On the Use of LOS to Define Feasible Heterogeneous Prospective Payments

Brigitte Dormont[*] and Carine Milcent[†]

September 12, 2005

Abstract

In many countries hospital payments are defined on the basis of per DRG prospective payment systems. The financing administrations generally set lump-sum payments per stay by computing the average costs per stay in each DRG. For that purpose, they generally use a sub sample of hospitals which are able to provide detailed information about their costs per stay. In actual practice, most countries use a hospital subsample whose representativeness is questionable. Indeed, participation in the cost database program is in general voluntary, and the participating hospitals must have accounting systems that enable them to provide detailed information about their costs. On the other hand, financing administrations generally have at their disposal extensive or even exhaustive information, at the stay-patient level, about DRG, length of stay, diagnoses and procedures implemented.

From a practical point of view, one drawback of lump-sum payments per stay is the lack

[*]University of Paris-Dauphine, France and Institute of Health Economics and Management, Lausanne, Switzerland. E-mail: brigitte.dormont@dauphine.fr

[†]PSE (CNRS-EHESS-ENPC-ENS), Paris, France and Institute of Health Economics and Management, Lausanne, Switzerland.E-mail: milcent@pse.ens.fr

of representativeness of the data used to compute the payments. From a more theoretical point of view, such payments do not take heterogeneity of patients and hospital for the same pathology into account.

Many studies have pointed to possible negative effects of careless implementation of a prospective payment system. In a preceding paper, we have proposed a payment system that creates incentives to increase hospital efficiency when hospitals are heterogeneous, without reducing quality of care (Dormont & Milcent, 2005). However, the implementation of our method of payment requires information about costs per stay for all the regulated hospitals. In this paper, we investigate the possibility of implementing such a method of payment by combining two sources of information: (i) a hospital sub-sample about costs per stay, (ii) an exhaustive database about length of stay, diagnoses and procedures implemented.

We use two database relative to stays for vaginal or C-sections deliveries in Swiss hospitals. The first database is a three dimensional nested database relative to 12,123 stays in 7 hospitals recorded over the years 1999-2001. This database provides information about costs par stay as well as about length of stay, diagnoses and procedures implemented. The second database concerns all the Swiss hospitals but does not provide any information about costs. It is a nested database relative to 29,495 stays in 80 hospitals recorded during the year 2002.

Our results show that the use of a hospital sub-sample leads to an overestimation of the average cost of C-sections deliveries (+ 12.0 % to + 32.4 %). On the other hand, the lump-sum for vaginal deliveries is underestimated by the use of the subsample (- 8.3 % to - 4.0 %). The implementation of our payment system, which gives hospitals incentives for efficiency but allows for heterogeneity and thus limits drawbacks such as patients selection

and lower quality, should lead to budget savings from 3 % to 13 %, depending on the DRG considered.

JEL classification: C23, H51, I18

Keywords : Hospital regulation, Prospective Payment System, Costs, LOS

1 Introduction

In many countries hospital payments are defined on the basis of per DRG prospective payment systems. The financing administrations generally set lump-sum payments per stay by computing the average costs per stay in each DRG. For that purpose, they generally use a sub sample of hospitals which are able to provide detailed information about their costs per stay.

The relevance of such payments is questionable for two main reasons. Firstly, using information on cost from a hospital subsample raises the question of the representativeness of the data used to compute the payments. Secondly, lump-sum payments per stay do not take heterogeneity of patients and hospitals into account. In this paper, we address these difficulties by combining two sources of information: (i) a hospital sub-sample about costs per stay, (ii) an exhaustive database about length of stay, diagnoses and procedures implemented.

We have at our disposal two database relative to stays for vaginal or C-sections deliveries in Swiss hospitals. The first database is a three dimensional nested database relative to 12,123 stays in 7 hospitals recorded over the years 1999-2001. This database provides information about costs per stay as well as about length of stay, diagnoses and procedures implemented. The second database concerns all the Swiss hospitals but does not provide any information about costs. It is a nested database relative to 29,495 stays in 80 hospitals recorded during the year 2002.

From a practical point of view, one drawback of lump-sum payments is the lack of representativeness of the data used to compute the payments. Combining our two sources of information allow us to improve the empirical approach.

From a more theoretical point of view, many studies have pointed to possible negative effects of careless implementation of a fully prospective payment system, namely patient selection, lower care quality and DRG "creep", i.e. the tendency to code patients in more costly DRG (Newhouse, 1996). In a preceding paper, we have proposed a payment system that creates incentives to increase hospital efficiency when hospitals are heterogeneous, without reducing quality of care (Dormont & Milcent, 2005). However, the implementation of our method of payment requires information about costs per stay for all the regulated hospitals. In this paper, we define heterogeneous and feasible payments by combining our two sources of information.

This article is organized as follows. In section 2 we address the issue of data representativeness and show the distortions arising from the use of a sub-sample. The issue of heterogeneity in hospital payment systems is discussed briefly in section 3. Our empirical specifications are presented in section 4, where we show how to use the exhaustive information about LOS to define feasible heterogenous payments. Our results are presented in section 5: we study the characteristics of the estimates obtained for hospital heterogeneity and for hospital inefficiency. Simulations of our payment method allow us to evaluate the budget savings which may result from its implementation. Section 6 concludes.

2 Average costs and data representativeness

We first address a very simple and practical question: what are the consequences of the use of a hospital subsample to set lump-sum payments? Then, we define a simple extension, which makes it possible to use a more extensive information about length of stays.

The practical question we address is in connection with the widespread situation of most countries. In actual practice, information about costs per DRG is recorded in a hospital subsample whose representativeness is questionable. Indeed, participation in the cost database program is in general voluntary, and the participating hospitals must have accounting systems that enable them to provide detailed information about their costs. On the other hand, financing administrations generally have at their disposal extensive or even exhaustive information, at the stay-patient level, about DRG, length of stay, diagnoses and procedures implemented.

In what follows, we show that our Swiss data are typical of this situation: a small subsample about costs and a much more extensive information about LOS. Information about hospitals stays presents the same pattern in most European countries.

2.1 The available information: our data

For our empirical work, we consider 4 DRGs relative to deliveries: DRG 370 and 371, i.e. C-section deliveries with low risk, with or without complications; DRG 372 and 373: vaginal deliveries with or without complications. We have at our disposal two databases relative to stays in Swiss hospitals. We have deleted the outliers from each database using thresholds computed by the financing administration.

The first database, named Swiss-DRG, is a three dimensional nested database relative to 12,123 stays in 7 hospitals recorded over the years 1999-2001. This database provides information about costs per stay as well as about length of stay, diagnoses and procedures implemented.

The second database, named OFS[1], is built from an exhaustive administrative source. It concerns all the Swiss hospitals but does not provide any information about costs. After cleaning the original information, we selected hospitals where stays for deliveries are observed. The resulting sample is a nested database relative to 29,495 stays in 80

[1] OFS stands for "Office Fédéral des Statistiques".

hospitals recorded year 2002. To implement our estimates, we had to select hospitals with at least 20 stays, resulting in a sample of 29 128 stays observed in 77 hospitals. However, we will refer to the exhaustive database for the OFS database in the following.

Tables 1 to 3 display the basic features of our data. Notice that the number of hospitals for which costs are recorded is very small. Moreover, these hospitals are situated in only two counties ("cantons") of Switzerland (over 26 "cantons" and "demi-cantons"): Vaud and Tessin. The results of table 3 show that the representativeness of the Swiss-DRG database is questionable. This is not surprising, given the small number of observed hospitals, and the fact that participation to the cost database is voluntary. In addition to the lack of representativeness, one drawback of this restricted sample is that it prevents us to implement refined econometric analysis in order to identify the inefficiency component in the hospital fixed effects, which mesure hospital heterogeneity (see section 4.1).

2.2 Average costs: the distortions arising from the use of a sub-sample

Table 4 displays several computations of the average cost per stay for each DRG. The Swiss financing administration uses the restricted sample where costs per stay are recorded, i.e. the Swiss-DRG database. Average costs computed from this database are denoted M2. One drawback of such computations is that the corresponding means are likely to be not representative of the "true" average costs, which would be computed if costs for all Swiss hospitals were observable. In actual practice, the financing administration uses a correction consisting in multiplying the costs of non teaching hospitals (6 over the 7 observed) by a coefficient equal to 1.24. The corresponding means are denoted M124 in the table.

A straightforward improvement could be the following: (i) using the hospital subsample to evaluate the link between the cost of a stay in a given DRG and the length of stays and patient characteristics; (ii) taking the information about LOS and patient characteristics (available in the exhaustive database) to predict the costs per stay for all hospitals; (iii) compute the corresponding average costs.

More precisely, we consider the following cost function:

$$C_{iht} = \gamma + D_{iht}\alpha + X'_{iht}\beta + \eta_h + \zeta_{iht} \ , \tag{1}$$

where D_{iht} is the length of stay i in hospital h in year t. $X'_{i,h,t}$ represents individual patient characteristics such as age and comorbidities (gender has a limited interest for deliveries). D_{iht} and $X'_{i,h,t}$ are observable in the exhaustive OFS database. ζ_{iht} is the disturbance. Our purpose is to estimate specification (1) using the Swiss-DRG database in order to build an unbiased prediction of the cost of the stays recorded in the OFS database. To avoid the estimates of parameters of being influenced by heterogeneity in cost levels between the specific hospitals of the restricted sample, we have introduced hospital specific effects η_h, with the constraint $\sum_h \eta_h = 0$. (Therefore, the estimated constant reflect the average level of the hospital specific effects which would have been estimated in a specification without constant.) The predicted cost is then the following:

$$\widetilde{C}_{ih} = E(C_{ih}|\, D_{ih},\ X'_{ih}) = \widehat{\gamma} + D_{ih}\widehat{\alpha} + X'_{ih}\widehat{\beta} \tag{2}$$

We assume that $\widetilde{C}_{ih}$ is an unbiased prediction of C_{ih} i.e. that $E(\widetilde{C}_{ih} - C_{ih}) = 0$. Table 4 displays the means M1 corresponding to these predicted costs. They are necessarily more representative of the average costs of all hospitals since their computation uses the observed length of stays and patient characteristics. Only the coefficients of prediction (2) are estimated from the Swiss-DRG database.

Comparison of means M1 and M2 shows that the use of a hospital sub-sample leads to an overestimation of the average cost for C-section deliveries and to an underestimation for vaginal deliveries. The differences are quite sizeable: + 32.4 % for DRG 370 (C-sections with complications) and - 8.3 % for DRG 372 (vaginal deliveries with complications). The use of a multiplier by the administration reinforces the overestimation of the cost of C-sections: differences between M1 and M124 are equal to + 42 % and +26.4 % for DRGs 370 and 371. On the other hand, the multiplier compensates for the underestimation of the cost of vaginal deliveries. The effect is quite spectacular as concerns DRG 373. The relative difference between M2 (average cost on Swiss-DRG database) and M1 (our estimates) is equal to - 4 %. It turns to +12.5 % between for M124 (corrected average cost on Swiss-DRG database) and M1! Given that this DRG is by far the most frequent (about 42 % of stays for deliveries), this correction is likely to have a sizeable impact on hospital expenditures.

On the whole, comparing the average costs M1 and M2 shows that the sub-sample (the Swiss-DRG database) is far from being representative of the average costs of deliveries in Switzerland. In addition, the distortion is not homogenous among DRGs: the use of M2 (or M124) instead of M1 results in incentives for an increasing use of C-sections. Thus, combining the extensive information provided by the OFS sample with estimates carried out on the Swiss-DRG database should contribute to an improvement of the evaluation of average costs. In what follows, we show that the joint use of the two databases could also makes it possible to define payments which allow for hospital heterogeneity.

3 Hospital payment systems: the issue of hospital heterogeneity

Each prospective payment system based on a lump-sum payment gives the hospital a perfect incentive for efficiency. However, the level of the lump-sum payment is difficult to establish: should it be unique or should it depend on hospital characteristics ? Many studies have underscored the great diversity in the conditions of care delivery for hospitals (teaching status, share of low income patients, local wage level, etc.). For instance, Pope (1990) shows that input prices can differ according to location, and that a hospital can be characterized by specific quality of services or severity of illnesses of admitted patients. These studies highlight the risks of a fully prospective payment system: patient selection and lower care quality.

In order to avoid these drawbacks, many authors have proposed payments that allow for patient and hospital heterogeneity. (Keeler (1990), Pope (1990), Ma (1994, 1998), Ellis (1998), Laffont and Tirole (1993)). It is also possible to consider extensions which introduce endogenous levels of number and quality of treatments (Ma (1994), Ellis (1998), Chalkley and Malcomson (2000)). In a preceding paper, we have considered an extension of Shleifer's basic model (1985), where the regulator is supposed to use the information available about observable sources of hospital cost heterogeneity and proposed a payment system that creates incentives to increase hospital efficiency when hospitals are heterogeneous (Dormont & Milcent, 2005).

However, the feasibility of our method of payment is questionable. Indeed, its implementation requires information about costs per stay for each regulated hospital. In this paper, we investigate the possibility of combining our two sources of information to address this practical issue. We show how to use the exhaustive information about LOS provided by the OFS database to define feasible payments which allow for hospital heterogeneity. As stated above, the practical question we address is in connection with the widespread situation of most countries, where information about costs is only available for a restricted sample of hospitals.

4 Empirical specifications

4.1 Setting payments on data relative to costs

Consider the simplest model, where the regulator sets a lump-sum payment per stay in a given DRG. The corresponding econometric specification is very simple: the cost is explained by a constant, γ. Denoting by $C_{i,h,\,t}$ the cost of stay i in hospital h, in year t, one has:

$$C_{iht} = \gamma + \nu_{iht} \ , \tag{3}$$

where ν_{iht} is the disturbance. It can be splitted in several components: a hospital specific effect, η_h, a hospital-year specific effect, ε_{ht} , and a random error term at the patient level $u_{i,h,t}$: $\nu_{iht} = \eta_h + \underbrace{\varepsilon_{ht} + u_{iht}}_{\zeta_{iht}}$. The random error term $u_{i,h,t}$ is assumed to be iid $(0, {\sigma_u}^2)$. It takes unobservable patient heterogeneity into account. $\varepsilon_{h,t}$ is a disturbance that is assumed to be iid $(0, \sigma_\varepsilon^2)$ and uncorrelated with $u_{i,h,t}$.

Using expression (3) to build hospital payments comes down to assuming that all the cost variability for a given DRG is due to inefficiency. As stated above, a careful regulator might prefer avoiding the possible negative effects of a fully prospective payment system by taking heterogeneity of patients and hospitals into account. In this case, the cost function corresponds to specification (1):

$$C_{iht} = \gamma + D_{iht}\alpha + X'_{iht}\beta + \eta_h + \zeta_{iht} \ , \text{ where } \zeta_{iht} = \varepsilon_{ht} + u_{iht}.$$

Considering an observable hospital characteristic such as its status (teaching, private not-for-profit or private-for-profit) is of no interest in our case. Indeed, previous econometric tests led us to the conclusion that hospital specific effects have to be specified as fixed instead of random (Dormont & Milcent, 2004). For identification purposes, all observable hospital characteristics which do not vary over time must therefore be removed from the specification.

Hospital specific effects η_h are related to time-constant observable and unobservable hospital heterogeneity. Estimating these effects allows us to evaluate, *ceteris paribus*, the difference in average cost between hospital h and the other hospitals. η_h can be seen as the result of three components: inefficiency, justifiable heterogeneity and care quality. Inefficiency refers to *long term* moral hazard: the hospital management can be permanently inefficient. Justifiable heterogeneity can be linked to the hospital's infrastructure and the existence of economies of scale or of scope. In principle, the goal of the regulator is to reduce inefficiency only.

The disturbance $\varepsilon_{h,t}$ is defined as the deviation, *ceteris paribus*, for a given year, of hospital h's cost in relation to its average cost. As shown in Dormont & Milcent (2005), it is entirely attributable to inefficiencies: it is an indicator of the effect on costs of transitory moral hazard.

In accordance with the principles of a prospective payment system, the regulator announces the following payment rule: $P^{rule}_{iht} = E(C_{iht})$, where $E(C_{iht})$ is the inconditional expectation of the costs *observed at the end of the year*. Indeed, when the payment rule in announced, $E(C_{iht})$ does not correspond to the average of observed costs. At the beginning of the year, the regulator cannot observe the costs which would result from the hospitals' efforts towards more efficiency. Therefore, effective payments are set *ex post.* An accurate estimation of the *ex post* payments is given by the expectation of costs linked to an efficient activity.

When the regulator wants to take patient and hospital heterogeneity into account the payment rule corresponds to a cost expectation which is defined conditionally on patient and hospital heterogeneity: $P_{iht}^{rule} = E(C_{iht}|\, X'_{iht}\ ,\ \eta_h)$. Using specification (1), the payment rule is given by:

$$P_{iht}^{rule} = \widehat{\gamma} + E(D_{iht})\widehat{\alpha} + X'_{iht}\widehat{\beta} + \widehat{\eta}_h \tag{4}$$

Notice that the proposed payment includes $\widehat{\eta}_h$: it takes all unobservable hospital heterogeneity into account, despite it is composed of inefficiency, as well as legitimate heterogeneity. This feature of the payment is due to our data limitation: the number of observed hospitals is too small to allow for the use of a stochastic cost frontier approach in order to identify separately the inefficiency component of η_h. Therefore, it seems to us advisable to pay for η_h, which includes also legitimate heterogeneity and care quality. As shown by (Dormont & Milcent, 2005), this payment rule can lead to substantial budget savings because it provides incentives to reduce costs linked to transitory moral hazard $\varepsilon_{h,t}$.

A prospective payment rule should include only variables which cannot be manipulated by the hospital's manager. However, the length of stay is an important explanatory variable of the cost. To set a prospective payment rule, we thus consider an unconditional expectation $E(D_{iht})$ as in expression (4). We can also consider a conditional expectation, such as $E(D_{iht}|\, X'_{iht}\ ,\ \delta_h)$, which can be evaluated from the estimation of a specification explaining the LOS, where δ_h is a hospital specific effect.

To evaluate the *ex post* payments corresponding to rule (4), one has to estimate the costs linked to an efficient activity.

4.2 Using exhaustive information about LOS to define feasible heterogenous payments

One main drawback of payment rule (4) is the lack of feasibility: its definition relies on the estimation of hospital specific effects η_h. Therefore, its implementation requires information about costs for each regulated hospital. In what follows, we propose a method to extend our heterogeneous payments to all the hospitals observed in the OFS database. As for our computations of average costs in section 2, the idea is to combine the information on costs available in the restricted hospital sub-sample with the extensive OFS database, which displays information about LOS and patient characteristics.

This new payment rule uses the hospital specific effect δ_h of a specification explaining the length of stay. One additional advantage of this payment rule is that it is now possible to identify the component of δ_h , thanks to the large number of hospitals observed in the OFS database.

4.2.1 Definition of feasible heterogenous payments

Consider D_{ih} the length of stay i in hospital[2] h.

$$D_{ih} = X'_{ih}b + \delta_h + \mu_{ih} \tag{5}$$

X'_{ih} are patient characteristics and μ_{ih} is the disturbance. δ_h is a hospital specific effect.

[2] Only one year is available for the OFS database. Our data have two dimensions: stay and hospital.

One has:

$$\delta_h = \delta + \theta_h + \lambda_h \tag{6}$$

The decomposition of δ_h is not a direct transposition of the analysis developped as regards the hospital specific effects η_h of the cost function. We consider a constant δ to express the fact that the smallest LOS must not be considered as a medical optimum for deliveries. Laws were passed in the USA (Postpartum discharge laws, from 1995 to 1998) to limit very short hospital stays for deliveries, often called *drive-through deliveries.* Typical laws use as reference a minimal level of 48h for normal vaginal deliveries and 96h for C-sections (Chamorand, 1996, Liu *et al.*, 2004). Specifying a constant δ comes down to define a reference equal to the average LOS in Swiss hospitals (controlled for patient characteristics). Medical guidelines could also be used as reference.

θ_h represents legitimate heterogeneity regarding LOS: it can vary significantly between hospitals, depending on the share of low-income patients, patient preferences and average distance to hospital (especially in Swizterland).

Inefficiency is specified with a component $\lambda_h \geq 0$: we assume that inefficiency tends to lengthen stays. Indeed, many Swiss hospitals have been financed through a *per diem* payment, which gives strong incentives to lengthen the stay. Moreover, 16 % of the observed hospitals are private-for-profit: a longer LOS is a mean to increase the invoice for accomodation.

From the estimation of (1) on the Swiss-DRG database, one can define a predicted cost (2):

$$\widetilde{C}_{ih} = \widehat{\gamma} + D_{ih}\widehat{\alpha} + X'_{ih}\widehat{\beta}$$

for the stays of hospitals observed in the OFS database. We propose to use the following payment rule:

$$P_{ih}^{rule} = \widehat{\gamma} + E(D_{ih}| X'_{ih}, \theta_h) \widehat{\alpha} + X'_{ih}\widehat{\beta} \tag{7}$$

This payment rule comes down to announce that hospitals will be reimbursed for each stay at the level of the conditional expectation of the costs observed at the end of the year. Expression (7) is equivalent to: $P_{ih}^{rule} = E(C_{ih}| X'_{ih}, \theta_h)$. In other words, the expectation is conditional on patient characteristics and on the hospital heterogeneity which has been identified as legitimate.

The use of a large sample of hospitals makes it possible to carry out a stochastic cost frontier analysis to identify the component of δ_h which is linked to inefficiency. Therefore, the regulator can announce that he/she is able to evaluate permanent inefficiency λ_h. The payment rule, together with this information on the regulator's ability, should lead the hospitals to join the efficient LOS frontier:

$$D_{ih}^{eff} = X'_{ih}b + \delta_h - \lambda_h \tag{8}$$

The *ex post* payments are then defined by:

$$P_{ih}^{ex\ post} = \widehat{\gamma} + D_{ih}^{eff}\widehat{\alpha} + X'_{ih}\widehat{\beta} \tag{9}$$

4.2.2 Estimation

Our econometric estimates will allow us to evaluate the magnitude of the inefficiency component λ_h and to simulate the implementation of payment rule (7) on our data.

To identify the inefficiency component, we used a SCF approach (Greene, 2004). This approach relies on a parametric specification: the disturbance is split into two components, a normal one, related to statistical noises and a half normal component, related to inefficiency. In our case, we assume:

$$\delta_h = \delta + \theta_h + \lambda_h \text{ with } \theta_h \sim N(0, \sigma_\theta^2), \text{ and } \lambda_h = |\ell_h|, \ \ell_h \sim N(0, \sigma_\ell^2) \tag{10}$$

In a first step, we estimate (5), where the δ_h are specified as fixed effects and where μ_{ih} is supposed to be iid (0, σ_μ^2). Given that the number of stays observed per hospital is large enough, they can be consistently estimated by OLS[3]. In the second step, we use first-step estimates $\widehat{\delta}_h$ and consider the SCF specification, assuming (10) to estimate (6) by the maximum likelihood estimator. This allows us to identify the component λ_h. From expression (8), it is then very easy to estimate the efficient level of length of stay:

$$\widehat{D}_{ih}^{eff} = X_{ih}'\widehat{b} + \widehat{\delta}_h - \widehat{\lambda}_h \tag{11}$$

From this result we can deduce an estimate of the *ex post* payments by replacing D_{ih}^{eff} by $\widehat{D}_{ih}^{eff}$ in (9).

5 Results

Graph 1 to 4 display a representation of hospital heterogeneity in LOS behavior and of the magnitude of inefficiency for each DRG. More exactly, we give the values of the hospital

[3] We also considered an alternative specification, allowing to inefficiency at the patient hospital level. In this case, the first step is also a maximum likelihood estimation of a SCF specification. More exactly, we assume $\mu_{ih} = \phi_{ih} + \kappa_{ih}$ with $\kappa_{ih} \sim N(0, \sigma_\kappa^2)$
and $\phi_{ih} = |\varphi_{ih}|, \varphi_{ih} \sim N(0, \sigma_\varphi^2)$. All estimates led to an unsignificant inefficiency ($\sigma_\varphi^2 = 0$). This result in maintained when considering a truncated normal or an exponential distribution instead of a half-normal one. It is in accordance with the idea that inefficiency is at the level of hospital and not at the level of stays.

effects $\widehat{\delta}_h$ and of $\widehat{\delta}_h - \widehat{\lambda}_h$, i.e. of the hospitals effect minus the inefficiency component. The observations are at the hospital level and have been sorted by increasing $\widehat{\delta}_h$.

The descriptive statistics given in table 1 have shown that the average LOS in the OFS database is equal to 8.5 days (respectively, 7.9 days) for C-section deliveries with complications (respectively, without complications) and to 5.8 days (respectively, 5.4 days) for vaginal deliveries with complications (respectively, without complications). In each case, the highest duration is observed for the DRG with complications. The estimates of fonction (5) show that the LOS levels and the difference in average LOS observed between C-section and vaginal deliveries are mainly explained by the coefficients of the patient's age: they are equal to 0.20 and 0.19 for C-sections versus 0.12 and 0.13 for vaginal deliveries. Given a very homogenous average age of patients equal to about 30 years, the pure effect of age is then equal to 6.1 and 5.9 for C-sections and to 3.5 and 3.9 for vaginal deliveries.

Given these values, we can see that the hospital heterogenity is very large for DRG 370 (C-sections with complications): from one day to 4.5 days. In addition, this heterogeneity is almost entirely due to inefficiency: the curve relative to $\widehat{\delta}_h - \widehat{\lambda}_h$ is flat and slightly above one day. The variability is comparable for DRG 371 (C-sections without complications): from 0.5 day to 4 days. But inefficiency has a small impact on these disparities: the two curves are rather parallels and there is still a sizeable variability for $\widehat{\delta}_h - \widehat{\lambda}_h$.

Similar patterns are observed in graphs 3 and 4, relative to vaginal deliveries. The disparities seem to be smaller. But not it is not the case, if we consider that average LOS are smaller for vaginal deliveries. Inefficiency λ_h explains much of the variability for the DRG 372, with complications. This is not the case for vaginal deliveries without complications. The two curves are parallels for DRG 373: the magnitude of inefficiency is independent of the size of the hospital specific effect and most of the variability is due to what we have estimated as being "legitimate heterogeneity".

Finally, graphs 5 to 8 illustrate our payments' variability, in comparison with lump-sum payments which could be applied on the basis of the average costs computed in section 2.2.

The estimates allow us to simulate the implementation of our payment rule on the stays for deliveries in Swiss hospitals. Table 5 gives for each DRG the potential budget savings associated with payment rule (7). As explained above, the *ex post* payments, denoted P in the table, are obtained by combining (8) and (9). To evaluate the budget savings, we simply consider that relevant estimates of current average costs are given by M1, i.e. our estimates resulting from the combination of the Swiss-DRG database and the exhaustive database (see section 2.2).

Let us recall that these payments have the great advantage of taking unobservable hospital heterogeneity into account as regards LOS, which permits to reimburse high-quality care. In addition, our payment rule takes patient characteristics into account, thus limiting the incentives for patient selection. Moreover, the use of the exhaustive database makes it possible to observe a large number of hospitals and to identify inefficiency by a SCF estimation: then, it is possible to set a payment rule which allows for heterogeneity but eliminate inefficiency.

Our approach thus permit us to overcome partly the difficulties linked to the lack of an extensive information about cost. Notice, however, that we cannot take into account the part of the unobservable hospital cost-heterogeneity which would be uncorrelated with the behavior related to LOS (more exactly, with δ_h).

Table 5 shows that the implementation of our payment system, which gives hospitals incentives for efficiency but allows for heterogeneity and thus limits drawbacks such as patients selection and lower quality, should lead to budget savings from 3 % to 13 %, depending on the DRG considered. The estimated potential savings are smaller for DRGs without complications, such as DRG 371 and 373.

6 Conclusion

Data limitations often lead to the use of a non representative hospital subsample to set lump-sum payments. This practical question is faced by most countries. Indeed, information about costs per DRG is in general recorded in a hospital subsample, whereas financing administrations have at their disposal an exhaustive information about DRG, length of stay, diagnoses and procedures implemented at the stay-patient level.

We show that it is possible to improve the payment definition by combining the subsample relative to costs with an information about LOS and patient characteristics collected for all the hospitals. The empirical application is carried out on samples relatives to stays for deliveries in Swiss hospitals. Our results show that the classical approach, i.e. the use of a hospital sub-sample, leads to an overestimation of the average cost for C-section deliveries and to an underestimation for vaginal deliveries. In addition to a lack of representativeness, this creates distortions which can induce incentives for an increasing use of C-sections. The method we propose is very simple and should improve the relevance of the computed average costs per DRG.

Another important issue is the taking into account of hospital heterogeneity. From a theoretical point of view, it is possible to define payments which allow for hospital and patient heterogeneity. But the feasibility of such payments is questionable, since their implementation relies on the observability of costs in each regulated hospital. Combining our two database, we define a payment system which is (i) feasible, (ii) allows for hospital and patient heterogeneity, (iii) provides incentives to reduce costs. Simulations provide an evaluation of the budget savings associated with the implementation of such payments: they vary from 3 % to 13 %, depending on the DRG considered. This payment system could be implemented at the present time in most countries, since it is feasible with the information that is currently avalailable.

Our empirical approach produces another useful result: estimates of efficiency indicators at the hospital level. Many studies propose a hospital ranking based on a DEA analysis or on SCF estimates of the hopital's cost function. These approaches are in general global and consider all stays in all DRGs. However, the DEA method is non parametric: its results do not allow a direct identification of the sources of inefficiency. The results of the SCF approach are also difficult to understand when it is applied to a cost function specified for stays in all DRGs. Indeed, the hospital production function is multiproduct with a very high number of products: the stays in each DRG. Differences in efficiency are likely to be explained by potential economies of scope that are difficult to identify correctly, because the number of product combinations to consider is huge. Our estimates provide an evaluation of efficiency for each DRG. This result is quite different and might help understanding more aggregate results.

7 References

Chalkley, M. and J. M Malcomson, 2000. Government purchasing of health services, in: Culyer A.J. and Newhouse J.P. editors, Handbook of Health Economics, Vol. 1A (North Holland, Amsterdam), Chapter 15, 847-890.

Chamorand R. 1996. Le système de santé américain. Notes et Etudes documentaires, *La Documentation Française,* n°5300

Dormont, B. and C. Milcent, 2004. The sources of hospital cost variability. Health Economics, Vol. 13: 927-939

Dormont, B. and C. Milcent, 2005. How to Regulate Heterogeneous Hospitals ? Journal of Economics and Management Strategy, vol 14, n°3: pp 591–621

Ellis, R. P., 1998. Creaming, Dumping, skimping: Provider competition on the intensive and extensive margins, Journal of Health Economics, vol 17: pp 537-555

Greene, W. 2004. Distinguishing between Heterogeneity and Inefficiency: Stochastic Frontier Analysis of the WHO's Panel Data on National Health Care Systems, Health Economics, Vol 13, Issue 10: pp 959-980

Keeler E. B., 1990. What proportion of hospital cost differences is justifiable ?, Journal of Health Economics 9(3), 359-365.

Laffont J. J. and J. Tirole, 1993. A theory of incentives in procurement and regulation, MIT Press

Liu, Z., Dow W. H. and Norton E. C. (2004) Effect of Drive-Through Delivery Laws on Postpartum Length of Stay and Hospital Charges, Journal of Health Economics 23, 129-155.

Ma, A. C. T., 1994. Health care payment systems: cost and quality incentives, Journal of Economics and Management Strategy, vol 3, n°1: pp 93-112

Ma, A. C. T., 1998. Health care payment systems: cost and quality incentives- Reply, Journal of Economics and Management Strategy, vol 7, n°1: pp 139-142

Newhouse J. P., 1996. Reimbursing Health Plans and Health Providers : Efficiency in Production Versus Selection , Journal of Economic Literature, Vol. XXXIV : 1236-1263.

Pope, G., 1990. Using hospital-specific costs to improve the fairness of prospective reimbursement, Journal of Health Economics, vol 9, n°3: pp 237-251

Shleifer, A., 1985. A theory of Yardstick Competition, Rand Journal of Economics, vol 16: pp 319-327

Table 1. The data

	Swiss-DRG database (7 hospitals ; 1999-2001) 12 123 stays			Swiss-DRG database (7 hospitals ; 2001) 3 437 stays		OFS database (80 hospitals ; 2002) 29 495 stays	
DRG	**# stays**	**Average Cost**[1]	**Average LOS**	**# stays**	**Average LOS**	**# stays**	**Average LOS**
370	344	9 910	8.67	118	8.75	2 168	8.51
371	2 009	7 575	7.55	632	7.38	8 848	7.91
372	2 362	4 060	5.51	785	5.31	6 095	5.82
373	7 408	3 823	5.25	1 902	5.13	12 384	5.42

1 : Swiss Francs (CHF)

Table 2. The OFS database

		Total	Of which : For profit	Of which : Teaching
Stays	Number	29 495	2 160	7 927
	%	100	7.3	26.9
Hospitals	#	80	13	5
	%	100	16.3	6.3

Table 3 Testing for differences in costs and LOS

		Swiss-DRG database		OFS database
		Cost	LOS	LOS
DRG 370	Teaching hosp.	+ 21.9	- 14.3	- 2.1*
	Non teaching hosp.	-	-	-
	For profit	-	-	+ 8.6
	Not for profit	-	-	-
DRG 371	Teaching hosp.	+ 12.1	- 20.8	- 5.0
	Non teaching hosp.	-	-	-
	For profit	-	-	+ 2.3
	Not for profit	-	-	-
DRG 372	Teaching hosp.	- 24.8	- 13.9	- 8.0
	Non teaching hosp.	-	-	-
	For profit	-	-	+ 18.9
	Not for profit	-	-	-
DRG 373	Teaching hosp.	- 26.9	-17.1	- 7.2
	Non teaching hosp.	-	-	-
	For profit	-	-	+16.3
	Not for profit	-	-	-

* : Not significant (5%)

Table 4. Average costs (CHF)

		DRG	370	371	372	373
OFS database *(Hospitals/ N_h=20)*	(1)	M1 Average cost computed using OFS database *(Using the estimate of E(C/D, X))*	7 483	6 764	4 428	3 983
Swiss-DRG database	(2)	M2 Average cost computed using Swiss-DRG database	9 910	7 575	4 060	3 823
		Relative difference with respect to M1	+ 32.4 %	+ 12.0 %	- 8.3 %	- 4.0 %
	(3)	M124 Average cost computed using Swiss-DRG database, with correction 1,24	10 627	8549	4407	4481
		Relative difference with respect to M2	+7.2 %	+ 12.8 %	+ 8.5%	+ 17.2 %
		Relative difference with respect to M1	+ 42.0 %	+ 26.4 %	-0.5 %	+ 12.5 %

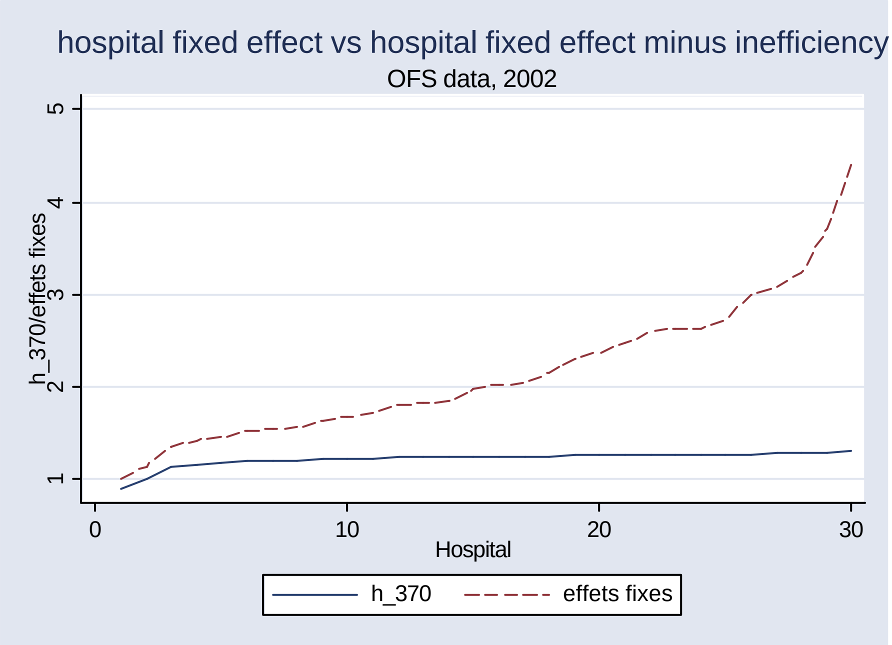
Swiss-DRG database *: 7 hospitals, 12 123 stays, years 1999-2001*
OFS database *: 80 hospitals, 29 495 stays, year 2002 , OFS* database */ N_h=20: 77 hospitals, 29 128 stays.*

Table 5. Payments and potential budget savings (CHF)

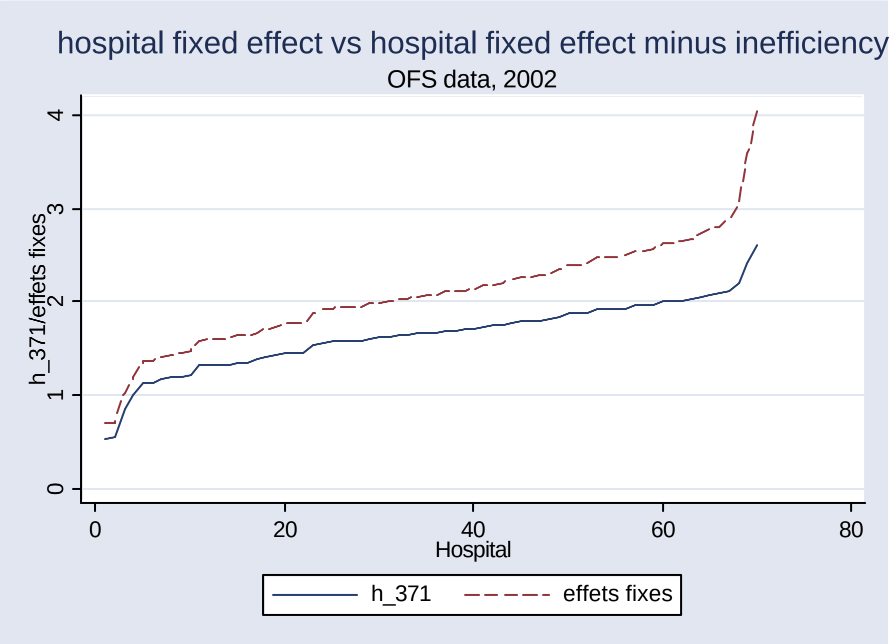
	DRG	370	371	372	373
(1)	M1 Average cost computed using OFS database *(Using the estimate of E(C/D, X))*	7 483	6 764	4 428	3 983
(2)	Heterogenous Payments P Average	6471	6477	4134	3 852
	Potential budget savings (= Relative difference with respect to M1)	- 13.5%	- 4.2 %	- 6.6 %	- 3.3 %

Swiss-DRG database *: 7 hospitals, 12 123 stays, years 1999-2001*
OFS database *: 80 hospitals, 29 495 stays, year 2002 , OFS* database */ N_h=20: 77 hospitals, 29 128 stays.*

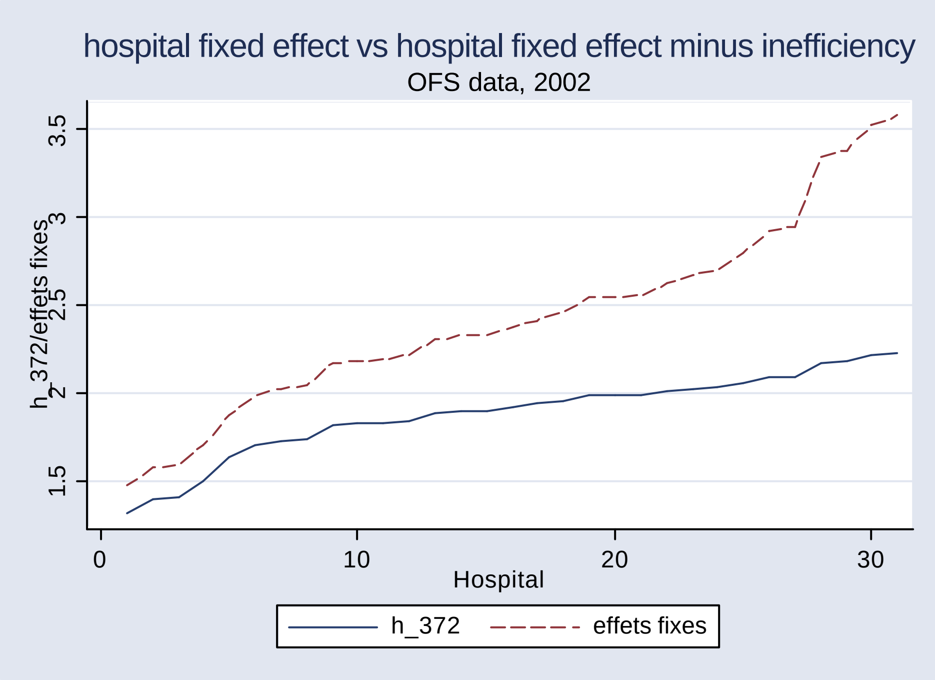
Graphique 1 : DRG 370



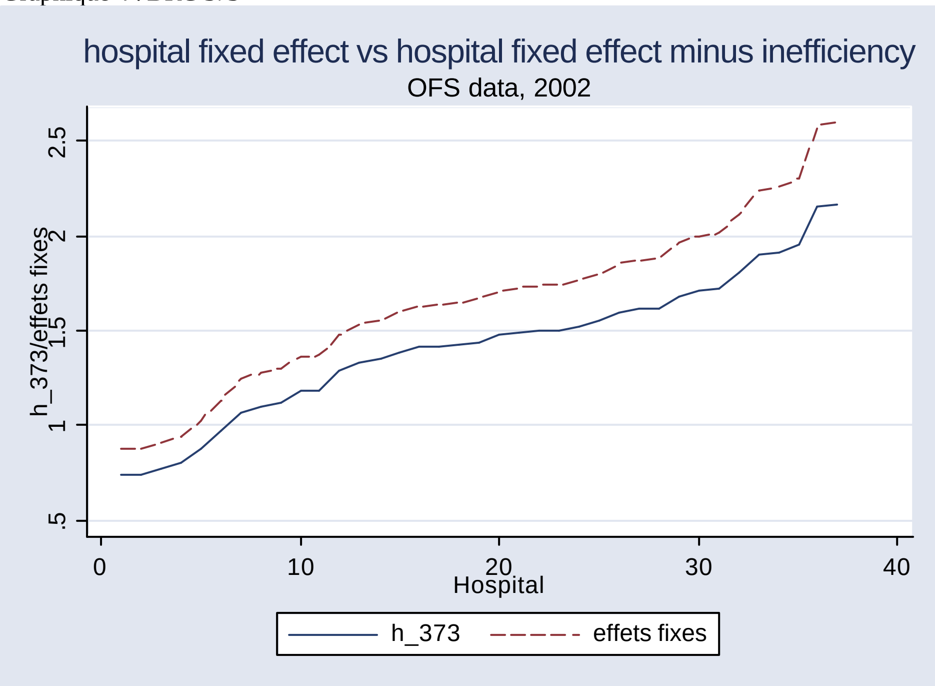
Graphique 2 : DRG 371



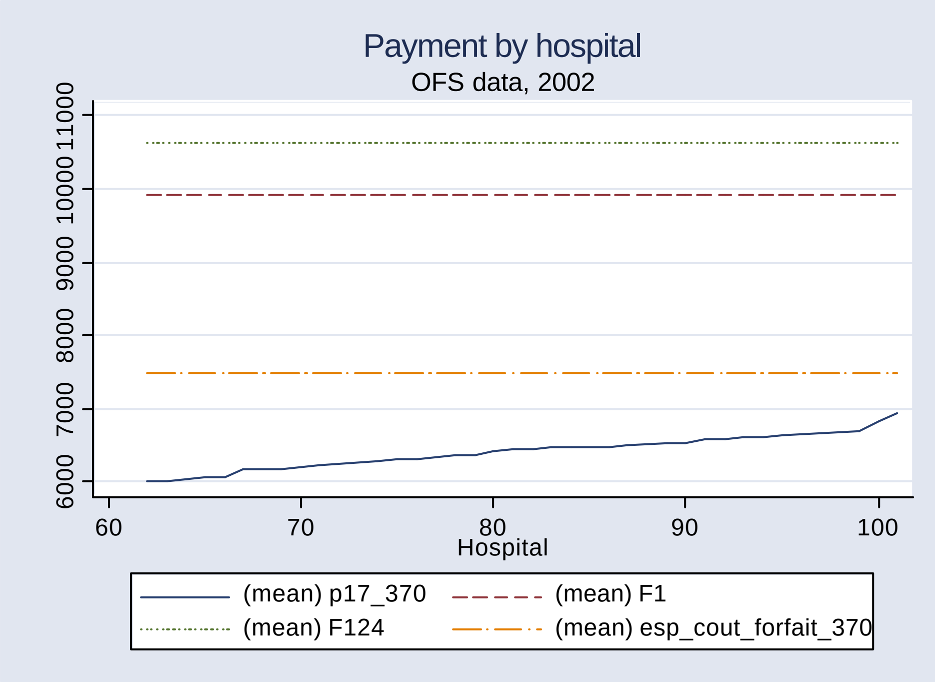
Graphique 3 : DRG 372



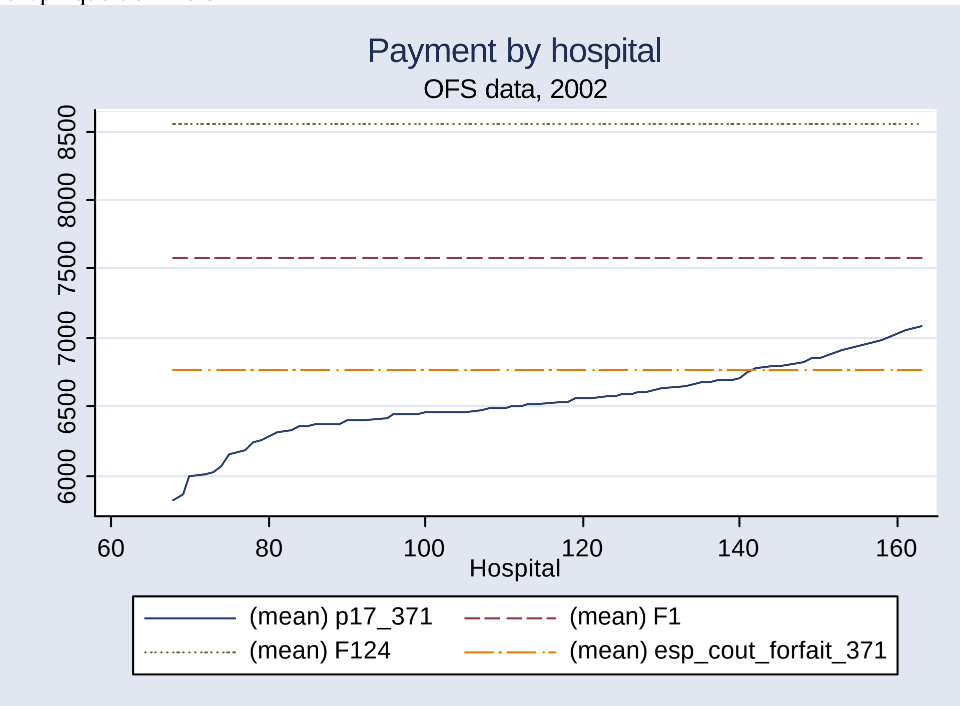
Graphique 4 : DRG 373



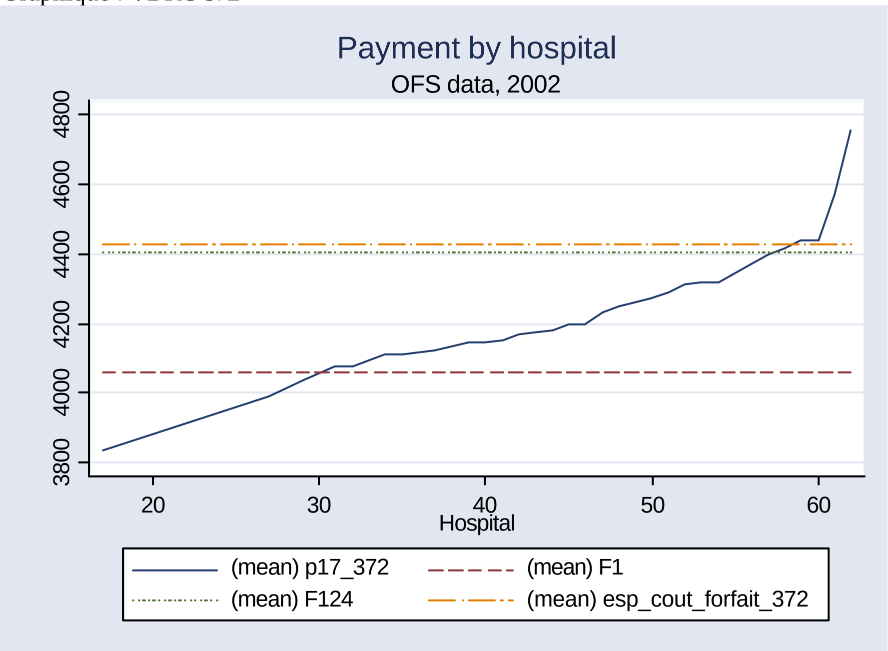
Graphique 5 : DRG 370



Graphique 6 : DRG 371



Graphique 7 : DRG 372



Graphique 8 : DRG 373