Working Paper No. 52

Accounting for Heterogeneity in the Measurement of Hospital Performance

Philippe K. Widmer, Peter Zweifel and Mehdi Farsi

December 2011

Accounting for Heterogeneity in the Measurement of Hospital Performance

Philippe K. Widmer[a,*], Peter Zweifel[a], Mehdi Farsi[b]

[a]*Department of Economics, University of Zurich, Switzerland*
[b]*Department of Management, Technology and Economics, ETH Zürich, Switzerland*

Abstract

With prospective payment of hospitals becoming more common, measuring their performance is gaining in importance. However, the standard cost frontier model yields biased efficiency scores because it ignores technological heterogeneity between hospitals. In this paper, efficiency scores are derived from a random intercept and an extended random parameter frontier model, designed to overcome the problem of unobserved heterogeneity in stochastic frontier analysis. Using a sample of 100 Swiss hospitals covering the years 2004 to 2007 and applying Bayesian inference, significant heterogeneity is found, suggesting rejection of the standard cost frontier model. Estimated inefficiency decreases even below the 14 percent reported by Hollingsworth (2008) for European countries. Accounting for unobserved heterogeneity would make hospitals rated below 85 percent efficiency according to the standard model gain up to 12 percentage points, serving to highlight the importance of heterogeneity correction in the estimation of hospital performance.

Keywords: Hospital efficiency, unobserved heterogeneity, Bayesian inference, Switzerland, Stochastic Frontier Analysis

JEL: C11, C33, D24, I18

---

* Corresponding author: Philippe Widmer, Department of Economics, University of Zurich, Hottingerstrasse 10, CH-8032 Zurich, Switzerland. Phone: +41 (0)44 634 22 70, Fax: +41 44 634 49 87, E-mail: philippe.widmer@soi.uzh.ch

1. Introduction

Performance-based prospective hospital payment has recently been introduced in several countries. It has greatly increased the importance of accurately measuring firm-specific performance, defined here in terms of operating costs per casemix-adjusted patient case. The challenge for policy makers is to pay for efficiency; however, this calls for filtering out differences that are caused by inefficiencies rather than heterogeneities due to exogenous influences.

In response to this need, there has been a growing body of research into the determinants of cost variability between hospitals (see Hollingsworth, 2008, Jacobs et al., 2006, and Worthington, 2004 for overviews of the literature). Specifically, the meta-analysis by Hollingsworth (2008) finds evidence of inefficiency in the hospitals of the United States and several European countries, amounting to a potential for cost reduction of 18 percent and 14 percent, respectively. Compared to these estimates, Switzerland is on the high side with Steinmann and Zweifel (2003), based on a Data Envelopment Analysis (DEA), coming up with 30 percent. In a comparison with Germany (the land of Saxony), Steinmann et al. (2004) once more found Swiss hospitals to be relatively inefficient. Using a Stochastic Frontier Analysis (SFA), Farsi and Filippini (2006) put the potential of cost reduction to 20 percent on average, which would translate into 7 percent of Switzerland's total health care expenditure (55 bn CHF) in 2007. However, these estimates do not account for heterogeneity in production technology, which is likely to be particularly marked due to Swiss federalism.

Efficiency scores from the articles cited above are simply defined as the ratio of observed cost to a value on the estimated single technology cost frontier (Farrell, 1957).[1] In the case of SFA introduced by Meeusen and van den Broeck (1977) and Aigner et al. (1977), this ratio is given by

$$\frac{C_{it}}{C_S(X_{it}; \alpha, \beta)} = e^{u_i + v_{it}}, \tag{1}$$

with C_{it} is the (arithmetic) cost of hospital $i = 1, ..., N$ at time $t = 1, ..., T$, $C_S(X_{it}; \alpha, \beta)$ is an estimated minimum cost for outputs and input prices comprised in a $NT \times (K + 1)$ matrix X, α is the unknown intercept, β is a $K \times 1$ vector of unknown slope parameters of the cost function, u_i is a random term with positive values only reflecting inefficiency, and v_{it} a conventional random error (see Section 2 below for details). Any difference in technology is captured in the composite error term $u_i + v_{it}$, which could bias the estimates of the inefficiency scores u_i (see also Greene, 2004a). This is particularly the case for Switzerland, where hospitals have to operate in different regulatory environments, causing

---

[1] Simple cost ratios, often used for policy purposes, are not sufficient because they neglect both economies of scale and heterogeneity of technology.

them to provide health care services using different technologies. Controlling for heterogeneity may therefore lead to inefficiency estimates that are more in line with those of the United States and the European countries.

There have been several approaches for dealing with this problem. The first is to introduce fixed effects in the SFM or in the distribution of u_i (see e.g. Worthington, 2004). Since the choice of the dummy variables must rely on observable characteristics of the hospital, this solution is limited to 'observable' heterogeneity, leaving potential for 'unobservable' heterogeneity to bias estimated inefficiency scores. Provided panel data available, a true random effects model can be estimated (see e.g. Farsi and Filippini, 2008, Farsi et al., 2006, Greene, 2005b, and Greene, 2004b). This is a special case of the Random Intercept Frontier Model (RIFM) to be presented below, which enables cost frontiers to vary between hospitals. Still, the RIFM is not without limitations because it only allows the intercept α to be stochastic and assumes heterogeneity to be homoscedastic. Additional flexibility is afforded by the Random Parameter Frontier Model (RPFM), a generalization of the SFM and the RIFM, which additionally allows the slopes β vary between hospitals. Implementation of RPFM until recently was hampered by the requirement of large computational power and panel data of sufficient quality. Improvement on both counts have rendered them feasible in the meantime (see e.g. Widmer, 2010, Huang, 2004, Orea and Kumbhakar, 2004, and Tsionas, 2002).

In this paper, we analyze the influence of unobserved heterogeneity between Swiss hospitals using a SFM, a RIFM, and an extended RPFM for SFA. Section 2 contains additional details for these models as well as for the Bayesian approach adopted in the model specification and parameter estimation. The database of about 100 Swiss hospitals covering the years 2004 to 2007 is presented in Section 3. Estimation results confirm the existence of unobserved heterogeneity in Swiss hospitals, suggesting rejection of the SFM. On average, the SFM overestimates Swiss hospital inefficiency by about 6 percent. Section 4 considers the implications of this study for hospital managers and policy makers, and concludes.

2. Modeling Unobservable Heterogeneity

It is common practice to define heterogeneity as time-invariant cost variation that is exogenous in the sense that it cannot be manipulated by management at least in the short run. This definition is adopted here.[2] As a benchmark, the specification of the cost function in the presence of observable heterogeneity is presented first (Section 2.1); this provides the point of departure for the modeling of

---

[2]   On the longer run, the choice of technology can be influenced by hospital management. The presence of inferior technology becomes a component of management inefficiency in this case.

unobservable heterogeneity in Section 2.2, which likely is the more important component given that the quality of (largely non-profit) hospital management is not easily measured.

2.1. The Standard Frontier Model with Observable Heterogeneity

When heterogeneity is perfectly observable, the SFM can be augmented in a way permitting to estimate consistent efficiency scores (Pitt and Lee, 1981). Let there be repeated observations $t = 1, ..., T$ for all hospitals $i = 1, ..., N$, and let heterogeneity Z_i be completely reflected by a $N \times L$ matrix of observable, time-invariant characteristics. Using a Cobb-Douglas cost function in logs for simplicity, the cost frontier can be specified as

$$C_{it} = \alpha + \gamma' Z_i + \beta' X_{it} + \gamma' X_{it} Z_i + u_i + v_{it}, \quad i = 1, ..., N, \quad t = 1, ..., T, \tag{2}$$

$$u_i \overset{iid}{\sim} f_N^+[0, \sigma_u^2] \text{ with } Cov(u_i; \alpha, X_{it}, Z_i, v_{it}) = 0, \tag{3}$$

$$v_{it} \overset{iid}{\sim} f_N[0, \sigma_v^2] \text{ with } Cov(v_i; \alpha, X_{it}, Z_i, u_i) = 0, \tag{4}$$

with parameters defined in the Introduction section. Note that observable heterogeneity enters the cost function in two ways. First, the intercept α is augmented by a hospital-specific term $\gamma' Z_i$. This is called separable heterogeneity because it captures cost variability that is unrelated to the hospital's technology, such as size of its catchment area.[3] Second, the slopes β pertaining to outputs and factor prices are individualized by $\gamma' X_{it} Z_i$. This is non-separable heterogeneity reflecting differences in technology such as the amount and vintage of capital in use.

In this model, firm-specific inefficiency u_i is assumed to be time-invariant, uncorrelated with the deterministic part of the cost function as well as with random noise v_{it}, and to follow a one-sided distribution supported on the interval $[0, \infty)$ such as the half normal, truncated normal, exponential, gamma or Weibull distribution. Alternatively one could model inefficiency as a time-invariant fixed effect (Schmidt and Sickles, 1984), which allows for correlation with production technology.[4] But because heterogeneity is assumed to be time-invariant here, this variant will not be pursued.

Given eqs. (2) to (4) as an option, cost variability in logarithms is given by

$$C_{it} - C_H(X_{it}, Z_i; \alpha, \beta, \gamma) = u_i + v_{it}, \tag{5}$$

---

[3]   Greene (2005a) permits correlation between separable heterogeneity and production technology in his true fixed effects model.

[4]   The fixed effect model has at least two drawbacks: (1) it only measures relative inefficiency, (2) no time-invariant technology parameters are allowed in the cost function.

where $C_H(X_{it}, Z_i; \alpha, \beta, \gamma)$ is individualized for each hospital as shown in eq. (2). It enables the hospital regulator to separate cost variability related to differences in technology from inefficiency u_i and random noise v_{it}. The resulting inefficiency scores are no longer biased by heterogeneity and can be used for performance-based reimbursement.

However, this result does not hold as soon as heterogeneity is not completely observable but has an unobservable component Z_i^*, resulting in terms $\gamma' Z_i^*$ and $\gamma' X_{it} Z_i^*$ in the error term of eq. (2). If the regulator uses the conventional benchmark C_H of eq. (2), the residuals become

$$C_{it} - C_H(X_{it}, Z_i; \alpha, \beta, \gamma) = \gamma' Z_i^* + \gamma' X_{it} Z_i^* + u_i + v_{it}. \qquad (6)$$

The existence of (time-invariant) unobservable heterogeneity Z_i^* now causes bias in the measurement of hospital performance in two ways:

- Rather than estimating true inefficiency u_i, eq. (2) will estimate an artificially augmented inefficiency term $\widetilde{u_{it}} = u_i + \gamma' Z_i^* + \gamma' X_{it} Z_i^*$;

- To the extent that unobserved heterogeneity Z_i^* is correlated with either observed heterogeneity Z_i or outputs and factor prices X_{it} estimates of technology parameters α and β are biased as well since $Cov(\widetilde{u_{it}}, Z_i) \neq 0$, $Cov(\widetilde{u_{it}}, X_{it}) \neq 0$.

With both inefficiency scores and technology parameters biased, prospective payment runs the risk of rewarding some hospitals for being seemingly efficient while punishing others for being seemingly inefficient.

U.S. experience with prospective payment suggests that unobservable heterogeneity could be substantial. If performance-based reimbursement took into account all relevant determinants of hospital cost, one would expect hospitals to discard technologies giving rise to characteristics that are not paid for and to move to the cost-efficient level. However, Keeler (1990) found that U.S. hospitals still have a great deal of unexplained cost variability although prospective payment had been in place since 1983. One reason is that changes in hospital technology are particularly costly, causing hospitals to be slow in adopting new technologies. Evidently, more advanced estimation techniques are necessary to disentangle latent heterogeneity from inefficiency for prospective hospital payment to have the desired efficiency-enhancing effects.[5]

---

[5] As stated by Newhouse (1996), one option to overcome this problem is to use non fully prospective reimbursement systems, that are not fully prospective.

2.2. The Random Parameter Model with Unobservable Heterogeneity

The discussion of the preceding subsection led to the conclusion that in the case of hospitals, part of their technological heterogeneity is unobservable for years to come. Therefore, reimbursement arguably should take into account both observable and unobservable heterogeneity. One way to achieve this is the RPFM (the RIFM allows only the intercept rather than all parameters to be random and will not be expounded separately below). A RPFM estimates an individual cost function for each hospital, admitting of both observable and unobservable heterogeneity (see e.g. Widmer, 2010, Greene, 2004b, Huang, 2004, Orea and Kumbhakar, 2004, and Tsionas, 2002). To save on notation, the case of unobservable heterogeneity only is presented below. No ex-ante information on heterogeneity is needed, except for the assumption that it is time-invariant and normally distributed over individual hospitals. This is achieved by introducing a $[(K+1) \times 1]$ vector of time-invariant random variable $w_i \sim N[0, \sigma_w^2]$ that changes the intercept of the cost function to become $\alpha_i = \alpha + w_i$ (separable heterogeneity) and to the slope parameters to become $\beta_i = \beta + w_i$ (non-separable heterogeneity), resulting in

$$C_{it} = (\alpha + w_i) + (\beta + w_i)' X_{it} + u_{it} + v_{it}, \quad \text{or} \tag{7}$$

$$C_{it} = \alpha_i + \beta_i' X_{it} + u_{it} + v_{it}. \tag{8}$$

In the special case where w_i captures all existing heterogeneity, the RPFM can be transformed back into a SFM by substituting w_i by $\gamma' Z_i$ in eq. (7).

However, this specification is somewhat restrictive because it assumes both the intercept and the slopes of the cost function to be time-independent. This neglects the fact that new medical technology affects the whole hospital industry in very much the same way (such as the introduction of SCAT scanners). Denoting these changes by a vector of time dummies M_t, eq. (8) can be generalized to read,

$$C_{it} = \alpha_{it} + \beta_{it}' X_{it} + u_{it} + v_{it} \quad \text{with}$$

$$\alpha_{it} = \bar{\bar{\alpha}} + \delta' M_t + w_i \quad \text{and} \quad \bar{\alpha}_t = \bar{\bar{\alpha}} + \delta' M_t;$$

$$\beta_{it} = \bar{\bar{\beta}} + \delta' M_t + w_i \quad \text{and} \quad \bar{\beta}_t = \bar{\bar{\beta}} + \delta' M_t. \tag{9}$$

This specification allows to disentangle inefficiency from unobservable heterogeneity both variable and time-invariant. Separable heterogeneity is captured by the random intercept $\alpha_{it} = \bar{\bar{\alpha}} + \delta' M_t + w_i$, where $\bar{\bar{\alpha}}$ is the mean intercept over all hospitals. Non-separable heterogeneity in technology parameters is captured by a $(K \times 1)$ vector of hospital-specific parameters $\beta_{it} = \bar{\bar{\beta}} + \delta' M_t + w_i$.

To derive individual effects, assume w_i to follow a multivariate normal distribution

$$w_i \sim f_{MN}\left[\bar{w}, \Sigma\right], \quad \text{with } \Sigma \sim f_W \begin{bmatrix} \sigma^2_{w_\alpha} & \sigma_{w_\alpha, w_\beta} \\ \sigma_{w_\alpha, w_\beta} & \sigma^2_{w_\beta} \end{bmatrix}, \tag{10}$$

with $\bar{w}$ equal to zero. Σ is Wishart distributed with a $[(K+1) \times (K+1)]$ positive definite covariance matrix $S = (\sigma^2_{w_\alpha}, \sigma^2_{w_\beta}, \sigma_{w_\alpha, w_\beta})$, denoting unobserved heterogeneity between hospitals. For $\Sigma = 0$ no time-invariant heterogeneity exists and the random parameter model simplifies to a SFM with no variation in β_{it} and α_{it}.

For the Bayesian estimation to be performed in Section 3, the posterior distribution for the random parameter model needs to be derived. It is given by

$$p(\alpha, \bar{\bar{\alpha}}, \beta, \bar{\bar{\beta}}, \delta, u, \Sigma, \sigma_v^{-2}, \sigma_u^{-2}; C, X, M) \propto p(\bar{\bar{\alpha}}, \bar{\bar{\beta}}, \delta, \Sigma, \sigma_v^{-2}, \sigma_u^{-2})$$

$$\times \prod_{i=1}^{N} \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma_v^2}} exp\left[-\frac{1}{2\sigma_v^2}\left(C_{it} - [\alpha_{it} + \beta'_{it}X_{it} + u_{it}]\right)^2\right]$$

$$\times \prod_{i=1}^{N} (2\pi)^{-K/2} |\Sigma|^{-1/2} exp\left[-\frac{1}{2}(w_i - \bar{w})' \Sigma^{-1}(w_i - \bar{w})\right]$$

$$\times \prod_{i=1}^{N} \prod_{t=1}^{T} p(u_{it}, \sigma_u^{-2}), \tag{11}$$

where probability distributions $p(\bar{\bar{\alpha}}, \bar{\bar{\beta}}, \delta, \Sigma, \sigma_v^{-2}, \sigma_u^{-2})$ for the priors remain to be specified in Section 3.1. The likelihood function in eq. (11) is as in Griffin and Steel (2007), normally distributed with σ_v^2 as the variance of the random noise $v_{it} = C_{it} - [\alpha_{it} + \beta'_{it}X_{it} + u_{it}]$ and with α_{it} and β_{it} as in eq. (9). The last term of eq. (11) points to a gain in flexibility over classical maximum likelihood applications, where a joint density function of the random noise v and the inefficiency term u is specified. Here, only random noise enters the likelihood function, while inefficiency is estimated along with the other parameters of the cost function.

The random intercept α_{it} and technology parameters β_{it} are estimated at three levels. At the first level, the overall influences on hospital costs $(\bar{\bar{\alpha}}, \bar{\bar{\beta}})$ are determined, corresponding to the first factor following the proportionality sign of eq. (11); the second-level estimates of time-specific effects $(\bar{\alpha}_t, \bar{\beta}_t)$ defined in eq. (9) and the third-level estimates of individual values $(\alpha_{it}, \beta_{it})$ are derived from the multivariate normal distribution shown in eq. (11). Finally, inefficiency $p(u, \sigma_u^{-2})$ given by the last factor of eq. (11) is estimated at two levels. The first-level estimate corresponds to the population mean. The second-level estimates yield firm-specific inefficiency scores u_{it}.

Note that estimates of the unknown parameters can be derived by the marginal posteriors of eq. (11). However, it is not always possible to compute the posteriors analytically. Therefore, iterative Monte Carlo Markov Chain (MCMC) simulation is used, which involves iterative sampling from posterior parameter densities. Here, we use WINBUGS to derive the estimates (see Ntzoufras, 2009 for an introduction to Bayesian analysis with WINBUGS).

The model formulation of eq. (9) is quite general, containing many other variants of RPFM cited as special cases. For example, if M_t turn out to be zero for all covariates, it reduces to the one described by Greene (2004b) where all parameters are allowed to vary between hospitals but are constant over time. Furthermore, if additionally only technology parameters contain heterogeneity, it reduces to the one of Tsionas (2002), with a common intercept for all hospitals. If only the intercept controls for heterogeneity, it corresponds to the true random effects model of Greene (2005a) where all cost functions have the same slopes, adjusting only for separable heterogeneity. This is similar to the applied RIFM but with heterogeneity that is constant over time. Finally, if additionally w_i is zero for all parameters, the RPFM reduces to a SFM with no heterogeneity in the cost function.

3. Empirical Application to Swiss Hospitals Using Bayesian Inference

In this section, we analyze the effect of unobserved heterogeneity on the performance of Swiss hospitals with three SFA formulations, a Random Parameter Frontier Model (RPFM) outlined in Section 2.2, a Random Intercept Frontier Model (RIFM) as a special case of the RPFM, and a Standard Frontier Model (SFM).

3.1. Data and Econometric Specification of the Cost Function

The data used in this study are provided by the annual reports of the Federal Office of Public Health. They include 333 Swiss hospitals for the time period 2004 to 2007, comprising information on 5 university hospitals (K111), 23 central hospitals (K112), 27 large regional hospitals (K121), 46 medium regional hospitals (K122), 46 small regional hospitals (K123), 28 specialized surgery hospitals (K231), and sundry hospitals, viz. psychiatric and rehabilitation clinics. In total, 127 of these 333 units are private, non-subsidized hospitals.

In the interest of comparability, the sundry category was discarded. After purging the data from missing values and outliers, an unbalanced panel of 405 observations of sufficient quality is available. Variables are defined as follows:

VC : Variable operational expense per year, in thousands of CHF (VC);

X_1 : CMI-adjusted inpatient cases (PCASES);

X_2 : Revenue from outpatients (OUTP);

X_3 : Price of labor, average wage per employee (PL);

X_4 : No. of beds (BEDS);

S_1 : No. of internship categories (INTERN);

S_2 : No. of specialties (SPEC);

S_3 : Dummy=1 for subsidized public hospitals (SUB);

S_4 : Share of inpatients with supplementary insurance, in percent (INSUR);

M_t : Year dummies, t=2005, 2006, 2007 (base = 2004);

Z_l : Hospital group dummies, l=K111, K112, K121, K122, K231 (base = K123).

Table 1: Descriptive Statistics

Variable	Mean	Min	Max	K111	K112	K121	K122	K123	K231
$VC^{1)}$	109,296	3,925	953,586	790,609	190,380	88,122	44,830	19,538	36,387
$PCASES$	7,731	497	52,143	43,046	14,651	7,093	3,796	1,545	2,741
$OUTP^{1)}$	20,499	0	186,174	142,331	42,845	14,753	6,826	2,387	6,698
$PL^{1)}$	101	34	188	106	103	100	100	99	108
$BEDS$	201	12	1,170	893	383	207	108	55	69
$INTERN$	18	0	134	118	32	15	10	5	3
$SPEC$	35	4	86	67	50	36	33	24	18
$SUB^{2)}$	87	0	100	100	100	100	88	80	37
$INSUR^{2)}$	25	3	100	18	19	20	25	27	47

[1)] in 1,000 CHF, 1 CHF=0.8 USD (2004 exchange rates).
[2)] in percent, SUB=100 means that 100 percent of all hospitals are subsidized.

Summary statistics are shown in Table 1 for all six hospital groups retained. They suggest that technological heterogeneity between Swiss hospital groups indeed influence cost. University hospitals (K111) for example have the highest variable costs ($VC = 790, 609$); this can be attributed to their high values of the two major outputs ($PCASES = 43,046$ and $OUTP = 142,331$) and possibly the fact that they are all subsidized ($SUB = 100$ percent) while having a small share of patients with supplementary insurance ($INSUR = 18$ percent). However, they also have the most internship programs ($INTERN = 118$) and specialties ($SPEC = 67$). Specialized hospitals (K231) on the other hand are on average small hospitals with fewer internship programs ($INTERN = 3$) and specialties ($SPEC = 18$) but are mostly non-subsidized ($SUB = 37$ percent) while having a high share of supplementary insured patients ($INSUR = 47$ percent).

With these data, we can specify a cost function where variable cost (VC) depends on two output categories ($PCASES$, $OUTP$), one input price for labor (PL), one capital stock ($BEDS$), and four structural variables ($INTERN$, $SPEC$, SUB, $INSUR$). Although some of them could be interpreted as observable indicators of technology (in particular, the number of specialties offered), they are treated as a category of their own here. Intercept and technology parameters are reflected by a linear function of dummies for different hospital groups (Z) and different time periods (M).[6] The underlying Cobb-Douglas cost function (subscripts $i = 1, ..., N$ and $t = 1, ..., T$ are dropped for simplicity) therefore reads,

$$
lnVC \quad = \quad \beta_0 + \sum_{m=1}^{4} \beta_m lnX_m + \sum_{z=1}^{3} \beta_z S_z + u + v, \quad \text{with}
$$

$$
\beta_{k,\ k=\{0,\ m,\ z\}} \quad = \quad \bar{\bar{\beta}}_k + \sum_{l=1}^{6} \gamma_{l,k} Z_l + \sum_{\tau=1}^{3} \delta_{\tau,k} M_\tau + w. \tag{12}
$$

In order to conduct Bayesian inference from the posterior given in Section 2.2, prior distributions need to be specified. The values for the hyperparameters are chosen in a way to imply relatively vague but proper priors. In particular, the priors are assumed to be independent,

$$
p(\bar{\bar{\alpha}}, \bar{\bar{\beta}}, \gamma, \delta, \Sigma, \sigma_v^{-2}, \sigma_u^{-2}) \quad = \quad p(\bar{\bar{\alpha}}) p(\bar{\bar{\beta}}) p(\gamma) p(\delta) p(\Sigma) p(\sigma_v^{-2}) p(\sigma_u^{-2}). \tag{13}
$$

Here, $p(\bar{\bar{\alpha}}) = f_N[0, \theta_{\bar{\bar{\alpha}}}]$, $p(\bar{\bar{\beta}}) = f_N[0, \theta_{\bar{\bar{\beta}}}]$, $p(\gamma) = f_N[0, \theta_\gamma]$, and $p(\delta) = f_N[0, \theta_\delta]$ have a normal distribution with mean zero and a diffuse prior for their corresponding precision θ. The precision of the likelihood function has a gamma distribution $p(\sigma_v^{-2}) = f_G[\mu, \theta_{\sigma_v^{-2}}]$ with diffuse shape and scale parameters. Inefficiency is assumed to be half normally distributed $p(u, \sigma_u^{-2}) = f_N^+[0, \sigma_u^{-2}]$ with $\sigma_u^{-2} = f_G[5, (5 * log(\overline{eff})^2)]$. This specification is in line with Griffin and Steel (2007) and Koop et al. (1997), permitting to impose a priori information with regard to mean efficiency, $\overline{eff} = exp(-\bar{u})$. Following the formulation of Griffin and Steel (2007), $\overline{eff} = 0.875$ is assumed for prior efficiency. Finally, the precision of the random parameters is specified as a Wishart distribution $p(\Sigma) = f_W[S]$ in accordance with eq. (9) with diffuse prior for the covariance matrix S.

To obtain posterior estimates, MCMC algorithms were run for 100,000 iterations, with the first 50,000 discarded as a burn-in phase. Different assumptions for priors and starting values converged to roughly the same values, suggesting that convergence to the posterior distribution was achieved.

---

[6] Note that average length of stay is not included variable. its expected value enters the casemix adjustment of *PCASES*. Therefore, deviations from expected value can be interpreted as reflecting management inefficiency u.

3.2. Econometric Results of the Cost Functions and their Cost Variability

Table 2 presents estimated means and standard errors of the cost function for the Standard Frontier Model (SFM), the Random Intercept Model (RIFM), and the Random Parameter Model (RPFM). The RPFM is presented with estimates for $\bar{\bar{\alpha}}$ and $\bar{\bar{\beta}}$. Results of second level-estimates of the technology parameters and the intercept are shown in Table 5 of the Appendix; they point to a cost shift over time, but without affecting the slope parameters. The three variants in Table 2 can be assessed using the DIC information criterion (Spiegelhalter et al., 2002). The lower the DIC-value, the better the goodness of fit of the estimated cost function, indicating that the RPFM has the best model fit, followed by RIFM, with SFM definitely behind.

Table 2: Econometric Results

	SFM		RIFM		RPFM	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Constant	2.176	(0.279)	3.334	(0.340)	4.948	(1.109)
PCASES	0.674	(0.037)	0.554	(0.045)	0.485	(0.115)
OUTP	0.025	(0.004)	0.015	(0.004)	0.013	(0.011)
PL	0.254	(0.058)	0.199	(0.049)	0.384	(0.110)
BEDS	0.256	(0.039)	0.311	(0.044)	−0.148	(0.127)
INTERN	0.004	(0.000)	0.001	(0.001)	0.008	(0.006)
SPEC	0.002	(0.001)	0.001	(0.001)	0.001	(0.007)
SUB	0.141	(0.036)	0.191	(0.055)	0.231	(0.529)
INSUR	0.004	(0.001)	0.004	(0.001)	0.003	(0.007)
σ_v^2	0.013		0.003		0.003	
σ_u^2	0.025		0.005		0.004	
σ_α^2	–		0.021		0.001	
λ_u	0.647		0.180		0.539	
λ_α	–		0.721		0.095	
DIC	−463.470		−1100.700		−1112.800	

Nevertheless, the three specifications produce fairly minor variations in technology parameters. Estimates also have the expected sign, with the only exception of *BED*. Being an indicator of capital stock, it should have a negative sign, which only obtains in the RPFM.

However, the main interest of this research revolves about technological heterogeneity that may not be accounted for in the SFM and its influence on estimated inefficiency scores. The relevant estimates are σ_v^2, σ_u^2, and σ_α^2, which stand for the variance of random noise, inefficiency, and separable heterogeneity, respectively; their relative importance (expressed as a share of total error variance) is given by λ_u and λ_α. For the SFM, total cost variability is 0.038 ($\sigma_v^2 = 0.013$, $\sigma_u^2 = 0.025$), with most variation in the inefficiency term. According to $\lambda_u = 0.647$, about 65 percent of cost variability is due to inefficiency, a share comparable to the literature cited in the Introduction section. Next, the RIFM with $\sigma_v^2 = 0.003$, $\sigma_u^2 = 0.005$, and $\sigma_\alpha^2 = 0.021$ confirms the existence of separable heterogeneity. It

also indicates a reduction of cost variability by 23 percent (from 0.038 to 0.029), due to its ability to capture heterogeneity in the random intercept. Variability due to inefficiency even declines by 80 percent (from 0.025 to 0.005) and from 65 percent to 18 percent in relative terms. Most of the cost variability between Swiss hospitals can now be attributed to unobserved separable heterogeneity ($\sigma_\alpha^2 = 0.021$), accounting for 72 percent of total cost variability. Still, non-separable heterogeneity is likely to exist, biasing coefficients and inefficiency scores.

The RPFM confirms this concern. As to bias, the capital indicator *BEDS* now has the expected negative sign (although insignificant) while the coefficient of *PL* attains the high value suggested by the argument that hospital costs are mainly driven by labor cost. As to cost variability, it again decreases markedly by 72 percent (from 0.029 RIFM to 0.008 RPFM). Most of the cost variability that cannot be attributed to casemix – equal to 0.042 and estimated by the residual sum of squares of a single regression with *VC* as a dependent variable and *PCASES* as the independent variable – can now be explained by the estimated cost function. Interestingly, while $\sigma_v^2 = 0.003$ and $\sigma_u^2 = 0.004$ change little from RIFM, the variance of separable heterogeneity diminishes drastically from 0.021 to 0.001, with its relative importance falling from 72 percent to a more plausible 9 percent. Consequently, the relevance of inefficiency increases to 54 percent, comparable to the SFM value.

Table 3: Variance-Covariance Matrix of the Wishart Distribution

	PCASES	OUTP	PL	BEDS	INTERN	SPEC	SUB	INSUR
PCASES	0.505*							
OUTP	−0.338	0.372*						
PL	0.257	−0.196	0.297*					
BEDS	−0.216	0.171	−0.132	0.257*				
INTERN	−0.125	0.088	−0.17	0.091	0.362*			
SPEC	−0.074	0.057	−0.035	0.043	0.001	0.132*		
SUB	0.208	−0.144	0.09	−0.091	0.060	−0.018	0.279*	
INSUR	0.389	−0.318	0.267	−0.238	−0.159	−0.081	0.159	0.584*

∗ Significant at the 95% confidence level.

Indeed, the RPFM attributes most of the cost variability to the technology parameters, as can be seen from the covariance matrix of the Wishart distribution in Table 3. The variances on the diagonal show that heterogeneity is strongly related to the output indicator *PCASES* and the structural variable *INSUR*, in spite of the fact that *PCASES* already adjusts for heterogeneity through a casemix index. At least in the Swiss case, this raises doubts about the relevance of the casemix index used to adjust for cost variability in prospective payment. As to the off-diagonal entries, the negative correlation between *BEDS* and *PL*, although insignificant, points to capital and labor being complements in the hospital sector.

3.3. Efficiency Scores

In many countries, public authorities finance major parts of hospital investment and decide about the opening, closing down, and restructuring of hospitals. For these decisions, efficiency scores may provide guidance. However, as shown in Section 3.2, accounting for heterogeneity has an impact on the cost variability attributed to inefficiency (σ_u^2, λ_u). Therefore, a comparison of mean efficiency scores, their distribution, and development over time between the three models is of considerable interest.

Efficient hospitals are on the estimated cost frontier ($\widehat{u_{it}} = 0$, $\widehat{eff} = 1$), those with inefficiency above the frontier ($\widehat{u_{it}} > 0$, $\widehat{eff} < 1$). Since the u_{it} are in logarithms, one has

$$\widehat{eff}_{it} = exp(-\widehat{u_{it}}), \tag{14}$$

with $\widehat{u_{it}}$ simulated from a half normal distribution.

Table 4: Efficiency Values by Model Type and Year

	Average	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis
SFM	0.89	0.96	0.73	0.04	−1.07	4.23
RIFM	0.95	0.98	0.84	0.02	−2.78	17.70
RPFM	0.95	0.98	0.90	0.01	−2.06	12.84
2007						
SFM	0.87	0.94	0.75	0.04	−0.90	3.37
RIFM	0.95	0.97	0.84	0.02	−3.74	20.56
RPFM	0.95	0.97	0.90	0.01	−3.09	16.53
2006						
SFM	0.88	0.95	0.75	0.04	−0.90	3.69
RIFM	0.95	0.97	0.86	0.02	−2.69	15.55
RPFM	0.96	0.97	0.94	0.01	−0.54	3.57
2005						
SFM	0.90	0.96	0.73	0.04	−1.46	6.25
RIFM	0.95	0.98	0.87	0.01	−1.77	11.61
RPFM	0.95	0.98	0.91	0.01	−1.35	8.60
2004						
SFM	0.91	0.96	0.78	0.03	−1.41	5.76
RIFM	0.95	0.98	0.89	0.01	−0.96	5.77
RPFM	0.95	0.97	0.92	0.01	−1.13	5.42

Starting with the top of Table 4, the first thing to note is that mean efficiency scores are 0.89 or higher, putting the potential for cost reduction at 11 percent or less. This figure is much closer to the 14 percent reported by Hollingsworth (2008) for other European countries and cited in the Introduction section. However, efficiency scores derived from the SFM are markedly lower than their RIFM and RPFM counterparts. Unobserved heterogeneity therefore does lower estimated efficiency

scores, causing the potential for cost reduction to be overstated. Over the observation period, the average SFM score is 0.89, suggesting a cost reduction potential of 11 percent. Using the RIFM that corrects for separable heterogeneity, one arrives at a mean score of 0.95, a value comparable to Farsi and Filippini (2008) who estimated a true random effects model. Thus, prospective payment based on a SFM would overestimate the potential for cost reduction by no less than 6 percent, causing financial distress to at least some cost-efficient hospitals who happen to be stuck with inferior technology e.g. due to old buildings. Turning to the RPFM, which distinguishes non-separable from separable heterogeneity, one does not find a change away from RIFM mean scores. However, the minimum value is now 0.90 rather than 0.84, accompanied by a decrease in (negative) skewness and kurtosis.

Figure 1: Efficiency Estimates of the SFM, RIFM, and RPFM, Years 2004-7
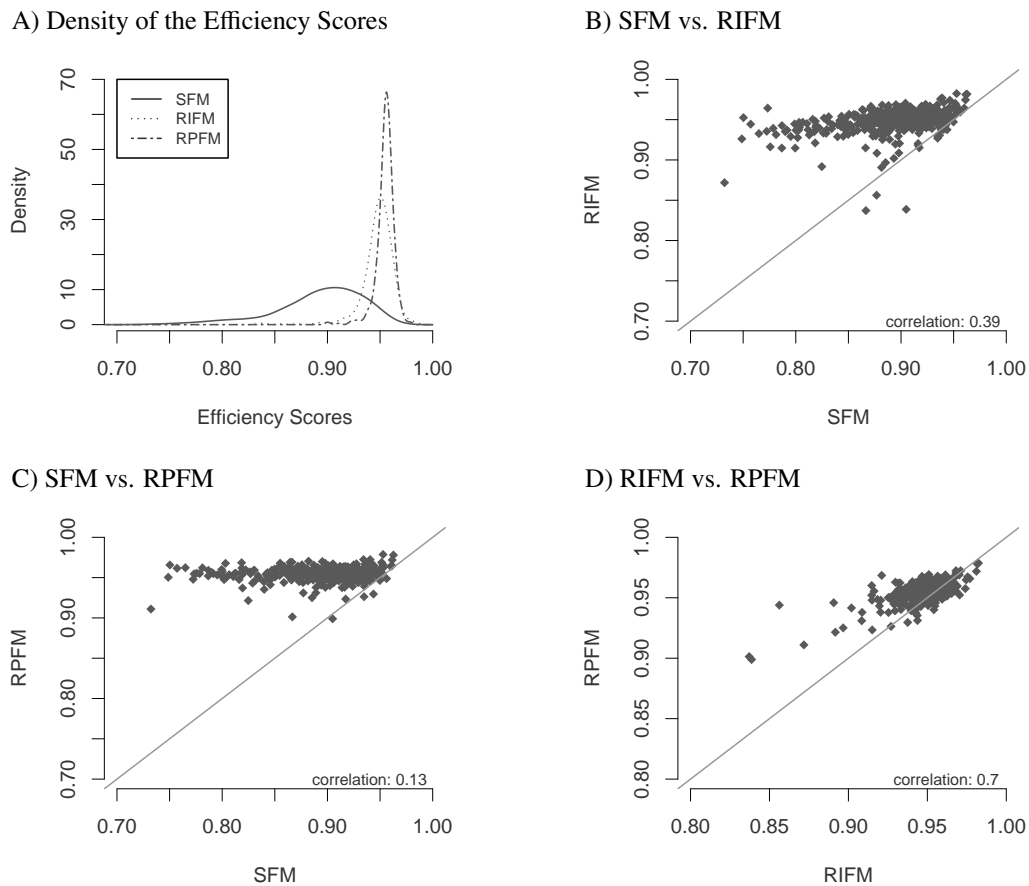
While these differences are evident from panel A of Figure 1, a comparison of individual efficiency scores is even more telling. Panel B of Figure 1 reveals that hospitals that would have been rated below 85 percent efficiency according to SFM gain up to 12 percentage points when the RIFM is applied instead. Panel C shows that this gain may even reach 15 points when the more general RPFM is used. Finally, the comparison between RIFM and RPFM in panel D of Figure 1 indicates that hospitals

with a RIFM score below 0.92 still would benefit from a transition to RPFM, although the gain rarely exceeds 5 percentage points. Therefore, at a given point of time and for a majority of Swiss hospitals, it clearly matters whether or not unobservable heterogeneity is taken into account in performance measurement.

Still, the three models might agree when it comes to development over time. Returning to Table 4, one notices that mean SFM efficiency scores have decreased over time, from 0.91 in 2004 to 0.87 in 2007. In sharp contrast, the RIFM and RPFM measures remained constant. Under the impression of SFM estimates, regulators would therefore have concluded that prospective payments should be cut to squeeze increasingly important cost reductions out of the hospital sector. Yet the evidence points to an increased importance of unobservable heterogeneity (see also the hikes in skewness and kurtosis of RIFM and RPFM scores especially in 2007). Such an increase is credible in view of the fact that in response to sluggish economic growth, hospital renovation projects were postponed or downsized. Failure to control for unobservable heterogeneity thus risks to punish more and more harshly those hospitals that are hampered by outdated technology.

4. Concluding Remarks

With prospective reimbursement of hospitals becoming increasingly common, measurement of hospital performance has been gaining in importance. However, measurement can be biased if it fails to account for technological heterogeneity, part of which is unobservable. This paper seeks to shed some light on the importance of both observable and unobservable heterogeneity by estimating a standard cost frontier model, a random intercept cost frontier model, and a random parameter cost frontier model. Bayesian inference is applied to the data of about 100 Swiss hospitals between 2004 and 2007.

Results confirm the existence of unobserved heterogeneity causing some of the cost variability even though inpatient cases are casemix-adjusted. This means that the standard frontier model, which does not control for heterogeneity, is insufficient for deriving unbiased performance measures. In the case of Switzerland, the biases may be substantial. Whereas Hollingsworth (2008) reports a potential for cost reductions of 14 percent for Europe, Farsi et al. (2006) put it to 20 percent for Swiss hospitals covering the years 1997 to 2002. In this paper, a comparable standard frontier model suggests 11 percent for 2004 to 2007, dropping to 5 percent when both observable and unobservable heterogeneity are controlled for. An element-wise comparison reveals that hospitals rated 85 percent efficient and less (using the standard method) would gain up to 12 percentage points. Therefore, quite a few hospitals, although highly efficient in fact, would end up in financial distress if regulators

were to cut reimbursement rates in an attempt to enforce the cost reductions indicated by the standard frontier model. This underlines the importance of accounting for both observable and unobservable heterogeneity in the estimation of hospital cost frontiers.

However, this analysis is not without limitations. First, there is the risk of misspecification causing bias in the benchmark values used to characterize the posterior distribution governing Bayesian inference. Specifically, the cost of capital is missing from the equation; but in addition, the indicator for unit cost of labor is an aggregate over skill categories ranging from physicians to orderlies, and just counting the categories of internships offered likely constitutes a poor measure of educational services provided. Second, the distinction between observable and unobservable heterogeneity remains somewhat arbitrary; for instance, if measures of the vintage of hospital capital stock in terms of buildings and medical technology were available, a greater part of cost variability would be attributed to the observable component, likely causing the estimated influence of total heterogeneity on cost to be reduced. This point relates to a third weakness, which is that management inefficiency continues to be measured as a residual rather than by direct indicators. Therefore, by minimizing the contribution of this residual to cost variation, random parameter models might end up going too far in exonerating hospital management.

This said, the evidence presented here does suggest that the standard cost frontier model is insufficient for measuring hospital performance due its failure to take technological heterogeneity into account at all. While the evidence is limited to a sample of Swiss hospitals, the reasons for heterogeneity are of a general nature. They also apply to other heavily regulated or public sectors such as energy, education, and banking, underlining the importance of specifying cost frontier models that yield unbiased efficiency scores.

Table 5: Econometric Results of the Random Technology Parameters

	RIFM	RPFM								
	CONSTANT	CONSTANT	PCASES	OUTP	PL	BEDS	INTERN	SPEC	SUB	INSUR
$\bar{\bar{\beta}}$	3.106	4.948	0.485	0.013	0.384	−0.148	0.008	0.000	0.231	0.003
	(0.32)	(1.11)	(0.115)	(0.011)	(0.110)	(0.127)	(0.006)	(0.007)	(0.529)	(0.007)
$K111$	0.693	0.439	0.170	−0.157	−0.013	−0.023	0.007	0.014	0.480	0.000
	(0.14)	(3.04)	(0.780)	(0.317)	(0.745)	(0.935)	(0.019)	(0.017)	(3.008)	(0.051)
$K112$	0.273	0.865	−0.250	0.066	−0.166	0.282	−0.004	0.002	0.915	−0.008
	(0.07)	(2.56)	(0.190)	(0.096)	(0.300)	(0.411)	(0.010)	(0.013)	(2.552)	(0.014)
$K121$	0.187	0.992	−0.231	0.028	−0.295	0.241	−0.007	0.010	1.024	0.007
	(0.06)	(2.57)	(0.251)	(0.096)	(0.254)	(0.394)	(0.006)	(0.011)	(2.574)	(0.013)
$K122$	0.097	0.105	−0.220	−0.012	−0.183	0.474	−0.005	0.000	0.601	0.002
	(0.04)	(1.58)	(0.157)	(0.012)	(0.145)	(0.205)	(0.008)	(0.009)	(0.955)	(0.010)
$K231$	0.082	0.676	−0.247	0.151	−0.129	0.046	−0.010	−0.005	0.682	0.002
	(0.07)	(2.16)	(0.236)	(0.105)	(0.276)	(0.229)	(0.080)	(0.019)	(0.987)	(0.014)
M_{05}	0.034	0.462	0.078	−0.004	−0.148	−0.065	0.000	0.000	−0.054	0.000
	(0.01)	(0.33)	(0.047)	(0.005)	(0.077)	(0.047)	(0.001)	(0.001)	(0.044)	(0.001)
M_{06}	0.129	1.148	−0.074	0.005	−0.152	0.069	0.000	−0.001	−0.071	0.000
	(0.01)	(0.43)	(0.059)	(0.009)	(0.105)	(0.063)	(0.001)	(0.001)	(0.062)	(0.001)
M_{07}	0.155	0.902	−0.130	0.004	−0.041	0.108	0.000	0.000	−0.057	0.000
	(0.01)	(0.45)	(0.062)	(0.010)	(0.108)	(0.064)	(0.001)	(0.001)	(0.064)	(0.001)

References

Aigner, D., Lovell, C. A. K., Schmidt, P., 1977. Formulation and estimation of stochastic frontier production function models. Journal of Econometrics 6 (1), 21–37.

Farrell, M., 1957. The measurement of productive efficiency. Journal of the Royal Statistical Scociety 120 (3), 253–281.

Farsi, M., Filippini, M., 2006. An analysis of efficiency and productivity in Swiss hospitals. Swiss Journal of Economics and Statistics (SJES) 142 (I), 1–37.

Farsi, M., Filippini, M., 2008. Effects of ownership, subsidization and teaching activities on hospital costs in Switzerland. Health Economics 17 (3), 335–350.

Farsi, M., Filippini, M., Greene, W., 2006. Application of panel data models in benchmarking analysis of the electricity distribution sector. Annals of Public and Cooperative Economics 77 (3), 271–290.

Greene, W., 2004a. Distinguishing between heterogeneity and inefficiency: stochastic frontier analysis of the world health organization's panel data on national health care systems. Health Economics 13 (10), 959–980.

Greene, W., 2004b. Interpreting estimated parameters and measuring individual heterogeneity in random coefficient models. Tech. rep.

Greene, W., 2005a. Fixed and random effects in stochastic frontier models. Journal of Productivity Analysis 23, 7–32.

Greene, W., 2005b. Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. Journal of Econometrics 126 (2), 269–303.

Griffin, J., Steel, M., 2007. Bayesian stochastic frontier analysis using WinBUGS. Journal of Productivity Analysis 27, 163–176.

Hollingsworth, B., 2008. The measurement of efficiency and productivity of health care delivery. Health Economics 17 (10), 1107–1128.

Huang, H.-C., 2004. Estimation of technical inefficiencies with heterogeneous technologies. Journal of Productivity Analysis 21 (3), 277–296.

Jacobs, R., Smith, P. C., Street, A., 2006. Measuring Efficiency in Health Care. Cambridge University Press.

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

Koop, G., Osiewalski, J., Steel, M. F. J., 1997. Bayesian efficiency analysis through individual effects: Hospital cost frontiers. Journal of Econometrics 76 (1-2), 77–105.

Meeusen, W., van den Broeck, J., 1977. Efficiency estimation from Cobb-Douglas production functions with composed error. International Economic Review 18 (2), 435–44.

Newhouse, J. P., 1996. Reimbursing health plans and health providers: Efficiency in production versus selection. Journal of Economic Literature 34 (3), 1236–1263.

Ntzoufras, I., 2009. Bayesian modeling using WinBUGS. Wiley Series in Computational Statistics. Wiley, Hoboken, NJ.

Orea, L., Kumbhakar, S. C., 2004. Efficiency measurement using a latent class stochastic frontier model. Empirical Economics 29 (1), 169–183.

Pitt, M. M., Lee, L.-F., 1981. The measurement and sources of technical inefficiency in the indonesian weaving industry. Journal of Development Economics 9 (1), 43 – 64.

Schmidt, P., Sickles, R. C., 1984. Production frontiers and panel data. Journal of Business & Economic Statistics 2 (4), 367–74.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., van der Linde, A., 2002. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society Series B 64, 583–640.

Steinmann, L., Dittrich, G., Karmann, A., Zweifel, P., 2004. Measuring and comparing the (in)efficiency of German and Swiss hospitals. The European Journal of Health Economics 5 (3), 216–226.

Steinmann, L., Zweifel, P., 2003. On the (in)efficiency of Swiss hospitals. Applied Economics 35 (3), 361–370.

Tsionas, E. G., 2002. Stochastic frontier models with random coefficients. Journal of Applied Econometrics 17 (2), 127–147.

Widmer, P. K., 2010. Does prospective payment increase hospital efficiency? evidence from a natural experiment. Working paper, Socioeconomic Institute.

Worthington, A., 2004. Frontier efficiency measurement in healthcare: A review of empirical techniques and selected applications. Medical Care Research and Review 61 (2), 1–36.