Int J Health Econ Manag (2015) 15:73–97
DOI 10.1007/s10754-014-9159-4

RESEARCH ARTICLE

Payment schemes and cost efficiency: evidence from Swiss public hospitals

Stefan Meyer

Received: 12 December 2013 / Accepted: 24 November 2014 / Published online: 2 December 2014

Abstract This paper aims at analysing the impact of prospective payment schemes on cost efficiency of acute care hospitals in Switzerland. We study a panel of 121 public hospitals subject to one of four payment schemes. While several hospitals are still reimbursed on a per diem basis for the treatment of patients, most face flat per-case rates—or mixed schemes, which combine both elements of reimbursement. Thus, unlike previous studies, we are able to simultaneously analyse and isolate the cost-efficiency effects of different payment schemes. By means of stochastic frontier analysis, we first estimate a hospital cost frontier. Using the two-stage approach proposed by Battese and Coelli (Empir Econ 20:325–332, 1995), we then analyse the impact of these payment schemes on the cost efficiency of hospitals. Controlling for hospital characteristics, local market conditions in the 26 Swiss states (cantons), and a time trend, we show that, compared to per diem, hospitals which are reimbursed by flat payment schemes perform better in terms of cost efficiency. Our results suggest that mixed schemes create incentives for cost containment as well, although to a lesser extent. In addition, our findings indicate that cost-efficient hospitals are primarily located in cantons with competitive markets, as measured by the Herfindahl–Hirschman index in inpatient care. Furthermore, our econometric model shows that we obtain biased estimates from frontier analysis if we do not account for heteroscedasticity in the inefficiency term.

Keywords Prospective reimbursement · Inpatient payment systems · Cost efficiency · Stochastic frontier analysis · Heteroscedastic frontier models

JEL Classification C23 · D22 · I11

Introduction

Continuously raising healthcare costs have become a matter of public concern not only in the USA, but in almost all the industrialisedcountries of Western Europe and throughout the

S. Meyer (✉)
Department of Health Economics, Faculty of Business and Economics (WWZ),
University of Basel, Peter Merian-Weg 6, 4002 Basel, Switzerland
e-mail: stefan.meyer@unibas.ch

world. In Switzerland, healthcare expenditure (HCE) has risen by 46 per cent over the last decade, reaching 10.9 % of GDP in 2010 (FSO 2011). It is widely accepted that income, technological progress, demographic ageing, and inefficiencies are the main driving forces behind this trend. However, as the first three factors are hardly controllable by regulators, the latter factor has become the focus of attention (Herr et al. 2011). A recent OECD study by Wranik (2012) confirms that the Swiss system is far from exploiting its full potential in terms of efficient health production: With regard to technical efficiency in the production of health, for instance, Switzerland ranked only 13th out of the 21 developed OECD countries in the sample.

In Switzerland, a substantial part of the national healthcare budget is spent on inpatient care. On average, expenditure on inpatient services in OECD countries account for about one third of total HCE, and amounted to 34.9 % in 2009 (OECD 2013). In comparison, the Swiss spent more than 45.5 % of their annual healthcare budget on hospital care in that year. This exceptionally high share was surpassed only by the Netherlands (45.8 %) and Italy (46.0 %). In Switzerland, therefore, even moderate inefficiencies in the provision of inpatient care will have a considerable effect on total healthcare spending.

Reimbursement schemes, which define the way healthcare services are paid for, are known to affect the behaviour of healthcare providers and thus also the amount and quality of service, access to healthcare, and cost efficiency (see Langenbrunner et al. 2009). In fact, in most industrialised countries, recent changes in provider reimbursement have aimed at curbing HCE by improving cost efficiency in inpatient care. Using data on Swiss public hospitals, we aim at providing new insights into the relationship between reimbursement schemes and cost efficiency. The four types of payment systems employed in Swiss public hospitals differ in terms of the amount of the per diem and the fixed per-case rate. While there are two flat-rate systems and one pure per diem scheme, one in two Swiss hospitals is paid on the basis of a hybrid system, which combines both elements of reimbursement. Consequently, we can simultaneously compare the performance effects of four different payment schemes. While the local governments (cantons) can decide upon the payment scheme, the basic rules of healthcare financing and health insurance coverage are determined at the national level. Therefore, contrary to international surveys, the problem of unobserved heterogeneity among different observation areas is alleviated. Moreover, many studies have used time-series data to look at efficiency effects of a change in the payment scheme (e.g., from per diem to flat-rate payments). This approach, however, is hardly able to completely distinguish the payment change effect from other exogenous shocks and time trends. To exploit the advantages of both cross-section analysis and time series, we use longitudinal data on a sample of hospitals over the period 2004–2009. Since we do not have data on the payment schemes of private hospitals, we only focus on state-owned and publicly-funded hospitals.[1,2]

We expect flat-rate systems (e.g., DRG) to perform best in terms of cost efficiency, *ceteris paribus*, as these hospitals cannot raise revenue by increasing treatment intensity or prolonging the length of stay (Aas 1995). On the other hand, per diem and fee-for-service schemes do not promote efficient production, as higher treatment costs are, at least partially, reimbursed. Mixed systems, though, are supposed to offer moderate incentives for cost containment, as they combine elements of both ordinary systems.

[1] We focus on state-owned and publicly-funded hospitals only, as it is not compulsory for private hospitals to implement the payment scheme introduced by the government. However, as far as we know, most private clinics also employ the cantonal system. Unfortunately, no reliable payment-scheme data on private acute care clinics was available until 2010.

[2] In 2009, 15.4 % of all acute-care patients in Switzerland were admitted to one of the 36 private general hospitals.

To analyse and isolate the effect of payment schemes on cost efficiency, a two-stage econometric model is applied. In a first step, we estimate a hospital cost frontier using stochastic frontier analysis (SFA). In a second stage, the impact of the payment scheme on hospital efficiency is analysed. We use the econometric model originally proposed by Battese and Coelli (1993), which allows for a simultaneous estimation of the two steps. This approach provides more efficient estimates of the parameters of the inefficiency-effects variables than those obtained using a two-stage estimation procedure (Coelli 1996). To disentangle payment scheme effects from cantonal heterogeneity, we include several explanatory variables at a regional level. Finally, studies indicate that SFA is prone to deliver biased coefficients due to heteroscedastic errors (Caudill and Ford 1993; Caudill et al. 1995; Hadri 1999). We try to overcome this drawback by parameterizing the variance of the two error components that are estimated by SFA.

We show that, compared to per diem, hospitals which are reimbursed by flat payment schemes perform better in terms of cost efficiency. Our SFA results suggest that mixed schemes create incentives for cost containment as well, although to a lesser extent. In addition, our econometric model shows that we obtain biased estimates if we do not account for heteroscedasticity in the inefficiency term.

The remainder of this paper is organised as follows. Section “hospital payment in Switzerland” outlines hospital payment in Switzerland. Section “payment schemes and cost efficiency” introduces a theoretic model, derives our hypothesis and gives an overview of empirical literature. Section “methodology” explains our estimation strategy and describes the data set. Section “results” presents the results of the stochastic frontier analysis. Section “concluding remarks” summarises and draws a conclusion.

Hospital payment in Switzerland

In Switzerland, all residents are required to purchase basic health insurance (BHI), which covers a exclusive range of treatments (benefit basket) detailed in the Swiss Federal Law on Health Insurance. People are free to buy their health insurance from any authorised private insurance company that offers BHI. The insurers must accept every individual who applies for coverage. Even though the benefit package is fixed, the insurers can propose optional health insurance contracts which provide lower premiums in exchange for higher deductibles or managed care contracts (OECD/WHO 2006).

While health insurance is regulated at the national level, the 26 Swiss cantons are responsible for the provision of healthcare services. Furthermore, they co-finance hospital care and decide upon how private insurers pay for inpatient and outpatient services by introducing a payment scheme. Cantons finance their healthcare expenditures by taxes.

Outpatient care, on the one hand, is mainly financed by insurance companies. In 2004, all Swiss cantons agreed to implement a national fee-for-service system that specifies medical fees for outpatient services provided by GPs and specialists. Inpatient services, on the other hand, are co-financed by the private health insurers and the cantons. According to the Swiss Federal Law of Health Insurance (KVG), the cantons have to pay at least 50 % of the eligible costs of public hospitals. In 2009, insurance companies reimbursed about 41.5 % of total costs in acute care, leaving 58.5 % to be paid by the public sector (FSO 2009). In general, cantons pay hospitals by means of global budget systems and (partially) cover generated deficits at the end of a financial year. In contrast to the unified tariffs in the outpatient sector, however, the cantons have implemented different payment schemes that govern the reimbursement

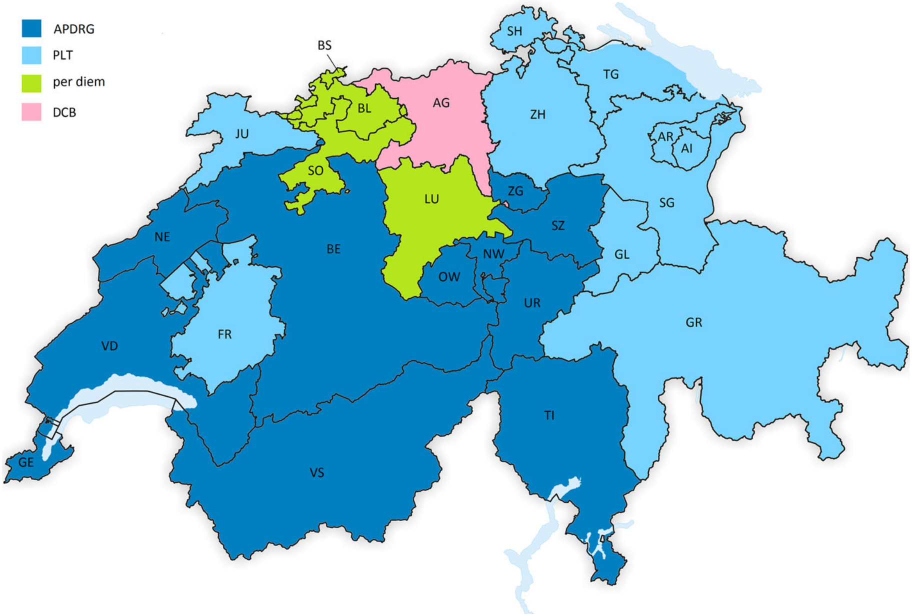

Fig. 1 Inpatient payment schemes in Switzerland, 2009

of hospitals by insurers. In 2009, private health insurance companies were subject to four different reimbursement schemes in the 26 cantons of Switzerland (see Fig. 1).

In the late 1990s, while the vast majority of public hospitals were still reimbursed on the basis of flat per diem payments, some cantons started to compensate acute-care clinics more prospectively. [3] In 2000, a hybrid system (PLT) [4] was implemented in the canton of Zurich. Other cantons quickly adopted the PLT system in the inpatient sector. By 2009, health insurance companies in 10 cantons had started to align payments on the basis of PLT. While PLT still entails a fixed amount per patient day for nursing and catering services, a substantial part of the money is paid prospectively per admission (flat). To account for differences in resource use associated with treatment, the flat amount differs according to the hospital department the patient is assigned to (e.g., general medicine, surgery, radiology, cardiology, etc.). Therefore, the system can be described as being partially prospective, consisting of a per day rate and a fixed amount per case.

In 1998, All Patient Diagnosis Related Groups (APDRG),[5] a prospective payment and classification system developed in the United States in the late 1980s, was tentatively implemented in selected Swiss hospitals. As this system proved to be an effective and practical alternative, other cantons gradually adopted this payment method. By 2006, almost half of the acute care hospitals in Switzerland got their services reimbursed by DRGs. APDRG hospitals receive a fixed payment per case depending on the diagnostic category the patient is

[3] The word *prospective* here refers to a payment system in which the hospital cannot influence income generated when treating a certain patient.

[4] *Prozess-Leistungs-Tarifierung* (German, "process- and performance-based pricing").

[5] APDRG was widely applied in the United States and, subsequently, updated versions of APDRG were adopted in various European countries, such as Spain and Portugal. Moreover, the system influenced the development of national DRG schemes, such as those of France and Australia (see Kobel et al. 2011).

Table 1 Switching cantons and payment schemes by canton and year

	2004	2005	2006	2007	2008	2009
Switching cantons						
AI	per diem	per diem	per diem	PLT	PLT	PLT
BE	DCB	DRG	DRG	DRG	DRG	DRG
GE	per diem	per diem	per diem	DRG	DRG	DRG
NE	PLT	PLT	DRG	DRG	DRG	DRG
VS	PLT	DRG	DRG	DRG	DRG	DRG
Number of cantons by payment scheme						
per diem	6	6	6	4	4	4
PLT	11	10	9	10	10	10
DRG	7	9	10	11	11	11
DCB	2	1	1	1	1	1
All	26	26	26	26	26	26

AG Aargau, *AI* Appenzell Innerrhoden, *AR* Appenzell Ausserrhoden, *BE* Berne, *BL* Basel-Landschaft, *BS* Basel-Stadt, *FR* Fribourg, *GE* Geneva, *GL* Glarus, *GR* Graubunden, *JU* Jura, *LU* Lucerne, *NE* Neuchâtel, *NW* Nidwalden, *OW* Obwalden, *SG* St. Gallen, *SH* Schaffhausen, *SO* Solothurn, *SZ* Schwyz, *TG* Thurgau, *UR* Uri, *VD* Vaud, *VS* Valais, *ZG* Zug, *ZH* Zurich, *TI* Ticino

grouped in. Unlike the PLT system, the highly diversified APDRG scheme offers more than 600 diagnostic groups. A specific case weight (CW) is assigned to every single group, reflecting the nationwide average cost of treating a patient in this category. Reimbursement rates are then calculated by multiplying the CW by a hospital-specific base rate, which is agreed upon by insurance companies and the hospital. Due to the great number of DRGs, resource use can be accounted for much more accurately compared to the flat department payments under PLT. Moreover, the DRG system is fully prospective, as there is no per diem rate.

The canton of Aargau and Berne employed a different prospective approach using department case-based payments (DCB). Similar to the PLT system, the insurer pays a fixed amount per patient according to the hospital department involved. As in the APDRG scheme, however, there is no additional per diem rate to cover nursing and catering services. Thus, this payment scheme is fully prospective as well.[6]

Table 1 lists the five cantons that implemented a new payment scheme for public hospitals during the period of study. Four cantons replaced their existing scheme by APDRG, while the canton of Appenzell Innerrhoden swapped its per diem scheme for PLT in 2007. In 2009, 22 cantons employed flat or hybrid payment schemes, while four cantons (BS, BL, SO, LU) continued to use per diem payments.

Payment schemes and cost efficiency

Theory and hypothesis

In a per diem approach, the apparent incentive is to raise the number of hospital days, by either increasing admissions or by raising the length of stay (LOS). According to Aas (1995),

[6] It should be noted that the largest public hospital in the canton of Aargau employed a slightly different payment scheme based on patient pathways (MIPP). However, as the two systems in question do not differ in the way they offer financial incentives to the hospital, we do not distinguish between them.

the incentive to lengthen a patient's period of stay is likely to be stronger than the incentive to raise admissions, since there is also motivation to lower resource use per day. Nevertheless, even if hospitals are interested in keeping costs per day down, per diem systems do not offer reasonable incentives to reduce inputs per case. Exactly the opposite is true. Aas (1995) argues that hospitals tend to subsidise the high costs during the first days of stay by low cost final days. Consequently, the LOS and thus average costs per case are likely to be relatively high under per diem.

Flat payment systems, on the other hand, simultaneously create the incentives to raise admissions while minimising costs per case (Langenbrunner et al. 2009). As providers usually have more control over resource use per case than over the total number of admissions, the former incentive is typically stronger. Given a certain level of quality, hospitals can contain costs per case by substituting towards cheaper inputs or by reducing total resource use per admission (e.g., reducing the LOS).

In addition, Ellis and Miller (2008) argue that incentives for cost containment are also affected by the accuracy of the flat payment system. Its effect is ambiguous, though. On the one hand, a more accurate payment scheme tends to lower cost variation within a payment category. Hence, there are greater monitoring problems and more incentives to *game the system* by classifying patients into a higher payment category (upcoding or DRG creep). As DRG creep typically increases compensation per patient, financial pressure and therefore incentives to work efficiently are reduced. On the other hand, however, highly diversified systems can help decrease incentives for *cream skimming*, a behaviour where hospitals try to admit the more lucrative patients (less severe cases). As hospitals must compete for these profitable patients by either increasing the intensity or quality of treatments, cost efficiency may worsen.

To illustrate the responses of the providers to the financial incentives created by per diem and flat-rate systems, we introduce a simple model. When a patient with a certain type of illness is admitted, the hospital can partially control the duration of stay by investing in cost efficiency measures (e.g., through better coordination of treatment activities). In order to keep things simple, we assume that the hospital maximises profit. Accordingly, the hospital's profit (π) from treating a patient can be written as

$$\pi(e) = F + p \cdot t(e) - C(t(e)) - \gamma(e), \tag{1}$$

where F is the flat payment per case, p is the per diem rate, and C are the total costs of treating a patient for t days. For simplicity reasons and without loss of generality, we make the assumption of a cost function with constant marginal costs, setting $C(t(e)) = C_F + c \cdot t(e)$.[7] One could think of $C_F \geq 0$ as the fixed costs incurred for the surgical intervention or main medical treatment, while $c \cdot t(e)$ captures all the variable costs that largely depend on the length of the inpatient stay (e.g., accommodation costs, meals, nursing, follow-up care, daily medication, etc.)

In inpatient care, the LOS is mainly determined by the severity of illness of the patient. This fact is implemented in our model by assuming that the hospital cannot set the duration of treatment directly. It can, however, decide upon its level of cost efficiency effort, $e \geq 0$, which is supposed to shorten the LOS at a non-increasing rate, i.e. $t_e < 0$ and $t_{ee} > 0$.[8] Hospitals which do not engage in efficiency measures at all may benefit from a significant drop in LOS by establishing a low level of efficiency (e.g., by improving efficient allocation of time slots

[7] The implications of the model still hold if the marginal costs are increasing in t, i.e., $C_{tt} > 0$.

[8] Implicitly, we solve the general case of $\pi(e) = F(\theta_i) + p \cdot t(\theta_i, e) - C(t(\theta_i, e)) - \gamma(e)$, where $\theta_i \in \{\underline{\theta}, \bar{\theta}\}$ is the severity of illness of the patient i.

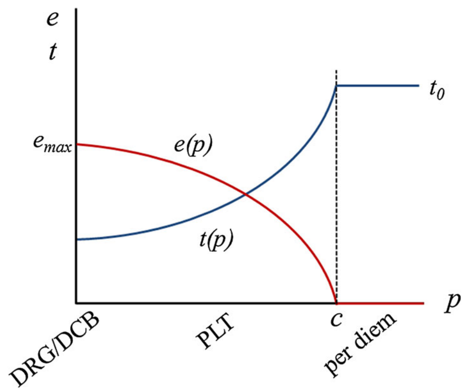

Fig. 2 Efficiency effort and payment schemes. e : Level of efficiency; t : Length of stay; p : per-diem rate; $e(p)$: Hospital efficiency as a function of the per-diem rate; $t(p)$: Length of stay as a function of the per-diem rate; e_{max} :Maximum level of efficiency that is achieved under flat payment; t_0 : Maximum length of stay that is realised under per diem

and operating theatres or implementing process management). Additional efficiency efforts made by highly-efficient facilities may still lead to a decrease in LOS. However, the marginal effect is likely to be smaller since the hospitals has to meet the strict quality standards set by the regulator.[9]

We refer to the maximum LOS as t_0, which occurs if the hospital does not engage in any cost efficiency measures at all, i.e. $e = 0$. Improved cost efficiency cannot be achieved for free, though. The function $\gamma(e)$ captures the corresponding loss of utility as well as the hospital's opportunity cost of time ($\gamma_e > 0$; $\gamma_{ee} > 0$). Differentiating (1) with respect to e and setting equal to zero yields the following first order condition, which determines the hospital's profit-maximising amount of cost efficiency effort (e^*):

$$\frac{\partial \pi}{\partial e} = (p - c)t_e(e^*) - \gamma_e(e^*) = 0 \tag{2}$$

We can easily see from (2) that if the hospital is paid on the basis of per diem ($p \geq c$), it will always be optimal to set the cost efficiency effort to zero. In this situation, all costs are reimbursed by the payer. The hospital, in turn, does not bear any risk for cost overruns at all. Hence, such payment methods do not offer any incentives for cost containment, since establishing a certain level of efficiency is costly.

In order to be able to compare the per diem scheme with our other payment systems in question, we have to show how the efficiency level (e) is affected by a change in the per diem rate (p). By applying the implicit function theorem to (2) and rearranging, we obtain

$$\frac{\partial e}{\partial p} = \frac{t_e}{(c - p)t_{ee} + \gamma_{ee}} < 0 \quad \forall\, p \in \{0, c\}\,. \tag{3}$$

Starting from a flat payment scheme with $p = 0$, (3) suggests that moving to a hybrid system with $0 < p < c$ is associated with a loss of efficiency. Eventually, when we approach a per diem setting ($p \geq c$), the level of efficiency effort drops to zero. Moreover, under reasonable assumptions with regard to the disutility function of the efficiency effort (γ) and the LOS function (t), the model indicates that the loss of efficiency is concave in p.[10]

Figure 2 summarises the relationship between the payment mode and the level of cost containment effort. In addition, the figure also illustrates how the LOS depends on the employed

[9] It is assumed that the quality of care delivered by the hospital can be verified by the regulator. Consequently, any decrease in the period of stay is achieved by efficiency measures only, while the quality of care is not affected. We further assume that the patient's opportunity costs of being in hospital do not depend on the LOS.

[10] The according proof can be found in the Appendix 1.

payment method. On the one hand, an increase in p is associated with a non-decreasing loss of efficiency effort, which approaches zero where p equals—or exceeds—marginal costs (c). On the other hand, the corresponding loss of efficiency goes hand in hand with a convex raise in the LOS, which reaches its maximum (t_0) at $p = c$.[11]

With regard to our empirical part, we expect cost efficiency to be correlated with the payment system in place. There are three different cases to consider. First, our model suggests that the two *flat systems* (DRG and DCB) with $p = 0$ can implement the highest level of cost efficiency. Accordingly, the hospital increases its level of efficiency effort (e) until the marginal costs of e, γ_e, equal the marginal cost savings achieved by the decrease in LOS, $c \cdot t_e$. This situation corresponds to the point e_{max} in Fig. 2. Second, if the hospital is subject to a *hybrid scheme* (PLT) with $0 < p < c$, incentives are diluted to some extent. This is due to the fact that the hospital's marginal savings from investing in efficiency effort then drop to $(c - p)t_e < c \cdot t_e$. The according level of efficiency effort (e^*) chosen by the hospital lies on the graph $e(p)$ in the interval $(0, c)$. Finally, the *per diem scheme* with $p \geq c$ does not create any incentives for cost containment. In this case, the hospital does not undertake any cost-saving activities (i.e., $e^* = 0$), as the marginal savings generated by reducing t are actually negative, amounting to $(c - p) < 0$.

Previous evidence

Several studies have used traditional parametric and non-parametric approaches to determine efficiency effects of hospital payment reforms. Regarding empirical evidence, changes in payment have been associated with improved technical efficiency in some European countries. Hospitals in Portugal (Dismuke and Sena 1999, 2001), Sweden (Gerdtham et al. 1999) and Norway (Biørn et al. 2003; Hagen et al. 2009) showed a higher level of technical efficiency following the implementation of DRG-based payment. Many other studies, on the other hand, have provided rather mixed evidence. According to Sommersguter-Reichmann (2000), the introduction of a prospective payment scheme in Austria had no effect on hospital-level efficiency. An earlier U.S. study by Borden (1988) did not find evidence for any efficiency gains in acute care hospitals in the state of New Jersey. Comparing different efficiency measures across 93 DRG and non-DRG hospitals from 1981 to 1984, the author concluded that DRGs had not had any positive effect on technical efficiency among hospitals. Similarly, Chern and Wan (2000) analysed the technical efficiency scores of 80 hospitals in the U.S. state of Virginia over a 10-year period. Contrary to the authors' hypothesis, the implementation of the PPS did not have any visible effect on hospital-level efficiency. Nevertheless, there is one noteworthy U.S. study which has found significant evidence with regard to differences in efficiency levels. Rosko and Mutter (2010) compared 543 small U.S. hospitals in rural areas that were subject to either a prospective payment system (PPS) or cost-based reimbursement. Accordingly, the SFA estimates revealed considerable differences in cost inefficiencies across hospital categories: While inefficiencies in PPS hospitals amounted to 10.3 %, on average, cost reimbursement was associated with significantly higher cost inefficiencies (15.9 %).

The relatively weak evidence indicates that efficiency studies with a in-country longitudinal perspective face considerable challenges. First, simple time-series data analysis is hardly able to find any effect if the time horizon is relatively short. A long-term study, on the other hand, may fail to establish any effect of the payment reform, as other exogenous shocks (e.g., other reform projects) occur in the interim (Street et al. 2011). Second, a pure time-series

[11] With regard to our data, a preliminary analysis showed that the average LOS in DRG and PLT hospitals was significantly lower than in facilities which applied per diem rates ($p < 0.01$), even after controlling for individual, hospital type, canton, and time fixed effects (see Table 6 in the Appendix 1).

approach is unlikely to completely distinguish the payment change effect from time fixed effects which affect all providers (e.g., demand and supply shocks). Hence, even though the policy change has had an impact on hospital performance, it cannot be quantified due to this methodological shortcomings.

With regard to popular *indicators* of efficiency, the average costs per case and the average LOS, empirical findings strongly support the efficiency hypothesis. A few studies have provided ample evidence for the relationship between the payment method and the LOS. For instance, Lave and Frank (1990), who examined 1,670 U.S. hospitals, found that flat payment systems lead to significant decreases in the LOS for medical, surgical and psychiatric patients. The average LOS in the inpatient sector fell by 15 per cent in the first three years following the implementation of the Medicare PPS. Likewise, Kahn et al. (1990) showed that for some diagnoses the average LOS dropped by even 24 % after the PPS was employed in U.S. hospitals. The authors used a nationally representative sample of 14,012 Medicare patients hospitalised in the periods 1981–1982 and 1985–1986. Sloan (1991) was able to isolate the effect of the payment scheme more clearly by comparing hospitals that switched to Medicare DRGs early on with those that continued to use per diem payments. On the basis of this quasi-experiment, he showed that the average LOS of Medicare patients dropped by 14.6 % in DRG hospitals, compared to a decrease of 7.9 % in the latter group. Hsiao et al. (1986) summarised the effect of two consecutive payment reforms in the U.S. state of New Jersey. In 1976, the budget regime was replaced by a more constraining per diem scheme. Only four years later, however, New Jersey adopted the national DRG-based system. While the annual decrease in the average LOS was −0.6 % during the budget regime between 1971–1976, the trend reversed under per diem (+0.7 % p.a.). Then, under DRG reimbursement, the average LOS fell again, amounting to −2.0 % per year.

Other countries have reported similar drops in LOS after moving to a prospective system. In a cross-country analysis, Moreno-Serra and Wagstaff (2010) estimated system-wide impacts of payment reforms in 28 former communist countries between 1990 and 2004. While the LOS was not affected in countries that shifted from line-item budgeting to fee-for-service or per diem, the authors found that the introduction of flat case payments reduced the average LOS by 4 %, approximately. Using hospital discharge data, Giammanco (1999) studied nearly 700,000 inpatient cases in Italy in 1995. Differentiating between ownership categories, Giammanco revealed significant decreases in the average LOS, amounting to −2.7 days in private hospitals (from 10.4 to 7.7) and −2 days in publicly-owned hospitals (from 8.1 to 6.1). Similar findings have been reported, for instance, in Germany (Hensen et al. 2007), England (Farrar et al. 2009), Norway (Biørn et al. 2003; Hagen et al. 2009), Sweden (Anell 2005) and Austria (Theurl and Winner 2007). The only Swiss study that examined the impact of hospital payment on the LOS found no significant differences between APDRG-hospitals and hospitals operating under a different scheme (Widmer and Weaver 2011). In both hospital categories, the average LOS decreased from 8.7 in 2001 to 7.4 in 2008.

A few authors have also demonstrated that shorter hospital stays due to prospective reimbursement also affect healthcare costs (Rosko 1984; Rosko and Broyles 1987; Helbing et al. 1990; Coffey and Louis 2001). In an extensive study, Davis et al. (1985) investigated in depth how the U.S. Medicare system changed inpatient utilisation, hospital capacity, and costs per case. The authors point out that the average costs per admission continued to increase after the implementation of Medicare in 1983, but at a much slower rate. While the annual increase in per case costs was 13.9 % between 1975 and 1983, the rate fell sharply to 7.8 % in 1984.

As with the efficiency studies, though, methodological shortcomings in the research cited above are common: With few exceptions, most studies have failed to apply sophisticated statistical methods (e.g., difference-in-difference), which are suitable for the evaluation of the causal effect of payment reforms. Therefore, it cannot be established to what extent the DRGs have contributed to the ongoing trend of shorter hospital stays—and the slower rate of cost growth in some countries.

Methodology

Stochastic frontier analysis

SFA was developed independently by Aigner et al. (1977) and Meeusen and Van Den Broeck (1977), introducing an econometric approach to frontier analysis with a composed error structure.[12] Accordingly, departures from the best practice frontier may be either stochastic (random shocks) or deterministic (inefficiencies). We employ SFA to estimate a stochastic cost frontier for Swiss public hospitals. Compared with production frontiers, which are used to estimate technical efficiency, cost frontiers suffer less from endogeneity problems, as they do not treat input quantities as exogenous right-rand variables.

To estimate the cost frontier consistently, we apply the pooled cross-section model to panel data suggested by Battese and Coelli (1993, 1995).[13] Their two-stage approach has been used by a growing number of economic studies on efficiency of healthcare providers (see, e.g., Rosko 1999, 2001, 2004; Smet 2007; Herr 2008; Rosko and Mutter 2010).

In our econometric model, we assume that the kernel of the stochastic cost frontier takes the multiple-output Cobb-Douglas form.[14] Expressed in logs, the Cobb-Douglas frontier can be written as

$$\ln TC_{it} = \alpha + \sum_{m} \beta_m \ln y_{it}^m + \sum_{n} \beta_n \ln p_{it}^n + \sum_{k} \beta_k s_{it}^k + v_{it} + u_{it}, \quad (4)$$

where TC_{it} are total costs incurred by hospital i at time t, y_{it}^m are the outputs produced by i, p_{it}^n are the input factor prices faced by hospital i, s_{it}^k are hospital characteristics that may influence total expenditure, and α and β are the technology parameters to be estimated. v_{it} is the normally distributed two-sided random-noise component with variance σ_v^2, and u_{it} is the non-negative inefficiency component of the idiosyncratic composed error term $\varepsilon_{it} = v_{it} + u_{it}$. Furthermore, since prices of all input factors are available, the Cobb-Douglas cost frontier must be linearly homogeneous in input prices, $\sum_n \beta_n = 1$.

According to Battese and Coelli (1995), the stochastic inefficiency effect u_{it} in (4) is specified as

$$u_{it} = z_{it}\delta + w_{it}, \; u_{it} \sim N^+(z_{it}\delta, \sigma_u^2), \quad (5)$$

where $z_{it} = (1, z_{it}^1, \ldots, z_{it}^L)$ is a vector of exogenous factors which directly impact inefficiency, δ is a vector of parameters to be estimated, and w_{it} are unobservable iid random

[12] There are two well-known frontier techniques that can be used to estimate efficiency scores at a firm level, SFA and data envelopment analysis (DEA). DEA is a non-parametric approach originally proposed by Charnes et al. (1978). However, econometricians have repeatedly criticised DEA due to its inability to separate variations in efficiency from random variations (Newhouse 1994).

[13] As Schmidt and Sickles (1984) mentioned, cross-sectional stochastic frontier models give rise to serious difficulties. For instance, Jondrow et al. (1982) noted that the variance of the conditional distribution of cost inefficiency does not go to zero when the sample size increases. As a result, we cannot obtain consistent efficiency estimates for a particular hospital even though the (whole) error term is estimated consistently.

[14] Alternatively, a translog cost function could be assumed. However, given our relatively small sample size, the translog specification would result in a considerable loss of degrees of freedom. In addition, the translog specification includes second-order terms and is therefore prone to multicollinearity (Farsi and Filippini 2008).

variables which are obtained by truncation of the normal distribution with mean zero and unknown variance, σ_u^2.

In general, σ_u^2 and σ_v^2 are assumed to be constant across observations and time, indicating homoscedastic error terms. However, in the presence of heteroscedastic errors in the two-sided stochastic error term (v_{it}), the standard errors in the model will be biased. Moreover, if the one-sided inefficiency term, u_{it}, is subject to heteroscedasticity, the coefficients obtained by SFA will be biased as well. Caudill and Ford (1993) investigated the effects of heteroscedasticity in the one-sided error, u_{it}, on parameter estimates in a frontier production function. They found that heteroscedasticity leads to biased estimates, particularly, when the model is estimated by the method of maximum likelihood. Moreover, Caudill et al. (1995), estimating a cost frontier in a Monte Carlo study, reported that the inefficiency measures were also affected by heteroscedasticity. In fact, not accounting for heteroscedasticity in the estimation also led to the overestimation of inefficiency for small facilities and the underestimation of inefficiency for large firms. Hadri (1999) extended the model proposed by Caudill et al. (1995) by introducing heteroscedasticity, not only in the one-sided error term, but also in the two-sided error term, v_{it}. According to the author, "ignoring this (...) will lead to inconsistent maximum likelihood estimators, and the usual test will no longer be valid" (Hadri 1999, p. 359).

Following Caudill et al. (1995) and Hadri (1999), we account for heteroscedasticity in the two error terms by parameterizing the variances of u_{it} and v_{it} as $\sigma_{uit}^2 = \exp(d_{it}\eta)$ and $\sigma_{vit}^2 = \exp(d_{it}\phi)$, respectively. d_{it} is one or more exogenous variables related generally to characteristics of firm size, while η and ϕ are vectors of unknown parameters to be estimated. Since firm-level data are used in frontier functions and firms vary widely in size, size-related heteroscedasticity is likely involved in the one-sided and two-sided error (Hadri et al. 2003).

To estimate the parameters of the stochastic frontier and the model for the cost inefficiency effects, the method of maximum likelihood is applied. The appropriate likelihood function and its partial derivatives with respect to the parameters of the model are given in the Appendix of Battese and Coelli (1993).[15]

Data

We use administrative data drawn from public and publicly-funded acute-care hospitals and specialised surgical clinics for the period 2004–2009 ($T = 6$). The primary sources for hospital-level data are the Annual Survey of Hospitals and the Hospital Medical Report. Both surveys are conducted annually by the Federal Statistical Office (FSO) and cover all public and private hospitals in Switzerland. An extract of both datasets is publicly accessible, including essential data on hospital category, inputs, inpatient and outpatient outputs, and costs. We first dropped all private hospitals from the sample. Public hospitals that reported complete information for at least two out of six years were included in our analytical file. On average, there are five observations per hospital in the sample. The number of observations range from 109 in 2004 to 98 in 2009 (see Table 2).

During this period, many existing hospitals merged or were incorporated into larger systems. The most significant drop in observations was registered in 2006. As a results, the panel is slightly unbalanced. After removing any observations with less than 750 inpatient days from our file, we obtained a panel of 122 Swiss hospitals, comprising a total sample size of 606. These observations cover a total of 4.9 million inpatient admissions, amounting to 58 % of all recorded hospital cases in Switzerland between 2004 and 2009 (FSO 2009).

15 As most statistical software packages do not provide the single-stage approach, we estimate the cost frontier using STATA commands proposed by Belotti et al. (2012).

Table 2 Number of hospitals in the sample by payment scheme and year

	2004	2005	2006	2007	2008	2009	All years
per diem	18	18	12	10	11	11	80
PLT	64	63	48	49	50	50	324
DRG	23	26	25	32	31	30	167
DCB	4	4	6	7	7	7	35
All	109	111	91	98	99	98	606

To estimate our efficiency model, we use data on cantonal payment schemes taken from the Swiss Conference of Health Ministers (CHM). However, as there are considerable time gaps in the CHM statistics, we use additional data on payment schemes provided by cantonal health departments. Furthermore, we take account of the fact that several cantons changed their payment systems during the observation period. In fact, though, only 7 hospitals in our sample were subject to a change in the payment system over the study period. Finally, we add three exogenous variables reflecting the market concentration in inpatient care, the market density in the outpatient sector, and the managed care penetration in the 26 cantons. All the regional data are drawn from publicly available sources and the FSO.

Table 3 reports summary statistics of the variables used in the analysis. Inpatient discharges ($CASES$) and outpatient revenue ($OUTP$) will be included as outputs in the cost function. As Rosko (2001) noted, the resource use varies considerably between different patient categories. We therefore adjust $CASES$ by a hospital-level Case Mix Index (CMI), which is calculated on the basis of APDRG. To clarify that the number of discharges has been adjusted, we add the subscript CMI to the discharges variable, $CASES_{CMI-Adjusted}$. The CMI accounts for all observable differences in patient morbidity across hospitals. Unfortunately, reliable information on outpatient production (e.g., adjusted outpatient visits) is hardly available. At the very least, the statistics contain some data on total outpatient revenue, which we use as a proxy for outpatient visits (Bjørn et al. 2003; Farsi and Filippini 2006).[16]

Labour and capital are recognised as inputs in the cost function model.The price of labour (PL) is approximated by the average annual salary per full-time-equivalent employee. To gain a proxy for the price of capital (PK), we divide total expenses on depreciation and interest by the number of hospital beds. For future research, a more complete specification of input prices is imperative. However, since data availability in our study is limited, we follow past practices (Grannemann et al. 1986; Zuckerman et al. 1994; Rosko 2001). We account for linear homogeneity in input prices by setting $\beta_{PK} + \beta_{PL} = 1$.[17]

As many authors before, we use hospital-specific measures as proxies for the prices of labour and capital. These measures, though, reflect the hospitals' choices about the average skill-mix of its employees and the amount and mix of capital equipment (Zuckerman et al. 1994). As a consequence, our estimates may be biased and inconsistent. Unfortunately, recent literature does not provide satisfactory instrumental variables. Nevertheless, to obtain some reassurance with regard to the endogeneity issue, we perform a Hausman test on the cost

16 However, outpatient revenue only gives some idea of outpatient costs, generally defined as average costs times quantity. Therefore, efficiency estimates may be biased, since they do not contain any information on cost efficiency in the outpatient wards of the hospitals in our dataset.

17 In our model, this restriction can be dealt with by subtracting $\ln(PL)$ from both sides of the equation. As a result, we obtain the restricted model $\ln(TC/PL) = \alpha + \sum \beta_m \ln y^m + \beta_{PK} \ln(PK/PL) + \sum \beta_k s^k + v + u$.

Table 3 Variable definitions and sample means by payment scheme

Variable	Description	per diem	PLT	DRG	DCB	All
Cost frontier variables						
TC	Total costs (CHF, in thousands)	197,765	103,724	149,505	125,459	130,010
$CASES_{CMI-Adjusted}$	Number of morbidity-adjusted discharges (adjusted by the APDRG case mix index)	11,613	7,041	9,036	8,576	8,283
$OUTP$	Revenue from outpatients (CHF, in thousands)	36,021	20,494	26,528	29,838	24,746
PL	Price of labour (CHF, in thousands)	102.28	99.87	98.71	103.94	100.10
PK	Price of capital (CHF, in thousands)	181.49	166.24	172.82	175.22	170.59
$WARDS$	Number of hospital wards and services	33.53	34.56	40.62	35.86	36.17
$INTERN$	Weighted number of internship categories	27.84	16.00	24.32	27.83	20.54
RQ	Target reservation quality of the hospital	1.87	1.61	1.78	1.53	1.69
$UNIVERSITY$	Binary variable (1,0) for university hospital	0.11	0.02	0.09	0.00	0.05
$SPECIALIST$	Binary variable (1,0) for specialised surgical clinic	0.15	0.07	0.01	0.00	0.06
Additional Z variables						
HHI	Herfindahl–Hirschman index in the market for inpatient care in the canton (calculated on the basis of beds in acute care)	0.24	0.16	0.25	0.11	0.19
$PHYS$	Number of GPs & specialists per 1,000 canton residents	2.66	1.92	1.96	1.51	2.00
$MCARE$	Share of people with managed care contracts in the canton	0.17	0.19	0.14	0.26	0.17
N		80	324	167	35	606

frontier after instrumenting the price of capital (Hausman 1978). The according test statistic fails to reject the hypothesis of no endogeneity in input prices ($p = 0.689$).[18]

The number of hospital wards ($WARDS$) account for observable product mix differences among hospitals. We further control for the degree of teaching activity ($INTERN$) by adding the adjusted number of internship categories offered in a hospital. We do not have access to sophisticated measures of quality of care. Nonetheless, we can at least control for some target level of in-house quality. Following Folland and Hofler (2001), we add a measure of reservation quality (RQ) to our cost function, which is based on the queuing model by Joskow (1980).[19] To control for other unobservable differences across hospital categories, we add

[18] We use four dummy variables as instruments for the cost of capital: TYPE1 (large general hospital), TYPE2 (medium general hospital), LEMAN (situated in the Cantons of Geneva, Vaud, or Valais) and ZH (situated in the Canton of Zurich). The $F(4, 592)$ statistic of the first stage regression amounts to 16.946. The Hausman $F(1, 594)$ statistic equals 0.161 ($p = 0.689$), which rejects the H_0 hypothesis of exogeneity.

[19] Joskow (1980) argues that hospitals set a target RQ, which can be estimated on the basis of the occupancy rate. Thereby, a higher reserve margin (low occupancy) indicates that the hospital aims to maintain a high

two indicator variables. $UNIVERSITY$ captures the unobserved level of research activities conducted in Swiss university hospitals, while $SPECIALIST$ is supposed to account for the different cost structure of specialised clinics (see (Farsi and Filippini 2006)). Finally, to account for time trends and technological change, we include a time-trend variable ($YEAR$) equal to 1 in 2004, 2 in 2005, and so sequentially up to 6 in 2009.[20]

The type of payment scheme, which we hypothesise to influence provider efficiency, is included in the regression model in (5). The influence of the payment method is tested by means of two dummy variables. PLT takes the value 1 if a hospital employs the mixed reimbursement scheme, and 0 otherwise. From a theoretic perspective, it is unlikely that the two flat-rate systems, DCB and DRG, exhibit different incentives for cost containment. We therefore construct an indicator variable, $FLAT$, that is equal to 1 if a hospital is either subject to a DRG or a DCB scheme. This step also ensures a certain level of model accuracy, since our sample only covers a total of seven DCB hospitals (i.e., 35 observations). The per diem scheme serves as the base category. Furthermore, a set of three covariates is included to account for regional differences. First, we calculate the Herfindahl–Hirschman Index (HHI) in acute care for 26 cantons. To obtain the market share of each hospital, we divide its supply of acute-care beds by the total number of beds in the canton. By canton, we then sum up the squared market shares of all hospitals in order to compute HHI. If the local market is relatively competitive (low HHI), hospitals face little market power. Consequently, the may be forced to contain cost by improving their cost efficiency. Second, the GP and specialists density ($PHYS$) serves as a proxy for the importance of the outpatient sector. The supply of primary healthcare may have nontrivial impacts on the demand for inpatient services. Still, efficiency scores are only supposed to be affected by the physician density if the changes in inpatient demand directly affect cost efficiency (e.g., through the price level). However, if the physician density only changes patient mix, we are not likely to see any correlation, since we account for the CMI in our cost function. Third, the variable $MCARE\%$ measures managed care penetration in the health insurance market. Since the premiums of managed care contracts are lower in general, the average inpatient prices and thus cost efficiency may differ across cantons.[21] To account for (non-linear) time effects, which may affect the efficiency of all institutions, we also add year fixed effects (2004–2009). Finally, to capture unobserved heterogeneity (fixed effects) across the regions of Switzerland, we also include seven region dummy variables.[22]

level of reservation quality (e.g., through a low average waiting time). By choosing a large RQ, the hospital is setting aside staffed beds as a reserve capacity available in case of unusually strong demand, hence the term reservation quality (Folland and Hofler 2001). We calculate $RQ_{it} = (B_{it} \times 365 - N_{it})/\sqrt{N_{it}}$, where B_{it} is bed supply and N_{it} is the number of patient days of hospital i at time t.

[20] Alternatively, we also estimate the model including time dummies instead of assuming a log-linear time trend. Since the sign and significance of the coefficients are not affected, we do not report the results of the dummy model.

[21] In Switzerland, the two main forms of mandatory insurance with a limited choice of healthcare providers are the health maintenance organisations (HMOs) and preferred provider organisations (PPOs). In PPOs, enrollees select a GP, who then acts as a gatekeeper for medical specialist care and inpatient care. Unless patients are in an emergency situation, they need a specific referral from the GP to the specialist or a hospital.

[22] These relatively homogeneous regions are used for statistical purposes by the FSO. The 7 regions are: Leman (GE, VD, VS), Mittelland (BE, SO, FR, NE, JU), Northwest (BL, BS, AG), Zurich (ZH), Eastern (SG, TG, AI, AR, GL, SH, GR), Central (UR, SZ, OW, NW, LU, ZG), Ticino (TI).

Results

Table 4 summarises the main results of the maximum-likelihood estimation of (4) and (5). We estimate three different specifications, comprising the model with heteroscedastic errors

Table 4 Parameter estimates for the cost function

$\ln TC$	(i) HEC V		(ii) HEC UV		(iii) HEC UV/FE	
	β	SE	β	SE	β	SE
Cost frontier						
$\ln CASES_{CMI-Adjusted}$	0.806***	(0.018)	0.849***	(0.018)	0.830***	(0.020)
$\ln OUTP$	0.056***	(0.014)	0.053***	(0.014)	0.057***	(0.014)
$\ln PK$	0.372***	(0.026)	0.377***	(0.025)	0.379***	(0.032)
$\ln WARDS$	0.002***	(0.000)	0.002***	(0.000)	0.002***	(0.000)
$\ln INTERN$	0.001***	(0.000)	0.001*	(0.000)	0.001**	(0.000)
RQ	0.044***	(0.007)	0.045***	(0.007)	0.042***	(0.007)
$UNIVERSITY$	0.073**	(0.033)	0.049	(0.030)	0.039	(0.031)
$SPECIALIST$	−0.069***	(0.024)	−0.072***	(0.025)	−0.074***	(0.029)
$YEAR$	0.011***	(0.004)	0.012**	(0.005)	0.011**	(0.005)
Effects on inefficiency						
$FLAT$	−0.370**	(0.162)	−0.284***	(0.073)	−0.265***	(0.102)
PLT	−0.127**	(0.064)	−0.101***	(0.038)	−0.135***	(0.052)
HHI	0.438***	(0.158)	0.327***	(0.074)	0.488***	(0.149)
$PHYS$	0.024	(0.041)	0.003	(0.027)	−0.017	(0.033)
$MCARE\%$	−0.322	(0.364)	−0.311	(0.209)	0.745*	(0.406)
$Y2005$	−0.057	(0.079)	−0.045	(0.051)	−0.047	(0.046)
$Y2006$	0.228	(0.133)	0.148	(0.064)	0.087	(0.056)
$Y2007$	0.186*	(0.126)	0.121**	(0.067)	0.025	(0.057)
$Y2008$	0.182	(0.143)	0.123*	(0.079)	−0.045	(0.070)
$Y2009$	0.167	(0.166)	0.125	(0.099)	−0.154	(0.107)
$ZURICH$					0.279**	(0.138)
$TICINO$					0.162	(0.120)
$CENTRAL$					0.044	(0.067)
$LEMAN$					0.247	(0.187)
$MITTELLAND$					0.282***	(0.097)
$NORTHWEST$					0.132	(0.111)
$\sigma^2_{uit} = \exp(d_{it}\eta)$		no		yes		yes
$\sigma^2_{vit} = \exp(d_{it}\phi)$		yes		yes		yes
$\gamma = \sigma^2_u/(\sigma^2_u + \sigma^2_v)$		0.866		0.899		0.875
Log-likelihood		357.264		366.730		386.537

$N = 606$; (i) HEC V parametrises heteroscedasticity only in the two-sided stochastic error term (v_{it}); (ii) HEC UV also accounts for heteroscedastic inefficiency effects (u_{it}); (iii) HEC UV/FE is similar to specification (ii). In addition, however, it controls for unobservable local characteristics by including 7 region dummies. The 7 regions are (cantons in parentheses): Leman (GE, VD, VS), Mittelland (BE, SO, FR, NE, JU), Northwest (BL, BS, AG), Zurich (ZH), Eastern (SG, TG, AI, AR, GL, SH, GR), Central (UR, SZ, OW, NW, LU, ZG), Ticino (TI); Cluster robust standard errors are given in parentheses; $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$; The sign of the efficiency variables are to be read as effects on *inefficiency* in v_{it} (HEC V), a model which accounts for both heteroscedastic u_{it} and v_{it} (HEC UV), and a augmented HEC UV model, which also allows for region fixed effects (HEC UV/FE). In our case, three indicators for hospital size ($\ln CASES_{CMI-Adjusted}$, $UNIVERSITY$ and $SPECIALIST$) serve as explanatory variables of σ_{ui}^2 and σ_{vi}^2. The coefficient of $CASES_{CMI-Adjusted}$ is negative in both cases, indicating that the variance of cost efficiency and random noise is less distinct among large hospitals ($p < 0.05$). The value of γ ranges between 86.6 and 89.9 %, which suggests that a substantial proportion of the total error variance in (4) is due to the stochastic inefficiency term, u_{it}. Consequently, estimating the model by SFA seems to be far more appropriate than using simple ordinary least squares.[23]

Comparing the *inefficiency effects* in column (i) and (ii), one can see that the coefficients and the corresponding standard errors tend to be smaller when heteroscedasticity in the inefficiency term is allowed for. This observation, however, is not true for the coefficients in the *cost frontier*: The parameter estimates are about evenly split between those that are larger than the estimates in (i) and those that are smaller. Caudill et al. (1995) found similar effects in their Monte Carlo study. The authors concluded that "in the multiproduct case, the [...] function is *twisted* by the heteroscedasticity" (Caudill et al. 1995, p.108). The inefficiency effect of $FLAT$ further diminishes when the region fixed effects (iii) are included, while the one of PLT increases again. This finding suggests that (ii) overestimates (underestimates) the true effect of FLAT (PLT) due to unobservable differences at the regional level. As regard the inefficiency estimates, the Spearman's statistic indicates that the rank correlation of inefficiency (u_{it}) is only moderate between HEC V and HEC UV (0.865), while being relatively high between HEC UV and HEC UV/FE (0.986). The mean inefficiency in (i) is relatively low (0.128), then rises significantly to 0.153 in the heteroscedastic frontier (ii), and comes down again to 0.142 in the model with fixed region effects (iii). Caudill et al. (1995), for instance, found even more pronounced changes in mean inefficiencies. Their analysis showed that the average inefficiency estimates were about 50 % higher at the mean when heteroscedasticity was taken into account.

The remaining discussions are based on the results of model (iii), which is considered superior. Not only can we correct for heteroscedastic errors in u_{it} and v_{it}, but we also account for all unobservable region characteristics that may over- or underestimate our inefficiency coefficients.

As expected, the coefficients of the two output variables in the cost function are positive. The coefficient of $\ln CASES_{CMI-Adjusted}$ can be treated as the cost elasticity of adjusted discharges. Its value is significantly smaller than one ($\beta = 0.83$). This implies that there might be economies of scale, which have not been fully exploited yet. Outpatient revenue only accounts for a relatively small share of total costs. The coefficient of $\ln PK$ suggests that capital expenditure accounts for about 38 % of total costs, leaving 62 % to labour. Furthermore, the coefficients of $\ln WARDS$ and $\ln INTERN$ indicate that diversification and teaching activity are positively related to hospital costs. This finding is consistent with recent studies using the same data (Widmer 2011). The positive coefficient of RQ confirms the costliness of maintaining a high reserve capacity. With regard to the hospital type, the positive coefficient of $UNIVERSITY$ suggests that university hospitals experience a shift in the cost function of +3.9 % (e.g., due to research activity). Moreover, another shift in costs is observed for specialised clinics, which seem to operate at significantly lower costs than general hospitals (−7.4 %). The estimated coefficient of the time variable ($YEAR$) is significantly different from zero, suggesting a change in the technology of production during

[23] If γ was zero, the variance of the inefficiency effects would be zero as well. Then, we would simply include the efficiency variables z_{it} in our cost function and estimate (4) by using ordinary least squares.

the observation period. With progress in medical technology, hospitals may use increasingly more advanced methods, which result in higher costs (Farsi and Filippini 2006). However, it should be noted that $YEAR$ may also capture other time-trend effects (e.g., changes in quality of reporting DRG cases).

The most interesting results, however, are obtained with the inefficiency system in (5). Taking a closer look at the payment scheme coefficients, we easily see that, compared to per diem, all three systems are associated with decreased cost inefficiency ($p < 0.01$). These results are consistent with our expectations concerning the financial incentives of prospective payment. In fact, the fully prospective systems, DRG and DCB, tend to affect CE more strongly than the hybrid system, PLT. Still, the PLT system, although being only partially prospective, seems to offer more incentives for cost containment than flat per diem payments. The difference between the two coefficients, $FLAT$ and PLT, is not significant, though. Expressed in marginal effects, $FLAT$ is associated with a reduction in cost inefficiency of 30.3 % (6.7–59.2 %), relative to the per diem system (Confidence intervals in parentheses). Likewise, the marginal effect of the mixed scheme (PLT) corresponds to a decrease in inefficiency of 14.5 % (3.4–23.7 %).

The Herfindahl–Hirschman index (HHI) in the market for inpatient care is positively related to inefficient production ($p < 0.01$). This result suggests that cost-efficient hospitals are likely located in competitive healthcare areas. Facilities which operate in highly competitive areas experience little market power. This, in turn, may force them to lower expenses, as they lose bargaining power when negotiating prices with health insurance companies. Although expected, this finding stands in stark contrast to many other studies that found a negative correlation between market concentration and inefficiency (see, e.g., Chirikos 1998; Rosko 1999, 2001).

The physician density, $PHYS$, is not correlated with hospital efficiency. This result fails to support the findings of an earlier study by Chirikos (1998), who took physician density (and population density) as proxy measures of demand. Analysing a panel of 186 Florida hospitals between 1982 and 1993, the author found that "inefficient facilities [hospitals] are more likely to be located in more population- and physician-dense areas. More elaborate multivariate models of efficiency values (...) confirm these differentials" (Chirikos 1998, p. 890). In general, though, population and physician density are highly correlated in most healthcare systems. Urban areas likely show a high concentration of GPs and specialists. Therefore, the finding by Chirikos (1998) may also be the consequence of multicollinearity in the model.

The estimated coefficient of $MCARE$ is slightly positive and suggests that cost inefficiency is more pronounced in cantons where the penetration of managed care contracts in BHI is higher. This effect, however, is reversed in the other specifications of our model. Nevertheless, this (weak) finding—being in contrast to the results of previous studies—suggests that managed care penetration is associated with decreased efficiency, although rather weakly (Rosko 2004). Still, this correlation may only be an artefact that is driven by reverse causality: High hospital costs due to inefficient production may raise the demand for managed care contracts in the canton in question.

The coefficients of the year fixed effects indicate that no linear time trend occurred during the observation period. Table 5 shows the mean cost inefficiency by payment scheme between 2004 and 2009. Comparing efficiency across the three systems, the pooled estimates actually confirm the findings of our econometric model. For instance, the per diem hospitals performed worst between 2004–2009 (19.8 %), followed by the hybrid PLT scheme (14.7 %). At the same time, hospitals which applied a fully prospective scheme, DRG or DCB, were on

Table 5 Mean inefficiency estimates by payment scheme and year

Year	per diem	PLT	FLAT	Overall	N
2004	0.185	0.133	0.117	0.138	109
2005	0.188	0.137	0.105	0.137	111
2006	0.192	0.156	0.105	0.144	91
2007	0.220	0.145	0.110	0.139	98
2008	0.208	0.159	0.110	0.146	99
2009	0.214	0.160	0.117	0.150	98
2004–2009	0.198	0.147	0.111	0.142	606

average the most cost-efficient (11.1 %).[24] While the cost inefficiency of hospitals subject to per diem (+2.9 %) or PLT (+2.7 %) slightly increased, facilities with flat-rate schemes managed to maintain their relatively low level of inefficiency (+0.0 %). Regarding the overall means across years, the inefficiency scores remained rather constant until 2007. This lateral movement was followed by a moderate upward trend that lifted the mean inefficiency score to 15.0 % in 2009. The pooled mean inefficiency of 14.2 % is somewhat lower than the estimates of two earlier studies (Steinmann et al. 2004; Farsi and Filippini 2006), while exceeding the results delivered by Farsi and Filippini (2008). These differences though are mostly explained by the set of variables that comprise the cost frontiers. For instance, Farsi and Filippini (2008), who reported a mean cost inefficiency of only 9.3 % (1998–2003), included the average LOS in the cost frontier. Therefore, inefficiencies that occur due to long stays are not captured by the inefficiency term.[25]

We run two robustness checks to verify our findings. In order to see whether the payment scheme effects are driven by the hospitals that were subject to a change in reimbursement, we perform a regression on a sub-sample of hospitals. To do so, we run specification (iii) again, using only those hospitals that showed no change in reimbursement over the study period ($N = 568$). The findings, however, indicate that the exclusion of the switching hospitals does not affect the sings and significance level of $FLAT$ and PLT (see Table 7 in Appendix 1). The mean inefficiency score decreases from 14.2 % to about 14.0 %, suggesting that the non-switching hospitals are slightly more efficient than the switching ones. The mean inefficiency estimates by payment scheme for the sub-sample can be found in Table 8 in Appendix 1.

As a second robustness check, we estimate a variable cost frontier. Some authors have argued that in the short term public hospitals cannot alter their stock of capital (Cowing and Holtmann 1983). Consequently, they propose estimating a variable cost frontier, while treating capital as a quasi-fixed input. In this case, total variable costs (Total costs—capital costs) serve as the left-hand variable, while the number of hospital beds ($BEDS$) is added to the frontier as a proxy for the stock of capital. While the signs and significance level of $FLAT$ and PLT do not change, the absolute size of the two coefficients decreases substantially (see Table 9 in Appendix 1). The lower values can be explained by the fact that variable cost frontiers give an impression of the level of *short-term* inefficiency (variable cost inefficiency), which is only one component of (total) cost inefficiency. According to the model estimates, the mean variable cost inefficiency amounts to 6.1 %.

[24] All differences in the pooled inefficiency scores (2004–2009) are significant at the 1 percent level.

[25] Using our data, the mean inefficiency decreases from 14.2 to 8.9 % if we consider the average LOS a cost frontier variable.

In summary, the two robustness checks mainly back our findings with regard to the effects of the different payment schemes. Since the absolute size of the payment scheme effects and the mean inefficiency estimates change only little within the restricted sample, we conclude that our results are not driven by changes in the payment policy of certain cantons.

Concluding remarks

By means of the SFA approach proposed by Battese and Coelli (1995), we have studied the impact of prospective payment schemes on cost-efficiency scores of acute care hospitals in a national setting. We used an unbalanced panel of 122 public and publicly-financed hospitals in Switzerland during the period 2004–2009. The four payment schemes that we analysed were per diem payments, two flat-rate schemes (DRG, DCB), and a mixed system (PLT). Controlling for hospital characteristics, local market conditions, and a time trend, we have shown that the hospitals which were reimbursed by flat payment schemes performed best in terms of cost efficiency. Furthermore, our results suggest that even hybrid schemes create incentives for cost containment, since PLT hospitals work more efficiently than per diem facilities. The difference in the estimated effect of the flat systems and the PLT scheme is not significant, though. In addition, we have demonstrated that SFA estimates are biased if heteroscedasticity in the one-sided inefficiency term (u_{it}) is not accounted for appropriately (Caudill and Ford 1993; Caudill et al. 1995; Hadri 1999).

Unlike previous studies, we were able to simultaneously analyse and isolate the efficiency effects of four different payment systems. Our findings are in line with the basic economic theory on the financial incentives of reimbursement schemes in healthcare (Ellis and McGuire 1986). In a prospective payment setting, a hospital can hardly influence the revenue side. There are therefore strong incentives to contain costs per case by reducing the treatment intensity and the LOS (Aas 1995). Our results indicate that even partially prospective reimbursement directly affects hospital behaviour. Consequently, to incentivise per diem hospitals, it may be sufficient to implement a moderate per case payment, while the per diem amount is adjusted accordingly to fit average costs per case.

Even though we have delivered new insights into the cost efficiency effects of inpatient payment schemes, there are several limitations to consider. First, we only focussed on public and publicly-financed hospitals. Along with the relatively small sample size, it is uncertain whether our findings apply to private hospitals as well. Moreover, as we used a Swiss dataset, it remains questionable whether our results are applicable to other healthcare systems. Second, as Swiss hospitals do not report reliable quality indicators, we cannot account for quality differences among them. Evidence suggests that ignoring quality differences may not be a serious problem (Mutter et al. 2008). However, if the quality of treatment is actually correlated with cost efficiency, we are likely to get biased inefficiency estimates. Further research may be able to close this gap by including data on different quality measures. Third, although we tried to control for observable and unobservable differences in local healthcare markets, there still might be unobservable heterogeneity among the 26 cantons that we could not account for. As a result, we may overestimate the efficiency effect of the payment scheme variables, as the estimated coefficients partially reflect these regional differences. Last but not least, the cost frontier approach gives rise to inevitable endogeneity problems. Apparently, inpatient and outpatient output is—at least partially—chosen by the hospital. Nevertheless, it seems plausible that hospital production is mostly demand-driven. There is, however, a more problematic endogeneity issue. As many authors before, we used hospital-specific measures as proxies for the prices of labour and capital. These measures, though, reflect the hospitals' choices about the average skill-mix of its employees and the amount and mix of capital equipment. As a consequence, our estimates may be biased and inconsistent. Unfortunately, recent literature does not provide satisfactory instrumental variables. With regard to instrumenting factor prices, Rosko and Mutter (2008) point out that in their study the predictive power of the proposed instruments was rather weak. According to Bound et al. (1995), poor instruments can even lead to worse results than accepting the bias due to endogenous covariates. In addition, the Hausman test we ran on the cost frontier cannot reject the hypothesis of exogeneity in input prices.

By 2012, the four payment systems that we analysed were replaced by a national DRG regime (SwissDRG). As a result, all the public and private hospitals in the 26 Swiss cantons are now paid according to a nationwide DRG catalogue. To maintain and further improve the high level of quality in inpatient care, a nationwide monitoring programme was established along with SwissDRG. As our findings indicate, the implementation of a prospective payment system may be a major step towards a more efficient inpatient sector in Switzerland. Future research will have to put greater emphasis on a dynamic perspective in order to provide a better understanding of how former non-DRG and APDRG hospitals respond to this new payment scheme.

Appendix A

Table 6 Regression of the average LOS on payment scheme variables

Variables	β	SE
CMI	0.935***	(0.046)
DRG	−0.252***	(0.032)
PLT	−0.217***	(0.043)
$COMPINS$	−0.382***	(0.118)
$YEAR$	−0.026***	(0.005)
$Constant$	1.254***	(0.085)
Type FE	Yes	
Region FE	Yes	
R^2	0.593	

$N = 606$; Cluster robust standard errors are given in parentheses; $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$; $COMPINS$ is the percentage of patients with supplementary insurance; Hospital types are: university hospital, general hospital (level 1, 2, and 3), geriatric hospital, and paediatric hospital

Table 7 Parameter estimates for the non-switching hospitals

$\ln TC$	HEC UV/FE	
	β	SE
Cost frontier		
$\ln CASES_{CMI-Adjusted}$	0.815***	(0.024)
$\ln OUTP$	0.064***	(0.015)
$\ln PK$	0.386***	(0.032)
$\ln WARDS$	0.002***	(0.000)
$\ln INTERN$	0.001*	(0.001)
RQ	0.048***	(0.007)
$UNIVERSITY$	0.074**	(0.032)
$SPECIALIST$	−0.086***	(0.030)
$YEAR$	−0.011**	(0.005)
Constant	−1.441***	(0.136)
Effects on inefficiency		
$FLAT$	−0.289**	(0.129)
PLT	−0.125**	(0.057)
HHI	0.495***	(0.162)
$PHYS$	−0.035	(0.043)
$MCARE\%$	0.882*	(0.450)
Constant	−0.200	(0.231)
Region fixed effects		Yes
Year fixed effects		Yes
$\sigma^2_{uit} = \exp(d_{it}\eta)$		Yes
$\sigma^2_{vit} = \exp(d_{it}\phi)$		Yes
$\gamma = \sigma^2_u/(\sigma^2_u + \sigma^2_v)$		0.876
Log-likelihood		369.095

$N = 568$; Cluster robust standard errors are given in parentheses; $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Appendix B

In order to show that the efficiency function $e(p)$ is concave, it is sufficient to prove that

$$\frac{\partial^2 e}{\partial p^2} < 0 \quad \forall\, p \in \{0, c\}\,.$$

Differentiating (3) with respect to p and defining $A = (c - p)t_{ee} + \gamma_{ee}$, one obtains

$$\frac{\partial^2 e}{\partial p^2} = \frac{\frac{\partial e}{\partial p}[A \cdot t_{ee} - t_e t_{eee}(c-p) - t_e\gamma_{eee}] + t_e t_{ee}}{A^2}.$$

Using (3) and rearranging yields

$$\frac{\partial^2 e}{\partial p^2} = \frac{2}{A^2} t_e t_{ee} - \frac{t_e^2}{A^3}[(c-p)t_{eee} + \gamma_{eee}],$$

which is strictly negative for all $p \le c$, given that $\gamma_{eee} \ge 0$ and provided that t_{eee} does not fall below a certain negative value, $t_{eee} \ge \frac{2At_{ee} - \gamma_{eee}t_e}{(c-p)t_e}$. Hence, under reasonable assumption with

Table 8 Mean inefficiency estimates by payment scheme and year

Year	per diem	PLT	FLAT	Overall	N
2004	0.168	0.134	0.116	0.135	102
2005	0.168	0.138	0.106	0.134	104
2006	0.188	0.154	0.106	0.143	86
2007	0.216	0.142	0.109	0.138	92
2008	0.204	0.157	0.110	0.145	92
2009	0.207	0.156	0.117	0.148	92
2004–2009	0.188	0.146	0.111	0.140	568

Table 9 Parameter estimates for the variable cost function

$\ln TVC$	HEC UV/FE β	SE
Cost frontier		
$\ln CASES_{CMI-Adjusted}$	0.135***	0.021
$\ln OUTP$	0.025***	0.009
$\ln PL$	0.400***	0.021
$\ln BEDS$	0.824***	0.026
$\ln WARDS$	0.002***	0.000
$\ln INTERN$	0.000	0.000
RQ	−0.008**	0.004
$UNIVERSITY$	0.138***	0.027
$SPECIALIST$	0.010	0.025
$YEAR$	0.004	0.005
Constant	0.713***	0.105
Effects on inefficiency		
$FLAT$	−0.063***	(0.020)
PLT	−0.054***	(0.019)
HHI	0.093***	(0.035)
$PHYS$	−0.041**	(0.016)
$MCARE\%$	−0.083	(0.105)
Constant	0.181***	(0.044)
Region fixed effects		yes
Year fixed effects		yes
$\sigma^2_{uit} = \exp(d_{it}\eta)$		yes
$\sigma^2_{vit} = \exp(d_{it}\phi)$		yes
$\gamma = \sigma^2_u/(\sigma^2_u + \sigma^2_v)$		0.669
Log-likelihood		643.064

$N = 606$; Cluster robust standard errors are given in parentheses; $*p < 0.10$, $**p < 0.05$, $***p < 0.01$; TVC : Total variable costs; $BEDS$: Number of hospital beds regard to the curvature of $t(e)$ and $\gamma(e)$, the model indicates that the loss of cost efficiency effort is concave in p. □

References

Aas, I. H. M. (1995). Incentives and financing methods. *Health Policy*, *34*, 205–220.

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

Anell, A. (2005). Swedish healthcare under pressure. *Health Economics*, *14*, 237–254.

Battese, G. E., & Coelli, T. J. (1993). A stochastic frontier production function incorporating a model for technical inefficiency effects. Working Papers in Econometrics and Applied Statistics 69 Department of Econometrics. Armidale: University of New England.

Battese, G. E., & Coelli, T. J. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, *20*, 325–332.

Belotti, F., Daidone, S., Ilardi, G., & Atella, V. (2012). Stochastic frontier analysis using Stata. CEIS Research Paper 251 Tor CEIS: Vergata University.

Biørn, E., Hagen, T. P., Iversen, T., & Magnussen, J. (2003). The effect of activity-based financing on hospital efficiency: A panel data analysis of DEA efficiency scores 1992–2000. *Health Care Management Science*, *6*, 271–283.

Borden, J. P. (1988). An assessment of the impact of diagnosis-related group DRG-based reimbursement on the technical efficiency of New Jersey hospitals using data envelopment analysis. *Journal of Accounting and Public Policy*, *7*, 77–96.

Bound, J., Jaeger, D. A., & Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogeneous explanatory variable is weak. *Journal of the American Statistical Association*, *90*, 443–450.

Caudill, S. B., & Ford, J. M. (1993). Biases in frontier estimation due to heteroscedasticity. *Economics Letters*, *41*, 17–20.

Caudill, S. B., Ford, J. M., & Gropper, D. M. (1995). Frontier estimation and firm-specific inefficiency measures in the presence of heteroscedasticity. *Journal of Business & Economic Statistics*, *13*, 105–111.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*, 429–444.

Chern, J.-Y., & Wan, T. (2000). The impact of the prospective payment system on the technical efficiency of hospitals. *Journal of Medical Systems*, *24*, 159–172.

Chirikos, T. N. (1998). Identifying efficiently and economically operated hospitals: The prospects and pitfalls of applying frontier regression techniques. *Journal of Health Politics, Policy and Law*, *23*, 879–904.

Coelli, T. J. (1996). *A guide to FRONTIER version 4.1: A computer program for stochastic frontier production and cost function estimation*. CEPA Working Papers 7 Department of Economics. Armidale: University of New England.

Coffey, R. M., & Louis, D. Z. (2001). Fünfzehn Jahre DRG-basierte Krankenhausvergütung in den USA. In M. Arnold, J. Klauber, & H. Schellschmidt (Eds.), *Krankenhaus-Report 2000* (pp. 33–47). Stuttgart, Germany: Schattauer.

Cowing, T. G., & Holtmann, A. G. (1983). Multiproduct short-run hospital cost functions: Empirical evidence and policy implications from cross-section data. *Southern Economic Journal*, *49*, 637–653.

Davis, K., Anderson, G., Renn, S. C., Rowland, D., Schramm, C. J., & Steinberg, E. (1985). Is cost containment working? *Health Affairs*, *4*, 81–94.

Dismuke, C., & Sena, V. (1999). Has DRG payment influenced the technical efficiency and productivity of diagnostic technologies in Portuguese public hospitals? An empirical analysis using parametric and non-parametric methods. *Health Care Management Science*, *2*, 107–116.

Dismuke, C., & Sena, V. (2001). Is there a trade-off between quality and productivity? The case of diagnostic technologies in Portugal. *Annals of Operations Research*, *107*, 101–116.

Ellis, R., & Miller, M. (2008). Provider payment methods and incentives. In K. Heggenhougen (Ed.), *International encyclopedia of public health* (pp. 395–402). San Diego, CA: Academic Press/Elsevier.

Ellis, R. P., & McGuire, T. G. (1986). Provider behavior under prospective reimbursement: Cost sharing and supply. *Journal of Health Economics*, *5*, 129–151.

Farrar, S., Yi, D., Sutton, M., Chalkley, M., Sussex, J., & Scott, A. (2009). Has payment by results affected the way that English hospitals provide care? Difference-in-differences analysis. *BMJ: British Medical Journal*, *339*, 1–8.

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

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

Folland, S. T., & Hofler, R. A. (2001). How reliable are hospital efficiency estimates? Exploiting the dual to homothetic production. *Health Economics*, *10*, 683–698.

FSO (2009). Calcul du cout de l'hospitalisation LAM al en 2009. http://www.bfs.admin.ch/bfs/portal/de/index/infothek/erhebungen_quellen/blank/blank/kg/02.Document.149170.xls.

FSO (2011). Kosten des Gesundheitswesens nach Leistungserbringern. http://www.bfs.admin.ch/bfs/portal/de/index/themen/14/05/blank/key/leistungserbringer.Document.21547.xls.

Gerdtham, U.-G., Löthgren, M., Tambour, M., & Rehnberg, C. (1999). Internal markets and health care efficiency: A multiple-output stochastic frontier analysis. *Health Economics*, *8*, 151–164.

Giammanco, M. D. (1999). The short-term response of hospitals to the introduction of the DRG based prospective payment system: Some evidence from Italy. *Giornale degli Economisti*, *58*, 27–62.

Grannemann, T. W., Brown, R. S., & Pauly, M. V. (1986). Estimating hospital costs: A multiple-output analysis. *Journal of Health Economics*, *5*, 107–127.

Hadri, K. (1999). Estimation of a doubly heteroscedastic stochastic frontier cost function. *Journal of Business & Economic Statistics*, *17*, 359–363.

Hadri, K., Guermat, C., & Whittaker, J. (2003). Estimation of technical inefficiency effects using panel data and doubly heteroscedastic stochastic production frontiers. *Empirical Economics*, *28*, 203–222.

Hagen, T. P., Veenstrab, M., & Stavem, K. (2009). Efficiency and patient satisfaction in Norwegian hospitals. HERO On line Working Paper Series 2006:2 Oslo University, Health Economics Research Programme.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, *46*, 1251–1271.

Helbing, C., Latta, V. B., & Keene, R. E. (1990). Hospital outpatient services under Medicare, 1987. *Health Care Financing Review*, *11*, 147–158.

Hensen, P., Beissert, S., Bruckner-Tuderman, L., Luger, T. A., Roeder, N., & Müller, M. L. (2007). Introduction of diagnosis-related groups in Germany: Evaluation of impact on in-patient care in a dermatological setting. *European Journal of Public Health*, *18*, 85–91.

Herr, A. (2008). Cost and technical efficiency of German hospitals: Does ownership matter? *Health Economics*, *17*, 1057–1071.

Herr, A., Schmitz, H., & Augurzky, B. (2011). Profit efficiency and ownership of German hospitals. *Health Economics*, *20*, 660–674.

Hsiao, W. C., Sapolsky, H. M., Dunn, D. L., & Weiner, S. L. (1986). Lessons of the New Jersey DRG payment system. *Health Affairs*, *5*, 32–45.

Jondrow, J., Lovell, C. K., Materov, I. S., & Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, *19*, 233–238.

Joskow, P. L. (1980). The effects of competition and regulation on hospital bed supply and the reservation quality of the hospital. *The Bell Journal of Economics*, *11*, 421–447.

Kahn, K. L., Rubenstein, L. V., Draper, D., Kosecoff, J., Rogers, W. H., Keeler, E. B., et al. (1990). The effects of the DRG-based prospective payment system on quality of care for hospitalized Medicare patients. *JAMA, The Journal of the American Medical Association*, *264*, 1953–1954.

Kobel, C., Thuilliez, J., Bellanger, M., & Pfeiffer, K.-P. (2011). DRG systems and similar patient classification systems in Europe. In R. Busse, A. Geissler, W. Quentin, & M. Wiley (Eds.), *Diagnosis-related groups in Europe: Moving towards transparency, efficiency and quality in hospitals European Observatory on Health Systems and Policies Series* (pp. 37–58). Maidenhill, UK: Open University Press.

Langenbrunner, J. C., Cashin, C., & O'Dougherty, S. (2009). Designing and implementing health care provider payment systems. How-to manuals. Number 13806 in World Bank Publications. Washington, DC: The World Bank.

Lave, J. R., & Frank, R. G. (1990). Hospital supply response to prospective payment as measured by length of stay. *Advances in Health Economics and Health Services Research*, *11*, 1–25.

Meeusen, W., & Van Den Broeck, J. (1977). Efficiency estimation from Cobb–Douglas production functions with composed error. *International Economic Review*, *18*, 435–444.

Moreno-Serra, R., & Wagstaff, A. (2010). System-wide impacts of hospital payment reforms: Evidence from Central and Eastern Europe and Central Asia. *Journal of Health Economics*, *29*, 585–602.

Mutter, R. L., Rosko, M. D., & Wong, H. S. (2008). Measuring hospital inefficiency: The effects of controlling for quality and patient burden of illness. *Health Services Research*, *43*, 1992–2013.

Newhouse, J. (1994). Frontier estimation: How useful a tool for health economics? *Journal of Health Economics*, *13*, 317–322.

OECD (2013). OECD health data: Health expenditure and financing. OECD Health Statistics (database). /content/data/data-00349-en. doi:10.1787/data-00349-en.

OECD, WHO. (2006). *OECD reviews of health systems*. Paris, France: OECD Publishing.
Rosko, M. D. (1984). The impact of prospective payment: A multi-dimensional analysis of new jersey's share program.Journal of Health Politics. *Policy and Law*, *9*, 81–101.
Rosko, M. D. (1999). Impact of internal and external environmental pressures on hospital inefficiency. *Health Care Management Science*, *2*, 63–74.
Rosko, M. D. (2001). Cost efficiency of us hospitals: A stochastic frontier approach. *Health Economics*, *10*, 539–551.
Rosko, M. D. (2004). Performance of U.S. teaching hospitals: A panel analysis of cost inefficiency. *Health Care Management Science*, *7*, 7–16.
Rosko, M. D., & Broyles, R. W. (1987). Short-term responses of hospitals to the DRG prospective pricing mechanism in New Jersey. *Medical Care*, *25*, 88–99.
Rosko, M. D., & Mutter, R. L. (2008). Stochastic frontier analysis of hospital inefficiency: A review of empirical issues and an assessment of robustness. *Medical Care Research and Review*, *65*, 131–166.
Rosko, M. D., & Mutter, R. L. (2010). Inefficiency differences between critical access hospitals and prospectively paid rural hospitals. *Journal of Health Politics, Policy and Law*, *35*, 95–126.
Schmidt, P., & Sickles, R. C. (1984). Production frontiers and panel data. *Journal of Business & Economic Statistics*, *2*, 367–374.
Sloan, F. (1991). Erfahrungen mit dem diagnosespezifischen Entgelt von Krankenhausleistungen in den USA: Das DRG-Experiment. In G. Neubauer & G. Sieben (Eds.), *Alternative Entgeltverfahren in der Krankenhausversorgung* (pp. 177–205). Gerlingen, Germany: Bleicher volume 24 of Beiträge zur Gesundheitsökonomie.
Smet, M. (2007). Measuring performance in the presence of stochastic demand for hospital services: An analysis of Belgian general care hospitals. *Journal of Productivity Analysis*, *27*, 13–29.
Sommersguter-Reichmann, M. (2000). The impact of the Austrian hospital financing reform on hospital productivity: Empirical evidence on efficiency and technology changes using a non-parametric input-based Malmquist approach. *Health Care Management Science*, *3*, 309–321.
Steinmann, L., Dittrich, G., Karmann, A., & Zweifel, P. (2004). Measuring and comparing the (in)efficiency of German and Swiss hospitals. *European Journal of Health Economics*, *5*, 216–226.
Street, A., O'Reilly, J., Ward, P., & Mason, A. (2011). DRG-based hospital payment and efficiency: Theory, evidence, and challenges. In R. Busse, A. Geissler, W. Quentin, & M. Wiley (Eds.), *Diagnosis-Related Groups in Europe: Moving towards transparency, efficiency and quality in hospitals European Observatory on Health Systems and Policies Series* (pp. 93–114). Maidenhill, UK: Open University Press.
Theurl, E., & Winner, H. (2007). The impact of hospital financing on the length of stay: Evidence from Austria. *Health Policy*, *82*, 375–389.
Widmer, M., & Weaver, F. (2011). *Der Einfluss von APDRG auf Aufenthaltsdauer und Rehospitalisierungen. Auswirkungen von Fallpauschalen in Schweizer Spitälern zwischen 2001 und 2008*. Obsan Bericht 49 Schweizerisches Gesundheitsobservatorium Neuchatel.
Widmer, P. K. (2011). *Does Prospective Payment Increase Hospital (In)Efficiency? Evidence from the Swiss Hospital Sector*. Working Paper Series 53 University of Zurich, Department of Economics.
Wranik, D. (2012). Healthcare policy tools as determinants of health-system efficiency: Evidence from the OECD. *Health Economics, Policy and Law*, *7*, 197–226.
Zuckerman, S., Hadley, J., & Iezzoni, L. (1994). Measuring hospital efficiency with frontier cost functions. *Journal of Health Economics*, *13*, 255–280.