Working Papers in Health Policy and Management

Volume 3

Andrew Street, David Scheller-Kreinsen,

Alexander Geissler and Reinhard Busse

Determinants of hospital costs and performance variation:

Methods, models and variables for the EuroDRG project

May 2010

Department of Health Care Management

Technische Universität Berlin

Andrew Street,[1] David Scheller-Kreinsen,[2] Alexander Geissler[2] and Reinhard Busse[3]

Determinants of hospital costs and performance variation: Methods, models and variables for the EuroDRG project

Working Papers in Health Policy and Management

Volume 3

May 2010

[1] Andrew Street is a professor of Health Economics and director of the Health Policy team in the Centre for Health Economics (CHE) at the University of York, UK.

[2] David Scheller-Kreinsen and Alexander Geissler are research fellows in the Department of Health Care Management at the Berlin University of Technology, Germany.

[3] Reinhard Busse is a professor and head of the Department of Health Care Management at the Berlin University of Technology, Germany.

EDITOR-IN-CHIEF / HERAUSGEBER DER SCHRIFTENREIHE
Prof. Dr. med. Reinhard Busse

EDITOR / REDAKTEUR
Matthew Gaskins

PUBLISHER AND DISTRIBUTOR / VERLAG UND VERTRIEB
Universitätsverlag der Technischen Universität Berlin
Universitätsbibliothek
Fasanenstr. 88 (im VOLKSWAGEN-Haus), D-10623 Berlin
Tel.: +49 30 314 76131; Fax.: +49 30 314 76133
e-mail: publikationen@ub.tu-berlin.de
http://www.univerlag.tu-berlin.de/

ISBN 978-3-7983-2124-3
ISSN 1869-6287

The EuroDRG project

Payment mechanisms represent one of the fundamental building blocks of any health system, introducing powerful incentives for actors in the system and important technical design challenges. Inpatient case payments, mainly referred to as diagnosis-related groups (DRGs), are used nowadays as a payment mechanism with ambitious aims: they seek to reimburse providers fairly for the work they undertake. Moreover, they intend to encourage efficient delivery and to discourage the provision of unnecessary services. Thereby are thought to overcome some of the drawbacks of more traditional hospital reimbursement systems. A case payment system that fulfils these hopes requires carefully balanced incentives as well as a methodologically sound system. Especially, DRGs need to accurately reflect the resources and costs of treating a given patient.

Fierce debate among practitioners, researchers and the public indicates that case payments still pose considerable technical and policy challenges, and many unresolved issues in their implementation remain. For example the HealthBASKET, project showed that DRG systems differ greatly between European Member States. One of the key conclusions of HealthBASKET was that structural components may play an even more important role than heterogeneity of treatment patters in cost variations within an episode of care. In this context, many European DRG systems may be heading in the wrong direction by concentrating almost exclusively on refining the medical classification of DRGs. In addition, over the last decade the Europeanization of health service markets generated pressure on national reimbursement systems. Increasing patient mobility caused growing health system interconnectedness. The EuroDRG project scrutinizes these challenges. Part one concentrates on the complexities of case payments for hospitals in national contexts. Special emphasis is put on identifying those factors, which are crucial for (1) calculating adequate case payments, (2) examining hospital efficiency within countries and across Europe fairly and (3) studying the relationship between costs and the quality of care provided in hospitals. The project uses comparative analyses of DRG systems across 10 European countries embedded in various types of health systems (Austria, Estonia, Finland, France, Germany, the Netherlands, Poland, Spain, Sweden, Ireland, UK). The second part of the project seeks to identify pan-European issues in hospital case payment by conducting cost analysis across European countries. Furthermore, the systemic factors, which are crucial for successful policy design in a slowly emerging pan-European hospital market, will be identified. Special emphasis is placed on (1) identifying ways to calculate these payments in an adequate fashion, (2) examining hospital efficiency within and across European countries, and (3) identifying factors that affect the relationship between the costs and quality of inpatient care.

Acknowledgements

The review presented in this working paper was generated in the context of the project "EuroDRG – Diagnosis-related groups in Europe: towards efficiency and quality" (www.eurodrg.eu), which is funded by the European Commission under the Seventh Framework Programme.

Contents

1 Introduction

The use of case payments intends to provide incentives for cost containment, efficiency and – at least in its theoretical formulation – to reward quality (Shleifer 1985). Although diagnosis-related groups (DRGs) have been used for years in many countries, hospital costs do not appear to have converged; indeed, rather large variations in hospital cost structures persist (Busse et al. 2008b). This empirical observation, various theoretical considerations, and recent policy initiatives to extend the scope of DRG systems to private sections of the hospital market, such as in England or in France, have led many to rethink the idea that a uniform payment for specific health care services can be applied across hospitals.

Instead, researchers are increasingly recognizing that differences in costs per DRG can be due to differences in the following categories of explanatory variables:

(1) patient mix (e.g. age, gender, socio-economic status, severity),

(2) diagnostics and treatment (both regarding the use/mix of new and emerging technologies as well as the intensity, e.g. physician or nursing time – “hospital/medical decision variables”),

(3) costs per unit of staff or technology (determined either by the hospital managers or being beyond their control – “structural variables”) and

(4) the factors that are included in the costs calculations (hence, the observed variations in costs would then be explained through the way costs are calculated, and what might be structurally left out).

The new paradigm in academia is therefore to establish a fair playing field (Mason et al. 2008b), which takes cost variation due to factors beyond the control of hospital managers into account. The underlying rationale is that the funding formula has to take account of “legitimate” cost heterogeneity (due to patient, treatment or structural variables) in order to avoid incentivizing unintended consequences in the hospital sector such as patient selection, lower quality of care, creaming, skimping and dumping (Dormont and Milcent 2004).

The purpose of this review is to examine the international evidence for existing instruments, systems or concepts for measuring and explaining differences in hospital costs. Costs here are defined as costs to the hospital to produce services rather than cost to purchasers ("price"), which is revenue from the point of the hospital.

The intention is to review but go beyond consideration of patient case-mix and "hospital/medical decision variables". The key question is to assess the (in)adequacyof existing DRG systems with regard to their (risk-)adjustment and to identify the structural or regional criteria that might need to be considered in analyzing hospital costs.

The review can be divided into two parts. The first will inform the analytical models applied in the analysis of patient-level cost data and in analyzing patient-level length of stay (as a proxy for cost) of the EuroDRG project. The objectives of these two approaches are the same, namely:

- To calculate costs (or length of stay) per case for each defined episode of care per country;
- To estimate cost (or length of stay) functions for each episode of care and joint functions including all care episodes utilizing the full set of explanatory variables across countries;
- To explore which adjustment factors explain variation in hospital costs (or length of stay) per country;
- To elaborate how far the DRG classification of each country is taking this empirically analysed variation in costs (or length of stay) into account, and adequately capturing the "legitimate" heterogeneity.

Empirical studies of variation in hospital costs fall into two camps: those based on analysis of the costs of individual patients and those – the vast majority – that analyse costs reported at hospital level. In this review, we consider how patient-level and hospital-level data are related and outline the approaches to analyzing these data. The second part of the review, which concentrates on hospital-level analysis, appears in the appendix to maintain a short and concise core text. This part of the review considers general specification choices as well as methods of efficiency analysis.

Define c_{ik} as the cost of treating patient $i, \quad i=1...I$, in hospital $k, \quad k=1...K$. It is straightforward to add together the costs for each individual patient in each hospital to arrive at the hospital's total cost:

$$TC_k = \sum_{i=1}^{I} c_{ik} \tag{1}$$

For those countries for which individual patient cost data are available, an analysis of variation in individual patient costs c_{ik} will be undertaken. We consider how to formulate models of patient costs in section 2.1.

In some countries, patient-level cost data are unavailable, but length of stay can be used as a proxy for costs, at least for those episodes of care that are delivered in the inpatient setting. The fact that the distributions of cost and length of stay differ means that we have to adopt slightly different estimation procedures, and these are outlined in section 2.2.

For completeness, in section 2.3 we also indicate how patient-level cost functions are related to hospital-level cost functions typically analysed in the empirical literature.

The primary question of interest is: why do hospital costs or lengths of stay vary among patients? Broadly speaking, costs or lengths of stay vary for two reasons:

- Firstly because patients have very different characteristics or are diagnosed and treated differently (the first two sets of variables named above)
- Secondly because of the characteristics of the hospital where their treatment is delivered (both within and beyond the control of hospital managers).

The research in the EuroDRG project is designed in a way to identify and disentangle these reasons. We discuss the type of patient characteristics and hospital characteristics that should be accounted for in section 3.

2 Conceptual and methodological framework

2.1 Analysis of patient-level costs

If patient-level cost data are available, analysis of why costs differ among patients can be undertaken by specifying a hierarchical or multi-level model, which recognizes that patients (level 1) are clustered within hospitals (level 2). The model might take the following general form:

$$c_{ik} = \alpha + \sum_{n=1}^{N} \beta_n \mathbf{x}_{ik} + \omega_k + \varepsilon_{ik} \qquad (2)$$

Where c_{ik} is the cost of patient *i* in hospital *k*, $\mathbf{x}_{ik}$ is a vector of *n* variables that describe patient *i* in hospital *k*, regarding their "patient variables", i.e. medical (diagnosis, severity, co-morbidities) and socio-demographic (age, gender, socio-economic status etc.) characteristics as well as "hospital/medical decision variables", i.e. diagnostics and treatment. These variables will be described at greater length in section 3.1. ω_k captures the hospital influence on costs over and above the patient characteristics while ε_{ik} is the standard disturbance.

Typically the distribution of patient costs is positively skewed, with some patients having much higher costs than the majority. So as to avoid these high cost patients exerted undue influence on the estimated relationships, we convert costs into logarithmic form, which serves to normalize the overall distribution of costs. Moreover, because there are only two levels to the hierarchy (patients clustered in hospitals), equation (1) can be estimated using standard fixed effects or random effects panel data models. Which model is most appropriate for the particular dataset being analysed can be established by a Hausman test (1978).

Whether estimated by applying a fixed or random effects model ω_k are termed "hospital effects" and can be interpreted as a measure of hospital performance, higher values implying that this hospital's costs are above average after taking into account the characteristics of the patients being treated (Hauck et al. 2003) and the choice and intensity of the diagnostics and treatment itself. These estimates ($\hat{\omega}_k$) can themselves be analysed to explore reasons why some hospitals appear to have higher costs than others. For instance, costs may vary according to the volume of activity if there are economies of scale or may be different in, say, teaching or specialist hospitals. We shall consider these characteristics in section 3.2 but, for the moment, we summarize them as a vector $\mathbf{z}_{km}$ of *m* variables measuring hospital characteristics (both under and beyond the control of hospital managers).

One way to account for hospital characteristics is to adopt a single-step process, using a model developed by Battese and Coelli (1995). This might take the following form:

$$c_{ik} = \alpha + \sum_{n=1}^{N} \beta_n \mathbf{x}_{ik} + \omega_k \left[\sum_{m=1}^{M} \gamma_m \mathbf{z}_k + \mu_k \right] + \varepsilon_{ik} \quad (3)$$

It is not always feasible to estimate this model, particularly the fixed effects specification is employed and some hospitals have unique characteristics. This makes it impossible to distinguish the influence of the characteristic from the hospital fixed effect.

An alternative is to specify an Estimated Dependent Variable model, as used in Laudicella et al. (2009), where $\hat{\omega}_k$ is regressed against the hospital characteristics:

$$\hat{\omega}_k = \alpha_o + \sum_{m=1}^{M} \gamma_m \mathbf{z}_k + \mu_k \quad (4)$$

This model forms the standard approach in the EuroDRGs project. For certain countries the Battese & Coelli approach may also be explored.

A major advantage of using patient-level data is that it is not necessary to use the legal entity (e.g. hospital) as the unit of analysis. Instead, data can be clustered to whatever level is deemed appropriate to the question. Most importantly, the use of patient-level data enables consideration of specific production lines, where there is likely to be greater standardization in the types of patients treated and activities undertaken. For instance, Laudicella et al. (2009) consider patients clustered within obstetrics departments and Kristensen et al. (2009) examine costs of hospital care for patients admitted with diabetes. The predecessor project to EuroDRG, HealthBASKET, used so-called vignettes, which in effect were patients with well-defined medical and socio-demographic characteristics, for whom the influence of differences in diagnostics/treatment as well as hospital and country characteristics were analyzed across nine countries (Busse et al. 2008a). Vignettes included hip replacement, stroke, acute myocardial infarction (AMI), child delivery, appendectomy and cataract surgery.

In the EuroDRG project, the intention is to analyse costs for patients classified to the same episode of care (building on the vignettes approach, but allowing variability in medical and socio-demographic variables). Define an episode of care as j, with $j = 1...J$ so that there are a total of *J* episodes of care dealt with in a hospital. This means that, instead of estimating equation (1) for ***all*** patients in each hospital, it is estimated separately for each episode of care of interest:

$$c_{ik} = \alpha + \sum_{n=1}^{N} \beta_n \mathbf{x}_{ik} + \omega_k + \varepsilon_{ik}, \quad \forall j \quad j = 1...J \tag{5}$$

This enables a better capture of the relevant diagnostics/treatment variables (e.g. type of hip implant, bare-metal stent vs. drug-eluting stent in AMI, open vs. laparoscopic appendectomy, normal delivery vs. C-section) in $\mathbf{x}_{ik}$ as well as structural variables (such as existence of catheter laboratory, stroke unit or neonatal intensive care unit) in ω_k.

2.2 Analysis of patient-level length of stay

For those countries where patient-level cost data are unavailable, length of stay will be analysed as a proxy for cost. The estimation procedure needs to be modified to reflect the fact that the distribution of costs differs from that of length of stay. In essence, length of stay cannot be considered a continuous variable, so the linear regression model should not be applied. Instead, it is better to use count data models that recognize that length of stay tends to take a limited set of quite low values with many patients having fairly short lengths of stay and relatively few staying for longer periods. Moreover, for some episodes of care, there might be a large proportion of patients who do not stay overnight (hence, being recorded as having a zero length of stay).

The main count data models are the Poisson regression model and the negative binomial regression model, with the latter more likely to provide a better fit to the length-of-stay data that we will be examining. For episodes of care with a high proportion of zero values, the zero-inflated negative binomial model might be appropriate. Good introductions to count data models are provided by Cameron and Trivedi (1986), Long (1997) and Jones (2000).

2.3 Analysis of hospital-level costs

As mentioned, most empirical studies of hospital costs take the hospital as the unit of analysis, mainly because data are reported at this level of aggregation. This is problematic, however, because a common production function cannot be assumed. Potentially this renders analysis invalid, because hospitals cannot be considered comparable.

When only TC_k is observed analysis of why some hospitals have different total costs than others involves specifying a cost function of the general form:

$$TC_k = \alpha + Q_k + \sum_{n=1}^{N} \beta_n \mathbf{X}_k + \sum_{m=1}^{M} \gamma_m \mathbf{z}_k + \varepsilon_k \tag{6}$$

Where Q_k measures the number of patients treated in hospital *k*. This number should be weighted in order to capture differences in (legitimate) costs; the most obvious choice for doing so – using DRG weights – has, however, the disadvantage of ex-ante accepting these weights as appropriate. There are the usual specification choices associated with estimating this function, such as whether to consider costs in natural units or in logarithmic form, and about the form in which explanatory variables appear. A review of these specification choices is included as section 2 of the Appendix.

Note that it is a simple matter to convert equation (5) from a total cost function to an average cost function simply by dividing total costs by total activity:

$$\frac{TC_k}{Q_k} = AC_k = \alpha + \sum_{n=1}^{N} \beta_n \mathbf{X}_k + \sum_{m=1}^{M} \gamma_m \mathbf{z}_k + \varepsilon_k \tag{7}$$

It is also possible to conduct efficiency analysis of hospital costs, using techniques such as stochastic frontier or data envelopment analysis (Jacobs et al. 2006). These techniques do not materially impact on the specification of the explanatory model, so we summarize these methods in Appendix 3.

In specifications (6 & 7) the vector $\mathbf{X}_k$ differs from the vector $\mathbf{x}_{ik}$ in the patient-level equation (1). Although the set of *n* variables are similar, rather being measured for each individual patient as in $\mathbf{x}_{ik}$, each variable in $\mathbf{X}_k$ is an aggregated summary of the patients in each hospital. For instance, rather than each patient's age, age is specified as some summary (e.g. the average) of all patients in the hospital. It is evident, therefore, that hospital-level analysis makes use of much less information about patients than is available in the patient-level analysis. Consequently, the results are less robust.

Note, however, that the vector of variables $\mathbf{z}_{km}$ capturing hospital characteristics is common to the two forms of analysis.

We now consider in more detail what variables are contained in these vectors.

3 Explanatory factors

3.1 Patient-level variables

The over-riding reason that costs differ among patients is because they suffer from very different conditions. Diagnosis-related groups (DRGs) are designed to classify patients into mutually exclusive groups of patients, with the patients in each group having the same expected "legitimate" cost (whether this is actually the case will clearly depend on both the number of different DRGs used as well as the accuracy of defining them and calculating their relative weights). Countries have developed their own classification systems but, for expositional purposes, define a DRG as j, with $j = 1...J$. If DRGs capture all the variation in costs that are due to patient characteristics, it would be possible to analyse variation in the costs of patients in different hospitals by specifying a simple model of the general form:

$$c_{ik} = \alpha + \boldsymbol{\beta}_j^{'}\mathbf{D}_{ik} + \omega_k + \varepsilon_{ik} \qquad (8)$$

Where $\mathbf{D}_{ik}$ is a vector of dummy variables specifying the DRG to which the patient is allocated, u_k is variation in cost that is attributable to the hospital in which the patient is treated and ε_{ik} captures random variation in patient costs.

DRGs, though, cannot account for all variation in costs among patients – there will be other characteristics of patients other than their DRG and the hospital where they are treated that explain their costs.

Some of these characteristics will be of general relevance, such as the patient's age, gender or socio-economic status. Hvenegaard et al. (2009) consider the additional explanatory power, over and above that offered by the DRG classification, in explaining variation in the costs of patients undergoing vascular surgery. In this study, such variables prove to add little additional information. This is consistent with the health insurance literature, where age and gender explain little of the individual risk of incurring health expenditure.

Other patient characteristics are more important, and are likely to vary according to the condition under consideration. For example Hvenegaard et al. (2009) consider a set of characteristics relevant to vascular patients including their smoking status, diabetes, and ASA score.

The HealthBASKET project has explored in depth the relevance of diagnostics and treatment (on the patient level) as well as certain structural variables: For hip replacement, the choice of cemented vs. non-cemented hip implants as well as the number of hospital beds were significant explanatory factors (Stargardt 2008). For stroke, receiving thrombolysis and length of stay were significant explanatory variables (Epstein et al. 2008). For acute myocardial infarction (AMI), these were the use of PTCA and stenting, length of stay, and the setting (i.e. urban/rural) of the hospital (Tiemann 2008). For normal child delivery, the variables length of stay and labour costs for nurses/midwives were significant (Bellanger and Or 2008). For appendectomy, open surgery (vs. laparoscopic surgery) was associated with significantly lower costs, whereas a larger number of hospital beds led to higher costs (Schreyögg 2008). For cataract surgery, the number of minutes used for the intervention had the highest explanatory power (Fattore and Torbica 2008). Similarly

Laudicella et al. (2009) – besides patient variables such as birth weight and whether or not the baby was still-born – consider a range of diagnostic and procedural variables in their analysis of the costs of patients admitted to obstetrics specialties, together with a set of hospital variables specific to obstetrics, including the number of babies delivered.

3.2 Hospital characteristics

The choice about what hospital characteristics, $\mathbf{z}_{km}$, to take into account depends on the purpose or perspective of the analysis. Broadly, the literature can be divided into two camps: technological functions derived from the theory of the firm and so-called behavioural functions designed to control for the exogenous constraints that prohibit cost minimization among hospitals.

A standard neo-classical framework involves explaining costs in relation to the level of production (e.g. number of patients treated) and the price of different inputs – notably labour, equipment and capital. If the research interest is in analysing costs from the perspective of hospitals themselves, this framework might be appropriate. The explanatory variables measure the input choices made by – or endogenous to - the hospitals themselves.

Early studies of hospital costs specified a behavioural function in a somewhat *ad hoc* manner, including “fairly indiscriminately all variables for which a causal relationship to hospital costs is hypothesized and data are available” (Breyer 1987). But recently more thought has gone into the specification, in recognition that, in most public health systems, some hospitals are required to operate in more adverse circumstances than others. These circumstances are often termed “environmental factors” and make achievement of a given level of cost more difficult. In effect these factors constrain the hospital from operating on the technically feasible minimum cost frontier.

The analytical objective is to control for these constraining factors in comparing costs between hospitals. This involves identifying a set of variables that reflects feasible production possibilities within a constrained environment (Smith 1981). There is often considerable debate as to what factors are considered constraints. In the short run, many factors are outside the control of hospitals. In the longer term a broader set of factors is potentially under hospital control, but the extent and nature of this control will vary depending on the nature of the context.

In what follows, we consider in more detail the type of variables that might be incorporated in the vector $\mathbf{z}_{km}$.

3.2.1 Teaching status

Teaching hospitals are often thought to have higher costs than non-teaching hospitals. There have been numerous studies devoted to estimating the cost differential and almost all econometric analyses of hospital costs include teaching status as a dummy variable. We do not attempt an exhaustive summary of this empirical research, except to say that the literature rarely explains why costs might differ in these two types of institution. There are two main explanations of differential costs.

The first reason is not to do with the teaching function of the hospital *per se*, but to perceived inadequacies in DRG system. Teaching hospitals are thought to treat patients of a greater severity than non-teaching hospitals. In theory, the DRG system should be able to account for these differences. In practice, the classification may be imperfect. If teaching hospitals systematically attract more severe patients within each DRG, their costs will be higher. So the "teaching status" variable is simply picking up additional patient heterogeneity, rather than it having anything to do with the teaching function itself.

The second reason is due to how teaching is funded. As part of their training medical students are required to accompany consultants around the hospital in order to observe and treat patients. Thus the provision of teaching may introduce delays to the treatment process if, say, consultants spend longer reviewing each patient prior to surgery or prior to discharge so that the medical students can learn from the review process. In essence, the provision of patient care and medical education is a joint and complementary process and it is not straightforward to disentangle these two components.

If funding fails to determine accurately the costs of teaching, the error will be captured by the “teaching status” variable in the analysis of patient-related hospital costs. In contexts where medical education is underfunded relative to patient care, teaching hospitals will have higher costs than non-teaching hospitals. In other contexts, funding for medical education might be more generous, to the extent that it subsidizes the provision of patient care. In such circumstances, costs of treatment in teaching hospitals will appear lower than in non-teaching hospitals.

3.2.2 Ownership

The famous question of whether “there is a dime’s worth of difference” (Sloan et al. 2001) between the performance of for-profit and not-for-profit or public and private hospitals has yet to be solved. Many theoretical and empirical investigations have examined differences with regard to organizational goals, behaviour and outcomes. Econometric analyses tend to include ownership as a dummy variable either being specified as private versus public or as for-profit versus not-for-profit (in the following we will refer only to for-profit and not-for-profit differences as the arguments for both taxonomies are very similar). As for teaching status, we do not attempt to summarize the findings of all existing studies, but rather focus on the implicit conceptual underpinnings, which are rarely spelled out explicitly.

There are two main predictions with regard to cost differences between for-profit and not-for-profit. The first argues that for-profit hospitals operate at lower costs and are more efficient than not-for-profit ones. The second argues the opposite: not-for-profit hospitals operate at lower cost per case.

Three main reasons are put forward to explain why for-profit hospitals should operate at lower costs. (1) The property rights model suggests that for-profit hospitals produce at lower costs as well-established control rights provide owners with a higher propensity to invest in cost-cutting innovations since they stand to gain personally from the resulting profits. (2) It is argued that in public and not-for-profit hospitals the lack of monitoring by private owners and capital markets means that managers are less tightly controlled and therefore have more opportunities to shirk. (3) Related but slightly different is the argument that quality and institutional reputation in the community rather than profit maximization are predominant principles for evaluating managers in the public sector and these determine managers' career advancement. As a consequence managers in public and not-for-profit hospitals face strong incentives to invest time and resources to maximize reputation enhancing technologies or representative infrastructure rather than focusing on profits or efficiency.

The second prediction, based on a number of empirical observations including systematic reviews and meta-analyses (Hollingsworth 2003; Vaillancourt 2003) , is that "public rather than private provision of health care seems more efficient for hospitals" (Hollingsworth 2008). Building on a mixture of theoretical and empirical findings, this strand of the literature argues that the finding of "higher profits" among for-profit hospitals are neither due to true costs nor due to ownership structure per se. Rather they are due to not-for-profit hospitals systematically attracting more complex or severe patients and of the inadequacy of DRGs to reflect actual costs and greater (unobserved) severity of these patients. In such circumstances, costs of treatment in not-for-profit hospitals will appear higher than in for-profit hospitals. Accounting for these imperfections reveals that public hospitals operate at least as efficiently as they have no ability or incentives to cherry-pick less complex patients.

Contract failure models add that as the quality and adequacy health care services are hard to determine for patients (Hansmann 1980) as health care markets are characterized by information asymmetry in favour of suppliers, i.e. hospitals. In this framework, for-profit hospitals face incentives to increase profit at the cost of lower quality of care, while not-for-profit hospitals, being characterized by a nondistribution constraint, have no incentive to lower quality in favour of profits (Wörz 2008: 37-38). As a consequence this set of models predict that revenue and profits are higher in for-profit hospitals while efficiency (if taking account of quality when measuring costs or ideally being measured in terms of health outcomes) will be lower. There is empirical evidence from a meta-analysis supporting this view (Devereaux et al. 2002).

Institutional approaches further stress that hospitals with different ownership regimes may have different costs because they are subject to different taxation regimes, may have different access to capital, may have differential influence on labour costs, their location of production, or the amount of (inadequately compensated) emergency care provided (Mason et al. 2008a). Whether these circumstances mean that for-profit or not-for-profit hospitals produce at higher cost then depends on the context specific configurations of the regulatory framework.

In contexts where patient and organizational variables fail to adequately reflect differences in factors constraining the behaviour of for-profit and not-for profit hospitals, the error will be captured by the dummy variable “ownership status”.

If for-profit hospital face a less constraining environment or benefit from more effective incentive structures relative to not-for-profit hospitals, for-profit hospitals will have lower costs than not-for-profit hospitals. If not-for-profit hospitals face a more favourable regulatory environment (e.g. generous tax breaks) relative to for-profit hospitals, not-for-profit hospitals will have lower costs than for-profits.

3.2.3 Volume of activity: economies of scale

Larger hospitals or units are often thought to produce at lower costs than smaller hospitals or units. The basic argument is that larger hospitals or units benefit from economies of scale, experiencing decreasing costs per output as volume increases. Larger hospitals or units might produce at lower cost due to the following reasons: specialized equipment often comes in units of some minimum size, so that larger output makes usage more efficient. Benefits from division of labour may arise because larger workforce facilitates restricting the range of services performed by each individual worker allowing for standardization. Also a larger workforce and technological infrastructure may allow compensating more easily for illnesses or departure of workers or technological problems.

Beyond a particular size, diseconomies of scale may arise, perhaps because a larger hospital or unit size leads to overly high expenditures for overhead, bureaucratic forms of organization, complex interdependencies implying problems of coordination and cooperation.

There have been numerous studies devoted to these issues, e.g. (Posnett 2002) and almost all econometric analyses of hospital costs include some measure of size as a dummy variable. To scrutinize the effect to different hospital sizes one can include the number of patients (e.g. in thousands) treated in hospital or department *j* in the model. These outlined arguments may also be consistently formulated to justify a variable that captures the volume of a specific episode of care. Hence one should add a variable that measures the number of patients falling in the definition of the respective episode of care treated in the hospital or department *j* measuring volume by episode of care.

If one of the respective variables is positive, this suggests that average costs after controlling for patient characteristics are higher in large departments or hospitals or in departments producing more of the services relevant to the respective episode of care. Additionally, a measure for the number of whole time equivalent general or specialist physician per 100 patients can be incorporated if reported. Different job and staff categories can be weighted according to their respective wages.

Whether size should be considered truly environmental factors depends to on the regulatory context. In some countries as in Germany hospital capacity is negotiated between state-level regulators and hospital management and then fixed for a certain period of time. Hence, in the short-run capacity, is exogenous, while in the long run it might be adjusted.

3.2.4 Range of activity: specialization versus economies of scope

The range of activity offered by the hospital may also influence costs, but it is difficult to predict the direction of influence. On the one hand, if there are economies of scope, increasing the range of activity leads to lower costs. Economies of scope refer to a situation in which the joint production of two or more products can be achieved at lower costs than the combined cost of producing each product individually. General hospitals offering a wide range of services may produce at lower costs in circumstances in which the production of different but related outputs jointly is cheaper than the production of outputs separately (Panzar and Willig 1981). These cost advantages can be generated through shared use of resources such as technology, workers, or general overhead such as spreading fix costs of operating rooms or intensive care units over multiple different but related operations. So, for example, it might be less expensive to situate the accident and emergency department, fracture clinic, and trauma and orthopaedic wards in close proximity rather than having them in different locations, as this means that equipment be shared and staff in different departments can work more effectively together particularly when patients need to transferred between each setting.

On the other hand, it is often argued that hospitals which specialize have lower costs than general hospitals (Dranove 1987). One reason for specialization is to protect or ring-fence resources for specific purposes, when otherwise there would be competing claims on their use (Harris 1977; Kjekshus and Hagen 2005; Street et al. 2009). Another argument is that specialization allows expertise to develop and flourish. Specialist hospitals may be better at evaluating practice over a more limited range of activities and clinical outcomes have been shown improve as doctors undertake more of the same procedure.

To analyse the effect of different range of services the number of specialties offered in a hospital can be used as a proxy. If positive, this variable would suggest that average costs after controlling for patient characteristics are higher in hospitals with more specialities. Another approach is to define a specialization index, capturing the distribution of activity across DRGs (Daidone and D'Amico 2009). In a similar vein, Laudicella et al. (2009: 8) include variables that measure the amount of neonatal and antenatal care as a proportion of all activity in the obstetrics department.

3.2.5 Quality

One reason that costs may differ among hospitals is that their quality differs. The relationship between quality and costs, however, is unclear. As outlined in the section on ownership, some models envisage that higher quality implies higher resource consumption, while other models suggest that higher quality can lead to more efficient delivery of hospital care (Wörz 2008). Moreover, the relationship between quality and ownership is a frequent issue (Currie et al. 2003; Devereaux et al. 2002; Vaillancourt 2003). However, the evidence is inconclusive with regard the conduciveness of different forms of ownership on quality.

For the analyses in the EuroDRG project it is moreover important to consider that quality is multi-dimensional and difficult to measure as it is difficult to establish both the direction and degree of influence of quality on costs.

Preferably measures of quality will be available for each patient, in which case influence can be established by including the measures in the $\mathbf{x}_{ik}$ vector of variables explaining patient costs in equation (3). Various measures of quality can be considered including waiting times, in-hospital or 30-day mortality, infections, readmissions, and changes in health status.

In the absence of patient-level measures of quality, some studies describe the process of care delivery or structure of the department or hospital in which care is delivered. These measures might describe, say, whether clinical guidelines have been adopted, the staffing complement, or the availability of particular types of equipment (see section 3.2.6). Whatever the measure, there is an assumption that investment in process or structural dimensions of quality is positively correlated with the quality of care delivered to individual patients.

3.2.6 Technological equipment

The relationship between hospital costs and technological equipment has been on the research agenda for several decades. Many of the investigations were and some are still inspired by the idea of non-price competition among hospitals (Joskow 1983; Luft et al. 1986). The latter argue that, as in most western health systems patients are not sensitive to hospital prices (since they are insured) and the quality of care delivered tends to be hard to judge, hospitals can feel motivated to (over-) invest in “high technology” as a competitive strategy to attract (profitable) patients. This may result in a “medical arms race”, which can cause higher hospital expenditure and may increase costs per case.

At the same time it has been stressed that technology investments in hospitals may encompass product, process and organizational innovation that can be hypothesized to influence hospitals in very different ways (Zweifel and Breyer 1997). Technological equipment may increase total costs per case but may also decrease total costs per case. Similarly process and organizational innovation can be hypothesized to increase or decrease costs per case. Given the theoretical background it seems to be hard to justify any general statement about the expected relationship between hospital costs and technology or technical equipment.

The multidimensional character of technology is one of the reasons why general technological indices - in its simplest form the sum of the number of services each hospital offers from a list of possible services - are criticized as they are not able to account for the heterogeneity of technology (Spetz and Maiuro 2004). Moreover, an underlying hypothesis suggesting a specific relation between technology on the one side and hospital costs on the other side can hardly be specified, which undermines well-grounded modelling.

A common alternative is the use of specific technology indices, which are modelled from a vector of known efficacious technologies at the level of the hospital (Pitterle et al. 1994; Prince 1998). However, also the use of this category of explanatory variables is problematic when used to explain cost variation between hospitals. The main reason is that technologies or technological equipment to be included are commonly selected in a more or less arbitrary way, while at the same time empirical studies confirms that variable omission and variable selection with regard to hospital technology substantially affect the results of parametric hospital cost functions (Cremieux and Ouellette 2001).

The episode of care approach seems well suited to avoid the main drawbacks of the outlined approaches. Firstly, as the costs of specific indications or are to be explained, the range of technologies and technological equipment to be considered is limited and decisions for including certain types of technical equipment can be more readily justified by reference to the treatment patterns of the episode at hand. The arbitrariness of the variable selection seems hence less problematic – though it remains a matter of degree. Moreover, practical experience within hospitals and episode specific costing studies may be used to formulate well-grounded hypotheses about the relationship between specific types of technical equipment and hospital costs. This makes it easier to justify the inclusion of specific variables approximating technological equipment during the modelling process.

To scrutinize the effect of diverging sets of technological equipment on hospital costs, dummy variables for the existence of e.g. a catheter laboratory, a stroke unit or a neonatal intensive care unit could be included in the model.

3.2.7 Geographical variation in input costs

Other factors that make the achievement of a given level of cost or production more difficult are geographical variations in input costs. So far the empirical evidence about the significance of differences in input costs is limited and only few DRG systems attempt to adjust reimbursement for geographical variation. Policy examples are the US Medicare wage index, which adjusts for differences in hospital wage levels by a factor reflecting the relative hospital wage level in the geographic area of the hospital compared to the US-wide average hospital wage level (Centre for Medicare and Medicaid Services 2007) and the English “Payment by Results” (PbR) where public hospitals receive additional payments to account for geographical variations in the costs of land, buildings and staff through the so-called Market Forces Factor (MFF) (Mason et al. 2008b).

The reason for accounting for input cost differentials across hospitals when comparing cost per episode of care is that most public hospitals cannot locate where they wish and therefore need to pay local factor prices rather than relocating to areas with lower costs. As a result, hospitals face locational constraints, which are outside their control and affect their production costs (Mason et al. 2008b)

To account for differences in wage levels beyond the control of hospitals one can adjust labour costs of each hospital by a scaling factor that represents variation in total (or average) costs due local labour market constrains. One way to do this is to divide the average compensation per hour in the hospital’s labour market (local, regional or state) by the average compensation per hour in the country. Contingent upon data availability the former can be performed for various job categories and summarized in a weighted index. The raw adjustor can then be multiplied by the hospital’s percentage of total costs attributable to salaries to determine final adjustors. Ultimately final adjustors with scores of less than 1.0 indicate low cost labour markets and those greater than 1.0 indicate high cost labour markets. For example, a hospital with an adjustor of 1.054 is expected to have total costs 5.4%

above the country average due to higher labour costs in its labour market. Similarly, a hospital with an adjustor of 0.98 is expected to experience 2% lower total costs due to the local labour market environment. Each hospital's costs per episode of care can then be adjusted accordingly. The respective variable in the regression captures differences in the labour prices faced by hospitals in different regions which are beyond their control.

Using similar approach, differences in capital and building costs can be captured across regions and countries provided that data requirements can be met.

3.2.8 Geographical location

It is common in many empirical analyses to include dummy variables that identify the location of the hospital. These might identify the geographical area or indicate whether the hospital is located in an urban or rural setting. The interpretation of the variables identifying geographical area will, of course, be country-specific.

In some contexts rural hospitals might have higher costs than urban hospitals. For example, when DRG funding was first introduced in the Australian state of Victoria rural hospitals received additional payments (Duckett 1994). There were two reasons for this. First, they were thought to face above ambulance and patient transport costs simply because they were transporting patients over longer distances. Second, because they were serving small communities rural hospitals were less likely to treat sufficient number of patients to achieve economies of scale. The payment, then, was justified to ensure access to health services for isolated communities.

In other contexts urban hospitals might appear to have higher costs than their rural counterparts. This might be because either they attract patients of above-average complexity and this is inadequately reflected by the DRG system or they face higher input prices which have not been adequately accounted for (see preceding section). It is difficult, though, to see why urban status *per se* drives higher costs. Rather the variable is likely to capture measurement error in other variables, particularly patient case-mix, hospital size and scope, and geographical variation in input prices.

Appendix

1 Examining hospital performance – the nature of the problem

It is often asserted that health care organizations face limited competitive pressures that would otherwise encourage them to innovate and adopt cost minimizing behaviour. There are various grounds for making this assertion. For instance, competitive behaviour may be less in evidence if health care is publicly funded/provided or in situations where health care organizations enjoy a geographical or specialist monopoly of supply. If competitive pressure is weak, there may be scope for better utilization of resources. At the same time common performance measures such as return on investment (ROI) or profitability seem inappropriate due to the fact that health care entities are often public or non-for profit and therefore do not have to generate profits or returns to investments.

Different classes of models are used by health economists to derive insights about hospital performance. Firstly, researchers use models, which attempt to explicitly model the production process in hospitals and compare (average or total) cost across hospitals. This group of models builds mainly on neoclassical microeconomic theory and produced rich and diverse sets of studies. Most popular approaches include hospital cost function (HCF) analysis operationalized via ordinary least squares (OLS) or corrected ordinary least squares (COLS) regression techniques. These were the main approaches during the 1980s and 1990s to derive insights about the hospital costs, scale and scope, as well as productivity and efficiency. Over the last decade the attention of most researchers shifted to performance analysis. Performance analyses aim to identify which organizations are doing better than others in either their overall operation or in specific areas of operation. This information may be used to stimulate better use of resources, either by encouraging organizations to act of their own volition or through the imposition of tailored incentives by a regulatory authority or funding agency.

Here the concept of efficiency is applied to consider the relative performance of organizations engaged in production, converting inputs into outputs. The fundamental building block of efficiency analysis is the ***production function***. The production function models the maximum output an organization could secure, given its level and mix of inputs. In very simple terms, the production process can be pictured as in the diagram below. The organization employs inputs (labour, capital, equipment, etc) and converts these into some sort of output. The middle box, where this production process takes place, is critical to whether some organizations are better than others at converting inputs into output.

Labour, intermediate and capital inputs → Organization of the production process → Outputs

The middle box is actually something of a ´black box´ because it is usually very difficult for outsiders to observe what goes on in the organization and how its production process is organized. Indeed, in some industries (such as the pharmaceutical sector) the production process might be a closely guarded secret and the source of competitive advantage.

This inability to observe the production process directly is a fundamental challenge for those seeking to analyse efficiency. Nevertheless, it is possible to think of a gold standard production process that describes the best possible way of organizing production, given the prevailing technology. This gold standard is termed the "production frontier", where the amount and combination of inputs is optimal. Any other scale of operation or input mix would secure a lower ratio of output to input. Organizations that have adopted this gold standard are efficient – they are operating at the frontier of the prevailing technological process. But organizations might be operating some way short of this gold standard: equipment might be outmoded, staff may be given to shirking, capital resources might stand idle periodically. These, and multiple other reasons, might explain inefficiency. To assess whether there is better scope for utilization of resources, insight can be gained by comparing organizations involved in similar activities. Rather than attempting to prise open the "black box", such comparative analysis concentrates on the extremes depicted in the diagram. Information about what goes in (inputs to the production process) and what comes out (outputs of the production process) allows comparison of input-output combinations of organizations that produce similar things. If an organization uses less input to produce one unit of output than another organization, the former is more productive. If we want to assess organizations that produce different amounts of output, we need to make judgements about whether there are economies of scale which, in turn, relies on understanding the gold standard production process. Organizations can then be judged in terms of their relative efficiency.

2 Specification choices (in parametric analysis of hospital costs and performance)

We concentrate on specifications choices to be made in the context of the parametric approaches such as hospital cost functions and the Stochastic Frontier Analysis (SFA) as these will be used in the EuroDRG project. Similar choices also have to be made in the context of the Data Envelopment Analysis (DEA). However, as the EuroDRG project will not apply these we only briefly introduce the DEA in very general terms in section 3 of the appendix and for the complexities of specification choices refer to the excellent introduction in Jacobs et al. (2006: 91ff).

2.1 Transformation of variables

One of the first choices researchers face is the form in which they incorporate variables in their model. Using variables in their natural units implies the assumption that the explanatory variables are related linearly to the dependent variable (Jacobs et al. 2006: 43). This approach assumes a constant rate of change in costs as the scale of activity changes. This assumption may not hold in practice as decreasing returns to scale or increasing returns to scale or scope can imply variable marginal cost. Taking logarithms of the variables in the model represents a way to model non-linear relationships. In this case, the interpretations of coefficients changes from natural units to percentage changes or elasticities. Other functional forms such as including higher–order powers of explanatory variables can also be applied. Their appropriateness can be scrutinized by applying a t-test.

2.2 Total vs. average cost function

The choice between estimating average and total cost functions depends to a large extent on the purpose and focus of the analysis. Conceptually economic models suggest that organizations can change their output or cost level in two ways: a) by making proportionate changes to the amount of all inputs used in the production process and b) by changing the composition of the input mix.

The effect of changing input composition depends on the scale properties of the function (also called the degree of homogeneity). If the function is homogeneous of degree 1, i.e. linearly homogenous, a proportionate change in use of input leads to a change of output by the same proportion. Under this condition of constant returns to scale, an average and a total cost function yield equivalent results (Jacobs et al. 2006: 45).

From a statistical point of view estimating an average function is preferable if the data being analysed are subject to heteroscedasticity, i.e. if the variance of the residual is not constant. The latter can be reduced by using an average cost function for the model or by estimating the ratio of output or cost to a deflating variable, where the residual is more likely to display homoscedasticity (Intriligator 1978). The disadvantage of using the ratio of output or cost to deflating variable is that there no statistical criteria to guide the selection of the deflating variable. One option is to making the dependent variable an average cost per unit of size. However, this approach is only appropriate in the case of constant returns to scale. Alternatively, one can deflate on an output – but in the context of multiple outputs the choice of the output for deflation substantially affects the estimates.

Before deflating a total cost function one should test for heteroscedasticity because the error term is transformed by deflation (Kuh and Meyer 1955). In addition, new bias may be entered by deflation if the variable used for deflation is multicollinear with one of the explanatory variables (Kuh and Meyer 1955). Total cost functions have the advantage of incorporating various types of output. Other approaches to deal with heteroscedasticity than estimating an average cost function in this context are estimating a logarithmic function (Maddala 1988) or using a robust estimator (White 1980).

2.3 Functional form

Various authors showed that the choice of the functional form of the production or cost function substantially affects the results of hospital cost or performance analysis (Folland and Hofler 2001; Rosko and Mutter 2008; Smet 2002; Street 2003). The two most common functional forms in parametric analyses are ad-hoc specifications and specifications, which in principal are informed by neo-classical propositions but accept assumptions to very different degrees. We will briefly outline the main features and implications of these in the following paragraphs.

(1) The most basic functional form are ad-hoc specifications, which intend to explain variations in costs per unit of output among hospitals. As indicators for output these early studies used mainly cost per case or cost per patient and as explanatory variables they included “fairly indiscriminately all variables for which a causal relationship to hospital costs is hypothesized and data are available” (Breyer 1987). The underlying explanatory model assumes that the cost determinants are linear and additively separable and are specified in an additive-linear functional form. These assumptions may not hold in practice. The additive-linear model for example would imply that an additional bed in the hospital raises costs by a fixed amount – independent of other factors such as hospital’s utilization, the spectrum of illnesses treatment or wage levels in the hospital.

The main advantage of this rather bold approach is that it can deal with the multi-product nature of hospitals, while keeping the number of variables manageable (Smet 2002). However, this strength comes at a cost – non-separable relationships between patient, organizational and or environmental variables cannot be accounted for and therefore estimates suffer from limited explanatory scope and power. While this form of hospital cost functions had considerable impact on public policy in the US and in the UK during the 1980s (Street 2003), most researchers recently concentrate on more flexible models.

(2) The second group – approaches following neoclassical proposition – enable researchers to investigate the variance of output following changes in the level or mix of inputs. Independent variables in this framework are inputs under the control of hospital managers. Within this group different approaches can be identified which accept various sets of restrictions with different implications. One can distinguish between (2a) closed functional form, (2b) flexible functional form and (2c) hybrid flexible functional form cost functions.

(2a) Closed functional form cost functions accept a large set of common but rather restrictive assumption used in neo-classical economics (Conrad and Strauss 1983; Cowing et al. 1983). These models, which we called "technological" cost functions in the first part of the review, build on the insights of duality theory (Shepard 1970) suggesting a one-to-one correspondence between production und cost function (Breyer 1987) and also implying a cost function which is non-decreasing and homogeneous of degree one in factor prices (Caves et al. 1980). A common form is the well-known Cobb-Douglas logarithmic form:

$$Y = F(L, K) = AL^{\alpha}K^{\beta} \qquad /\ln \rightarrow \quad lnY_i = A + \alpha lnL_i + \beta lnK_i + \varepsilon_i \qquad (1)$$

where:

Y *= total output*

A *= technology*

L *= labour*

K *= capital*

α *and* β *positive parameters*

The logarithmic form allows these parameters to be interpreted as elasticities. This approach entails relatively high risks of misspecification, since researchers often feel that they can hardly specify the true cost function. Moreover, the strict version of this approach requires that only factor input prices and quantities are incorporated as explanatory variables. The influence of structural variables such as market characteristics, and patient characteristics cannot be scrutinized. Hence, the estimates may be biased given that the underlying model disregards exogenous factors (Breyer 1987). Consequently, this group of hospital cost functions is hardly able to derive meaningful insight about potential "legitimate" hospital cost variation.

(2b) To circumvent these problems and to avoid misspecification, flexible functional forms, which try to approximate the true functional form using techniques such as local second-order Taylor approximation, were developed (Smet 2002). Thereby the main problems of both the ad-hoc and the strict neoclassical approaches are circumvented, i.e. the deterministic approach towards the relationship between the variables in the model. Most common variants of flexible functional forms are the quadratic, the transcendental (translog) and the generalized transcendental functional form. However, these models face other difficulties: due to their technical form they can no longer handle large number of outputs such as different cases or patients categories (Smet 2002). Moreover, the limited range of outputs that cost functions in this group can handle generates additional drawbacks. Calculating economies of scope and scale, i.e. deriving inputs about decreasing marginal cost as output levels increase or returns from specialization require that the function is scrutinized at the level of (multiple) outputs and not in the neighbourhood of approximations. Moreover, to estimate economies of scope and scale information about incremental costs is required. Since flexible functional form specifications need to reduce the number of outputs in the model, they often operate on the basis of some aggregate measure of output such as average or total costs. Since average and total cost functions fail to provide information about incremental costs and costs at zero output level respectively, their use precludes analyzing issues such as economies of scale and scope.

(2c) Hybrid functional form hospital cost functions allow relaxing some neoclassical assumptions, while still operating within the neoclassical framework. For example using hybrid functional forms allows supplementing factors of production as explanatory variables by other variables representing organizational constrains. Hence, the impact of other exogenous factors on performance and cost variation can be scrutinized, while the basic assumption of homogeneity in factor prices is maintained. These variables are not interpreted as contributing to the cost minimum, but rather explain deviations of observed costs from cost minimization.

Overall, we face an ambiguous trade-off: either to prioritize flexibility of the functional form and less restrictive assumptions with regard to nature of organizational and market characteristics or to focus on the ability of the model to account for the multi-product nature of hospitals. Increasingly, ad hoc specifications and closed functional forms are being considered of limited value due to restrictive assumptions and over simplistic specification. Flexible functional and hybrid forms represent the majority of recent applications (Rosko and Mutter 2008; Smet 2002; Street 2003).

Vignettes or episode approach such as used in the EuroDRG project constitute a third option that captures the benefits of flexible functional forms while not suffering from artificially ignoring the multi-product nature of hospitals.

2.4 Input variables

Most parametric analysis of hospital costs build explicitly or implicitly on a neo-classical model of the production process. In this framework factor markets are assumed to be competitive and free from monistic pressures (Smet 2002). Hence, hospital managers face given prices and can solely decide on the mix of inputs, the scale of input use and their application in the production process. Reliable price data for inputs, however, are often hard to collect or simply not available. Therefore, researchers often simply omit input prices in their models. This approach implies the assumption that input prices are identical across hospitals included in the sample or that the substitution rate for hospital technology is zero (Cowing et al. 1983). As a consequence, insights about the effects of differences in input prices, such as different wage levels, capital costs or differences in the cost of medical technology across hospitals, cannot be generated. Moreover, as we saw earlier the application of the strict neoclassical model implies that the hospital cost function is homothetic. Given the high degree of restrictiveness of these assumptions, recent analyses often relaxed the former and incorporated input prices in the model in the framework of more flexible functional forms.

Another central issue when considering the input side is, how inputs are disaggregated (Jacobs et al. 2006). Studies vary to a great extent with regard to the level of disaggregation: some studies use aggregate costs, while others try to disaggregate all relevant inputs. This variation of practice reflects on the one hand differences in the availability of data for input prices and quantities. On the other hand, choices with regard to the level of disaggregation also co-determine the focus of the analysis. Using total costs as the single measure of inputs implies the assumption that hospitals can deploy inputs freely without external constrains. This approach hence takes a long term perspective, assuming that hospitals can adopt an optimal mix of capital and labour. Disaggregation comes at the cost that models become less parsimonious, which captures some of the explanatory power of those factors in the production for which weighting is less clear (Jacobs et al. 2006: 30). Commonly disaggregated inputs are labour and capital. An important additional aspect is to account for services that hospitals have outsourced.

2.5 Outputs

As the model of the production process is the starting point of all analyses of hospital costs good measures for output are essential. Nevertheless, it is well known that defining outputs is problematic in the health sector, as health care is rarely demanded for its own sake but rather due to its positive contribution to health status. Given this background, hospital output should ideally be measured in terms of health outcomes. However, health outcomes are hard to measure and rarely systematically collected by hospitals. Researchers therefore tend to use activity indicators such as the number of cases treated, the number of hospital days per organization or the number of procedures performed as approximations for output. Using these indicators suffers from various limitations:

(1) Firstly, hospitals generating better health outcomes with less activity, for example by developing smart medical pathways are penalized. However, the quality of the services is hard to measure and therefore hard to incorporate in production or cost functions (Jacobs et al. 2006: 28).
(2) Secondly, given the multi-product nature of hospitals, the relative value or weight of each output has to be specified.

As we will see below in the sub-section on weights (2.7), parametric analyses tend to implicitly assume that the expenditure allocation within the hospital reflects societal values of different outputs, which might be questionable. Moreover, in the neo-classical production framework the level of output should be completely beyond the control of the hospital managers and physicians, i.e. exogenous. This assumption has been challenged from two angles:

(1) Firstly, it has been argued that hospital managers may manipulate the supply side by producing more if their hospital is characterized by a favourable cost structure or less if the opposite is the case.

(2) Secondly, many studies argue that physicians may induce demand, operating exogenous to the decisions of hospital management.

In both cases, estimates generated by a model that does not account for the former are biased. Nevertheless, exogenous determination of levels of output is often assumed given that case payment systems and the cost containment pressure by third-party payers weakened the relationship between gross-price charged and the quantity demanded (Smet 2002). Still, one of the defining features of well-suited models to determine hospital performance or sources of cost variation is that they incorporate well-designed measures of output and that they are capable to account for output being determined (partially) within hospitals via econometric techniques such as instrumental variables (Smet 2002).

2.6 Environmental variables

Any analysis of hospital costs needs to clarify how it approaches factors that constrain the production of health care. There is often considerable debate as to what factors should be considered constraints. In the short run, many factors are outside the control of the organizations and could be considered “legitimate” factors of variation. In the long-term a broader set of factors is potentially under the control of the organizations, but the extent and nature of this control will vary depending on the nature of the context.

Therefore the assumptions with regard to the equilibrium condition, i.e. the state (long- vs. short-term) in which managers can determine the level of all inputs, are important. Some authors assume that hospital managers are essentially able to set all inputs at the cost-minimizing level, include a price for capital and therefore estimate a long-run cost function. Others argue that only easily adjustable inputs can be employed at the cost-minimization level and estimate short run cost functions (Smet 2002).

Econometric techniques offer some orientation by providing tools to test whether hospitals operate in short or long run equilibrium using the so-called envelope condition (Braeutigam and Daughety 1983). The latter, however, can only be performed with reliable and detailed data for inputs, which are often not available. Estimating a long-run hospital cost function without testing the appropriateness of the long-term equilibrium assumption may lead to biased estimates and misspecification of functional form. Most researchers therefore agree that in the absence of high quality measures of capital input, short-run hospital cost functions should be estimated (Cowing et al. 1983; Smet 2002; Vita 1990).

Having agreed on the priority given to short-run models, one needs to consider how to incorporate those factors, which are beyond the control of hospital mangers in the analysis. In the framework of the parametric analyses there are three main ways to account for environmental variables:

- Firstly, comparisons can be restricted to hospitals operating within similarly constrained environments (Everitt et al. 2001).
- Secondly, external constrains can be included directly as explanatory variables in the production model.
- Thirdly, risk adjustment techniques adjust organizational output for differences in constrains before they are deployed in the efficiency model (Jacobs et al. 2006: 35). In the light of the multi-product nature of hospitals this approach is often considered superior to the other approaches, since it adjusts each output only for those constrains that apply specifically rather than adjusting all outputs generally for external constrains.

Risk adjustment to account for the fact that output is heterogeneous with regard to certain dimensions is often operationalized via product mix descriptor variables, which can be grouped into three broad categories: (1) case-mix complexity, (2) quality, and (3) intra-diagnosis related group severity of illness (Rosko and Mutter 2008). While contributing to avoid specification bias, the use of product descriptor variables has three additional global effects, which need to be taken into account. Firstly, the addition of further explanatory variables is always likely to lead to a decline of the previously unexplained residual and thereby to reduce estimated average inefficiency or cost variation (Jacobs et al. 2006). However, it reduces parsimony and adds uncertainty with regard to the results. Finally, the data requirements for risk adjustment are demanding and some controversy exists about appropriate techniques (Jacobs et al. 2006: 36).

2.7 Weights

A key aspect, which policy makers need to consider when scrutinizing cost and output variation across individual or groups of organizations, is whether all organizations are meant to fulfil the same task and objectives (Smith and Street 2005). Ideally, any analysis that compares output or costs across organizations needs to clarify the underlying assumptions with regard to the relative value it allocates to different outputs and the way specific outputs are therefore included and weighted in the model (Smith and Street 2005).

Clearly, in the context of the multi-product hospital world, the former is very complex. Various approaches have been used to account for this complexity such as creating a single index of outputs, estimation of a cost rather than a production function or the use of distance functions (Coelli and Perelman 2000; Shepard 1970). Smith and Street (2005) point out that common to all approaches is that the weight for each output is calculated via the sample mean cost of producing an additional unit of output. This implies that weights are not allocated with reference to societal values, but rather estimated as a by-product of statistical estimation (Smith and Street 2005). As a consequence most models implicitly assume that the expenditure allocation decisions of managers and physicians within hospitals reflect the values that society places on different outputs (Smith and Street 2005). Moreover, all models attach zero weight to any output that is not explicitly specified in the production model. Hence any resources devoted to outputs, which do not enter the model, are hypothesized to contribute to inefficiency or illegitimate cost variation. Thirdly, assumptions with regard to the nature of the rate of return from additional units of output, i.e. constant versus variable returns, considerably affect the weights that are attached to different outputs.

2.8 Dynamic nature of production

Unsolved is the question how to deal with the dynamic nature of the production process as current performance depends on past inheritances. Furthermore, hospital managers also invest resources in future attainments (Smith and Street 2005). Adequate measures for organizational endowments are lacking. Given these limitations, estimating cost determinants or inefficiency scores by analysing panel data might be a more suitable way to account for dynamic effects in the production process (Bond 2002; Jacobs et al. 2006; Smith and Street 2005).

2.9 Modelling the error term

One of the fundamental innovations of the SFA is interpreting the residuals, i.e. the difference between observed hospital costs or outputs and that predicted, as comprising random error and inefficiency. These two components are assumed to be distributed differently: for the random error term a normal distribution is assumed, while the inefficiency part can take various distributional forms. However, for cross-sectional data no economic or statistical criteria are available to inform the choice of the distributional characteristics of the inefficiency term (Newhouse 1994; Schmidt and Sickles 1984). Researchers are commonly provided with four distributions which they can apply to the inefficiency error term: a half-normal, a truncated-normal, exponential or a gamma distribution (Greene 2004). The choice of the distribution generates different estimates of average inefficiency in the sample as a whole and of relative inefficiency of individual hospitals (Jacobs et al. 2006; Rosko and Mutter 2008; Zuckerman et al. 1994). Overall, the significance of the impact of different distributions on relative and average inefficiency (Jacobs et al. 2006; Rosko and Mutter 2008; Zuckerman et al. 1994) seems to be limited, especially when comparing it to the difference resulting from decomposing inefficiency into two components versus interpreting the entire residuals as inefficiency using COLS (Jacobs et al. 2006: 57).

2.10 One versus two stage estimation

Identifying correlate variables of inefficiency such as firm-specific or environmental variables is highly relevant for the identification of "legitimate" sources of hospital performance variation. The analysis of the former can be conducted in one or two stages (Rosko and Mutter 2008). The first generation of SFA studies use two-stage approaches in which inefficiency estimates are calculated in a first step and in a second stage these estimates are regressed against a set of factors thought to correlate with inefficiency (Pitt and Lee 1981). One stage approaches simultaneously estimate the stochastic frontier and the inefficiency model, which allows estimating the effect of hospital specific and environmental variables directly (Battese and Coelli 1995; Coelli et al. 2005; Hjalmarsson et al. 1996; Wang and Schmidt 2002). Research indicates that two-stage estimation leads to substantially biased results (Li and Rosenman 2001; Rosko and Mutter 2008; Wang and Schmidt 2002) and should hence be interpreted with great caution and where possible ought to be checked against results from one stage estimation.

3 Efficiency analysis

3.1 The main approaches to efficiency analysis: SFA and DEA

The two main techniques to measure relative efficiency are Stochastic Frontier Analysis (SFA), which is a form of econometric model, and Data Envelopment Analysis (DEA), which is a form of linear optimization. We introduce and compare each of these techniques briefly.

SFA is a special form of the standard econometric model that builds on production theory. The standard model takes the general form:

$$Y_k = \alpha + \sum_{m=1}^{M} \beta_m \mathbf{z}_{mk} + \varepsilon_k \qquad (2)$$

where Y_k is the total amount of output for each organization $k = 1...K$, α is a constant and we have a set of $m = 1...M$ explanatory variables $\mathbf{z}_m$. These explanatory variables might measure the amount of labour or capital input used by each organization or the constraints faced by the organization in attaining the production frontier. β_m captures the influence of a particular explanatory variable z_m on output. If β_m is positive, the explanatory variable has a positive influence on output. The error term ε captures the imperfect relationship between these explanatory variables (i.e. the model) and observed output.

SFA has been developed to address a fundamental measurement problem: inefficiency is unobservable. How can one measure something that is not there: the lack of effort that an organization ought to exert, for instance? SFA tries to solve this problem in two stages.

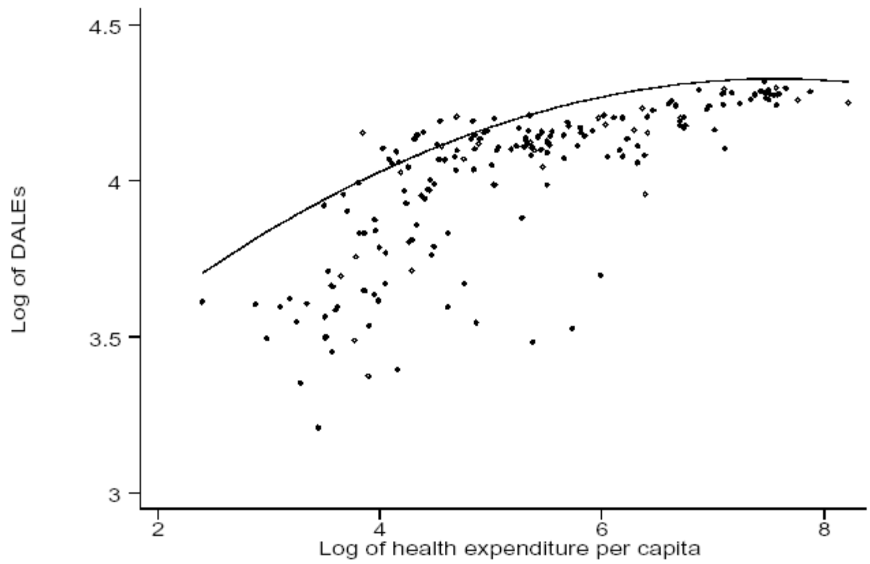
First, it specifies the form and location of the production frontier. This involves comparing different organizations producing the same things, to see which organization(s) are doing best. This involves specifying a production model where $Y_k = f(\mathbf{z}_{mk})$. We shall consider this specification in more detail shortly.

Second, the production of the remaining organizations is compared to the frontier, so that unobservable inefficiency is estimated by the distance from the frontier. However, as we noted in equation (1), true (unobserved) performance is also measured imperfectly, captured previously by the error term ε. But in the context of inefficiency, the term ε captures ***both*** measurement error and inefficiency. The SFA model disentangles these components, so that $\varepsilon = u + v$, where u captures inefficiency and v captures measurement or modelling error. The standard equation, then, is re-specified as:

$$Y_m = \alpha + \sum_{m=1}^{M} \beta_m \mathbf{z}_{mk} + u_k + v_k \qquad (3)$$

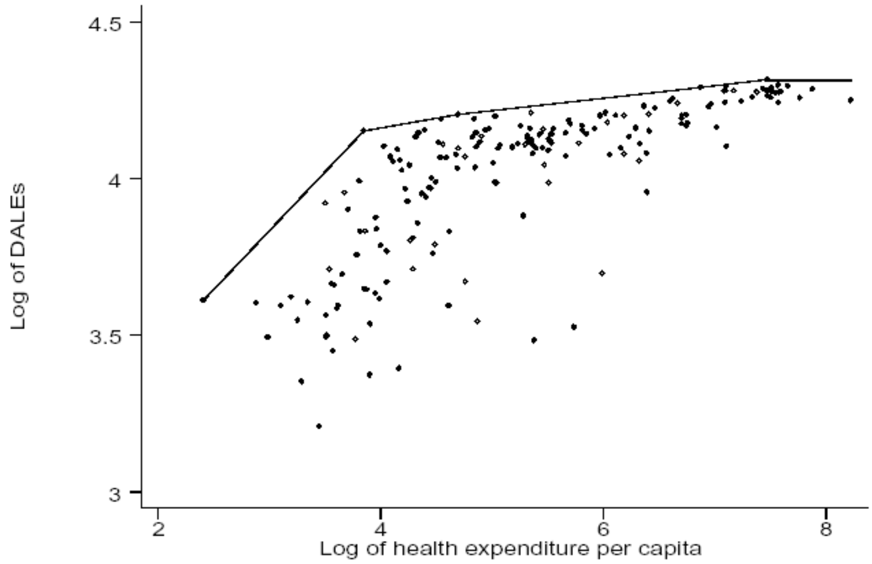
Figure 1 provides a graphical illustration of the process, taken from the World Health Organization's attempt the measure the relative performance of different countries in producing health (World Health Organization, 2000). Here we observe Disability Adjusted Life Expectancy (DALE) as the output of production and health expenditure per capita as the input. Each dot represents a unit of observation – each is a country in this example. The SFA frontier runs through the cloud of observations but not through the uppermost data points. Some points lie above the frontier because of measurement error v. For points below the frontier, the vertical distance comprises a combination of inefficiency u and measurement error v.

Figure 1: An example of a stochastic frontier going through a cloud of observations



DEA derives from an engineering tradition that presumes accurate data, so does not allow for measurement error. This greatly simplifies the production model which, in its most basic form, is specified as the ratio of output to input. This conforms to the simple notion that an organization that employs less input than another to produce the same amount of output can be considered more efficient. Those observations with the highest ratios of output to input are considered efficient, and the efficiency frontier is constructed by joining these observations up in the input-output space. The frontier thus comprises a series of linear segments connecting one efficient observation to another. Inefficient organizations are “enveloped” by the efficiency frontier in DEA. The inefficiency of the organizations within the frontier boundary is calculated relative to this “envelope”, as can be seen in Figure 2, using the same dataset as in Figure 1, but employing DEA instead of SFA. Thus, while careful specification of the production model is required for SFA, when applying DEA the location (and the shape) of the efficiency frontier is determined by the data.

Figure 2: An example of a DEA frontier enveloping a cloud of observations



DEA models can be run for both constant returns to scale (CRS) and a more flexible variable returns to scale (VRS) (as shown in Figure 2) which may be appropriate when not all organizations can be considered to be operating at an optimal scale.

3.2 Comparing DEA and SFA

As described earlier, DEA and SFA take different approaches to establishing the location and shape of the production frontier, and to determining where each organization is located in relation to the frontier. Here we briefly summarize these differences. This is important because these differences are the reasons why the techniques produce different estimates of the relative efficiency of individual organizations.

SFA controls for supposed influences on output, and contends that unexplained variations in output are due – in part, at least – to inefficiency. Unlike standard econometric models, therefore, where interest lies in the explanatory variables, SFA models extract organization-specific estimates of inefficiency from the unexplained part of the model. The implication is that standard econometric tools to test model specification cannot be applied to SFA models both because of the interpretation placed on ε_k and because organization-specific rather than average estimates are required. Thus un-testable judgments need to be made about the adequacy of stochastic frontier models and the inefficiency estimates they yield (Smith and Street 2005).

In applying DEA the location and shape of the frontier are established by the data. The outermost observations, those with the highest level of output given their scale of operation, are deemed efficient. Referring back to figures 1 and 2, DEA would consider observations lying at the upper extreme of the data cloud to be fully efficient, whereas SFA may consider that these organizations exhibit some degree of inefficiency.

The consequence of plotting the frontier through the outermost observations is that DEA is highly flexible, the frontier moulding itself to the data. The drawback, though, is that the frontier is sensitive to organizations that have unusual types, levels or combinations of inputs or outputs. These will have a scarcity of adjacent reference observations (usually termed "peers"), perhaps resulting in sections of the "frontier" being inappropriately positioned.

While DEA might be thought to win out over the SFA method in terms of flexibility, this is offset by how the technique interprets any distance from the frontier. There are two key differences. Firstly, DEA assumes correct model specification and that all data are observed without error while SFA allows for the possibility of modelling and measurement error. Consequently, even if the two techniques yield an identical frontier, SFA efficiency estimates are likely to be higher than those produced by DEA.

Secondly, DEA uses a selective amount of data to estimate each organization's efficiency score. DEA generates an efficiency score for each organization by comparing it only to peers that produce a comparable mix of outputs. This has two implications.

(1) If any output is unique to an organization, it will have no peers with which to make a comparison, irrespective of the fact that it may produce other common outputs. An absence of peers results in the automatic assignation of full efficiency for the organization under consideration.

(2) When assigning an efficiency score to an organization not lying on the frontier, only its peers are considered, with information pertaining to the remainder of the sample discarded. In contrast, SFA appeals to the full sample information when estimating relative efficiency.

In addition to making greater use of the available data, this facet of the SFA approach will make the sample's efficiency estimates more robust in the presence of outlier observations and to the presence of atypical input/output combinations. But this advantage over DEA is mainly a matter of degree – the location of (sections of) the DEA frontier may be determined by outliers, but outliers also exert influence on where the SFA frontier is positioned. Moreover, there are no statistical criteria for sorting these unusual observations into outliers or examples of best practice (Smith and Street 2005).

References

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

Bellanger MM, Or Z (2008). What can we learn from a cross-country comparison of the costs of child delivery? *Health Econ.,* 17, (1 Suppl): 47–57.

Bond S (2002). Dynamic panel data models: a guide to microdata methods and practice. *CeMMAP working papers 09/02*, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

Braeutigam RR, Daughety AF (1983). On the estimation of returns to scale using variable cost functions. *Economic Letters,* 11, (1–2): 25–31.

Breyer F (1987). The specification of a hospital cost function. A comment on the recent literature. *J.Health Econ.,* 6, (2): 147–157.

Busse R, Schreyogg J, Smith PC (2008a). Variability in healthcare treatment costs amongst nine EU countries – results from the HealthBASKET project. *Health Econ.,* 17, (1 Suppl): 1–8.

Busse R, Schreyogg J, Smith PC (2008b). Variability in healthcare treatment costs amongst nine EU countries – results from the HealthBASKET project. *Health Econ.,* 17, (1 Suppl): 1–8.

Cameron AC, Trivedi PK (1986). Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests. *Journal of Applied Econometrics,* 1, (1): 29–53.

Caves D, Christensen LR, Tretheway MW (1980). Flexible cost functions for multi-product firms. *Review of Economics and Statistics,* 62, 477–481.

Coelli TJ, Perelman S (2000). Technical Efficiency of European Railways: A Distance Function Approach. *Applied Economics,* 32, (15): 1967–1976.

Coelli TJ, Rao D, O'Donnell C, Battese GE (2005). *An introduction to efficiency and productivity analysis*, New York, Springer.

Conrad RF, Strauss RP (1983). A Multiple- Output, Multiple- Input Model of the Hospital Industry in North Carolina. *Applied Economics,* 15, (3): 341–352.

Cowing TG, Holtmann AG, Powers S (1983). Hospital cost analysis: a survey and evaluation of recent studies. *Adv.Health Econ.Health Serv.Res.,* 4, 257–303.

Cremieux PY, Ouellette P (2001). Omitted variable bias and hospital costs. *J Health Econ.,* 20, (2): 271–282.

Currie G, Donaldson C, Lu M (2003). What does Canada Profit from the For-Profit Debate on Health Care? *Canadian Public Policy,* 29, (2): 227–251.

Daidone S, D'Amico F (2009). Technical efficiency, specialization and ownership form: evidences from a pooling of Italian hospitals. *Journal of Productivity Analysis,* 32, (3): 203–216.

Devereaux PJ, Choi PT, Lacchetti C, Weaver B, Schunemann HJ, Haines T, Lavis JN, Grant BJ, Haslam DR, Bhandari M, Sullivan T, Cook DJ, Walter SD, Meade M, Khan H, Bhatnagar N, Guyatt GH (2002). A systematic review and meta-analysis of studies comparing mortality rates of private for-profit and private not-for-profit hospitals. *CMAJ.,* 166, (11): 1399–1406.

Dormont B, Milcent C (2004). The sources of hospital cost variability. *Health Econ.,* 13, (10): 927–939.

Dranove D (1987). Rate Setting by Diagnosis Related Groups and Hospital Specialization. *RAND Journal of Economics,* 18, (3): 417–427.

Duckett SJ (1994). Reform of Public Hospital Funding in Victoria. *Australia Studies in Health Service Administration 77*, School of Health Service Management, University of New South Wales: Kensington.

Epstein D, Mason A, Manca A (2008). The hospital costs of care for stroke in nine European countries. *Health Econ.,* 17, (1 Suppl): 21–31.

Everitt BS, Landau S, Leese M (2001). *Cluster Analysis*, London, Arnold.

Fattore G, Torbica A (2008). Cost and reimbursement of cataract surgery in Europe: a cross-country comparison. *Health Econ.,* 17, (1 Suppl): 71–82.

Folland ST, Hofler RA (2001). How reliable are hospital efficiency estimates? Exploiting the dual to homothetic production. *Health Econ.,* 10, (8): 683–698.

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

Hansmann H (1980). The role of nonprofit enterprise. *Yale Law Journal,* 89, 835–901.

Harris JE (1977). The Internal Organisation of Hospitals: Some Economic Implication. *Bell Journal of Economics,* 8, 467–482.

Hausman JA (1978). Specififation Tests in Econometrics. *Econometrica,* 46, (6): 1251–1271.

Hjalmarsson L, Kumbhakar SC, Heshmati A (1996). DEA, DFA and SFA: A Comparison. *Journal of Productivity Analysis,* 7, 303–327.

Hollingsworth B (2003). Non-parametric and parametric applications measuring efficiency in health care. *Health Care Manag.Sci.,* 6, (4): 203–218.

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

Hvenegaard A, Street A, Sorensen TH, Gyrd-Hansen D (2009). Comparing hospital costs: what is gained by accounting for more than a case-mix index? *Soc.Sci.Med.,* 69, (4): 640–647.

Intriligator MD (1978). *Econometric Models, Techniques and Applications*, N.J.: Prentice-Hall, Englewood Cliffs.

Jacobs R, Smith P, Street A (2006). *Measuring Efficiency in Health Care*, Cambridge, New York, Cambridge University Press.

Jones AM (2000). Health Econometrics, In: Culyer, AJ and Newhouse, JP, eds. Handbook of Health Economics, Elsevier: 265–344.

Joskow PL (1983). Reimbursement policy, cost containment and non-price competition. *J Health Econ.,* 2, (2): 167–174.

Kjekshus LE, Hagen TP (2005). Ring fencing of elective surgery: does it affect hospital efficiency? *Health Serv.Manage.Res.,* 18, (3): 186–197.

Kristensen T, Laudicella M, Ejersted C, Street A (2009). Cost variation in diabetes care delivered in English hospitals.

Kuh E, Meyer JR (1955). Correlations and regression estimates when the data are ratios. *Econometrica,* 23, 400–416.

Laudicella M, Olsen KR, Street A (2009). What explains variation in the costs of treating patients in English obstetrics specialties? *CHE Research Paper 49*, Center for Health Economics, University of York: York, UK.

Li T, Rosenman R (2001). Cost inefficiency in Washington hospitals: a stochastic frontier approach using panel data. *Health Care Manag.Sci.,* 4, (2): 73–81.

Long JC (1997). *Regression Models for Categorical and Limited Dependent Variables (Advanced Quantitative Techniques in the Social Sciences)*, Sage Publications.

Luft HS, Robinson JC, Garnick DW, Maerki SC, McPhee SJ (1986). The role of specialized clinical services in competition among hospitals. *Inquiry,* 23, (1): 83–94.

Maddala GS (1988). *Introduction to Econometrics*, London, Collier Macmillan.

Mason A, Street A, Miraldo M, Siciliani L (2008a). Should prospective payments be differentiated for public and private healthcare providers? *Health Economics, Policy and Law,* 0, 1–21.

Mason A, Street A, Miraldo M, Sivey P, Siciliani L (2008b). Establishing a Fair Playing Field for Payment by Results. *CHE Research Papers 39*, Centre for Health Economics, University of York: York, UK.

Newhouse JP (1994). Frontier estimation: how useful a tool for health economics? *J.Health Econ.,* 13, (3): 317–322.

Panzar J, Willig R (1981). Economies of scope. *American Economic Review,* 71, 268–272.

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

Pitterle ME, Bond CA, Raehl CL, Franke T (1994). Hospital and pharmacy characteristics associated with mortality rates in United States hospitals. *Pharmacotherapy,* 14, (5): 620–630.

Posnett J (2002). Are bigger hospitals better?, In: McKee, M and Healy, J, eds. Hospitals in a changing Europe, Buckhingham, Philadelphia: Open University Press: 100–118.

Prince TR (1998). A medical technology index for community hospitals. *Health Care Manage.Rev.,* 23, (1): 52–63.

Rosko MD, Mutter RL (2008). Stochastic frontier analysis of hospital inefficiency: a review of empirical issues and an assessment of robustness. *Med.Care Res.Rev.,* 65, (2): 131–166.

Schmidt P, Sickles RC (1984). Production frontiers and panel data. *Journal of Business and Aconomic Studies,* 2, 299–326.

Schreyögg J (2008). A micro-costing approach to estimating hospital costs for appendectomy in a cross-European context. *Health Econ.,* 17, (1 Suppl): 59–69.

Shepard RW (1970). *Theory of Cost and Production Funktions*, Princeton, Princeton University Press.

Shleifer A (1985). A Theory of Yardstick Competition. *Rand Journal of Economics,* 16, 319–327.

Sloan FA, Picone GA, Taylor DH, Chou SY (2001). Hospital ownership and cost and quality of care: is there a dime's worth of difference? *J.Health Econ.,* 20, (1): 1–21.

Smet M (2002). Cost characteristics of hospitals. *Soc.Sci.Med.,* 55, (6): 895–906.

Smith A (1981). An Inquiry into the Nature and Causes of the Wealth of NationsGlasgow Edition of the Works and Correspondence of Adam Smith 2, Indianapolis: Liberty Press.

Smith P, Street A (2005). Measuring the efficiency of public services: the limits of analysis. *J.R.Statist.Soc.A,* 168, (2): 401–417.

Spetz J, Maiuro LS (2004). Measuring levels of technology in Hospitals. *The Quarterly Review of Economics and Finance,* 44, (3): 430–447.

Stargardt T (2008). Health service costs in Europe: cost and reimbursement of primary hip replacement in nine countries. *Health Econ.,* 17, (1 Suppl): 9–20.

Street A (2003). How much confidence should we place in efficiency estimates? *Health Econ.,* 12, (11): 895–907.

Street A, Sivey P, Mason A, Miraldo M, Siciliani L (2009). Are English treatment centres treating less complex patients? *Health Policy,* 94, (2): 150–157.

Tiemann O (2008). Variations in hospitalisation costs for acute myocardial infarction – a comparison across Europe. *Health Econ.,* 17, (1 Suppl): 33–45.

Vaillancourt RP (2003). Performance Evaluations of For- Profit and Nonprofit U.S. Hospitals Since 1980. *Nonprofit Management & Leadership,* 13, (4): 401–422.

Vita MG (1990). Exploring hospital production relationships with flexible functional forms. *J Health Econ.,* 9, (1): 1–21.

Wang H, Schmidt P (2002). One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. *Journal of Productivity Analysis,* 18, 129–144.

White HA (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica,* 84, 817–830.

Wörz M (2008). *Erlöse- Kosten- Qualität: Macht die Krankenhausträgerschaft einen Unterschied?*, Wiesbaden, Verlag für Sozialwissenschaften.

Zuckerman S, Hadley J, Iezzoni L (1994). Measuring hospital efficiency with frontier cost functions. *J.Health Econ.,* 13, (3): 255–280.

Zweifel P, Breyer F (1997). *Health Economics*, Oxford, Oxford University Press.