Herwartz, Helmut; Strumann, Christoph

Working Paper

On the effect of prospective payment system on hospital efficiency and competition for patients in Germany

Economics working paper / Christian-Albrechts-Universität Kiel, Department of Economics, No. 2011,03

Provided in cooperation with:
Christian-Albrechts-Universität Kiel (CAU)

C | A | U
Christian-Albrechts-Universität zu Kiel
Department of Economics
Economics Working Paper
No 2011-03

on the effect of prospective payment system on hospital efficiency and competition for patients in Germany

by Helmut Herwartz and Christoph Strumann

On the effect of prospective payment system on hospital efficiency and competition for patients in Germany

Helmut Herwartz* Christoph Strumann‡

March 3, 2011

Abstract

The introduction of hospital reimbursement based on diagnosis related groups (DRG) in 2004 has been a conspicuous attempt to increase hospital efficiency in the German health sector. In this paper changes of hospital efficiency, quantified as a Malmquist index decomposition in pure technical efficiency change, are analyzed for periods before and after the reform. We implement a two-stage semi-parametric efficiency model that allows for spatial interdependence among hospitals. The results reveal an enhancement in overall efficiency after the DRG introduction. Moreover, an increase in the magnitude of negative spatial spillovers among German hospital performance can be diagnosed. This result is in line with a rise of competition for (low cost) patients.

JEL-Classification: C21, D61, I11, I18

Keywords: Hospital efficiency, data envelopment analysis, spatial analysis, diagnosis related groups

*Institut für Statistik und Ökonometrie, Christian-Albrechts-Universität zu Kiel, Olshausenstr. 40, D-24118 Kiel, Germany, h.herwartz@stat-econ.uni-kiel.de

†Corresponding author. Institut für Statistik und Ökonometrie, Christian-Albrechts-Universität zu Kiel, Olshausenstr. 40, D-24118 Kiel, Germany, cstrumann@stat-econ.uni-kiel.de, tel.: +49 431 880 2379, fax: +49 431 880 7605

‡We thank Andrew Street, the participants of the XVth. Spring Meeting of Young Economists 2010 in Luxembourg, of the Jahrestagung des Vereins für Socialpolitik 2010 in Kiel and of the 15. Nachwuchsworkshop der DStatG 2009 in Merseburg for helpful comments and discussions on earlier versions of this manuscript. We also thank the employees of the Forschungsdatenzentrum der Statistischen Landesämter - Standort Kiel/Hamburg for their cooperation.

1 Introduction

The German health sector is characterized by a steady increase of hospital expenditures. It has doubled from 1991 to 2007, reaching almost 60 billion Euro in 2007 (Statistisches Bundesamt, 2008). This amounts to around 2.5% of the German gross domestic product (GDP). Until 2004 German hospitals have been reimbursed by per diem payments. This system has invoked incentives to hospitalize patients as long as possible, likely resulting in an inefficient use of resources. In December 1999 the left-wing government announced the introduction of a prospective payment system based on diagnosis related groups (DRG) in 2004 as an attempt to increase the efficiency of hospitals (e.g. Hensen et al., 2008 and Lungen and Lapsley, 2003). Under the DRG based financing system hospitals receive a fixed rate for each admission depending on a patient's diagnosis. If the costs for a particular case are lower (higher) than the reimbursement, the hospital realizes profits (losses). As a consequence, hospitals face an increased pressure on their financial performance and a higher risk of insolvency.

The effects of the reform on hospital efficiency in Germany have not been evaluated empirically yet. However, some consequences can be expected. Under the prospective payment system it is profitable to decrease the lengths of stay and to increase simultaneously the number of treated cases. Evidence for such behavioral adjustments of hospitals is given for dermatological hospital admissions by Hensen et al. (2008). Furthermore, Böcking et al. (2005) mention that hospitals preferably treat cases with high reimbursement rates and a low level of complexity. This implies relatively less resource usage in comparison with treating patients with the same diagnosis but higher level of complexity. A hospital which is successful in attracting so called low cost patients (e.g. due to reputation, bribery[1]) might show a better performance than its neighbor hospitals. Hence, a rise of competition for low cost patients could be reflected by an increased negative spatial interdependence of hospital efficiency.

The aim of this study is to subject two hypotheses about potential effects of the financial reform on overall hospital performance and spatial interdependence to empirical testing. In particular, firstly, we examine if hospitals have realized efficiency gains and, secondly, if stronger

[1]In Germany in summer 2009, there was an affair about bribery payments from several hospitals to primary care physicians for the admission of low cost patients.

negative spatial interdependence of hospital efficiency has emerged after the DRG reform. For this purpose a spatial two-stage semi-parametric efficiency model is implemented. In the first stage, hospital efficiency is quantified by means of the non-parametric *Data Envelopment Analysis* (DEA). To identify efficiency gains as a consequence of changes in the hospital incentive structure, we determine the Malmquist index decomposition in pure technical efficiency change, which is not affected by technological progress and scale adjustments (e.g. Burgess and Wilson, 1995 and Sommersguter-Reichmann, 2000). In a second stage, potential effects of the DRG reform on hospital efficiency change and spatial spillovers among hospital efficiency are investigated by means of a parametric spatially autoregressive model with spatially autoregressive disturbances (SARAR). An unbalanced cross-section of around 1500 German hospitals is analyzed over 12 years (1995 to 2006) covering the DRG announcement and introduction period. The results show an enhancement in overall efficiency. Moreover, an increase in the magnitude of negative spatial spillovers among German hospital performance can be diagnosed. This result is in line with an expected rise of competition for (low cost) patients invoked by the prospective payment system.

In Section 2, the two hypotheses about potential effects of the financial reform on overall hospital performance and spatial interdependence are put forth. Section 3 sketches the measurement of efficiency and efficiency change, the SARAR model and the empirical testing strategy and describes the data. Empirical results are discussed in Section 4. Section 5 concludes. An appendix delivers methodological details for Section 3.

2 Effects of the financial reform

The introduction of the DRG based financing system in 2004 has been intended to increase hospital efficiency by changing the incentive structure (e.g. Hensen et al., 2008, Böcking et al., 2005 and Lungen and Lapsley, 2003). The empirical evidence for efficiency gains after the DRG reform, obtained for various countries is, however, not fully conclusive. For the cases of Norway (Biörn et al., 2010 and Biörn et al., 2003), Portugal (Dismuke and Sena, 1999) and Finland (Linna, 2000), positive effects of the DRG introduction on hospital efficiency have been found.

In the same time no effect is detected for the case of Austria (Sommerguters-Reichmann, 2000). To examine the intended efficiency enhancement of German hospitals, we analyze the change of technical efficiency during the period from 1995 to 2006. Hence, the sample starts in a pre-reform period and covers both the announcement (at the end of 1999) and the introduction (2004) of the reform. If there is any effect of the reform, we expect efficiency gains after these particular dates. Hospital efficiency improvements should be interpreted as a response to the changed incentives, because of two reasons. Firstly, we use the Malmquist index decomposition in pure technical efficiency change, which is invariant to changes of technology and scale adjustments (e.g. Burgess and Wilson, 1995 and Sommersguter-Reichmann, 2000). Secondly, there have been no other major exogenous shocks affecting hospital efficiency during the sample period. We formalize the following hypothesis

- $H(A)$: The announcement (2000) or introduction (2004) of the DRG based financing system has been followed by improvements of hospital efficiency.

As Ellis (1998) points out in a theoretical equilibrium approach, under prospective payment, health providers dump the most severely ill patients and compete to attract low cost patients. The latter are characterized by a low level of complexity and an expected relatively short hospital stay. Empirical evidence for an implicit patient selection after the shift to a prospective payment system is found for the US by Norton et al. (2002) and Ellis and McGuire (1996). Hospitals which are successful in attracting low cost patients in the nearby area use relatively less resources in comparison with hospitals treating patients of the same area with the same diagnosis but higher levels of complexity. These cases might be characterized by prolonged hospital stays. Hence, the performance of two contiguous hospitals is expected to be negatively correlated if one of the two competes (more successful) in the described way. Then, strengthened competition for low cost patients could be reflected by an increased magnitude of negative spatial interdependence of hospital efficiency. This leads to the hypothesis

- $H(B)$: The announcement (2000) or introduction (2004) of the DRG based financing system has been followed by an increase in the magnitude of negative spatial interdependence of hospital efficiency.

3 Methodology

In this Section, we sketch the spatial two-stage semi-parametric efficiency model. In the first stage, DEA efficiency scores and the Malmquist index decomposition in pure technical efficiency change are determined. In the second stage a parametric SARAR regression model is implemented. Furthermore, this Section illustrates the empirical testing strategies for the hypotheses $H(A)$ and $H(B)$. Potential effects of the DRG reform on hospital efficiency gains ($H(A)$) are investigated by means of a regression of logarithmic pure technical efficiency change. We implement an unbalanced panel data model with time dummy variables, while controlling for hidden and observable heterogeneity across hospitals in form of fixed effects and explanatory variables, respectively. An enhancement of efficiency is identified by testing for increasing time effects. The second hypothesis ($H(B)$) is examined by means of spatial cross-sectional regressions of logarithmic DEA efficiency scores. Spatial spillover estimates are then tested for a decrease over time. Moreover, the data and the construction of relevant variables used for the analysis are described in this Section.

3.1 Hospital efficiency and the Malmquist decomposition

Hospital efficiency is estimated in a first step by means of the non-parametric DEA. In this framework, the production or cost function does not require an explicit specification. Thus, assumptions about profit-maximization or cost-minimization behavior, which might be inappropriate for (non-profit) hospitals (Zweifel et al., 2009), can be avoided. The input-based DEA efficiency score of hospital i, $\theta^C_{i,t_1|t_2}$, is obtained under the assumption of constant returns to scale through a comparison of its set of inputs and outputs of period t_1 to that of all hospitals in period t_2, where $t_1, t_2 \in \{t-1, t\}$. The measure denotes the radial distance of the i-th hospital at time t_1 to the frontier function at time t_2, which is determined from a linear combination of the best practicing (efficient) units in t_2. As shown by Kneip et al. (1998) the DEA efficiency scores are consistent estimates for the true efficiency scores (details are given in Appendix A).

The input-based Malmquist index of efficiency change from $t-1$ to t is the geometric mean of the change in efficiency under both frontier functions in $t-1$ and t. For the i-th hospital the index, along with its decomposition in efficiency change (EC) and technological change (TC), is given by

$$MI_{i,t} = \left[\frac{\theta^C_{i,t-1|t-1}}{\theta^C_{i,t|t-1}} \cdot \frac{\theta^C_{i,t-1|t}}{\theta^C_{i,t|t}}\right]^{1/2} = \underbrace{\frac{\theta^C_{i,t-1|t-1}}{\theta^C_{i,t|t}}}_{EC_{i,t}} \cdot \underbrace{\left[\frac{\theta^C_{i,t|t}}{\theta^C_{i,t|t-1}} \cdot \frac{\theta^C_{i,t-1|t}}{\theta^C_{i,t-1|t-1}}\right]^{1/2}}_{TC_{i,t}}. \quad (1)$$

In (1), $EC_{i,t}$ measures the movement over time of hospital i towards the frontier function and represents a change in efficiency (Färe et al., 1992). Moreover, $TC_{i,t}$ is the geometric mean of the change in efficiency under changing technology given the production bundles of $t-1$ and t, and indicates a shift in the constant returns to scale technology. The efficiency change component can be further decomposed in pure technical efficiency change (PEC) and scale efficiency adjustments (SEA)

$$EC_{i,t} = \underbrace{\frac{\theta^V_{i,t-1|t-1}}{\theta^V_{i,t|t}}}_{PEC_{i,t}} \cdot \underbrace{\left[\frac{\theta^V_{i,t|t}}{\theta^C_{i,t|t}} \cdot \frac{\theta^C_{i,t-1|t-1}}{\theta^V_{i,t-1|t-1}}\right]}_{SEA_{i,t}},$$

where $\theta^V_{i,t_1|t_2}$ is the respective efficiency score under variable returns to scale (Banker et al., 1984). The pure efficiency change, $PEC_{i,t}$, measures the relative efficiency enhancement and is invariant to changes in the technology and scale adjustments (Burgess and Wilson, 1995 and Sommersguter-Reichmann, 2000). For the ease of interpretation, we consider the inverse of $PEC_{i,t}$ as the pure technical efficiency change of hospital i from period $t-1$ to t

$$\gamma_{it} = \theta^V_{i,t|t} / \theta^V_{i,t-1|t-1}.$$

By construction, values above (below) unity indicate an improvement (regress) in efficiency.

DEA scores are constrained to the interval $(0, 1]$, with 1 indicating an efficient hospital. To avoid the censoring problem in the second stage regression analysis, we compute super efficiency scores by means of the tie-breaking procedure proposed by Andersen and Petersen (1993). For this purpose, the efficient units are ranked according to the amount by which their input vectors could be increased without becoming inefficient. To compare hospitals with distinct exogenously fixed input variables we account for non-discretionary input variables (Banker and Morey, 1986). Due to the deterministic nature of DEA, measurement errors for observations of the reference set can distort the estimated efficiency scores for all hospitals (e.g. Wilson, 1995). Similarly, hospitals performing particularly poor might also invalidate the second stage regression results. Therefore, we apply an outlier detection proposed by Johnson and McGinnis (2008) to identify hospitals having an outstandingly good or poor performance. Hospitals are treated as an efficient outlier if it is possible to double the inputs without becoming inefficient. An inefficient outlier is detected if a convex-combination of worst performing hospitals can produce the same level of output using half the inputs. Outlying hospitals are excluded from the analysis.[2]

3.2 The spatial regression model

The rationale about spatial interdependence among German hospital performance leads to the conjecture that spatial dependence might similarly characterize changes of hospital efficiency. Thus, for both variables, $\theta^V_{i,t|t}$ and γ_{it}, we implement a SARAR model to account for two distinct channels of spatial dependence simultaneously. On the one hand, negative spatial spillovers might occur due to the competition for low cost patients and, on the other hand, positive spatial dependence could be the result of similar unobservable factors of nearby observations. As mentioned above, γ_{it} is analyzed in an unbalanced fixed effects panel and $\theta^V_{i,t|t}$ in a cross-sectional model framework. Both models read in time t as

$$y_t = \lambda \boldsymbol{W}_t y_t + \boldsymbol{Z}_t \beta + e_t, \quad \text{with } e_t = \rho \boldsymbol{M}_t e_t + \epsilon_t,\ t = 1, \ldots, T, \tag{2}$$

where y_t is an $N \times 1$ vector comprising the variables of interest, i.e. the logarithm of pure technical efficiency change, $y_t = (\ln(\gamma_{1t}), \ldots, \ln(\gamma_{Nt}))'$, or the logarithm of DEA efficiency scores, $y_t = (\ln(\theta^V_{1,t|t}), \ldots, \ln(\theta^V_{N,t|t}))'$, $\boldsymbol{Z}_t$ is an $N \times K$ matrix of observations of K explanatory variables and β a $K \times 1$ vector of parameters. The pattern of spatial dependence is captured by the $N \times N$ spatial weights matrices $\boldsymbol{W}_t$ and $\boldsymbol{M}_t$ with zero diagonal elements and row normalized

[2] As it turns out, the empirical testing results of $H(A)$ and $H(B)$ are qualitatively similar for alternative threshold values for the outlier detection.

constants (such that each row sums to unity). The number of hospitals, N, varies with t, since some hospitals are not observed over all time periods. The spatial lag coefficient λ measures the direct effect of the weighted neighboring observations on the elements in y_t (Anselin, 1988). Spatial dependence due to similar unobservable factors of nearby observations is quantified by the spatial autocorrelation coefficient ρ. Both spatial parameters are restricted to be less than unity in absolute value. Finally, ϵ_t is an $N \times 1$ vector of location specific i.i.d. disturbances, $\epsilon_t \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I}_N)$, where $\boldsymbol{I}_N$ is the N-dimensional identity matrix. Adding time and individual effects the (unbalanced) model can be written in matrix notation as

$$y = \lambda \begin{pmatrix} \boldsymbol{W}_1 \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \boldsymbol{W}_T \end{pmatrix} y + \begin{pmatrix} \mathfrak{o}_1 & \cdots & \mathfrak{o}_1 \\ \iota_2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \iota_T \end{pmatrix} \delta + \alpha + \boldsymbol{Z}\beta + e, \quad e = \rho \begin{pmatrix} \boldsymbol{M}_1 \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \boldsymbol{M}_T \end{pmatrix} e + \epsilon, \quad (3)$$

where $y = (y_1', \ldots, y_T')'$, $\boldsymbol{Z} = (\boldsymbol{Z}_1', \ldots, \boldsymbol{Z}_T')'$, $e = (e_1', \ldots, e_T')'$ and $\epsilon = (\epsilon_1', \ldots, \epsilon_T')'$. The coefficients of the time dummy variables, δ_t, are collected in $\delta = (\delta_2, \ldots, \delta_T)'$, where $t = 1$ is the benchmark, $\mathfrak{o}_t$ and ι_t is an $N_t \times 1$ vector of zeros and ones, respectively, where N_t is the number of hospitals sampled in time t. The fixed effects are summarized in $\alpha = (\alpha_1', \ldots, \alpha_T')'$, where α_t is an $N_t \times 1$ vector comprising the individual effects of the N_t hospitals. These are dropped out by means of the within transformation. The panel and cross-sectional models are estimated by means of a Maximum Likelihood approach[3] (see Appendix B for a formal representation of the likelihood function).

3.3 Empirical testing strategy of DRG Hypotheses

While controlling for observable and hidden hospital heterogeneity, the period after the DRG announcement or introduction should be characterized by a significant rise of γ_{it} under $H(A)$.

[3]Simar and Wilson (2007) mention that in finite samples the estimated efficiency scores are biased and serially correlated in a complicated fashion. The convergence rate of $\theta^V_{i,t|t}$ depends on the number of inputs and outputs and is typically lower than the parametric convergence rate. Therefore the serial correlation and the bias itself, disappear asymptotically with the same rate as $\theta^V_{i,t|t}$ converges. Maximum Likelihood estimates of regressions of $\theta^V_{i,t|t}$ are consistent, but inference based on the inverse of the negative Hessian of the log-likelihood is generally invalid. To overcome this problem, we apply a bootstrap procedure suggested by Simar and Wilson (2007). However, the difference between the bootstrap and asymptotic results is negligible.

The hypothesis is examined by testing for significant increases of the means of the time dummy coefficients of 5 subperiods. In particular, to verify an announcement effect (AE), the pre-announcement period (SP_1: 1996 to 1999) is compared with the post-announcement-pre-reform period (SP_2: 2000 to 2003). An introduction effect is evaluated by means of two alternative strategies. The first approach (IE_1) takes the AE into account and compares SP_2 with the post-reform period (SP_3: 2004 to 2006). Secondly, the AE is neglected and the pre-reform period (SP_4: 1996 to 2003) is compared with SP_3 (IE_2). Finally, an overall effect (OE) is examined by comparing SP_1 with the post-announcement period (SP_5: 2000 to 2006). The empirical testing strategy can be summarized by the following hypotheses with the tested effects in parentheses

$$H_0^{\mathrm{A}_1} : \bar{\delta}_1 = \bar{\delta}_2 \qquad \text{vs.} \qquad H_1^{\mathrm{A}_1} : \bar{\delta}_1 < \bar{\delta}_2 \qquad (AE) \tag{4}$$

$$H_0^{\mathrm{A}_2} : \bar{\delta}_2 = \bar{\delta}_3 \qquad \text{vs.} \qquad H_1^{\mathrm{A}_2} : \bar{\delta}_2 < \bar{\delta}_3 \qquad (IE_1) \tag{5}$$

$$H_0^{\mathrm{A}_3} : \bar{\delta}_4 = \bar{\delta}_3 \qquad \text{vs.} \qquad H_1^{\mathrm{A}_3} : \bar{\delta}_4 < \bar{\delta}_3 \qquad (IE_2) \tag{6}$$

$$H_0^{\mathrm{A}_4} : \bar{\delta}_1 = \bar{\delta}_5 \qquad \text{vs.} \qquad H_1^{\mathrm{A}_4} : \bar{\delta}_1 < \bar{\delta}_5 \qquad (OE), \tag{7}$$

where $\bar{\delta}_1 = \frac{1}{4}\sum_{t=1996}^{1999} \delta_t$ (SP_1), $\bar{\delta}_2 = \frac{1}{4}\sum_{t=2000}^{2003} \delta_t$ (SP_2), $\bar{\delta}_3 = \frac{1}{3}\sum_{t=2004}^{2006} \delta_t$ (SP_3), $\bar{\delta}_4 = \frac{1}{8}\sum_{t=1996}^{2003} \delta_t$ (SP_4), $\bar{\delta}_5 = \frac{1}{7}\sum_{t=2000}^{2006} \delta_t$ (SP_5) and δ_t is the time dummy coefficient for the year t. The year 1996 serves as reference, i.e. $\delta_{1996} = 0$. The hypotheses are tested by means of one-sided t-tests based on the covariance matrix of estimated time effects.

Hypothesis $H(B)$ suggests an increase of the magnitude of negative spatial spillovers λ. In order to test $H(B)$ we firstly apply a cross-sectional spatial regression of logarithmic DEA efficiency scores. Average spatial spillover estimates of SP_1 to SP_5 are then compared with each other analogously to the procedure described for $H(A)$.

3.4 Data and variable construction

3.4.1 The data set

The data are drawn from two distinct sources. Hospital data are extracted from the annual hospital statistics collected by the statistical offices of the federal states ("Statistische Landesämter"). It includes basic hospital characteristics, e.g. forms of ownership, the number of beds, staff, patients, etc., and data on the cost structure of the hospitals, as total costs, payroll costs, material expenses etc. The district- and state-level data are obtained from the "Regionaldatenbank Deutschland - GENESIS", which is administered by the statistical office of North Rhine-Westphalia ("Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen"). Annual data cover the period from 1995 until 2006 and have been provided by the "Forschungsdatenzentrum der Statistischen Landesämter - Standort Kiel/Hamburg". In Germany university hospitals are generally in charge of the education of young medical doctors and research programmes. Thus, a comparison with other hospitals is rather difficult. Therefore university hospitals are not considered in the analysis. For each year, around 450 hospitals have missing values for relevant variables or data inconsistencies, like declaring costs of less than 100 Euro or having zero-values for beds, physicians etc. These hospitals are also excluded from the sample. Moreover, 0.3% (1996) to 1.9% (2001) and 0% (1995) to 0.9% (2003) of the hospitals are detected as efficient and inefficient outliers, respectively, and removed from the analysis. Finally, to facilitate the interpretation of the time effects, hospitals with less than 2 data points are not included in the panel model.

3.4.2 Inputs and outputs

The input variables controlled by the hospitals are the amount of material expenses (in 2005 prices) (*exp*), the number of employed physicians (*phys*), nurses (*nurses*) and non-medical employees (*nonmed*). Notably, the capacity of beds is imposed by the states for most hospitals and therefore a non-controllable instrument for these hospitals. Accordingly, the number of beds (*beds*) is treated as a non-discretionary input. For a hospital's output we take the number of cases weighted for the respective average resource usage (*wcases*, Herr, 2008), which is approximated by the nationwide average length of stay of patients treated in a particular clinical department (details can be found in Appendix C). As a second output variable the number of apprentices is considered (*appr*).[4]

[4] As a robustness check we have applied two further input specifications. Firstly, the number of employees are replaced by the expenses for labor. Secondly, in order to minimize measurement errors in the labor variables

3.4.3 Explanatory variables

To control for observable heterogeneity across hospitals, the following hospital- and district/state level variables are selected.

Hospital specific variables: A main finding of hospital efficiency studies for Germany is a lower efficiency in privately owned hospitals in comparison with their public counterparts (e.g. Herr, 2008, Steinmann et al., 2004, and Helmig and Lapsley, 2001). To control for private for profit and non-profit private hospitals respective dummy variables (*private* and *non-profit*) are included. We explore the impact on hospital efficiency of the market share, *ms*, which is obtained as the number of patients of a hospital relative to competitors located in the same district. Town and Vistnes (2001) and Dranove and Ludwick (1999) confirm that higher market shares reduce costs or raise profits due to improvements of the hospital's bargaining power. Hospitals usually do not adjust their working staff promptly in response to changes in the number of treated patients. Thus, hospitals with a relatively low occupancy rate (*occrate*) are expected to have an oversized staff that is unlikely to meet the current demand for inpatient care efficiently. The mortality rate (*mort*) is used as a proxy for poor quality (e.g. Propper et al., 2004). Differences in the hospitals' budgets are controlled by total expenses (in 2005 prices) per bed (*budget*). The management of hospitals with a more complex service-structure is likely to face additional difficulties to organize the production efficiently (e.g. Farsi and Filippini, 2008, and Lee et al., 2008). The degree of specialization (*spec*) is measured by an information theory index (Evans and Walker, 1972) in terms of differences between the national and hospital's proportions of cases belonging to several clinical departments.

District/state specific variables: Treatments of older people are likely to be more cost- and resource-intensive, because they are often accompanied with higher degrees of comorbidity and complications (Augurzky et al., 2006). Several authors (e.g. Herr, 2008, Chang, 1998) address the influence of the patients' age structure on hospital efficiency and find that higher

the number of full time equivalent employees are used instead of (crude) numbers of employees. However, for this measure data are only available for physicians and non-physicians. Thus, this input specification might neglect more heterogeneity across the hospitals in respect to their staffing mix in comparison with the initial specification of number of employed physicians, nurses and non-medical employees. In summary, the results regarding $H(A)$ and $H(B)$ are qualitatively very similar across all considered specifications.

proportions of older patients increase the inefficiency. The fraction of people aged over 65 years and living in a hospitals' district is considered as a demographic variable (*age*65). The degree of the district's urbanization is captured by the population density (population per square kilometer) (*popdens*) and exogenous socioeconomic factors are controlled by the district's GDP per capita (*gdp*).

In each federal state of Germany, a commission composed of members of the state government and health insurances creates the hospital requirement and financing plan ("Krankenhausbedarfsplanung") for providing inpatient care to the population in the hospitals' service area (Mörsch, 2010). Hence, hospitals which are in the same state are confronted with the same regional legal requirements. To account for this type of observable heterogeneity, we include state dummy variables, with North Rhine-Westphalia serving as reference. In the fixed effects specification any type of time invariant heterogeneity between the hospitals cancels out by the within transformation. According to § 4 of the Hospital Financing Act (*Krankenhausfinanzierungsgesetz-KHG*) the financial support in the German hospital sector is dualistic, i.e. operating costs are paid by insurance companies, while investments are funded by federal states (§ 9 KHG). Thus, financial stress in the federal states could reduce the financial means for investments and might influence the creation of the hospital requirement and financing plan. In the cross-sectional model, the state dummy variables take into account all kinds of variations between the federal states. To control for variations of the federal states' financial situation over time in the panel model, we further include the debts of the federal states per GDP (*debt*).

3.4.4 Spatial weights matrices

To address robustness of the empirical results, two alternative weights matrices are used to implement W_t and M_t in (2). The elements $w_{ijt} = w^*_{ijt}/\sum_{j=1}^{N_t} w^*_{ijt}$ and $m_{ijt} = m^*_{ijt}/\sum_{j=1}^{N_t} m^*_{ijt}$ are built on binary matrices, with $w^*_{ijt} = 1$ and $m^*_{ijt} = 1$, if the i-th and the j-th hospital are contiguous, respectively. The definition of contiguity differs across alternative weights matrices. The first concept, denoting W_d and M_d, is to define hospitals as contiguous to each other if they are located in the same district. For the second set of weights matrices, W_n and M_n, two hospitals are considered contiguous if they are either located in the same district, or if their respective districts of residence are neighbors.

4 Results

Firstly, the results of the unbalanced fixed effects panel regression model of pure technical efficiency change and diagnostic results for $H(A)$ are considered. Afterwards, we turn to the second hypothesis $H(B)$ and discuss the results of the cross-sectional regression model of DEA efficiency scores. In order to gain additional information about spatial dependence in the data, we also apply more parsimoniously parameterized model specifications for both, the panel and cross-sectional framework. A model neglecting spatial dependence is denoted by OLS, furthermore the spatial error model (SEM), $\lambda = 0$ and the spatial lag model (SLM), $\rho = 0$ are estimated. Descriptive statistics (means and standard deviations) for the variables included in the empirical analyses and the number of hospitals entering the first (DEA and efficiency change) and second modeling stage (regression analysis) are documented in Table 1. Since there are no substantial differences between the statistics of the explanatory variables of the panel and cross-sectional model, they are only given for the panel regression model. For the regression analysis some of the introduced regressors (*ms*, *mort*, *occrate*, *budget*, *gdp*, *popdens* and *dept*) are measured in natural logarithms.

4.1 Pure technical efficiency change panel model

Table 2 displays the results of the panel regression models explaining the pure technical efficiency change derived by DEA. First of all, the spatial regression models obtain higher log likelihood statistics in comparison with OLS model evaluation. The best fit is achieved by means of the SARAR model under W_d and M_n. The SEM and SARAR model yield significantly positive spatial error correlation estimates if the spatial error process is modeled by means of M_n. Spatial spillovers might have less importance to explain the pattern of hospital efficiency change. Applying the district spatial weights matrix for the spatial lag process the SARAR model

Table 1: Mean values and standard deviations of selected variables

	95-06	95-99	00-03	04-06
output variables				
weighted cases (in 1000)	5.47	5.30	5.65	5.47
	(5.13)	(4.56)	(5.46)	(5.42)
apprentices	45.5	54.4	36.6	44.9
	(70.9)	(71.7)	(66.8)	(73.3)
input variables				
number of beds	271	298	259	251
	(230)	(227)	(232)	(227)
number of physicians	50.8	49.1	48.4	56.2
	(60.9)	(53.1)	(60.8)	(69.9)
number of nurses	196	211	189	184
	(188)	(183)	(193)	(187)
number of non-medical staff	211	230	202	197
	(228)	(231)	(230)	(219)
material expenses (in 2005 prices and mio. €)	12.4	15.1	11.7	9.62
	(14.9)	(16.1)	(15.1)	(12.0)
DEA results				
technical DEA efficiency scores	51.8	54.2	54.5	44.9
	(22.2)	(23.3)	(20.2)	(21.6)
hospitals (first-stage: DEA)	20372	7569	7259	5544
pure technical efficiency change	1.01	1.15	0.93	0.98
	(0.61)	(0.66)	(0.68)	(0.36)
hospitals (second-stage: eff. change)	17955	5660	6855	5440
explanatory variables (panel model)				
market share	24.9	25.1	24.3	25.4
	(25.2)	(24.5)	(24.7)	(26.3)
specialization index	0.95	0.85	0.94	1.06
	(1.00)	(1.00)	(0.98)	(1.01)
mortality	2.65	2.80	2.64	2.50
	(2.29)	(2.27)	(2.43)	(2.10)
occupancy rate	78.1	80.3	78.9	74.6
	(11.6)	(7.6)	(8.9)	(16.6)
hospital budget per bed (in 2005 prices and 1000 €)	95.3	87.3	95.6	104
	(33.4)	(27.6)	(33.1)	(37.5)
plus 65 ratio (in%)	17.6	16.2	17.5	19.4
	(2.2)	(1.76)	(1.81)	(1.96)
federal state's GDP per capita (in 1000 €)	25.6	23.7	25.9	27.4
	(10.7)	(10.0)	(10.7)	(11.2)
population density (population per km^2)	832	883	813	800
	(1020)	(1039)	(1014)	(1002)
federal state's depts per GDP	18.9	17.5	18.3	21.2
	(10.4)	(7.65)	(10.3)	(12.5)
hospitals (second-stage: panel model)	16097	5297	6116	4684
hospitals (second-stage: cross-sectional model)	18221	7032	6404	4785

Standard deviations in parentheses.

Table 2: Regression results of unbalanced fixed effects panel models

	OLS	SEM$\|_{\lambda=0}$		SLM$\|_{\rho=0}$		SARAR			
		M_d	M_n	W_d	W_n	W_d&M_d	W_n&M_n	W_d&M_n	W_n&M_d
1997	0.734**	0.734**	0.732**	0.727**	0.598**	0.760**	0.729**	0.736**	0.600**
1998	0.099**	0.099**	0.096**	0.098**	0.081**	0.102**	0.096**	0.096**	0.081**
1999	0.002	0.002	−0.001	0.002	0.002	0.002	−0.001	−0.001	0.002
2000	0.098**	0.097**	0.091**	0.097**	0.080**	0.100**	0.091**	0.092**	0.080**
2001	0.009	0.008	0.002	0.009	0.009	0.008	0.002	0.002	0.009
2002	−0.012	−0.013	−0.021	−0.012	−0.004	−0.015	−0.021	−0.022	−0.004
2003	0.177**	0.177**	0.164**	0.176**	0.153**	0.181**	0.164**	0.164**	0.153**
2004	0.026	0.026	0.012	0.026	0.033*	0.024	0.013	0.012	0.033*
2005	0.221**	0.221**	0.203**	0.220**	0.193**	0.225**	0.203**	0.203**	0.194**
2006	0.374**	0.373**	0.353**	0.371**	0.320**	0.383**	0.352**	0.354**	0.320**
private	−0.011	−0.011	−0.009	−0.011	−0.010	−0.011	−0.009	−0.009	−0.011
non-profit	−0.000	0.000	−0.000	−0.000	−0.000	0.001	−0.000	−0.000	−0.000
ln(*ms*)	0.065**	0.065**	0.064**	0.065**	0.064**	0.066**	0.064**	0.064**	0.064**
spec	−0.088**	−0.088**	−0.086**	−0.088**	−0.087**	−0.087**	−0.086**	−0.086**	−0.087**
ln(*mort*)	−0.040**	−0.040**	−0.040**	−0.040**	−0.040**	−0.041**	−0.040**	−0.040**	−0.040**
ln(*occrate*)	0.150**	0.149**	0.152**	0.150**	0.150**	0.149**	0.152**	0.152**	0.149**
ln(*budget*)	−0.218**	−0.216**	−0.211**	−0.217**	−0.213**	−0.215**	−0.211**	−0.211**	−0.212**
*age*65	−0.019**	−0.020**	−0.016**	−0.019**	−0.017**	−0.020**	−0.016**	−0.016**	−0.017**
ln(*gdp*)	−0.070	−0.070	−0.061	−0.071	−0.071	−0.068	−0.061	−0.061	−0.071
ln(*popdens*)	0.038	0.037	−0.007	0.037	0.012	0.039	−0.007	−0.008	0.012
ln(*dept*)	−0.258**	−0.257**	−0.243**	−0.255**	−0.220**	−0.266**	−0.243**	−0.245**	−0.221**
ρ	-	0.017*	0.247**	-	-	0.055**	0.243**	0.251**	0.003
λ	-	-	-	0.010	0.185**	−0.039	0.004	−0.006	0.183**
$LOGLIKE^a$	148.5	150.2	171.6	149.1	168.0	151.5	171.6	171.8	168.0
R^2_{adj} (in %)	44.79	44.81	44.56	44.01	30.26	47.97	44.19	45.02	30.41
spatial correlation tests									
LM_E^d	3.231*	0.002	0.001	0.640	0.106	0.034	0.002	0.377	0.000
LM_L^d	1.158	0.354	0.266	0.001	0.365	0.001	0.274	0.009	0.825
LM_E^n	63.383**	53.783**	0.372	57.640**	1.248	55.217**	0.358	0.357	1.218
LM_L^n	49.968**	42.199**	0.001	44.796**	0.075	46.309**	0.000	0.009	0.078
average time effect estimates									
$\bar{\delta}_1$ (96-99)	0.209	0.209	0.207	0.207	0.170	0.216	0.206	0.208	0.171
$\bar{\delta}_2$ (00-03)	0.068	0.067	0.059	0.067	0.059	0.068	0.059	0.059	0.059
$\bar{\delta}_3$ (04-06)	0.207	0.207	0.189	0.206	0.182	0.210	0.189	0.190	0.182
$\bar{\delta}_4$ (96-03)	0.158	0.158	0.152	0.157	0.131	0.163	0.151	0.153	0.132
$\bar{\delta}_5$ (00-06)	0.128	0.127	0.115	0.127	0.112	0.129	0.115	0.115	0.112
test statistics									
$\bar{\delta}_1 < \bar{\delta}_2$ (AE)	15.448	15.278	14.388	15.290	12.180	15.533	14.343	14.452	12.191
$\bar{\delta}_2 < \bar{\delta}_3$ (IE$_1$)	−12.779**	−12.604**	−10.701**	−12.698**	−11.268**	−12.534**	−10.710**	−10.707**	−11.254**
$\bar{\delta}_4 < \bar{\delta}_3$ (IE$_2$)	−4.816**	−4.735**	−3.626**	−4.804**	−4.705**	−4.600**	−3.646**	−3.610**	−4.691**
$\bar{\delta}_1 < \bar{\delta}_5$ (OE)	6.402	6.351	6.651	6.318	4.608	6.574	6.609	6.702	4.622

Significance level: ** 5%; * 10%; ^{a} significance levels are given for log-likelihood ratio tests against OLS; AE: Announcement effect; IE$_1$: Introduction effect under consideration of AE; IE$_2$: Introduction effect under the assumption of no AE; OE: Overall effect; LM_E^d, LM_L^d, LM_E^n, LM_L^n denote the LM test for spatial error (E) and lag (L) dependence under the district (d) and neighborhood (n) weights matrix; 1996 serves as reference; R^2_{adj} is the adjusted degree of explanation; estimation based on 16097 observations.

yields insignificantly negative spatial spillover estimates. In contrast, under W_n the spatial lag estimates become positive and significant if the spatial error structure is modeled by means of W_d in the SARAR model or neglected in the SLM. However, as indicated by the respective (substantially lower) log likelihood statistics this might be explained by a false specification of the spatial error structure. Spatial dependence is also confirmed by means of a Lagrange Multiplier (LM) test for spatial error (LM_E) and lag (LM_L) dependence (Anselin, 1988). Under W_n and M_n, both tests are highly significant for OLS and the spatial models applying W_d and M_d. In summary, the results indicate spatial error dependence formalized by means of the neighborhood spatial weights matrix as most appropriate to describe spatial patterns. This finding underpins that hospitals which are in the same region have similar opportunities and constraints (e.g. market characteristics, composition of patients, regional legal requirements) implying spatial clustering of hospital efficiency change. However, the values of the regression coefficients do not vary strongly across the models.

The results of the empirical testing strategy for $H(A)$ reveal a negative effect of the DRG announcement on hospital efficiency change. Moreover, the negative announcement effect dominates the overall effect (of the reform announcement and introduction), which also appears to be negative. However, the DRG introduction (by itself) is found to have a significantly positive effect on hospital efficiency gains. The time effect estimates are, on average, significantly higher for the post-reform period (2004-2006) than for the pre-reform (1996-2003) and post-announcement-pre-reform (2000-2003) period.[5]

4.2 Cross-sectional efficiency

Log likelihood values for the cross-sectional regression models are reported in Table 3. The SARAR model under W_d and M_d obtains the highest statistics for almost all years and these are significantly higher than their OLS counterparts. In contrast to the efficiency change model,

[5]Noting that the increase in time effect estimates might be driven by poorly performing hospitals exiting the market we address this issue by means of estimating a binary response model (Probit and Logit). The results do not indicate a systematic and significant effect of a hospital's performance in period t on the probability to exit the market in $t+1$ over the considered time periods. Therefore, we do not believe in a selection bias which is responsible for the increased overall hospital efficiency.

the district based spatial weights matrix appears to be most appropriate to model the spatial error and lag structure of hospital efficiency.

Table 3: Comparison of cross-sectional models

year	obs.	OLS	SEM$\|_{\lambda=0}$		SLM$\|_{\rho=0}$		SARAR			
			W_d	W_n	W_d	W_n	W_d&M_d	W_n&M_n	W_d&M_n	W_n&M_d
1995	1355	−1003.4	−1002.9	−1003.1	−1003.3	−1003.2	−1002.5	−1003.1	−1003.1	−1002.7
1996	1359	−706.2	−706.0	−705.7	−705.9	−705.8	−705.9	−705.7	−705.6	−705.7
1997	1354	−373.6	−373.3	−373.3	−371.8*	−372.1*	−366.2**	−371.9	−371.8	−371.8
1998	1430	−365.2	−365.0	−365.0	−363.6*	−364.7	−359.0**	−364.6	−363.6	−364.6
1999	1534	−362.3	−361.4	−362.2	−360.2**	−361.7	−349.3**	−361.1	−360.1	−360.6
2000	1544	−438.0	−437.7	−437.6	−437.0	−437.8	−432.6**	−437.6	−436.9	−437.5
2001	1583	−646.9	−645.1*	−645.4*	−646.8	−644.9**	−640.7**	−644.8	−645.4	−642.7**
2002	1653	−477.8	−473.4**	−477.3	−473.7**	−475.8**	−447.1**	−474.4**	−473.1**	−470.6**
2003	1624	−516.6	−514.4**	−516.4	−512.1**	−514.9*	−490.2**	−514.3	−512.0**	−512.1**
2004	1632	−889.7	−887.1**	−889.5	−889.1	−886.8**	−877.1**	−885.3**	−889.0	−883.6**
2005	1604	−845.6	−844.9	−845.3	−842.8**	−841.2**	−832.6**	−838.8**	−842.4**	−840.2**
2006	1549	−702.4	−702.2	−701.9	−697.5**	−702.1	−684.8**	−700.7	−696.6**	−702.0

Significance level: ** 5%; * 10%.

In the following, we examine the regression variables to explain the variation in hospital efficiency. Similar to the panel model, the estimated coefficients do not vary markedly across the distinct model specifications. Therefore, we concentrate on the results of the best fitting model as indicated by the highest log likelihood statistics. Table 4 provides the regression results of the SARAR model under W_d and M_d. The estimation results reveal private hospitals to be less efficient than their public counterparts after 2001. This is in line with other empirical findings (Herr, 2008, Farsi and Filippini, 2008, and Helmig and Lapsley, 2001). Before 2002 private for profit hospitals seem to be more efficient than public hospitals. However, for most periods the estimates are not significant. Interestingly, there is no convergence in efficiency of private hospitals to the public counterparts after the introduction of DRG. This result is in contrast to the findings of Barbetta et al. (2007), who analyze hospital efficiency in Italy. Since profit incentives are no longer associated with an inefficiently long hospital stay (Herr, 2008), the result rather supports a conjecture of Helmig and Lapsley (2001). They argue that local governments sell the most inefficient hospitals to private investors, while holding the more efficient

Table 4: Regression results of cross-sectional SARAR(W_d,M_d) model

Regressor	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
intercept	3.676**	2.562**	2.192**	2.145**	1.792**	0.811**	0.791*	−0.548	−0.948**	−1.080**	−0.060	0.868*
SH	-	-	-	-	-	-	0.231**	0.084	0.088	0.114	0.049	0.126
HH	0.181*	0.073	0.031	0.010	−0.018	−0.011	−0.013	0.069	0.262**	0.215*	0.330**	0.226**
NI	0.235**	0.120**	0.027	0.001	0.009	0.014	0.049	−0.084*	−0.034	0.008	0.004	0.045
HB	−0.090	−0.056	−0.123	−0.009	−0.006	−0.063	−0.025	−0.067	−0.131	−0.114	−0.149	−0.158
HE	0.221**	0.190**	0.120**	0.082*	0.035	0.042	0.056	−0.089*	−0.091*	−0.041	−0.122**	0.002
RP	0.292**	0.217**	−0.016	−0.047	−0.049	−0.052	−0.026	−0.129**	−0.111**	−0.002	0.051	0.089
BW	0.304**	0.156**	−0.106**	−0.092**	−0.103**	−0.094**	−0.117**	−0.298**	−0.253**	−0.227**	−0.181**	−0.174**
BY	0.474**	0.339**	0.050	0.030	0.005	0.026	0.037	−0.217**	−0.150**	−0.179**	−0.154**	−0.080*
SL	0.356**	0.245**	0.157*	−0.015	0.054	0.037	0.173*	0.055	0.050	0.121	0.138	0.154
BE	−0.005	−0.050	0.046	0.102	0.060	0.023	0.232**	0.367**	0.416**	0.516**	0.499**	0.393**
BB	−0.196**	−0.111*	−0.067	−0.080	0.041	0.023	0.128*	0.078	0.148**	0.217**	0.184**	0.207**
MV	−0.016	0.094	0.145**	-	-	-	-	0.268**	0.240**	0.210**	0.182*	0.245**
SN	−0.116	−0.065	−0.025	−0.002	0.038	0.009	0.062	−0.037	−0.034	−0.025	−0.025	−0.028
ST	0.056	−0.002	0.024	−0.076	0.005	0.023	0.099	0.032	0.046	0.034	0.072	0.027
TH	−0.335**	−0.303**	−0.113*	−0.081	−0.022	−0.007	0.096	0.026	0.030	−0.044	−0.015	−0.059
private	−0.059	0.025	0.161**	0.102**	0.023	0.036	0.083**	−0.119**	−0.082**	−0.148**	−0.113**	−0.152**
non-profit	0.030	0.018	0.024	0.028	−0.004	−0.009	−0.001	−0.115**	−0.087**	−0.151**	−0.179**	−0.132**
$\ln(ms)$	0.004	0.010	0.157**	0.155**	0.161**	0.128**	0.129**	0.235**	0.238**	0.234**	0.223**	0.205**
spec	0.019	−0.052**	−0.098**	−0.121**	−0.136**	−0.162**	−0.175**	0.080**	0.075**	0.122**	0.112**	0.101**
$\ln(mort)$	0.019	−0.012	0.007	−0.021*	−0.050**	−0.055**	−0.043**	−0.023**	−0.037**	−0.023*	−0.017	−0.032**
$\ln(occrate)$	−0.174	−0.074	−0.355**	−0.114	−0.013	−0.020	−0.145**	0.427**	0.387**	0.171**	0.419**	0.535**
$\ln(budget)$	−0.399**	−0.336**	−0.179**	−0.257**	−0.308**	−0.199**	−0.173**	−0.375**	−0.336**	−0.287**	−0.448**	−0.547**
$age65$	−0.002	−0.000	−0.008	−0.009	−0.002	0.006	0.006	0.010	0.020**	0.016**	0.014*	0.011
$\ln(gdp)$	0.073	0.072	0.149**	0.166**	0.196**	0.157**	0.128**	0.262**	0.228**	0.303**	0.249**	0.254**
$\ln(popdens)$	0.018	0.022	0.023	0.015	0.048**	0.031**	0.068**	0.079**	0.082**	0.106**	0.103**	0.083**
ρ	0.074	−0.002	0.187**	0.176**	0.225**	0.154**	0.185**	0.290**	0.274**	0.228**	0.205**	0.237**
λ	−0.046	−0.016	−0.193**	−0.186**	−0.220**	−0.153**	−0.145**	−0.262**	−0.267**	−0.186**	−0.199**	−0.248**
$LOGLIKE$	−1002.5	−705.9	−366.2	−359.0	−349.3	−432.6	−640.7	−447.1	−490.2	−877.1	−832.6	−684.8
R^2_{adj} (in %)	14.24	12.68	34.94	37.21	39.49	35.97	33.72	45.74	42.79	31.45	33.37	35.52
N	1355	1359	1354	1430	1534	1544	1583	1653	1624	1632	1604	1549

Significance level: ** 5%; * 10%; SH: Schleswig-Holstein, HH: Hamburg, NI: Lower Saxony, HB: Bremen, HE: Hesse, RP: Rhineland-Palatinate, BW: Baden-Wuerttemberg BY: Bavaria, SL: Saarland, BE: Berlin, BB: Brandenburg, MV: Mecklenburg-Western Pomerania, SN: Saxony, ST: Saxony-Anhalt, TH: Thuringia.

ones. Furthermore, a positive relationship between market share and efficiency is found.[6] The relationship between specialization and efficiency has changed over time. After 2002, a specialized hospital is, on average, less inefficient in comparison with a non-specialized hospital. Due to the financial reform a trend towards specialization (Knorr, 2003) and merger (Rocke, 2003) is expected, which might lead to an increased market share of the involved hospitals. Thus, the results support these strategies as promising options to increase hospital performance. The occupancy rate has a significantly positive parameter estimate for most years, implying that hospitals which are fully stretched are less inefficient. Furthermore, high mortality and the budget size are correlated with higher inefficiencies, while hospital performance appears invariant with regard to the age structure of the district's population. Noting that the estimated coefficients are mostly in line with findings of related studies (e.g. Herr, 2008, Farsi and Filippini, 2008, Lee et al., 2008, Chang, 1998), we believe that the explanatory factors control appropriately for heterogeneity among hospital performance and offer the correct identification of potential spatial dependence patterns.

Table 5 displays spatial parameter estimates for distinct model specifications. Due to space considerations, SARAR model results are only shown for specifications with the same applied pattern for the spatial lag and error dependence, since log-likelihood statistics do not indicate an obvious priority to any specification, except to the choice of W_d and M_d. First of all, the results confirm the presence of negative spatial spillovers and positive spatial error correlation, irrespective of the spatial weights matrix. The spatial parameter estimates of the restricted models (SEM and SLM) are of a lower magnitude in comparison with the respective estimates of the SARAR model. This might be explained by the fact of having only one channel of spatial dependence in these models. Positive spatial error correlation mitigates the negative spatial lag estimates and vice versa. Under W_d and M_d, the SARAR model obtains significant spatial parameter estimates for almost all years. If the broader concept of spatial contiguity is applied,

[6] Augurzky et al. (2006) mention the importance of the geographic area where the market share is built up. In a rural area a higher market share can be the result of being the only provider of inpatient treatments leading potentially to an inefficient production of medical care, due to the lack of competitors. At the opposite, a higher market share in an urban area can be the result of an efficient performance. We incorporate an interaction of the variable ms and an agglomeration dummy variable. However, there is no considerable difference between the impact of rural or urban market share.

W_n and M_n, estimated spatial parameters are larger in absolute value, but insignificant until 2001.

Table 5: Spatial correlation estimates

year	$\hat{\rho}\vert_{\lambda=0}$ (SEM)		$\hat{\lambda}\vert_{\rho=0}$ (SLM)		$\hat{\rho}$ (SARAR)		$\hat{\lambda}$ (SARAR)	
	W_d	W_n	W_d	W_n	W_d&M_d	W_n&M_n	W_d&M_d	W_n&M_n
1995	0.031	−0.098	0.008	−0.070	0.074	−0.087	−0.046	−0.010
1996	−0.018	−0.154	−0.018	−0.115	−0.002	−0.108	−0.016	−0.037
1997	0.019	−0.102	−0.055*	−0.249	0.187**	0.152	−0.193**	−0.350
1998	0.009	−0.062	−0.051*	−0.128	0.176**	0.089	−0.186**	−0.188
1999	0.042	0.045	−0.056**	−0.165	0.225**	0.203	−0.220**	−0.276
2000	0.022	−0.129	−0.038	−0.072	0.154**	−0.125	−0.153**	−0.000
2001	0.057*	−0.305	−0.004	−0.296*	0.185**	−0.105	−0.145**	−0.227
2002	0.096**	0.148	−0.075**	−0.291*	0.290**	0.294*	−0.262**	−0.410**
2003	0.065**	0.042	−0.078**	−0.262*	0.274**	0.192	−0.267**	−0.356*
2004	0.070**	0.066	−0.025	−0.375**	0.228**	0.303*	−0.186**	−0.578**
2005	0.034	0.107	−0.062**	−0.494**	0.205**	0.360**	−0.199**	−0.738**
2006	0.012	0.132	−0.081**	−0.091	0.237**	0.326*	−0.248**	−0.325
average spatial spillover estimates								
$\bar{\lambda}_1$ (95-99)	-	-	−0.034	−0.145	-	-	−0.132	−0.172
$\bar{\lambda}_2$ (00-03)	-	-	−0.049	−0.230	-	-	−0.206	−0.248
$\bar{\lambda}_3$ (04-06)	-	-	−0.056	−0.320	-	-	−0.211	−0.547
$\bar{\lambda}_4$ (95-03)	-	-	−0.041	−0.183	-	-	−0.165	−0.206
$\bar{\lambda}_5$ (00-06)	-	-	−0.052	−0.269	-	-	−0.208	−0.376
test statistics								
$\bar{\lambda}_1 > \bar{\lambda}_2$ (AE)	-	-	0.762	0.846	-	-	2.462**	0.522
$\bar{\lambda}_2 > \bar{\lambda}_3$ (IE$_1$)	-	-	0.363	0.744	-	-	0.146	1.817**
$\bar{\lambda}_4 > \bar{\lambda}_3$ (IE$_2$)	-	-	0.859	1.293*	-	-	1.684**	2.280**
$\bar{\lambda}_1 > \bar{\lambda}_5$ (OE)	-	-	1.060	1.395*	-	-	2.857**	1.538*

Significance level: ** 5%; * 10%.

Finally, the diagnostic results of the SARAR model for $H(B)$ are discussed in detail. For all spatial specifications, the magnitude of average spatial spillover estimates increases over time. However, the test results differ across the spatial weights matrices. Under W_d an announcement effect of the DRG reform is identified while an effect of the DRG introduction is only detected if the AE is neglected (IE$_2$). This is in contrast to the results obtained under W_n. There is no evidence for an effect in response to the announcement, but to the DRG introduction, irrespective if the AE is taken into account (IE$_1$) or not (IE$_2$). However, under both spatial specifications an overall effect is detected. In summary the results convey the expected rise of competition for low cost patients invoked by the DRG reform.

The competition for low cost patients might have several effects. On the one hand, in order to attract patients, hospitals have to acquire reputation by quality of care, service, room facilities etc. On the other hand, hospitals treating patients with high complexities receive inappropriate cost reimbursements and may experience solvency problems (Böcking, 2005). To save costs they might decrease the quality of treatment. Several studies (e.g. Perelman et al., 2008, Picone, 2003) find a positive relationship between social deprivation and the length of hospital stay, e.g. due to higher complexities (Krieger et al., 1997). Thus, there might be a cost differential between underprivileged and well-off patients, which is not taken into account by the German hospital cost reimbursement. This typically yields an implicit patient selection of the hospitals with consequences on social equity in health (Perelman et al., 2008). To avoid such behavior, Perelman et al. (2008) suggest to integrate the impact of socio-economic status on length of stay to the cost reimbursement.

5 Conclusions

This study is the first approach that considers spatial interdependence of hospital efficiency in Germany for the period 1995 to 2006 that includes the announcement (2000) and introduction (2004) of the DRG based financing system. In particular, two hypotheses about potential effects of the financial reform on overall hospital efficiency gains and spatial interdependence of hospital performance are examined.

Accounting for observed and hidden hospital characteristics, we find an increased growth of efficiency after the DRG introduction. Noting that there have been no major exogenous shocks affecting hospital efficiency during the period of study, the results confirm the intention of the reform to improve the efficiency of the health care system. Furthermore, the results reveal two distinct channels of spatial interdependence of hospital performance, i.e. positive spatial error correlation and negative spatial spillovers. While the former could be explained by similar opportunities and constraints of nearby hospitals, the latter might occur in response to competition between the hospitals. Moreover, the increase in the magnitude of negative spatial spillovers is in line with an expected rise of competition for low cost patients invoked by the reform of the financing system.

The increase of efficiency after the introduction of the prospective payment system could be achieved by opportunistic practices of the hospitals. For instance, they could refer cases prematurely to other health care institutions (e.g. rehabilitation centers) or readmit the patients (Böcking et al., 2005). In order to account for such a behavior it might be important to incorporate in the efficiency measurement information about the quality of treatments and hospital stay. Future research should consist in constructing adequate quality adjusted efficiency measures. Another important issue is to analyze the competitive behavior of the hospitals. The increased negative spatial spillovers among hospital performance after the DRG reform might be explained by an increased competition for low cost patients, implying the practice of patient selection. The consequences for so called high cost patients are of particular interest in order to derive policies for targeting an equal access to inpatient care.

References

Anderson, P. and N. C. Petersen (1993). A procedure for ranking efficient units in Data Envelopment Analysis. *Management Science* 39(10), 1261-1264.

Anselin, L. (1988). Spatial econometrics: methods and models (Boston: Kluwer Academic Publishers).

Augurzky, B., D. Engel and C. Schwierz (2006). Who gets the Credit? Determinants of the probability of default in the German hospital sector. RWI Discussion Papers 0054, Rheinisch-Westfälisches Institut für Wirtschaftsforschung.

Banker, R. D. and R. C. Morey (1986). Efficiency analysis for exogenously fixed inputs and outputs. *Operations Research* 34(4), 513-521.

Banker, R. D., A. Charnes and W.W. Cooper (1984). Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis. *Management Science* 30, 1078-1092.

Barbetta, G. P., G. Turati and A. M. Zago (2007). Behavioral differences between public and private not-for-profit hospitals in the Italian national health service. *Health Economics* 16(1), 75-96.

Biorn, E., T. P. Hagen, T. Iversen and J. Magnussen (2003). The effect of activity-based financing on hospital efficiency: a panel data analysis of DEA Efficiency Scores 1992-2000. *Health Care Management Science* 6(4), 271-283.

Biorn, E., T. P. Hagen, T. Iversen and J. Magnussen (2010). How different are hospitals' responses to a financial reform? The impact on efficiency of activity-based financing. *Health Care Management Science* 13(1), 1-16.

Böcking, W., U. Ahrens, W. Kirch and M. Milakovic (2005). First results of the introduction of DRGs in Germany and overview of experience from other DRG countries. *Journal of Public Health* 13(3), 128-137.

Burgess, J. and P. Wilson (1995). Decomposing hospital productivity changes, 1985-1988: A nonparametric Malmquist approach. *Journal of Productivity Analysis* 6, 343-363.

Chang, H. (1998). Determinants of hospital efficiency: the case of central government-owned Hospitals in Taiwan. *Omega* 26(2), 307-317

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

Dranove, D. and R. Ludwick (1999). Competition and pricing by nonprofit hospitals: a re-assessment of Lynk's analysis. *Journal of Health Economics* 18(1), 87-98.

Ellis, R. P. (1998). Creaming, skimping and dumping: provider competition on the intensive and extensive margins. *Journal of Health Economics* 17(5), 537-555.

Ellis, R. P. and T. G. McGuire (1996). Hospital response to prospective payment: Moral hazard, selection, and practice-style effects. *Journal of Health Economics* 15(3), 257-277.

Evans, R. G. and H. D. Walker (1972). Information theory and the analysis of hospital cost structure. *The Canadian Journal of Economics* 5(3), 398-418.

Färe, R., S. Grosskopf, B. Lindgren, P. Roos (1992). Productivity changes in Swedish pharmacies 1980-1989: A non-parametric Malmquist approach. *Journal of Productivity Analysis* 3, 85-101.

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

Helmig, B. and I. Lapsley (2001). On the efficiency of public, welfare and private hospitals in Germany over time - A sectoral DEA-study. *Health Services Management Research* 14(4), 263-274.

Hensen, P., S. Beissert, L. Bruckner-Tuderman, T. A. Luger, N. Roeder and M. L. Müller (2008). Introduction of diagnosis-related groups in Germany: evaluation of impact on inpatient care in a dermatological setting. *The European Journal of Public Health* 18(1), 85-91.

Herr, A. (2008). Cost and technical efficiency of German hospitals: does ownership matter? *Health Economics* 17(9), 1057-1071.

Johnson, A. L. and L. F. McGinnis (2008). Outlier detection in two-stage semiparametric DEA models. *European Journal of Operational Research* 187(2), 629-635.

Kneip, A., B.U. Park and L. Simar (1998). A note on the convergence of nonparametric DEA estimators for production efficiency scores. *Econometric Theory* 14, 783-793.

Knorr, G. (2003). Probleme der Grundversorgungskrankenhäuser im DRG-System - Spezialisierung als Ausweg? *Krankenhaus* 9, 679-682.

Krieger, N., D. R. Williams and N. E. Moss (1997). Measuring social class in US public health research: Concepts, methodologies, and guidelines. *Annual Review of Public Health* 18, 341-378.

Lee, K., K. Chun and J. Lee (2008). Reforming the hospital service structure to improve efficiency: Urban hospital specialization. *Health Policy* 87(1), 41-49.

Linna, M. (2000). Health care financing reform and the productivity change in Finnish hospitals. *Journal of Health Care Finance* 26(3), 83-100.

Lungen, M. and I. Lapsley (2003). The reform of hospital financing in Germany: an international solution? *Journal of Health Organization and Management* 17(5), 360-372.

Mörsch, M. (2010). Bestandsaufnahme zur Krankenhausplanung und Investitionsfinanzierung in den Bundeslandern. Deutsche Krankenhausgesellschaft, Dezernat II, Berlin.

Norton, E. C., C. H. Van Houtven, R. C. Lindrooth, S.-L. T. Normand and B. Dickey (2002). Does prospective payment reduce inpatient length of stay? *Health Economics* 11(5), 377-387.

Perelman, J. A. Shmueli and M.-C. Closon (2008). Deriving a risk-adjustment formula for hospital financing: Integrating the impact of socio-economic status on length of stay. *Social Science & Medicine* 66, 88-98.

Picone G., R. M. Wilson and S.-Y. Chou (2003). Analysis of hospital length of stay and discharge destination using hazard functions with unmeasured heterogeneity. *Health Economics* 12(12), 1021-1034.

Propper, C., S. Burgess and K. Green (2004). Does competition between hospitals improve the quality of care? Hospital death rates and the NHS internal market. *Journal of Public Economics* 11(7-8), 1247-1272.

Rocke, B. (2003). Zur Theorie und Praxis der Kooperationen und Fusionen im Krankenhausbereich. *Krankenhaus* 7, 531-535.

Simar, L and P. W. Wilson (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Jounal of Econometrics* 136(1), 31-64.

Sommersguter-Reichmann, M. (2000). The impact of the Austrian hospital financing reform on hospital productivity: empirical evidence on efficiency and technology changes using a non-parametric input-based Malmquist approach," *Health Care Management Science* 3(4), 309-321.

Statistisches Bundesamt (2008). Krankenhaus- Landschaft im Umbruch. Begleitmaterial zur Pressekonferenz am 10. Dezember 2008 in Berlin.

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

Town, R. and G. Vistnes (2001). Hospital competition in HMO networks. *Journal of Health Economics* 20(5), 733-753.

Wilson, P.W. (1995). Detecting influential observations in data envelopment analysis. *Journal of Productivity Analysis* 6(1), 27-45.

Zweifel, P., F. Breyer and M. Kifmann (2009). Health Economics (Berlin: Springer).

A Calculation of DEA efficiency scores

The estimated super efficiency score, $\theta^V_{i,t|t}$, is obtained under the assumption of variable returns to scale (Banker et al., 1984) by solving the following linear program

$$
\begin{aligned}
\theta^V_{i,t|t} = \arg\min_{\theta^V_{i,t|t},\nu} \{\theta^V_{i,t|t} > 0| \sum_{l\neq i} \nu_{lt} q_{plt} \geq q_{pit} \quad &\forall \quad p \in \{1,...,s\} \\
\theta^V_{i,t|t} x^D_{jit} \sum_{l\neq i} \nu_{lt} x^D_{jlt} \quad &\forall \quad j \in \{1,...,m^D\} \\
x^N_{kit} \geq \sum_{l\neq i} \nu_l x^N_{klt} \quad &\forall \quad k \in \{1,...,m^N\} \\
\sum_{l\neq i} \nu_{lt} = 1,\ \nu_{lt} > 0 \quad &\forall \quad l = 1,...,N_t\},
\end{aligned}
$$

where q_{rit}, x^N_{kit} and x^D_{jit} denote output, non-discretionary and discretionary input variables of hospital i at time t. The numbers of outputs, non- and discretionary inputs, and reference hospitals at time t are s, m^N, m^D, and N_t, respectively.

B ML estimation

Model (3) can be written as

$$
\mathbf{BA}\widetilde{y} = \mathbf{B}\begin{pmatrix}\widetilde{\mathbf{1}} & \widetilde{\boldsymbol{Z}}\end{pmatrix}\begin{pmatrix}\delta \\ \beta\end{pmatrix} + \epsilon,
$$

where $\widetilde{y}$, $\widetilde{\mathbf{1}}$ and $\widetilde{\boldsymbol{Z}}$ are the time demeaned variables of y, $\mathbf{1}$ and $\boldsymbol{Z}$, respectively, where

$$
\mathbf{1} = \begin{pmatrix} \mathsf{o}_1 & \cdots & \mathsf{o}_1 \\ \iota_2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \iota_T \end{pmatrix},\ \mathbf{B} = \begin{pmatrix} \boldsymbol{B}_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \boldsymbol{B}_T \end{pmatrix},\ \mathbf{A} = \begin{pmatrix} \boldsymbol{A}_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \boldsymbol{A}_T \end{pmatrix}
$$

and $\boldsymbol{B}_t = \boldsymbol{I}_{N_t} - \rho\boldsymbol{M}_t$, $\boldsymbol{A}_t = \boldsymbol{I}_{N_t} - \lambda\boldsymbol{W}_t$. Assuming a multivariate normal distribution of the error terms, the log likelihood function is given by

$$
\ln L = \sum_{t=1}^{T}\left(-\frac{N_t}{2}\ln(2\pi\sigma^2) + \ln|\boldsymbol{A}_t| + \ln|\boldsymbol{B}_t| - \frac{\epsilon_t'\epsilon_t}{2\sigma^2}\right), \tag{8}
$$

where

$$\epsilon_t = \begin{cases} \boldsymbol{B}_t\left(\boldsymbol{A}_t\widetilde{y}_t - \widetilde{\boldsymbol{Z}}_t\beta\right) & \forall \quad t=1 \\ \boldsymbol{B}_t\left(\boldsymbol{A}_t\widetilde{y}_t - \widetilde{1}_t\delta_t - \widetilde{\boldsymbol{Z}}_t\beta\right) & \forall \quad t=2,\dots,T \end{cases}$$

and $\sigma^2 = \sum_{t=1}^{T}\left(\epsilon_t'\epsilon_t/N_t\right)$. The ML estimator is

$$\begin{pmatrix} \hat{\delta}_{ML} \\ \hat{\beta}_{ML} \end{pmatrix} = \left[\begin{pmatrix} \widetilde{\mathbf{1}}' \\ \widetilde{\boldsymbol{Z}}' \end{pmatrix} \widehat{\mathbf{B}}'\widehat{\mathbf{B}} \begin{pmatrix} \widetilde{\mathbf{1}} & \widetilde{\boldsymbol{Z}} \end{pmatrix}\right]^{-1} \begin{pmatrix} \widetilde{\mathbf{1}}' \\ \widetilde{\boldsymbol{Z}}' \end{pmatrix} \widehat{\mathbf{B}}'\widehat{\mathbf{B}}\widehat{\mathbf{A}}\widetilde{y},$$

where $\widehat{\mathbf{B}} = \begin{pmatrix} \widehat{\boldsymbol{B}}_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \widehat{\boldsymbol{B}}_T \end{pmatrix}$, $\widehat{\boldsymbol{B}}_t = \boldsymbol{I}_{N_t} - \hat{\rho}_{ML}\boldsymbol{M}_t$, $\widehat{\mathbf{A}} = \begin{pmatrix} \widehat{\boldsymbol{A}}_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \widehat{\boldsymbol{A}}_T \end{pmatrix}$ and $\widehat{\boldsymbol{A}}_t = \boldsymbol{I}_{N_t} - \hat{\lambda}_{ML}\boldsymbol{W}_t$.

C Construction of case mix weights

The more time the treatments of cases belonging to the j-th clinical department takes relative to all other treatments, the higher the weight, π_j, of the corresponding cases. Let c_{ij} be the number of cases in the j-th clinical department of the i-th hospital at time t[7]. Then, the weighted cases of hospital i at time t are calculated as

$$wc_i = \sum_{j=1}^{J} \pi_j \, c_{ij},$$

where $\pi_j = los_j/los_G$, $los_j = (\sum_{i=1}^{N} days_{ij}/c_{ij})/N$ is the mean length of stay for the cases belonging to the j-th clinical department over all hospitals and $los_G = (\sum_{j=1}^{J} los_j)/J$ is the mean length of stay over all clinical departments and all hospitals at time t.

[7] For ease of illustration the time index t is neglected.