Felder, Stefan; Tauchmann, Harald

Working Paper

Regional Differences in the Efficiency of Health Production: an Artefact of Spatial Dependence?

Ruhr Economic Papers, No. 112

Provided in Cooperation with:
RWI – Leibniz-Institut für Wirtschaftsforschung, Essen

Ruhr Economic Papers

Stefan Felder and Harald Tauchmann

Regional Differences in the Efficiency of Health Production: an Artefact of Spatial Dependence?

#112

UNIVERSITÄT DUISBURG ESSEN

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstraße 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
Hohenzollernstr. 1/3, 45128 Essen, Germany

Ruhr Economic Papers #112

Responsible Editor: Volker Clausen

ISSN 1864-4872 (online) – ISBN 978-3-86788-125-8

Ruhr Economic Papers

#112

Stefan Felder and Harald Tauchmann

Regional Differences in the Efficiency of Health Production: an Artefact of Spatial Dependence?

UNIVERSITÄT DUISBURG ESSEN

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

ISSN 1864-4872 (online)
ISBN 978-3-86788-125-8

Stefan Felder and Harald Tauchmann*

Regional Differences in the Efficiency of Health Production: an Artefact of Spatial Dependence?

Abstract
The inefficiency of health care provision presents a major health policy concern in Germany. In order to address the issue of efficiency comprehensively – i.e. at the level of the entire system of health care provision rather than individual service providers – empirical analyses are often based on data at the regional level. However, regional efficiencies might be subject to spatial dependence, rendering any analysis biased that aims at identifying the determinants of efficiency differentials. We address this issue by specifying a spatial autoregressive model to explain efficiency scores for German districts which we derive through data envelopment analysis. Regression results suggest that spatial dependence is not a dominant feature in the data. Hence, ignoring spatial interdependence is unlikely to severely bias results of efficiency analyses based on regional data. This holds, in particular, for the role of the states in the efficiency of health production. Significant heterogeneity among states is found in the data regardless of whether or not spatial dependence is accounted for.

JEL Classification: I12, R10

Keywords: Health production, data envelopment analysis, spatial autoregressive model

June 2009

---

* Stefan Felder, University of Duisburg-Essen; Harald Tauchmann, RWI. – The authors are grateful to Rüdiger Budde for generating the regional distance-matrix that is used for our empirical analysis and to Miriam Krieger for editorial assistance. – All correspondence to Harald Tauchmann, RWI, Hohenzollernstr. 1-3, 45128 Essen, Germany, e-mail: harald.tauchmann@rwi-essen.de.

1 Introduction

The efficiency of health care provision has in recent years become a major concern of health policy and applied health economics. Several empirical analyses address this issue at the level of individual health care providers such as hospitals (e.g. Staat, 2006 and Herr, 2008). However, rather than the inefficiency of individual providers, poor coordination e. g. between outpatient and inpatient care might in fact present the prime source of technical and economic inefficiency. This applies in particular to Germany, where a rigid separation of inpatient and outpatient care, free and direct access to medical specialists, and generous coverage by social health insurance are frequently blamed for generating inefficiencies such as the over-use of services, redundant medical treatments, and medical malpractice due to insufficient exchange of information.

These potential sources of inefficiency evidently cannot be addressed using data at the level of individual providers. Recent analyses based on regional data do comprehensively study quality, effectiveness and efficiency of health care provision in Germany, taking into account the interdependence of different health care sectors (e.g. Augurzky et al., 2009b; Schwierz & Wübker, 2008; Busse, 2009). However, by using regionally aggregated data, the implicit assumption - inherent in conventional regression analyses - of statistically independent observation units becomes almost indefensible.[1] Rather, health care efficiency at the regional level is likely to be shaped by different sources of spatial interdependence, as regional health care systems are not strictly separated. Patients can be provided with services in regions other than the one they reside in, thus net patient flows between regions will most evidently bias figures on regional efficiency. Typically, regions exhibiting net patient inflows fully account for inputs (e.g. medical infrastructure or services), while the corresponding outputs (e.g. reductions in mortality or morbidity) are partly assigned to the sending regions. While little is known about patient-flows in outpatient care in Germany, Augurzky et al. (2009a) provide detailed figures on flows for the hospital sector. Indeed, in some regions the share of patients treated within the region’s border is substantially lower than 50% (cf. Augurzky et al., 2009a).

Furthermore, health care infrastructure is likely to be regionally specialized. Medical specialists and specialized hospital departments not available in all regions will provide services to patients from neighbouring regions, even if net patient flows are small. Regions are thus likely to benefit in terms of reduced morbidity and mortality from efficient and high quality service in adjacent regions. Hence, when analyzing the quality and efficiency of health care provision at the regional level ignoring potential spatial interdependence may result in misleading empirical evidence in two ways. Firstly, efficiency figures may be biased, due in particular to net patient flows. Secondly, the determinants of regional efficiency differentials may be biased. The question of whether states[2] differ in the efficiency of health care provision has recently attracted much interest, as health inequality is a controversial issue in the public debate. Institutional factors do suggest such differences: Firstly, the states are responsible for investments in inpatient care such as the construction and closure of hospitals.

[1] To some degree Schwierz & Wübker (2008) allow for inter-regional correlation by carrying out a ‘multi-level analysis’. But in fact they account for correlation at the state level rather than spatial dependence.

[2] The term ‘state’ refers here to the German ‘Länder’, while ‘federal’ denotes the German ‘Bund’.

Secondly, the availability of outpatient care is organized on the state level, the number of resident doctors effectively restricted by public demand planning in many regions.[3] Moreover, empirical analyses (cf. Augurzky et al. 2009b) report substantial differences in state-level efficiency. This result, however, might represent an artefact of spatial dependence. Our analysis is concerned with this issue, focusing on whether state-specific heterogeneity can be identified even if spatial dependence is accounted for.

The empirical analysis follows a two step approach. Firstly, we calculate health care efficiency figures for Germany at the district level employing data envelopment analysis (DEA). These efficiency scores are likely to be biased as they ignore spatial interdependence. This first step, however, mimics conventional efficiency analysis and hence does not correct for such biases. The second step then addresses spatial interdependence using a Cliff-Ord (Cliff & Ord, 1973) modeling framework where the DEA efficiency scores derived in step one serve as dependent variable. The remainder of the paper is organized as follows: section 2 introduces the data; section 3 discusses our two step empirical approach; section 4 presents and discusses the estimation results; section 5 concludes.

2 The Data

The empirical analysis is based on district-level data from Germany. In 2004, Germany was sectioned into 449 districts ('Kreise'[4]) with an average of 180,000 inhabitants and 800 km^2 in area. Districts exhibit pronounced heterogeneity with respect to both population and size. This is partly due to the typical pattern of large and medium-size – sometimes even small – cities and towns constituting one district (urban district, 'Stadtkreis') and the surrounding countryside constituting another (rural district, 'Landkreis'). Thus, large rural districts often directly border on densely populated urban districts. The two largest German cities, Berlin and Hamburg, are urban districts as well. District populations hence range from less than 35,000 to almost 3.5 million people.

Comprehensive district-level data is provided by the Federal Statistical Office ('Statistisches Bundesamt')[5] as well as the Federal Office for Building and Regional Planning ('Bundesamt für Bauwesen und Raumordnung')[6]. We combine data from both sources and add information from the German Hospital Register ('Verzeichnis der Krankenhäuser und der Vorsorge- oder Rehabilitations-einrichtungen') that is also provided by the Federal Statistical Office. Specifically, we use (i) demographic information, i.e. population and deaths by gender and age, (ii) territory, from which we calculate population density, (iii) per capita income, (iv) district unemployment rates, (v) the number of physicians by medical specialty, (vi) the number of hospital beds by medical specialty, (vii) the share of hospital beds by type of ownership (private, public, non-profit). The state a district is located in is also considered in the regression analysis. The 'city states' Berlin, Hamburg and Bremen, each a single district in their own right[7], serve as reference. We focus on the year 2004, for which comprehensive information on these variables is available. Some districts were excluded due to

[3] Resident doctors' associations organized at state level, rather than state governments themselves, are responsible for the provision of outpatient care.

[4] Due to mergers, the number of districts has constantly been reduced.

[5] See database 'Statistik Regional'.

[6] See database 'INKAR'.

[7] Except for Bremen, which is divided into two urban districts, i.e. the 'City of Bremen' and 'Bremerhaven'.

missing data. Our analysis is thus based on a data set of 435 individual districts. Table 1 displays descriptive statistics for the variables included in the empirical analyses.

Table 1: Descriptive statistics (435 districts)

Variable	Mean	s.d.	min	max
standardized mortality	1.014	0.087	0.787	1.264
hospital beds (per 10.000 inhabitants)	67	12	2	228
general practitioners (per 10.000 inhabitants)	54	8	31	84
resident medical specialists (per 10.000 inhab.)	100	51	23	328
output-oriented technical efficiency (DEA score)	0.827	0.076	0.652	1.000
population (million people)	0.188	0.220	0.035	3.388
population density (1.000 people per km^2)	0.509	0.657	0.040	4.024
per capita disposable income (10.000 € per year)	1.413	0.185	1.085	2.304
unemployment rate	0.111	0.055	0.040	0.293
non-profit hospital ownership (share in beds)	0.279	0.329	0.000	1.000
private hospital ownership (share in beds)	0.141	0.258	0.000	1.000

Carrying out spatial regression analysis requires a spatial-weighting matrix that parameterizes the distance between spatial units. We use a matrix based on distances in terms of average travel time by car between two districts' centroids. We believe that travel time is far more important for patient flows and efficiency spillovers than geographical distance alone. We distinguish two variants of spatial-weighting matrices, one including squared inverse distances and another using a contiguity matrix design. The latter assigns a value of 1 for contiguous neighbours and a value of 0 for all other elements. 'Contiguous' is defined here as a travel time between two districts of less than 30 minutes, thus allowing for the inclusion of 'island districts' without any factually contiguous neighbours.

3 The Empirical Approach

3.1 Data Envelopment Analysis

The first step of our empirical analysis is to perform a 'conventional' efficiency analysis without taking into account any potential spatial interdependence. The purpose is to generate district-specific performance figures (efficiency scores) rather than to identify structural parameters of the underlying production technology. For this reason, we prefer data envelopment analysis to a stochastic frontier approach. Policy makers regularly demand such descriptive analyses from applied research. We therefore use DEA as a descriptive tool for aggregating information on inputs and outputs for health production into a single dependent variable capturing efficiency. In technical terms, data envelopment analysis represents a nonparametric linear programming approach to efficiency analysis (see e.g. Seiford & Thrall, 1990). In the present application, DEA does not impose the restriction of constant returns to scale in production.

The limited availability of data at the district level unfortunately limits the opportunities for formulating a rich model of efficiency that includes measures of morbidity as outcome variables. Instead, we use the number of deaths per year and district as a single output variable of health production capturing mortality.[8] Clearly, the raw number of deaths does not tell us much about the efficiency or quality of health care provision, as district populations substantially differ in size and demographic composition. We therefore standardize the number of deaths to national gender- and age-specific death rates.[9] Figure 1, a map of the districts in Germany, indicates substantial heterogeneity in standardized mortality across districts. While for some districts mortality is 21% lower than expected on the basis of demographic structure alone, others exhibit mortality rates as much as 26% above the expected level. Standardized mortality is highest in the north-east of the country – the former GDR – and lowest in the south-west - the state of Baden-Wurttemberg and southern Bavaria. As DEA requires at least one desired output rather than an unwanted 'bad', we use the inverse of standardized deaths as left-hand side variable (cf. Scheel, 2000).

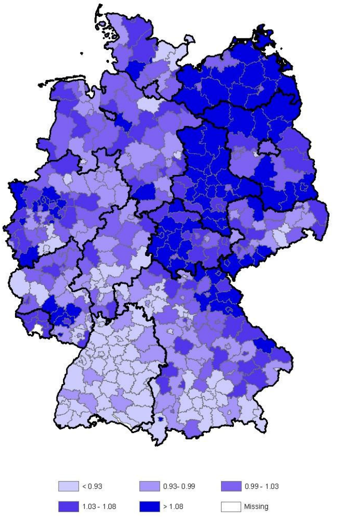

Figure 1: Standardized mortality, 2004
Source: Statistisches Bundesamt, own calculations

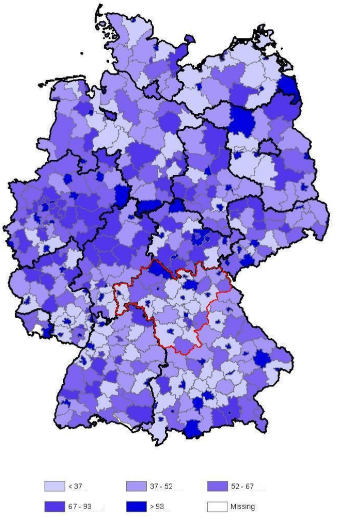

Figure 2: Hospital beds per capita, 2004
Source: Statistisches Bundesamt (Statistic Regional)

As comprehensive quantitative information on health care utilization (e.g. the number of physician consultations) is not available to us at the district level, we use medical infrastructure to measure the inputs to health production.[10] Specifically, hospital beds, the number of general practitioners, and the number of resident medical specialists - each per 10,000 inhabitants - enter the data envelopment analysis (see Figures 2 trough 4). Figures 2 and 4 show the typical pattern of hospitals and medical specialist located in cities or towns, with a much weaker supply of medical infrastructure in the countryside. This pattern is most distinct in Franconia (marked red in Figures 2 and 4) and some bordering regions, where the arrangement of districts adheres closely to the 'Stadtkreis-Landkreis' scheme. Though the distribution of general practitioners (Figure 3) exhibits some

8 This is not uncommon in applied health economics; see Hall and Jones (2007) for a recent example.

9 Unfortunately, the data available to us include precise age only up to 75 years. For the population older than this, figures are reported in age-classes, i.e. 75 to 80, 80 to 85, and > 85 years.

10 As with patient flows, publically available information on inpatient care is much better than information on outpatient care.

interesting regional features – e.g. remarkably low numbers for the western state of North Rhine-Westphalia – there is no distinct difference between urban and rural districts.

Figure 5 displays the corresponding DEA results in terms of output-oriented technical efficiency scores ($0 <$ DEA score ≤ 1). These scores can be interpreted as the factor by which mortality might be reduced, relative to other districts, without changing the inputs considered here[11]. On average, districts reach an efficiency level of 0.827 while the worst score is 0.652. The share of districts classified as fully efficient (unity efficiency score) is only 3%.[12] The regional pattern of efficiency scores to some degree mirrors the distribution of standardized mortality. That is, while the share of highly efficient districts is rather large in the south-west, many of the least efficient ones are located in the north-east of the country.

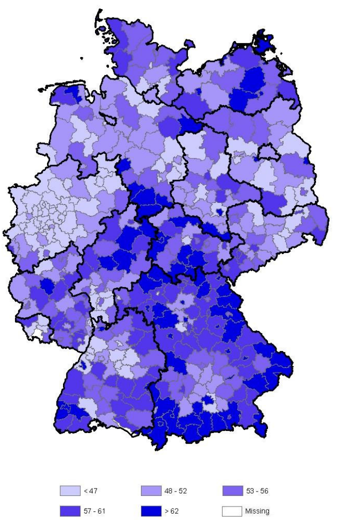

Figure 3: General practitioners per capita, 2004
Source: BBR (INKAR 2006)

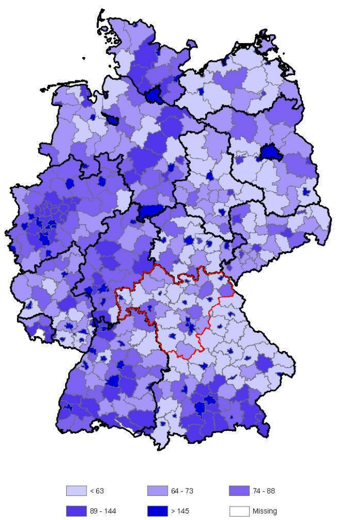

Figure 4: Resident medical specialists per capita, 2004
Source: BBR (INKAR 2006)

Interestingly, especially in the region of Franconia, DEA results exhibit a pattern of relatively less efficient urban districts surrounded by rural ones featuring high, in some cases even perfect, efficiency scores (see Figure 6 and Table 2).

Table 2: DEA efficiency scores for some Franconian urban and contiguous rural districts

	Bamberg	Bayreuth	Coburg	Hof	Schweinfurt	Wurzburg
urban district	0.821	0.782	0.854	0.675	0.846	0.815
rural district	0.914	1.000	0.975	0.735	1.000	1.000

[11] As we conceptualize DEA in terms of a descriptive tool rather than a causal model, this interpretation should not be overstretched.

[12] Since the share of fully efficient districts is very low, censoring from the right is not a major problem for the subsequent regression analysis that uses DEA scores as left-hand side variable.

This corresponds to the regional distribution of hospitals and medical specialists and supports our original hypothesis of a strong influence of net patient flows on district-specific efficiency scores. Inhabitants of rural Franconia thus receive medical treatment in urban district such as Bamberg, Bayreuth, Schweinfurt and Wurzburg, leading to a downward bias in the urban efficiency scores while the rural contiguous neighbours seem highly efficient. Indeed, applying Moran's I test to the raw[13] efficiency scores indicates a highly significant spatial correlation, regardless of the type of spatial weighting matrix used.

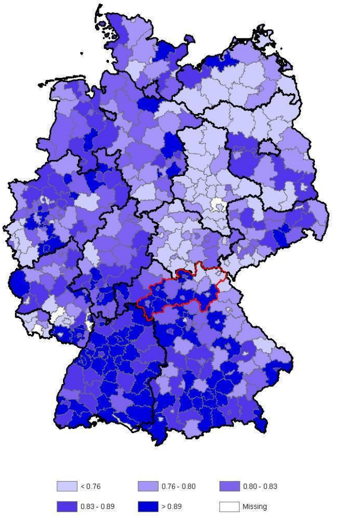

Figure 5: DEA efficiency scores, 2004
Source: Own calculations

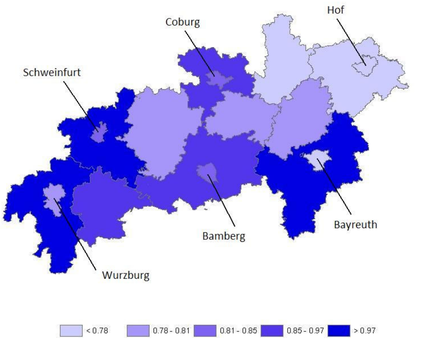

Figure 6: DEA efficiency scores, 2004 (enlarged detail)
Source: Own calculations

3.2 A Spatial Regression Model

If spatial interdependence occurs in health care provision, it must be considered in the regional analyses of health care efficiency so as to achieve unbiased and conclusive results. In a spatial regression model explaining regional efficiency, unobserved net patient flows might most appropriately be modelled with spatially negatively correlated errors. In contrast, potential spillover effects from one region to another may be captured by including the efficiency scores of neighbouring districts as right-hand side variables in the regression model. The spatial autoregressive model with autoregressive disturbances (see Kelejian & Prucha, 1998) accommodates both types of spatial interdependence. The basic structure of the model is given by:

$$y = X\beta + \lambda W y + u \tag{1}$$

$$u = \rho M u + \varepsilon \tag{2}$$

$$with \quad \mathrm{E}(\varepsilon_i) = 0, \quad \mathrm{var}(\varepsilon_i) = \sigma^2, \quad and \quad \mathrm{cov}(\varepsilon_i, \varepsilon_j) = 0 \quad \forall\, i \neq j$$

[13] That is, no covariates enter the initial regression from which residuals for calculating the test statistic are obtained.

Here, *y* indicates the $N \times 1$ vector of the dependent variables (DEA scores), where N denotes the number of observations (districts). As usual, u denotes the corresponding vector of error terms and X denotes the $N \times k$ regressor matrix, where β is a vector of coefficients subject to estimation. W and M represent $N \times N$ spatial weighting matrices, capturing the pattern of spatial dependence. Hence, the endogenous variable y enters the right hand side of the equation (1) via Wy that represents a $N \times 1$ vector of spatial lags of y_i, i.e. weighted sums of y_j, with $j = 1 \ldots N,\ j \neq i$. Hence, the efficiency of any district *i* potentially depends on the efficiency of any other district under consideration.

An appropriate spatial weighting matrix must satisfy certain requirements (see e.g. Kelejian& Prucha, 2008). In particular, the diagonal elements w_{ii} and m_{ii} must be zero and the intensity of spatial dependence must be restricted and decline sufficiently with increasing distance between two spatial units. Formally, each row and column sum must be uniformly bounded if the number of spatial units grows to infinity, and $(I - \lambda W)$ as well as $(I - \rho M)$ must be non-singular. Yet, W and M need not be symmetric, allowing for different travel times between two districts depending on the direction of travel.[14] The model further allows for the special case $W = M$, in which the same pattern of spatial dependence applies to the covariance structure among the error terms u_i and to the pattern of direct spillover between the dependent variables y_i. In practical applications, both W and M are typically exogenously given and not subject to estimation, except for rare cases where panel data are available and the number of cross-sectional units is small compared to the number of periods.

In applied work spatial weighting matrices are typically normalized, in order to allow for a straightforward interpretation of the model parameters. The most common approach is row standardization, where all matrix elements w_{ij} and m_{ij} are divided by the corresponding row sum, yielding uniform ones that equal unity. Other approaches are (i) minmax standardization, i.e. standardization by the minimum of the largest row and the largest column sum, and (ii) eigenvalue standardization, i.e. standardization by the modulus of the largest eigenvalue of the relevant matrix. Through minmax- as well as eigenvalue standardization heterogeneity in the total relevance of spatial dependence is left unchanged across observations. In contrast, by using row standardization, one implicitly assumes uniform total importance of spatial interdependence. Thus, the choice of row standardization might substantially affect estimation results. Kelejian& Prucha (2008) therefore argue against row-normalization – despite its popularity – unless this implicit assumption is clearly suggested by economic theory.

Finally, λ and ρ are unknown scalar coefficients that are estimated along with β. Clearly, no spatial autoregression – i.e. no direct spillover effects – is present for the case $\lambda = 0$, while for $\rho = 0$ the errors are spatially uncorrelated. In many applications, the unit-interval is regarded as the valid and relevant parameter space for λ and ρ, which has much intuitive appeal. Indeed, the valid parameter space depends on how the spatial weighting matrix is normalized (see Kelejian& Prucha (2008) for a detailed discussion). Nevertheless, for row-, minmax-, and eigenvalue-normalized

[14] To a very limited degree, this applies to our data.

matrices, any value within the unit-interval is valid, though some other values (smaller than -1) may also meet the relevant regularity requirements.

The spatial autoregressive model exhibits a close analogy to familiar autoregressive time-series models. Hence, (1) and (2) allow for inversion, yielding a reduced form, moving average representation of the model:

$$y = (I - \lambda W)^{-1}[X\beta + u] = (I + \lambda W + \lambda^2 W^2 + \dots)[X\beta + u] \tag{3}$$

$$u = (I - \rho M)^{-1}\varepsilon = (I + \rho M + \rho^2 M^2 + \dots)\varepsilon \tag{4}$$

from which the conditional mean $\mathrm{E}(y \mid X)$ is directly derived as:

$$\mathrm{E}(y \mid X) = (I - \lambda W)^{-1} X\beta = (I + \lambda W + \lambda^2 W^2 + \dots) X\beta \tag{5}$$

Thus, in the spatial autoregressive model a change in an exogenous variable x_i exerts an effect not only on region i but – potentially – on any other region j. Moreover, due to feedback effects from j to i, β does not represent the total effect on region i.

The spatial autoregressive model does not allow for straight forward estimation of equation (1) by OLS as the endogenous variable y enters the right-hand-side via the spatial lag, for which $\mathrm{E}(Wy'u) \neq 0$ holds. Thus, in order to avoid endogeneity bias, we follow the approach suggested by Kelejian & Prucha (1998). This three-stage estimator rests on initially estimating equation (1) using a conventional instrumental variables approach, with X, WX, W^2X, MX, M^2X, and WMX serving as instruments for the endogenous spatial lag Wy.[15] Secondly, based on the residuals obtained from the initial two-stage least-squares (2SLS) regression, Kelejian & Prucha (1998) derive a set of moment restrictions that allow for estimating ρ and σ^2 via non-linear least squares. Finally, with an estimate for ρ in hand, a Cochrane-Orcutt transformation can be applied and 2SLS estimation carried out again using the transformed data. Here, the transformed left-hand side variable is $\tilde{y} \equiv y - \hat{\rho}My$, while the transformed right-hand side variables are defined analogously.

4 Estimation Results

Table 3 displays results from the spatial autoregressive regression model explaining efficiency scores derived by data envelopment analysis. Several model variants are estimated using differently specified spatial weighting matrices. The reported variants all consider the case $W = M$, as theory in our application does not clearly argue in favour of a particular pattern of spatial dependence with respect to spatial spillovers and spatial error correlation. We consider an inverse-squared-distance matrix as our preferred specification. Alternatively, a contiguity matrix is used, with contiguous neighbours defined in terms of travel time less than 30 minutes between the two.[16] 30 minutes clearly represents an arbitrary choice. More generally speaking, the concept of 'contiguous neighbours' is hard to define in the case of irregular spatial units where location does not follow a

[15] Evidently, if $W = M$ holds, the number of available instruments is reduced.

[16] Additional variants were estimated, specifying either W or M as inverse-squared-distance matrix and choosing the contiguity matrix specification for the other. Results are available upon request.

grid-structure and where ‘economic integration’ is, what really matters. For districts, one might think of a joint border as the relevant criterion defining contiguity. However a joint border (or its absence) does not necessarily indicate connection (or distance) in economic terms. Thus, the preferred inverse-squared-distance design has more intuitive appeal. All reported results are based on eigenvalue-standardized matrices.[17]

Table 3: Regression results for the spatial auto-regressive model without state dummies

	OLS (no spatial dependence)		inverse-squared-distance		contiguity	
	estimate	s.e.	estimate	s.e.	estimate	s.e.
population	0.018	0.015	0.026	0.025	0.017	0.014
population density	-0.004	0.005	-0.011	0.010	-0.009	0.006
unemployment rate	-0.496***	0.076	-0.541***	0.138	-0.536***	0.077
income	0.133***	0.023	0.114***	0.040	0.121***	0.022
hospitals non-profit	-0.007	0.009	0.001	0.016	-0.003	0.009
hospitals private	0.003	0.011	0.006	0.019	0.004	0.011
constant	0.694***	0.038	0.718***	0.066	0.717***	0.038
spatial lag (λ)	-	-	0.013	0.029	-0.001	0.006
ρ	-	-	0.488**	0.200	0.203***	0.069

Notes: *, **, and *** indicate significance at the 0.1, 0.05 and the 0.001 level, respectively.

We start with a simple model variant that ignores any potential spatial dependence and estimates β by OLS. These estimation results do not indicate any role of hospital ownership for the efficiency in health production at the district level. By contrast, the unemployment rate and average net income are highly significant predictors of a districts’ efficiency. Moreover, both exhibit the expected signs, i.e. unemployment is negatively and income positively correlated with efficiency. These coefficients are unlikely to capture a pure causal effect. Rather, unobserved morbidity might enter the model via these two controls, as morbidity is typically inversely related to unemployment while wealthy people tend to be of better health. From a spatial point of view, income may also serve as proxy for mobility, i.e. well paid employees are likely to travel longer distances to go to work. Long distance commuters in fact represent prime candidates for generating net-patient-flows – especially with respect to outpatient care – as they tend to visit physicians at the place of work rather than the place of residence. This argument analogously applies to the unemployment rate, as unemployed individuals more likely consult physicians at the place of residence. Finally, neither population size nor population density exhibits a significant impact on efficiency.

Turning to the model specifications that consider spatial dependence, the value of the β coefficients are close to the initial OLS ones. Moreover, the spatial lag is clearly insignificant for either specification. However, the estimation results indicate substantial and statistically significant spatial error-correlation for both the variant based on an inverse-squared-distance matrix and the one using a contiguity matrix design. The relevant estimates for ρ are positive. This seems to contradict our hypothesis that unobserved net patient flows leads to negatively correlated error terms. Hence, other unobserved factors are likely to drive the error correlation.

[17] While minmax standardization yields similar results, row standardization provides substantially different patterns of estimated spatial dependence.

Table 4 presents the estimation results for an extended model that includes state dummies. The results indicate jointly significant deviations in efficiency across states. Yet, when individually compared to the reference category ’city states’, only Saxony is significantly more efficient than the reference.

Table 4: Regression results for specification with state dummies included

	OLS (no spatial dependence)		inverse-squared-distance		contiguity	
	estimate	s.e.	estimate	s.e.	estimate	s.e.
population	0.019	0.016	0.022	0.028	0.021	0.016
population density	0.003	0.005	-0.002	0.009	-0.002	0.006
unemployment rate	-0.762***	0.129	-0.799***	0.206	-0.808***	0.123
income	0.122***	0.024	0.112***	0.039	0.121***	0.023
hospitals non-profit	0.012	0.010	0.014	0.017	0.012	0.010
hospitals private	-0.001	0.012	0.000	0.018	0.000	0.011
Schleswig-Holstein	0.014	0.048	0.008	0.061	0.003	0.037
Lower Saxony	0.031	0.047	0.018	0.059	0.018	0.035
North Rhine-Westphalia	0.017	0.046	0.002	0.056	0.002	0.034
Hesse	0.025	0.047	0.006	0.060	0.008	0.036
Rhineland-Palatinate	0.021	0.049	0.001	0.062	0.007	0.037
Baden-Wurttemberg	0.065	0.047	0.050	0.059	0.049	0.035
Bavaria	0.031	0.048	0.013	0.061	0.017	0.036
Saarland	-0.052	0.048	-0.066	0.071	-0.066	0.043
Brandenburg	0.078	0.048	0.073	0.061	0.071**	0.036
Mecklenburg-West Pomerania	0.076	0.050	0.072	0.062	0.069*	0.038
Saxony	0.097**	0.048	0.086	0.060	0.089**	0.036
Saxony-Anhalt	0.054	0.049	0.040	0.062	0.043	0.037
Thuringia	0.047	0.049	0.033	0.061	0.039	0.036
constant	0.691***	0.065	0.707***	0.095	0.709***	0.057
spatial lag (λ)	-	-	0.032	0.022	0.005	0.005
ρ	-	-	-0.030	0.224	0.087	0.073
joint significance (p-value):						
states	0.000		0.023		0.000	
hospital ownership	0.472		0.716		0.482	

Notes: *, **, and *** indicate significance at the 0.1, 0.05 and the 0.001 level, respectively.

Unlike in the first models, spatial dependence does not appear to be a relevant issue, as the estimates for λ and ρ are statistically insignificant regardless of the type of spatial weighing matrix used. In addition, Moran’s I test applied to the original OLS-residuals does not indicate strong spatial error correlation, either. Hence, correlation among district-specific efficiency scores that in our first models is attributed to spatially correlated errors seems to actually originate from state fixed effects. As the distance between districts belonging to the same state is on average smaller than the distance between two arbitrary districts, disentangling spatial dependence from state effects generally is ambitious. Nevertheless, our results clearly point to state fixed effects rather than spatial error correlation being of prime importance for district-specific efficiency. Hence, state-level efficiency does affect efficiency on the district level. And, most importantly, our results do not support the hypothesis that state effects are an artefact of spatial dependence, as the finding of significant state dummies turns out to be robust to any specification of spatial inter-relation among districts.

However, standard errors indicate that the relevant estimates for λ and in particular ρ are quite imprecise (see Table 4). One might therefore argue that the available data are insufficient to reveal two distinct channels of spatial dependence simultaneously. Table 5 displays results for more parsimoniously parameterized models that allow either for spatial error correlation or for spatial spillovers, i.e. either λ or ρ is restricted to the value of zero. If the restriction $\rho = 0$ is imposed and inverse-squared-distances enter the spatial weighting matrix W, estimation results indeed indicate statistically significant and positive, though moderate, spillover effects. This result remains in line with our previous reasoning that the effectiveness of health care provision is likely to benefit from good services in neighbouring regions and vice versa. However, estimates for the remaining coefficients β are still more or less equal to those obtained from the simple OLS estimation. Moreover, the pattern of significance among the controls does not change. For any other restricted model variant, neither spatial spillovers nor spatial error-correlation are statistically significant, while the estimated coefficients are in line with the previously obtained ones. Hence, model variants that parameterize potential spatial dependence more parsimoniously do not point at any bias originating from ignoring spatial interdependence either.

Table 5: Regression results for restricted model variants

	inverse-squared-distance				Contiguity			
	$\rho = 0$		$\lambda = 0$		$\rho = 0$		$\lambda = 0$	
	estimate	s.e.	estimate	s.e.	estimate	s.e.	estimate	s.e.
Population	0.022	0.016	0.019	0.028	0.022	0.016	0.018	0.016
population density	-0.003	0.006	0.002	0.009	0.000	0.006	0.000	0.005
unemployment rate	-0.803***	0.127	-0.771***	0.211	-0.770***	0.127	-0.807***	0.124
income	0.112***	0.024	0.121***	0.039	0.120***	0.024	0.122***	0.023
hospitals non-profit	0.014	0.010	0.012	0.018	0.013	0.010	0.011	0.010
hospitals private	0.000	0.012	-0.001	0.019	0.000	0.012	-0.001	0.011
Schleswig-Holstein	0.007	0.048	0.013	0.063	0.009	0.047	0.006	0.037
Lower Saxony	0.018	0.047	0.030	0.060	0.026	0.046	0.021	0.035
North Rhine-Westphalia	0.001	0.046	0.015	0.057	0.009	0.045	0.006	0.033
Hesse	0.005	0.048	0.024	0.061	0.018	0.047	0.013	0.035
Rhineland-Palatinate	0.001	0.049	0.020	0.063	0.013	0.048	0.012	0.037
Baden-Wurttemberg	0.049	0.047	0.063	0.060	0.059	0.046	0.053	0.035
Bavaria	0.012	0.048	0.029	0.061	0.024	0.047	0.021	0.035
Saarland	-0.067	0.048	-0.054	0.072	-0.059	0.047	-0.062	0.043
Brandenburg	0.072	0.048	0.077	0.063	0.075	0.047	0.073**	0.037
Mecklenburg-West Pomerania	0.072	0.050	0.075	0.064	0.072	0.049	0.071*	0.038
Saxony	0.086*	0.048	0.096	0.061	0.092**	0.047	0.093***	0.036
Saxony-Anhalt	0.039	0.049	0.053	0.063	0.047	0.047	0.047	0.037
Thuringia	0.033	0.049	0.046	0.062	0.041	0.047	0.042	0.036
constant	0.709***	0.064	0.695***	0.098	0.697***	0.063	0.706***	0.057
spatial lag (λ)	0.032**	0.016	-	-	0.005	0.005	-	-
ρ	-	-	0.056	0.176	-	-	0.094	0.059
joint significance (p-value):								
states	0.000		0.063		0.000		0.000	
hosp. ownership	0.367		0.786		0.402		0.532	

Notes: *, **, and *** indicate significance at the 0.1, 0.05 and the 0.001 level, respectively.

Though our results argue against spatial dependence representing a dominant feature of regional efficiency, some regional patterns found in the data remain unexplained. This in particular holds for the striking pattern of highly efficient rural districts directly adjoined to less efficient urban ones pointing at negative spatial error correlation. The absence of such correlations in our results, even if income and unemployment – which to some degree might capture unobserved net patient flows – are dropped from the regressions, remains a puzzle. Moreover, the case of spatial spillovers, i.e. direct interdependence between regional efficiency scores, remains vague. While the majority of specifications do not yield significant coefficients for the spatial lag, a single one does. Yet, in this single case the degree of spatial spillovers is still very small in absolute terms.

5 Conclusions

This paper investigates the possible spatial dependence of district-specific efficiency of health production in Germany. If spatial dependence held, this might bias empirical analyses and result in misleading policy advice. Our initial descriptive analysis of efficiency at the district level reveals patterns that seem to indicate the presence of spatial error correlation and/or spatial spillovers. However, the subsequent regression analysis does not support this initial impression. Rather, significant spatial error correlation is only found if state effects are not controlled for. Thus, what appears as spatial dependence in the raw data might better be explained by state-level fixed-effects. This does not seem implausible, as several decisions concerning the supply of health care are made at the state level. Moreover, the result does not argue in favour of spatial spillovers representing a major determinant of regional efficiency of health production either.

We conclude that analyses based on district-level data are unlikely to suffer from severe bias if regional interdependence is not explicitly accounted for. This holds in particular for the finding of a significant influence that states exert on the efficiency of health care provision. In our analysis this result, which has been found previously, turns out to be robust to various ways of modelling spatial dependence. Hence, regional differences in the efficiency of health production, and especially the role of state-level policies, seem not to represent an artefact of spatial dependence. However, some striking spatial patterns found in the raw data remain unexplained by the present analysis, implying a need for further research in this field. And, evidently, our key result only applies to efficiency at the district level in Germany. Spatial dependence might play a more prominent role under a different definition of spatial units.

References

Anselin, L. (2002): Under the hood – issues in the specification and interpretation of spatial regression models, *Agricultural Economics* 27, 274-267.

Augurzky, B., S. Felder, H. Tauchmann & A. Werblow (2009b): Effizienzreserven im Gesundheitswesen, *RWI: Materialien* 49.

Augurzky, B., S. Krolop, R. Gülker, Ch.M. Schmidt, H. Schmidt, H. Schmitz, Ch. Schwierz & S. Terkatz (2009a): Krankenhaus Rating Report 2009 – Im Auge des Orkans, *RWI: Materialien* 53.

Busse, R. (2009): Health care and avoidable mortality in Germany, presentation at Annual Meeting of the German Association of Health Economists.

Cliff, A.D. & J.K. Ord (1973): Spatial Autocorrelation, London: Pion.

Hall, R.E. & C.I. Jones (2007): The value of health and the rise in health spending, *The Quarterly Journal of Economics,* 122(1), 39-72.

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

Kelejian, H.H. & I.R. Prucha (1998): A generalized spatial two-stage least squares procedure for estimation a spatial autoregressive model with autoregressive disturbances, *Journal of Real Estate Finance and Economics* 17, 99-121.

Kelejian, H.H. & I.R. Prucha (2008): Specification and estimation of spatial autoregressive models with autoregressive heteroskedastic disturbances, *CESifo Working Paper* # 2448.

LeSage, J. & R.K. Pace (2009): Introduction to Spatial Econometrics, CRC Press.

Scheel, H. (2000): Undesirable outputs in efficiency valuations, *European Journal of Operational Research* 132, 400-410.

Schwierz, Ch. & A. Wübker (2009): The Influence of the Health System on Avoidable Deaths from Ischaemic Heart Diseases in East versus West Germany, available at SSRN: http://ssrn.com/abstract=1289525.

Seiford, L.M. & R.M. Thrall (1990): Recent developments in DEA – the mathematical programming approach to frontier analysis, *Journal of Econometrics* 46, 7-38.

Staat, M. (2006): Efficiency of hospitals in Germany: a DEA-bootstrap approach, *Applied Economics* 38, 2255-2263.