Preface:
The aim of this Working Paper (WP) is to contribute to the current debate about the future configuration of the Danish hospital sector, in particular the issue of 'optimal' hospital size. From an economic perspective we find that the empirical evidence underpinning the planned hospital structure is ambiguous and partly lacking. The WP contributes to this debate with an empirical study of the question of economies of scale in the Danish hospital sector and estimates of an optimal hospital size.

Our WP is addressed to foreign and Danish economists with an interest in industrial organization and, in particular, the organisation of the (Danish) hospital sector. Furthermore, our WP serves as the back ground paper for an article submitted to a peer-reviewed health economic journal. In addition to the content of the submitted article, the WP contains details such as an appendix and a more detailed discussion of methodological issues and concepts.

Acknowledgments
An earlier version of this WP was presented at the Health Economics’ Study Group Joint UK & Nordic meeting at the Health Economic Research Unit, University of Aberdeen 27-29 August 2008. We thank the discussant, Professor Jørgen T. Lauridsen, for helpful comments and suggestions. We would also like to acknowledge helpful comments from our SDU colleague, associate professor Jacob Nielsen Arendt, and visiting fellow at the Danish Institute for Health Services Research Professor Andrew Street, the University of York. Finally, we would like to thank John Mason of Manuscripts.dk for linguistic revision and secretary Helle Møller Jensen for helping in formatting and compiling the WP.

This project was partly financed by the Danish VAT foundation (Det kommunale momsfond). All errors, omissions and opinions are the sole responsibility of the authors.

Troels Kristensen, Kim Rose Olsen and Kjeld Møller Pedersen
November 2008
Title: Economies of scale and optimal size of hospitals: Empirical results for Danish public hospitals

Abstract:
Context and aim: The Danish hospital sector is facing a significant rebuilding programme, driven by a political desire to concentrate activity in fewer and larger hospitals. Our aim is to analyse whether the current configuration of Danish hospitals is subject to scale economies that may justify such plans and to estimate an optimal hospital size.

Methods: We estimate cost functions using panel data on total costs, DRG-weighted casemix, and number of beds for three years from 2004-2006. A short-run cost function is used to derive estimates of long-run scale economies by applying the envelope condition.

Results: We identify moderate to significant long-run economies of scale when applying two alternative translog cost functions. However, using a quadratic functional form we identify constant economies of scale for the medium-sized sub-groups and decreasing economies of scale for the largest sub-groups. The optimal number of beds per hospital is estimated to be 275 beds per site. Sensitivity analysis to partial changes in model parameters yields a joint 95% confidence interval in the range 130 – 585 beds per site.

Conclusions: The results indicate that it may be appropriate to consolidate the production of small hospitals (<230 beds) on fewer and larger units.

Keywords: Economies of scale, optimal size, hospitals, cost function.


1. Introduction
In Denmark the number of somatic hospitals has decreased from 117 in 1980 to 52 in 2004. Part of this decrease is due to the fact that the concept or idea of a hospital has changed. Until the early 1990’s there was always a one-to-one relationship between a hospital as a management entity and its geographical location. During the past 15 years, however, hospitals at different locations have been merged so that many ‘hospitals’ today consist of several geographically distinct ‘production units’ being managed together. These new entities, consisting of several production units, are called management entities or conglomerate hospitals. This trend towards centralization is not uniquely Danish but is also found in, for instance, England, Norway and Sweden [1,2,3].

Hospital plans from the five Danish regions show that this development is expected to continue in the years to come [4]. They are planning a significant rebuilding program including green field investments at 5 new sites, significant extension and reconstruction of several existing hospitals and mergers or closures of several small hospitals¹.

Whether the increasing centralization of hospitals is to be seen as an advantage depends on a) whether there are economies of scale i.e. lower average costs and b) whether bigger hospitals lead to improved clinical outcomes [5]. Exploiting economies of scale may help to limit costs of health care outputs without compromising their quality and volume. On the other hand, hospitals may become so large that the cost of treatment will be higher due to diseconomies of scale. Furthermore, plans to concentrate further assume that the ‘optimal’ hospital size is bigger than in the current configuration.

From an economic perspective the evidence base underpinning centralization is weak, that is to say that there is a conspicuous absence of research and discussion of economies of scale in Danish hospital production. No econometric studies of economies of scale in hospitals have ever been undertaken in Denmark. In Europe, unlike in the U.S. where more literature exists on economies of scale, the economics of this trend towards larger hospitals have not been sufficiently analyzed. Notable exceptions are [6] and [7] along with a survey [5]. Evidence of the ‘optimal’ hospital size is important at a time when the hospital sector is facing major restructuring. Therefore, the aim of this study is to assess whether there are unexploited economies of scale in the current configuration and to estimate an ‘optimal’ hospital size. The present study

¹ The association of Danish Regions has published the rebuilding program for each region, see http://www.godtsygehusbyggeri.dk.
is limited to assessing economies of scale and ‘optimal’ hospital size for Danish hospitals in the period 2004-2006 from a hybrid econometric cost function perspective [8].

The unit of analysis is the hospital “production unit”, not the hospital management entity. This approach has become increasingly relevant due to the trend towards concentration in secondary healthcare. In relation to the rebuilding programme it is the geographical hospital “production unit” which is the relevant decision unit when deciding to build new hospitals to replace one or more former hospitals. It is not the management entities with satellite production units which may be located far from each other that are the relevant analytical unit. A distance of 30-50 km between units within the same management entity is quite common. Besides, using the production unit as the unit of analysis means that we can interpret the estimated economies of scale and hospital sizes in relation to the actual geographical hospital production units instead of multisited hospital management entities. In the following, consequently, the term hospital is reserved for freestanding “production units” in specific geographical sites rather than “hospital management entities” which consist of several production units at different sites.

Our presentation of earlier studies is restricted to those that use econometric cost functions that resemble the methods used in this study. However, this study differs from the majority of earlier studies in several ways – especially in its estimation of long run cost functions and ‘optimal’ hospital size. So far, this approach has not been used in European studies. The literature search revealed only a single Canadian study that has estimated an ‘optimal’ hospital size using the envelope condition [9]. All other earlier studies of ‘optimal’ size are based on scale estimates – excluding specific estimates of ‘optimal’ hospital size.

2. Earlier results
The empirical literature on economies of scale in hospitals is extensive, if all statistical techniques are included [5]. Despite the fact that the literature reflects different methods and covers many different countries the results are remarkably consistent, according to a recent survey of 103 studies by Aletras [10], i.e. these studies reveal constant economies – or even diseconomies – of scale for the average hospital with about 200-300 beds, see also Aletras et al. [11]. However, studies based on structural or hybrid econometric cost functions only represent about one fifth of these studies. According to [11] economies of scale were evident only for small hospitals with less than 200 beds and the ‘optimal’ size for acute hospitals ranged form 200 to 400 beds (based on the interpretation of scale estimates). For hospitals above 400-600 beds it was concluded that the average cost increases.

Studies after 1997 based on structural or hybrid econometric cost function do not confirm the above-mentioned consistency. In North America the application of panel data has shown economies of scale in Canada [13,9]. A third study based on cross-section data also indicated economies of scale [14]. Moreover, a study of acute care hospitals in California has revealed a minor trend towards economies of scale [15].

In contrast to, for example, [13] and [14], the present writers use casemix-adjusted output measures instead of particularly constructed casemix indexes to adjust for differences in patient mix and severity. Furthermore, this study differs from [9], for example, by including costs shifters to adjust the structural model for cost drivers that are specific to hospitals.

Finally, it is apparent that the studies described do not rely on the latest data. This study applies the latest data and data adapted for managerial decision-making and efficiency-measurement in the Danish hospital sector.
3. Methods

Using econometric assessment of economies of scale and ‘optimal’ size of hospitals, a number of choices need to be made such as unit of analysis, model of hospital production, model for cost functions, specification of cost and output variables, and estimation technique.

As elsewhere, the number of hospitals in Denmark has declined radically over the past two decades as a result of mergers and closures. This means that many hospitals have changed from being an institution located on a single geographical site to a management conglomerate of hospitals spread across several geographical sites, often with a degree of division of labour and hence specialization. When estimating economies of scale and “optimal” hospital size this is a challenge if the cost and output data for hospitals are aggregated at the conglomerate level. This study is based on data for hospital production sites to reveal knowledge relevant for policy making about the cost and production characteristics of physical production entities in the Danish hospital sector. If the management conglomerate sites had been used as the sole unit of analysis, then estimates could have been conducted only in relation to an ‘optimal management unit’.

The short-run cost function

Most studies argue that there is no evidence that hospitals operate in their long-run equilibrium, i.e. hospitals do not adjust all their inputs to their cost-minimizing levels [16]. From a theoretical point of view, therefore, it is appropriate to estimate short-run cost functions. The argument is that this approach allows hospitals to use possible non-optimal levels of the fixed inputs in the short run. Hence, hospitals are only assumed to use cost minimizing quantities of easily adjustable variable inputs, such as nurses, physicians and materials. Furthermore, cost function estimation by frontier estimators may account for deviations from the cost frontier (non-minimum cost functions).

To estimate the short-run cost function, three different functional forms have been applied to examine the sensitivity of results to the chosen functional forms, because they cannot be determined a priori from a theoretical model. The first two specifications are the translog model and Cobb-Douglas model. The third specification is the less common quadratic form which, like the translog model, belongs to the family of ‘flexible’ forms. The three functional forms can be expressed as a version of the family of ‘flexible’ functional forms, which are second order Taylor approximations to an unknown functional form.

\[ g(C_i) = \alpha_i + \sum_{j=1}^{K} \beta_{ij} x_{ij} + \sum_{j=1}^{K} \sum_{k=1}^{J} \beta_{ijk} f(x_{ij}) f(x_{ik}) + u_i \quad i = 1, \ldots, N; t = 1, \ldots, T \]  

(1)

where \( g \) is a real valued function of the total cost for somatic treatment \( C \) for hospital \( i \) \((i = 1, \ldots, N)\) in period \( t \) \((t = 1, \ldots, T)\), \( \alpha \) is a constant, \( f \) a real valued function of the cost determinants, \( x \) is a matrix of outputs \( Q_{m} \) \((m = 1, \ldots, M)\) and cost shifters. \( u \) is a vector of unknown parameters and \( u_i \) is a random disturbance term.

Firstly, (1) is equal to the quadratic form when \( f \) and \( g \) are chosen to be equal to \( x \).

Secondly, if \( f \) and \( g \) are chosen to be the natural logarithmic function, (1) yields the translog model. Thirdly, the Cobb-Douglas version is the special nested case of the translog model where squared and interaction terms in (1) are excluded. The exact model specifications are shown in the appendix. Since this study aims to allow \( x \) to contain cost shifters for hospital \( i \) in time period \( t \), the estimated models also belong to the family of hybrid cost functions [8].

This category of functional forms is now preferred to the more naïve structural functional form, see, for instance [17].

In equation (1) input prices are left out, because it is assumed that input prices do not vary significantly between the hospitals, i.e. union agreements are made on a national basis. The reason is that pay and work conditions are fairly uniform across the public hospitals in Denmark, since wages are negotiated nationally and wages mainly vary with the experience and level of education. The public regulation of the hospital’s purchases of medicines and other not pay-related inputs means that it is accepted that the hospitals have more or less uniform prices in these areas. Additional arguments for the assumption about constant prices are that it is a panel data study with a relatively short time dimension and that the literature indicates the difficulty of measuring the input prices for hospitals e.g. the price of capital [6,18].

Cost and output variables

The advantages of ‘flexible’ cost functions are that they do not prejudge the existence or degree of economies of scale. Unfortunately, this increased flexibility is obtained at the cost of there being more parameters to be estimated than in more restricted functional forms such as the classical Cobb-Douglas cost functions. A consequence is that the estimation of ‘flexible’ cost functions often results in multicollinearity problems, see e.g. [6,19], especially if attempts are made to disaggregate the outputs into more subgroups. This was also the case in this study. In other

---

3. Cost shifters controls for the fact that the cost of individual hospitals may be influenced by other external factors to the hospital management than the demand for output such as capital endowment, strikes, epidemics and whether conditions.
words the estimated coefficients became insignificant and unstable, and the signs changed in such hospitals, for instance, different managerial abilities and severity of illness. Therefore it is less likely that FE models differ from other variables as is often the case in cross-sectional behavioural equations \[25\]. Besides the fixed effect model accounts for technical inefficiency among the variable inputs, even though it does not disentangle the fixed effects from expected behavior.

Sensitivity analysis was conducted for the optimal hospital size by substituting the upper and lower bound of the 95% confidence interval with respect to the proxy for fixed capital \( K \). We have merged the multiple hospital outputs into two aggregated output measures, DRG value for inpatient and outpatient activity. This approach has the advantages that calculations of the optimal hospital size are conducted for individual hospitals. The Danish DRG system adjusts each output (discharge) for casemix and to a certain extent for severity through the DRG cost weights attached to each discharge. Therefore, this study uses measure the both cost and output measures. The Danish DRG system adjusts output measures and aggregation of outputs are important tasks.

The 'optimal' size of a hospital can be calculated from the short-run cost function by application of the envelope condition \[9, 27\]. Given the cost function \( C(K) \), this calculation can be expressed as:

\[
K^* = \frac{\partial C(K)}{\partial K} = 0
\]

\[
K = \frac{\partial C(K)}{\partial K} = 0
\]

The sake cost function has been estimated in two different ways. The first approach uses the long-run cost function, the envelope condition to calculate the long-run cost function \( C(K) \) which remains in \( 2 \) after differentiation with respect to \( K \). Substituting the long-run cost function into the short-run cost function yields the long-run cost function. Since \( 1 \) is a second order equation for the long-run cost function and the long-run quadratic cost function respectively:

\[
C(K) = aK^2 + bK + c
\]

\[
K^* = \frac{-b + \sqrt{b^2 - 4ac}}{2a}
\]

Finally, the small number of observations, even in panel data, is an important issue in all small countries. These circumstances make the choice of a limited amount of observations and the aggregation of outputs two important tasks. We have merged the multiple hospital outputs into two aggregated output measures, DRG value for inpatient and outpatient activity. This approach has the advantages that the long-run cost function is a function of outputs and included cost shifters (if cost functions are considered). The estimated short-run cost function is a translog cost function, and the envelope condition to calculate the long-run cost function set equal to zero defines the 'optimal' relationship between beds and outputs as defined in \( 2 \). The estimated short-run cost function and the envelope condition yield the long-run cost function. This means that the first order condition of the short-run cost function is set equal to zero. The long-run cost function is an exponential function of hospital size, see appendix for details. Sensitivity analysis was conducted for the optimal hospital size by substituting the upper and lower bound of the 95% confidence interval with respect to the proxy for fixed capital \( K \).  

Sensitivity analysis was conducted for the optimal hospital size by substituting the upper and lower bound of the 95% confidence interval with respect to the proxy for fixed capital \( K \). We have merged the multiple hospital outputs into two aggregated output measures, DRG value for inpatient and outpatient activity. This approach has the advantages that the long-run cost function is a function of outputs and included cost shifters (if cost functions are considered). The estimated short-run cost function is a translog cost function, and the envelope condition to calculate the long-run cost function set equal to zero defines the 'optimal' relationship between beds and outputs as defined in \( 2 \). The estimated short-run cost function and the envelope condition yield the long-run cost function. This means that the first order condition of the short-run cost function is set equal to zero. The long-run cost function is an exponential function of hospital size, see appendix for details. Sensitivity analysis was conducted for the optimal hospital size by substituting the upper and lower bound of the 95% confidence interval with respect to the proxy for fixed capital \( K \).
The second approach applies an alternative way, where the cost function is estimated directly without use of the envelope condition [6]. This approach, the 'direct approach', has been achieved by omitting the number of beds in the estimation of the cost function. Hence, in contrast to the 'envelope condition approach', the direct approach assumes that hospitals use an optimal amount of capital in terms of beds (K) in the short and long run, see for instance, [16]. In other words, the direct approach assumes fixed cost to become variable in the long run, in other words as a function of, for example, outputs. Besides, it assumes that fixed cost varies with output across the data set and that the hospitals are always endowed with an optimal amount of capital. The latter is a relatively restrictive assumption to be discussed later.

The degrees of freedom gained from dropping the beds variable (K) are used to include two output measures, DRG value of inpatients and outpatients instead of the total DRG value per hospital.

Derivation of economies of scale estimates

In accordance with [28], economies of scale are estimated in a way that shows the relative rise in costs when output is increased proportionally. Since the translog and Cobb-Douglas models are logged in all variables and the quadratic forms are unlogged this yields (3) and (4):

\[
SE_1 = \sum \frac{\partial g(Q)}{\partial (Q_m)}
\]  

(3)

\[
SE_2 = \sum \frac{C}{Q_m/m - \frac{C}{Q_m/m} - \frac{C}{Q_m/m}}
\]  

(4)

\(SE_1\) in (3) expresses the sum of first order partial derivatives of the cost function (1) with respect to each output \(Q_m\) in logs. The logarithmic transformations imply that each of these derivatives is an estimate of cost elasticities for each \(Q_m\).

\(SE_2\) in (4) measures the sum of cost elasticities with respect to output. Each of the cost elasticities in \(SE_2\) is calculated using the standard (unlogged) approach, because the quadratic form is in cost levels. In the translog and quadratic models, in which scale estimates by definition are flexible, the sub-group median hospital was used to calculate scale estimates for each of the defined size groups. The size groups were defined by quartiles. The smallest size group (1st quartile) consists of the 25% of hospitals, which has the smallest number of beds, while the other size groups, 2nd quartile, 3rd quartile and 4th quartile, include hospitals with a size in the respective quartiles. Both \(SE_1\) and \(SE_2\) express the multi-product analog of marginal cost divided by average cost. The exact model specifications are shown in the Appendix.

In equations (3) and (4), \(SE\) values less than 1 indicate economies of scale corresponding to cost increases, which are smaller than the proportional output increase. \(SE\) values larger than 1 show diseconomies of scale.

Data

The data comes from a national cost database developed by the National Board of Health [29]. The cost database is based on patient activity and cost information from most public hospitals and is also used to calculate Danish DRG tariffs. Total hospital costs are actual costs incurred in respective years adjusted for costs from shared facilities with other hospitals, such as laundry. They are used as the best available proxy for the total cost for somatic treatment. DRG values, or in other words the reimbursement received by hospitals, give the most appropriate picture of the value of hospital production.

There may be some inconsistencies for the DRG values for the three years 2004 to 2006, because the DRG grouper used for 2005 and 2006 was different from that used for 2004 (giving different input prices). This means that 2004 data is based on 2007 input prices while 2005 and 2006 data is based on real 2008 input prices. We assume, however, that the effect of this is negligible due to a low inflationary level. Variables used and descriptive statistics are shown in table 1.
Table 1 Descriptive statistics for Danish hospitals in the years 2004-2006

<table>
<thead>
<tr>
<th>Variable</th>
<th>Year</th>
<th>Description</th>
<th>Average</th>
<th>Std. dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent:</strong></td>
<td></td>
<td>in 1000 DKK (Danish currency)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>2004</td>
<td>Adjusted operational costs</td>
<td>530,934</td>
<td>648,465</td>
<td>21,581</td>
<td>3,591,319</td>
</tr>
<tr>
<td></td>
<td>2005</td>
<td></td>
<td>616,490</td>
<td>784,055</td>
<td>19,035</td>
<td>4,337,614</td>
</tr>
<tr>
<td></td>
<td>2006</td>
<td></td>
<td>439,160</td>
<td>449,149</td>
<td>16,761</td>
<td>1,890,084</td>
</tr>
<tr>
<td><strong>Independent:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q1</td>
<td>2004</td>
<td>DRG value inpatient</td>
<td>328,514</td>
<td>379,452</td>
<td>0</td>
<td>2,123,694</td>
</tr>
<tr>
<td></td>
<td>2005</td>
<td></td>
<td>361,908</td>
<td>477,284</td>
<td>0</td>
<td>2,516,725</td>
</tr>
<tr>
<td></td>
<td>2006</td>
<td></td>
<td>232,824</td>
<td>231,748</td>
<td>3,265</td>
<td>898,218</td>
</tr>
<tr>
<td>Q2</td>
<td>2004</td>
<td>DRG value outpatientb</td>
<td>209,637</td>
<td>270,439</td>
<td>4,909</td>
<td>1,180,949</td>
</tr>
<tr>
<td></td>
<td>2005</td>
<td></td>
<td>274,984</td>
<td>335,966</td>
<td>2,514</td>
<td>1,372,589</td>
</tr>
<tr>
<td></td>
<td>2006</td>
<td></td>
<td>226,344</td>
<td>247,555</td>
<td>3,555</td>
<td>1,000,748</td>
</tr>
<tr>
<td>Q1 + Q2</td>
<td></td>
<td>Total DRG value (Q1 + Q2)</td>
<td>538,152</td>
<td>592,734</td>
<td>22,968</td>
<td>3,304,644</td>
</tr>
<tr>
<td></td>
<td>2005</td>
<td></td>
<td>538,152</td>
<td>592,734</td>
<td>22,968</td>
<td>3,304,644</td>
</tr>
<tr>
<td></td>
<td>2006</td>
<td></td>
<td>538,152</td>
<td>592,734</td>
<td>22,968</td>
<td>3,304,644</td>
</tr>
</tbody>
</table>

Independent cost shifters: in no. of beds and percentage of hospitals

<table>
<thead>
<tr>
<th>Variable</th>
<th>Year</th>
<th>Description</th>
<th>Average</th>
<th>Std. dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>2004</td>
<td>Average number of staffed beds</td>
<td>281.6</td>
<td>250.9</td>
<td>25.6</td>
<td>1107.1</td>
</tr>
<tr>
<td></td>
<td>2005</td>
<td></td>
<td>265.1</td>
<td>259.5</td>
<td>9</td>
<td>1136.7</td>
</tr>
<tr>
<td></td>
<td>2006</td>
<td></td>
<td>176.0</td>
<td>152.1</td>
<td>9</td>
<td>517</td>
</tr>
</tbody>
</table>

*Unbalanced due to missing data for 2006. The numbers of observations in 2004-06 are 57, 55 & 31 respectively.

b Including the value of grey zone DRG activity

Data in table 1 shows that hospital production units on average had operating costs in the range DKK 530 to 616 million. The DRG values are measured in local currency, DKK. The total value of DRG production for each hospital is divided into two output categories: 1) the production value of inpatients and 2) the production value of outpatients, including both so called grey zone patients and emergency patients.

Grey zone patients are patients that the hospital staff both can choose to treat as outpatient or as inpatient (in connection with hospitalization). To avoid distortion of this substitution choice, a special grey zone DRG rate is used. The grey zone DRG rate is calculated as the weighted average between what it costs to perform same-day surgery or outpatient treatment, and the corresponding price for similar inpatient treatment.

The average number of beds per hospital production unit is in the range 265 to 281, but this average covers wide variation between production units (e.g. min. 9, max. 1136 in 2005). The average number of disposable beds per hospital is used as a proxy for the size of hospitals and fixed inputs.

Table 1 also shows that the percentage of public hospitals that were university hospitals was on average approximately 21% in the period 2004 to 2005. Finally, it should be noted that the data for 2006 is generally sparser than the data for the previous year. This is due to data being missing for some of the large units in 2006, i.e. unbalanced panel data. Psychiatric hospitals are excluded from this study. Danish psychiatric hospitals do not use the DRG system. In special hospitals, e.g. Frihild Klinikken in Braedstrup and Hammel Neurocenter, the production process is considered to be atypical. Therefore, six hospitals were excluded.
Results

Table 2 shows the results for the short-run cost models in the Cobb-Douglas, the translog and the quadratic model specification.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cobb-Douglas</th>
<th>Translog</th>
<th>Quadratic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.1423*</td>
<td>-0.1966**</td>
<td>-0.0297</td>
</tr>
<tr>
<td>Inpatients (DRG value)</td>
<td>0.4425***</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Outpatients (DRG value)</td>
<td>0.2736***</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total DRG value</td>
<td>0.6921***</td>
<td>1.4800***</td>
<td></td>
</tr>
<tr>
<td>of in- and outpatients</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. number of beds</td>
<td>0.0511</td>
<td>-0.0403</td>
<td>-1.2360**</td>
</tr>
<tr>
<td>(Total DRG value)</td>
<td></td>
<td>-0.0262</td>
<td>0.1266**</td>
</tr>
<tr>
<td>(Avg. number of beds)</td>
<td></td>
<td>-0.0967</td>
<td>0.8131***</td>
</tr>
<tr>
<td>Total DRG value*Avg.</td>
<td>-</td>
<td>0.0938</td>
<td>-0.5377***</td>
</tr>
<tr>
<td>number of beds</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>143</td>
<td>143</td>
<td>143</td>
</tr>
<tr>
<td>Number of hospitals</td>
<td>54</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>R2</td>
<td>0.6028</td>
<td>0.7177</td>
<td>0.7657</td>
</tr>
<tr>
<td>Within</td>
<td>0.9837</td>
<td>0.9778</td>
<td>0.9517</td>
</tr>
<tr>
<td>Between</td>
<td>0.7965</td>
<td>0.9663</td>
<td>0.9374</td>
</tr>
<tr>
<td>Overall</td>
<td>0.7965</td>
<td>0.9663</td>
<td>0.9374</td>
</tr>
<tr>
<td>F-test (5,78)</td>
<td>28.11***</td>
<td>103.53***</td>
<td>30.10***</td>
</tr>
<tr>
<td>Hausman ch2(5)</td>
<td>8.49***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*** P < 0.01, ** P <0.05%, * P <0.10

1 Model did not meet the assumptions of the Hausmann test

The beta estimate for the average number of beds changes sign and significance across the model specifications leaving the effect ambiguous. The university hospital dummy in table 1 was eliminated in the fixed effect model 2.

The regression results in table 2 are used to estimate the long-run cost function based on the envelope condition, shown together with the direct approach to long-run cost function in table 3. The results in table 2 are also used to estimate the scale elasticities shown in table 4.

Table 3 Regression and calculated result – long-run cost functions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Translog &amp; Envelope condition</th>
<th>Translog, FE (without beds)</th>
<th>Quadratic &amp; Envelope condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.1950</td>
<td>-0.1709***</td>
<td>1.4980</td>
</tr>
<tr>
<td>Total DRG value</td>
<td>0.6845</td>
<td>-</td>
<td>2.8092</td>
</tr>
<tr>
<td>Total DRG value</td>
<td>-0.0045</td>
<td>-</td>
<td>0.3524</td>
</tr>
<tr>
<td>Inpatients DRG value</td>
<td>-</td>
<td>0.5388***</td>
<td>-</td>
</tr>
<tr>
<td>Outpatients DRG value</td>
<td>-</td>
<td>0.4100***</td>
<td>-</td>
</tr>
<tr>
<td>Inpatients DRG value</td>
<td>-</td>
<td>0.0879**</td>
<td>-</td>
</tr>
<tr>
<td>Outpatients DRG value</td>
<td>-</td>
<td>0.1094***</td>
<td>-</td>
</tr>
<tr>
<td>Total DRG value*Avg.</td>
<td>-</td>
<td>-0.1483***</td>
<td>-</td>
</tr>
<tr>
<td>number of beds</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Number of hospitals</td>
<td>54</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>R2</td>
<td>0.6028</td>
<td>0.7177</td>
<td>0.7657</td>
</tr>
<tr>
<td>Within</td>
<td>0.9837</td>
<td>0.9778</td>
<td>0.9517</td>
</tr>
<tr>
<td>Between</td>
<td>0.7965</td>
<td>0.9663</td>
<td>0.9374</td>
</tr>
<tr>
<td>Overall</td>
<td>0.7965</td>
<td>0.9663</td>
<td>0.9374</td>
</tr>
<tr>
<td>F-test (5,78)</td>
<td>28.11***</td>
<td>103.53***</td>
<td>30.10***</td>
</tr>
<tr>
<td>Hausman ch2(5)</td>
<td>8.49***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*** P < 0.01, ** P <0.05%, * P <0.10

Table 3 shows the three long-run cost functions. The first version of the translog function and the quadratic function are calculated from the short-run cost functions in table 2 by substitution of equation (2) into equation (1). The second version of the translog model is a directly estimated, fixed-effect, long-run cost function. In this model, the total output vector has been divided into two output measures – inpatient and outpatient DRG value – and no cost shifters have been included to avoid collinearity. The beta estimates of the Cobb-Douglas and translog cost functions are elasticities, whereas the betas of the quadratic model show the absolute increases in costs. The difference in

1 In an earlier cross-section analysis the university hospital dummy was positively significant for each of the years 2004-2006. This indicates, as expected, that university hospitals incur higher costs, see the method section.
Overall, the models suggest that results of the flexible models depend on the functional form used, while the nested Cobb-Douglas and the translog models yield similar results. However, an important difference is that the translog and quadratic models allow us to make non-constant scale estimates. The results indicate that scale estimates are increasing with the size of the hospitals, which should be interpreted as an indication of declining economies of scale as the size of the hospitals becomes larger.

Table 4 shows estimates of economies of scale when applying the short-run cost and long-run cost functions respectively. Both short-run and long-run economies of scale are measured by conventional ray scale economies, which are the elasticity of cost taken along a ray that holds product mix constant. SE < 1 implies scale economies and SE > 1 implies diseconomies when outputs are changed proportionately.

All scale estimates are calculated for four size groups, measured by the number of beds on information at the shape of the cost curve. This policy implication of economies of scale for all size groups is an L-shaped cost curve where average cost decreases when hospital output increases. This means the cheapest way of operating a hospital system would be to build hospitals that are as large as possible. In the extreme, we would plan one super hospital per region of a single hospital for the entire country. Diseconomies of scale imply that the average cost curves must be U-shaped because average cost increases as a function of output [30].

The short-run estimates in table 4 indicate that results are dependent on the functional forms used. When applying a Cobb-Douglas functional form we identify constant economies of scale, while a quadratic form indicates increasing returns to scale as the size of the hospitals becomes larger. Overall, we identify significant to moderate long-run economies of scale when applying a translog functional form. However, using a quadratic functional form we identify constant economies of scale for the medium-sized sub-groups and decreasing economies of scale for the largest sub-groups.

The translog scale estimates for the largest size groups lie around 0.67 or very close to the value 1, equivalent to constant economies of scale in the long run. Thus, there is nothing sensitive to the scale estimate of the model. Besides, the directly estimated translog model indicates that results are sensitive to the scale estimate of the model. Hence, we use the direct approach (0.70) for the smallest to 1.02 for the largest hospital size group (0.69 for the largest hospital and 0.67 for the smallest), the variation shows up at larger when we use the direct approach for the smallest to 1.02 for the largest hospital size group (0.69 for the largest hospital and 0.67 for the smallest).

Table 4 shows estimates of economies of scale when applying the short-run cost and long-run cost functions respectively. Both short-run and long-run economies of scale are measured by conventional ray scale economies, which are the elasticity of cost taken along a ray that holds product mix constant. SE < 1 implies scale economies and SE > 1 implies diseconomies when outputs are changed proportionately.
Figure 1 shows the long-run (LR) scale estimates as a function of hospital size for the three different LR model specifications. The figure shows that LR scale estimates for the translog model using the envelope condition lie below 1 for all hospital sizes, whereas they start to exceed 1 for hospitals above around 400 beds in the direct translog LR model. The estimates based on the quadratic form show less correlation between hospital size and LR scale estimates even though a positive trend can be detected with increasing size of hospitals. This increased level of noise probably stems from the lack of compression of outliers in the unlogged model.

In contrast to table 3, which showed scale estimates for the median hospitals data in each size groups (representative units), figure 1 shows the short-run scale characteristics for all observed hospitals. The smallest quartile has estimates in the range of 9-50.9 beds, while the following three size groups (2nd quartile, 3rd quartile and 4th quartile) have observations in the intervals 50.9-229.0, 229.0-356.6 and 356.6-1136.7 beds.

Finally, figure 1 shows that there are three outlier observations that we did not find any arguments to exclude and that there are relatively few observations among the largest hospital production sites.

### Optimal hospital size

The estimation based on (1) and the calculation of an ‘optimal’ hospital size based on (2) yields 204.9 beds for the median Danish hospital in the translog model and 275.2 beds for the quadratic functional form.

In figure 2 the estimated optimal hospital size is shown as a function of present size. Using the 45 degree line as point of departure, the figure illustrates how the ‘optimal’ size of each hospital deviates from the present size (‘45º’ line). The results of both models indicate that small and medium-sized hospitals with less than 204 or 275 beds are too small, while the larger hospitals are too large. However, it is not evident whether, for example, ‘small is too small’ in the translog model, since optimal and actual sizes are not different, at least not statistically. Both results are in line with the above-mentioned literature review by Aletras & Jones, which points to optimal sizes
of hospitals as being between 200 and 400 beds. An example is a recent Canadian study, which estimated an 'optimal' value of 179.5 beds [9]. Applying the 'directly' estimated long-run cost function, it is not possible to derive the quadratic form that is not very frequently used in the quadratic form. However, the long-run cost function has the great advantage that the coefficients are more difficult to interpret in the quadratic form. On the other hand, the quadratic form has the advantages that it does not require logarithmic transformation of the dependent variable. The translog model [15] has been applied in many studies of economies of scale, for instance, in a number of studies of hospital efficiency [17]. The results indicate that economies of scale have been exploited for the largest quartile (357-1137 beds), and that these hospitals are facing constant economies of scale where no economic gains can be obtained by centralization. From an econometric perspective, the quadratic form for estimates of 'optimal' hospital size [33]. To avoid these transformation issues (i.e. the introduction of approximate square and interactions terms), to be included, so that the model can capture varying economies of scale. Therefore, as mentioned above and found in [36], it is not possible to include more than one or two different output measures without getting multicollinearity which, of course, lowers the structural model requires all parts of the flexible functional model, including the quadratic form has the advantages that it does not require logarithmic transformation of the dependent variable. The translog model, changes in the parameter estimate of the average number of beds within the 95% confidence interval can result in 'optimal' size estimates in the range 130.0 to 419.2 beds per hospital. Finally, including the intersection of possible partial changes in each of the estimated parameters included in the envelope condition (2). According to the estimated parameters indicate that the 'optimal' number of beds is estimated to be in the interval 130 to 419 beds per hospital. The small number of observations may be due to the fact that the coefficients are more difficult to interpret in the quadratic form. From another perspective, the translog and quadratic functional form have the great advantage that they are flexible forms, which minimizes the risk of misspecification. From another perspective, the translog and quadratic functional form have the great advantage that they are flexible forms, which minimizes the risk of misspecification. From another perspective, the translog and quadratic functional form have the great advantage that they are flexible forms, which minimizes the risk of misspecification.

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Sensitivity analysis for estimate of optimal hospital size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model parameters</td>
<td>Quadratic model (275 beds)</td>
</tr>
<tr>
<td>Average number of beds</td>
<td>130.0</td>
</tr>
<tr>
<td>Total DRG value * Average number of beds</td>
<td>235.2</td>
</tr>
</tbody>
</table>

This study shows that parametric estimates of economies of scale to the chosen model specification. Theory makes assumptions about first and second order effects of cost functions, marginal cost elasticity and the envelope condition. This study is based on the relatively well-known properties of parametric flexible cost functions guided by neoclassical production theory [33]. Since, due to the fact that input prices were assumed to be constant, this study was limited to estimating a second-best practice where it was impossible to test all cost function regularity conditions, see for instance, [6]. The translog cost function, which has become a standard approach, has been applied in many studies of hospital efficiency [17]. In the same way, the Cobb-Douglas function is a standard functional form, but it has been used to estimate economies of scale since the development of the more flexible functional form, such as the translog model [15]. The only functional form that is very frequently used in the quadratic form.
The above-mentioned approach, which uses only one output index, is debatable. From one point of view, the multi-dimensional output can only be aggregated if the original dimensions are broad terms of, for example, quality, patient characteristics and institutional conditions in order to avoid bias due to omitted variables. On the other hand, empirical studies show that quality, efficiency, and patient outcomes are all significant in hospital cost functions, leading to arguments that better quality comes at a cost [17].

Furthermore, the Cobb-Douglas model has been applied in the literature to measure hospital size. However, it could be claimed that measuring hospital size based on the Cobb-Douglas model for cost estimation is a simplified approach. The Cobb-Douglas model assumes that costs are due to factors such as the number of beds, labor, and materials. In this way, the Cobb-Douglas model can be used to estimate total costs, but it cannot be used to calculate 'optimal' scale and size as a function of size or activity, as can be seen from table 4 where the short-run scale estimates become constant. The implication is that the Cobb-Douglas model may be preferred in situations where only a relatively small number of degrees of freedom are available as in the case of small country case. There is no available data as in the 'small country' case. From another point of view, use of the Cobb-Douglas model based on the opportunity cost approach. However, it can be argued that this is exactly what the DRG system allows. An opportunity cost approach can be justified by the Danish National Board of Health's approach to costing. The more correct economic approach requires calculation of both amortization, interests on data capital and the market value of assets used for fixed cost in a hospital activity. In a long-run perspective, this is due to the fact that fixed costs are assumed to be driven by the number of beds per hospital.

The Cobb-Douglas model has been applied to try to avoid multicollinearity. The lack of flexibility in the Cobb-Douglas model implies that fewer parameters have to be estimated than in the flexible cost functions. For example, it can be used to calculate optimal size and scale as a function of size or activity, as can be seen from table 4. However, the Cobb-Douglas model may be preferred in situations where only a relatively small number of degrees of freedom are available as in the case of small country case. There is no available data as in the 'small country' case.

Furthermore, since we do not know the monetary capital costs, the meaning for the results in this study of the choice between the applied physical measure and the monetary measure of total cost is ambiguous. Furthermore, the explanation is that it is difficult to compare more years than the three used here. Danish Regions, the Ministry of Finance [37]. Another consequence of the above mentioned "small country" problem is that the number of observations is too small to conduct specifications tests to determine a "correct" functional form. Outliers will simply influence too much. Instead the results of several functional forms are presented to "measure" sensitivity to different functional forms.

A time wise bigger panel would give more degrees of freedom and minimize multicollinearity. The flexibility of the Cobb-Douglas model also shows that it cannot be used to calculate optimal size and scale as a function of size or activity, as can be seen from table 4. A time wise bigger panel would give more degrees of freedom and minimize multicollinearity. The flexibility of the Cobb-Douglas model also shows that it cannot be used to calculate optimal size and scale as a function of size or activity, as can be seen from table 4.

While the present study assumes that the quality of hospital output is homogeneous across the hospitals, it is claimed that this is otherwise a crude generalization. The reasoning behind this assumption is a combination of a lack of recognized and objective approaches to measuring hospital size. Furthermore, the explanation is that it is difficult to secure a sufficient amount of degrees of freedom to include more variables to adjust for quality.
It is not realistic to assume that hospitals can adjust all their inputs quickly as was the case in the 'direct approach' [44]. Most studies indicate that hospitals cannot adjust all inputs on an efficient frontier does not appear to exert much influence on the quality of the estimation of economies of scale.

5. Conclusion

Overall the study shows that when applying two alternative specifications of translog cost functions, there are significant to moderate long-run economies of scale for all size groups of Danish hospitals in 2004-2006. These results indicate that the medium-sized subgroups of hospitals are more efficient than the small subgroups. According to the scale results, it may be appropriate to concentrate hospital production, such as health services in larger hospitals in the Danish hospital system than was the case in 2004-2006. In other words, it is not known whether the unit cost will decline (be U-shaped) or it will be L-shaped when hospital size increases above 1200 beds. Overall, this study supports the hypothesis that there may be cost advantages (or no disadvantages) for the smallest sub-group in producing hospital services in large hospitals, because it is outside the range of data used here. Moreover, the 'optimal' number of beds per hospital is estimated to be 275 beds per site within a 95% confidence interval between 130 to 585 beds per hospital. This is roughly in line with international results.

The analysis concludes that the translog model should be preferred, because it is outside the range of data used here. However, policy conclusions should be drawn with caution, because the findings are solely based on panel data for the years 2004-2006. The findings reported here should be investigated further, for example, through the use of additional data, alternative estimation methods such as data envelopment analysis (DEA), and estimation of potential efficiency gains from consolidation.
Appendix I

(1) Short run Quadratic (1a), Translog (1b) and Cobb Douglas (1c) cost functions:

\[ C(Q, K) = \beta_0 + \beta_1 Q + \frac{1}{2} \beta_2 Q^2 + \beta_1 K + \frac{1}{2} \beta_3 K^2 + \beta_4 QK \]  
\[ \ln C(Q, K) = \beta_0 + \beta_1 \ln Q + \frac{1}{2} \beta_2 \ln Q^2 + \beta_1 \ln K + \frac{1}{2} \beta_3 \ln K^2 + \beta_4 \ln Q \ln K \]
\[ \ln C(Q, K) = \beta_0 + \beta_1 \ln Q + \beta_1 \ln Q + \beta_1 \ln K \]  

(2) The optimal hospital production unit size measured in terms of beds is calculated from the short run cost functions (1a-c) by application of the envelope condition:

\[ \frac{\partial C(Q, K)}{\partial K} = \beta_1 + \beta_2 K + \beta_4 Q = 0 \quad \Rightarrow \quad K = -\frac{\beta_2 Q}{\beta_4} \]  
\[ \frac{\partial \ln C(Q, K)}{\partial K} = \beta_1 + \beta_2 \ln K + \beta_4 \ln Q = 0 \quad \Rightarrow \quad \ln K = -\frac{\beta_4 \ln Q}{\beta_2} \quad \Rightarrow \quad K = e^{-\frac{\beta_4 \ln Q}{\beta_2}} \]  
\[ \frac{\partial \ln C(Q, K)}{\partial \ln Q} = \beta_1 = 0 \quad \Rightarrow \quad \text{unfeasible} \]  

Calculation of long run cost function

The long run cost function is calculated from the short run cost function (1a,b) by substitution of the optimal number of beds (2a, 2b respectively) derived by the envelope condition (2). For the long run Quadratic cost function this yields:

\[ C(Q_t) = \beta_0 + \beta_1 Q_t + \frac{1}{2} \beta_2 Q_t^2 + \frac{1}{2} \beta_1 \left( -\beta_2 \ln Q^2 + \beta_2 \ln Q - \beta_2 \ln Q + \beta_2 \right) + \beta_3 \left( -\beta_2 \ln Q + \beta_2 \right) \]

After mathematical reduction the long run cost function can be reduced to:

\[ C(Q_t) = \beta_0 + \beta_1 Q_t + \frac{1}{2} \beta_2 \left( \frac{\ln Q + \beta_2}{\beta_2} \right)^2 \]

We omitted the long run Translog cost function since the only differences from the above mentioned quadratic cost function is that the total DRG-production value \( Q_t \) is replaced by logged levels.

Appendix II

Calculation of long run economies of scale

The expression for long run economies of scale (3a, 4a) is calculated from the long run cost function (2a) and (3, 4 respectively).

(3) Long run economies of scale - Translog cost function

\[ SE_1 = \frac{\delta \ln C(Q_t)}{\delta \ln Q_t} = \beta_1 - \beta_2 \frac{\beta_1}{\beta_2} + \left( \beta_2 - \beta_4 \right) \ln Q_t \]  

By mathematical reduction expression (3a) can be reduced to the following:

\[ SE_1 = \beta_1 - \beta_2 \frac{\beta_1}{\beta_2} + \left( \beta_2 - \beta_4 \right) \ln Q_t \]

(4) Long run economies of scale - Quadratic cost function

\[ SE_2 = \frac{\partial C(Q_t)}{\partial \ln Q_t} \bigg/ \frac{C}{Q_t} = \frac{\beta_1}{\beta_2} - \beta_2 \frac{\beta_1}{\beta_2} + \left( \beta_2 - \beta_4 \right) \ln Q_t \]  

By mathematical reduction expression (4a) can be reduced to the following:

\[ SE_2 = \frac{\beta_1}{\beta_2} - \beta_2 \frac{\beta_1}{\beta_2} + \left( \beta_2 - \beta_4 \right) \ln Q_t \]


