Hospital productivity and the Norwegian ownership reform – A Nordic comparative study

Sverre A.C. Kittelsen
Jon Magnussen
Kjartan Sarheim Anthun
Unto Häkkinen
Miika Linna
Emma Medin
Kim Rose Olsen
Clas Rehnberg

Frisch Centre, Norwegian University of Science and Technology, SINTEF, STAKES, Karolinska Institutet, Danish Institute for Health Services

UNIVERSITY OF OSLO
HEALTH ECONOMICS RESEARCH PROGRAMME

Working paper 2008: 10

HERO

Hospital productivity and the Norwegian ownership reform

– A Nordic comparative study

Sverre A C Kittelsen† [1], Jon Magnussen[2], Kjartan Sarheim Anthun[3], Unto Häkkinen[4], Miika Linna[4], Emma Medin[5], Kim Rose Olsen[6] and Clas Rehnberg[5].

1. October 2008

Health Economics Research Programme at the University of Oslo
HERO 2008

We acknowledge the contribution of Mikko Peltola, Timo Seppälä, Annette Søberg Roed, Jes Søgaard, Anni Ankjær-Jensen and Kjersti Hernæs participating in the Nordic Hospital Comparison Study Group (http://info.stakes.fi/nhcsg/EN/index.htm) in the collection of data and discussion of study design and results. Some of these may later become co-authors. We further thank the Norwegian Research Council (NFR) and the Health Economics Research Programme at the University of Oslo (HERO), as well as the respective employers, for financial contributions.

† Corresponding author, sverre.kittelsen@frisch.uio.no, Tel +47-91843889, Fax +47-22958815.

[1] Ragnar Frisch Centre for Economic Research, Gaustadalléen 21, N-0349 Oslo, Norway
[2] Department of Public Health and Community Medicine, Norwegian University of Science and Technology, N-7489 Trondheim, Norway
[3] SINTEF Health Research, N-7465 Trondheim, Norway.
[4] Centre for Health Economics, STAKES, FI-00531 Helsinki, Finland.
[5] Medical Management Centre, Karolinska Institutet, SE-171-77 Stockholm, Sweden.
[6] Danish Institute for Health Services Research, DK-2100 Copenhagen, Denmark.

Health Economics Research Programme at the University of Oslo
Financial support from The Research Council of Norway is acknowledged.
ISSN 1501-9071 (print version.), ISSN 1890-1735 (online), ISBN 82-7756-196-2

Abstract:

In a period where decentralisation seemed to be the prominent trend, Norway in 2002 chose to re-centralise the hospital sector. The reform had three main aims; cost control, efficiency and reduced waiting times. This study investigates whether the hospital reform has improved hospital productivity using the other four major Nordic countries as controls. Hospital productivity measures are obtained using data envelopment analysis (DEA) on a comparable dataset of 728 Nordic hospitals in the period 1999 to 2004. First a common reference frontier is established for the four countries, enveloping the technologies of each of the countries and years. Bootstrapping techniques are applied to the obtained productivity estimates to assess uncertainty and correct for bias. Second, these are regressed on a set of explanatory variables in order to separate the effect of the hospital reform from the effects of other structural, financial and organizational variables. A fixed hospital effect model is used, as random effects and OLS specifications are rejected. Robustness is examined through alternate model specifications, including stochastic frontier analysis (SFA). The SFA approach in performed using the Battese & Coelli (1995) one stage procedure where the inefficiency term is estimated as a function of the set of explanatory variables used in the second stage in the DEA approach. Results indicate that the hospital reform in Norway seems to have improved the level of productivity in the magnitude of approximately 4 % or more. While there are small or contradictory estimates of the effects of case mix and activity based financing, the length of stay is clearly negatively associated with estimated productivity. Results are robust to choice of efficiency estimation technique and various definition of when the reform effect takes place.

Keywords: Efficiency, productivity, DEA, SFA, hospitals
JEL Classifications: D24, I12, C14

1. Background

Efficient use of resources is a common health policy goal across virtually all health care systems. To obtain this goal a mix of policy initiatives aimed at purchaser and providers are used. In broad terms the health care systems in Europe can be characterised as either insurance based or tax based. Policy initiatives in social insurance based systems have primarily been directed towards increased competition between insurers (Saltman et al., 2005), while health care policy in tax based systems to a larger extent have focused on the relationship between purchasers and providers. Lately, however, we have seen several reforms whose aim has been the organization of the purchaser level (Saltman et al., 2007). Notable examples are the introduction of primary care trusts in the UK, and the recentralization of the Norwegian and Danish systems. Parallel to this the use of activity based financing, usually via patient classification systems such as the diagnosis related groups (DRGs), seems to increase in some countries (UK, Denmark) and decrease in others (Norway[1]).

The multitude of models used to organize the purchaser level and to pay providers can be explained both by historical, cultural and political factors. It is nevertheless interesting to observe such a variety of solutions in the organization of a sector where the overall goals (productivity, quality, access and cost containment) hardly differ between countries. This variety is the background for this paper. Our focus is on the centralization of hospital ownership following the Norwegian hospital reform in 2002, and its subsequent effect on one of the health policy goals; *efficiency in the utilization of resources.* Thus our basic intention is

[1] The share of activity based financing in Norway increased from an initial level of 30 % in 1997 to 60 % in 2003. It has since been reduced to 40 %, the main argument being fear of cream-skimming.

by the way of a cross-national comparative study to provide answers to the question; *does centralization of the purchaser and provider role provide effects in terms of increased levels of productivity?* Analyses of health care reforms are often difficult because of the problems related to separating the effects of the reform from other possible explanations, such as developments in medical technology, public expectations and policy changes. Thus the setting of this project is the hospital care systems of Norway, Sweden, Finland and Denmark. While similar, there are still substantial differences with regards to the financing of hospital services, the degree of (de)centralization, management structures and the degree of political involvement (Linna et al., 2008). While the focus is on Norway, we however use Finland, Sweden and (to some degree) Denmark as a control group. In order to analyse the productivity effects of the Norwegian reform, it is not necessary to use output prices, nor is it necessary to make any assumptions on the technological possibility sets of different countries and years, but only to establish a common reference set for measuring productivity and use the relative rates of transformation along the frontier of the reference set as weights instead of prices. In order to isolate the reform effect from the effect of other changes that may influence the productivity development of Norwegian hospitals, it is however necessary to control for variables that change over time for individual hospitals.

There are a relatively large number of analyses of hospital efficiency and productivity, but only a few are based on cross-national data sets (Mobley & Magnussen, 1998; Dervaux et al., 2004, Steinmann et al. 2004, Linna et al., 2006, Kittelsen et al., 2007). Such analyses often find quite substantial differences in performance between countries. Such differences may be due to the dissimilar hospital structures and financing schemes discussed above, but may also result from methodological problems. Cross-national analyses are often based on data sets that only to a limited extent are comparable – in the sense that inputs and outputs are defined and measured differently across countries. It is, however, well known that the way we measure hospital performance may influence the empirical efficiency measures (Magnussen, 1996). We therefore provide robustness tests and alternate specifications of the reform effect on productivity.

2. The Norwegian Hospital reform

In a period where decentralisation seemed to be the prominent trend, Norway in 2002 chose to re-centralise the hospital sector. There were four main elements in this reform. Firstly, the central governments took over responsibility for all somatic and psychiatric hospitals and other parts of specialist care from the 19 counties. As a result approximately 100 000 employees or 60 000 man-years and almost 60 % of county councils budget were transferred from the counties to the state. Second, the Minister of Health Affairs, as the general assembly for the regional health enterprises, became responsible for overall general management of specialist care. Third, the central government kept the five health regions that were established in 1974 as the organizational unit for coordination and steering. This implied that the new organization could start out with up-to-date descriptions of supply side and demand side factors, and with already prepared plans for restructuring.

While these three elements of the hospital reform imply a centralization of the hospital sector; the fourth element of the reform represents a decentralization: Both the health regions and the hospitals were organized as health enterprises, which are separate legal entities to a large extent modelled on commercial companies. Five regional health enterprises were established covering each of the five health regions. A board elected by the Ministry of Health is the body that is formally responsible to the ministry. The regional health enterprises have the statutory responsibility for ensuring the provision of health services to inhabitants in their geographical area, and each regional health enterprise is the owner of most health care providers in its region. The argument for choosing enterprises and not the directorate model is related to the aim of having politicians on arm length distance. The hospitals and clinics were merged into 42 (local) enterprises. This number was later (2003) reduced to 32. Both the numbers of regional and local enterprises have been under consideration after reform. In 2007 the number of regional health enterprises was reduced to 4, and the number of local health enterprises is presently 28. Representatives from the regional enterprises are often leaders of the boards that govern the local enterprises. The system can be characterized as highly integrated.

The reform had three main aims; cost control, efficiency and reduced waiting lists. In an early discussion of reform effects Magnussen et al. (2007) concludes that cost control has not improved but that waiting lists are down and that efficiency seemingly did increase. Their conclusion was however not based on a rigorous analysis. Thus the aim of this paper is to add to our knowledge of the effects of the reform by providing a better analysis of the effects of the reforms on hospital productivity.

3. Data

Meaningful cross national analyses must be based on comparable data. Measures of hospital efficiency and productivity require accurate measures of inputs and outputs. It is well known that efficiency measures are sensitive to operationalisation of hospital output (Magnussen, 1996). In a cross national analysis where we also utilise panel data there are several challenges both related to defining inputs and outputs similarly, and related to how one should adjust for wage/price differences. We have chosen the following approach:

Table 1: Relative input price indices for physician and nurses wage costs (including social security and pension payments) and overall hospital input costs.

	1999	2000	2001	2002	2003	2004
Physicians wage cost index						
Denmark	-	0.889	0.933	0.975	1.009	1.041
Sweden	0.855	0.898	0.945	1.006	1.054	1.076
Finland	0.770	0.767	0.795	0.860	0.939	1.000
Norway	0.892	0.907	0.943	0.981	1.131	1.159
Nurses wage cost index						
Denmark	-	1.223	1.270	1.317	1.360	1.412
Sweden	1.119	1.177	1.239	1.266	1.323	1.361
Finland	0.850	0.868	0.897	0.927	0.965	1.000
Norway	1.173	1.267	1.250	1.368	1.413	1.464
Overall hospital input price index						
Denmark	-	1.124	1.158	1.199	1.230	1.267
Sweden	1.030	1.066	1.115	1.147	1.184	1.210
Finland	0.867	0.878	0.900	0.931	0.968	1.000
Norway	1.087	1.134	1.140	1.214	1.272	1.306

No information available for Denmark in 1999.

Inputs are measured as operating costs, which for reasons of data availability are exclusive of capital costs and costs associated with teaching and research. Costs are initially measured in nominal prices in each country's national currency. To harmonize costs between the four countries over time we have constructed three separate price indices; one for physicians, one for nurses and one for "other resources". The wage indices are based on official wage date and include all personnel costs, i.e. pension costs and indirect labour taxes. The index for "other resources" is based on the harmonized CPI from Eurostat, and is converted to EURO using a purchaser parity corrected price index from OECD. We assume that costs are distributed between the three inputs with 20 % on physicians, 50 % on nurses and 30 % on other resources. Finally we construct a Paasche-index using Finland in 2004 as reference point. Note that this represents an approximation, the index will only hold exactly in the case of the relative use of inputs is constant over time. Table 1 shows relative wages for physicians and nurses and the overall deflator used in the analysis, where in each case is relative to the level for Finland in 2004[2].

Table 2: Hospital observations by country and year.

	1999	2000	2001	2002	2003	2004	Total
Denmark				54			54
Sweden			39	49	49	51	188
Finland	37	39	39	38	39	38	230
Norway	42	42	43	43	43	43	256
Total	79	81	121	184	131	132	728

Outputs are measured by using the Nordic version of the Diagnosis related groups (DRGs). We define six broad output categories; inpatient medical care, day medical care, inpatient surgical care, day surgical care, other patient care and outpatient visits. Within each category patients are weighted with a common Nordic cost weight calculated as the weighted average of the national cost weights. Outpatient visits were not weighted. While national weights may change from year to year, we have chosen to use the weights calculated for the year 2002 on the whole data set. Hence we expect to reduce problems associated with changes in coding practice. Table 2 shows the distribution of hospitals between countries and years, while table 3 shows summary statistics.

We note that Swedish hospitals are larger than the hospitals in the other countries. In some cases this is due to data limitations[3]. Also note that Swedish hospitals have a lower level of

[2] Even after harmonizing the input price levels there might be differences in the accounting systems that might cause some uncertainty and impact the comparability of the data. To the extent that these differences are country-specific they may influence estimated productivity levels, but not the estimates of the effect of the Norwegian reform.

[3] Some of the Swedish and a few of the Norwegian units for analysis are administrative aggregates rather than hospitals. The variance around the mean will be less due to this.

day care patients, and like the Norwegian hospitals have also a lower level of outpatient activity[4].

Table 3: Average values for real costs (input) and service production (outputs) per hospital by country, and outputs per unit of real costs.

	Denmark	Sweden	Finland	Norway	All
Input:					
Costs in MEUR	71.35	134.30	69.28	70.95	86.81
Outputs:					
Surgical inpatients DRGs	6795	10675	8204	6926	8288
Medical inpatients DRGs	9074	13470	9113	8825	10134
Surgical daypatients DRGs	0	881	1298	1474	1156
Medical daypatients DRGs	0	373	104	203	200
Other DRGs	1370	994	646	816	849
Outpatients	123611	141382	120308	63290	103368
Outputs per MEUR:					
Surgical inpatients DRGs	5.0	10.7	12.7	8.5	9.8
Medical inpatients DRGs	6.6	13.6	14.1	10.8	11.9
Surgical daypatients DRGs	0.0	0.9	2.0	1.8	1.4
Medical daypatients DRGs	0.0	0.4	0.2	0.2	0.2
Other DRGs	1.0	1.0	1.0	1.0	1.0
Outpatients	90.2	142.3	186.2	77.5	121.7
Observations	54	188	230	256	728
Years	1	4	6	6	6

4. Methods

Technical productivity

Efficiency and productivity are often used interchangeably. Strictly speaking productivity denotes the ratio of inputs and outputs, while efficiency is a relative measure comparing actual to optimal productivity. Most productivity indexes rely on prices to weigh several inputs and/or outputs, but building on Malmquist (1953), Caves et al. (1982) recognised that one can instead use properties of the production function, i.e. rates of transformation and

[4] The reporting of day-care and outpatient visits from the Swedish hospitals present local variations and has large scarcities. As for 2004, 23 percent of the data was missing. As for 2001- 2003 the share of missing data was even larger.

substitution along the frontier of the production possibility set, for an implicit weighting of inputs and outputs.

This analysis departs from Farrell (1957) who defined *technical efficiency* as:

$$\mathrm{TE}_i = \mathrm{Min}\left\{\theta \middle| (\theta \mathbf{x}_i, \mathbf{y}_i) \in T\right\} \qquad (1)$$

Where $(\mathbf{x}_i, \mathbf{y}_i)$ is the input/output vector for an observation *i*, and *T* is the technology or production possibility set. For an input/output-vector $(\mathbf{x}, \mathbf{y})$ to be part of the production possibility set, we need to be able to produce **y** using **x.**

If there are variable returns to scale, Farrell's measure of technical efficiency depends on the size of the observation, so that we can account for (dis)economies of scale. Even if we were interested in technical efficiency, our data does not allow us to estimate the scale properties since there are some instances in Sweden and Norway where we do not have each hospital as an observational unit, only administrative aggregates of hospitals. Our approach here is instead to use a measure of *technical productivity* by rescaling inputs and outputs[5]:

$$\mathrm{TP}_i = \mathrm{Min}_{\theta,\lambda}\left\{\theta \middle| (\theta \mathbf{x}_i, \mathbf{y}_i) \in \lambda T\right\}, \qquad (2)$$

where the homogenous envelopment of the technology λT contains all input-output combinations that are a proportionate rescaling of a feasible point in the technology set T .

Furthermore, it is not necessary to assume that the technologies of different countries and time periods are identical in order to compare productivity, as long as one has a common reference set. While it is normal to use a specific time period as a reference (as in Berg et al., 1992), or

[5] While this is formally identical to a ”CRS technical efficiency” measure, our interpretation here is instead that the reference surface is a homogenous envelopment of the underlying technology. This is the same assumption normally used in Malmquist indices of productivity change (see e.g. Grifell-Tatjé and Lovell, 1995), but we prefer here to analyse the effect of the reform on productivity levels rather than indices.

to construct geometric averages of indices based on consecutive time periods (as in Färe et al., 1994), the approach followed here is to use information from all technologies as the reference:

$$\overline{T} = \bigcup_{t}\bigcup_{c} T^{tc} \tag{3}$$

where the technology T^{tc} is for year *t* and country *c*. The reference set (3) is not itself a technology, only an envelopment of technologies, as is the rescaled set $\lambda\overline{T}$. The consequence of using $\lambda\overline{T}$ instead of T is to estimate *productivity* by weighing inputs and outputs by the relative slopes on the frontier that represents homogenous envelopment of all observed technologies, rather than estimate technical *efficiency* by weighing inputs and outputs by the relative the slopes on the frontier of a specific technology or its estimate.

DEA estimates of technology

In order obtain empirical measures we utilise the method known as data envelopment analysis (DEA), as developed in among others Charnes et al. (1978). DEA is built on three major assumptions:

Firstly we assume that the observed combinations of inputs and outputs are possible. A sufficient condition for this is that there are no measurement errors. Even though this assumption is implausible in its extreme form, measurement errors will be of less importance when the goal is to analyze the development of productivity over time. Secondly we assume free disposal, i.e. one can always use more of an input without reducing production or produce less without increasing the use of inputs. Finally we assume convexity, i.e. linear combinations of observations are also possible. When these three are combined an estimate of the production possibility set in a country *c* and year *t* is given as:

$$\hat{T}^{tc} = \left\{ (\mathbf{x},\mathbf{y}) \middle| \mathbf{y} \le \sum_{j\in N^{tc}} \lambda_j \mathbf{y}_j, \mathbf{x} \ge \sum_{j\in N^{tc}} \lambda_j \mathbf{x}_j, \sum_{j\in N^{tc}} \lambda_j = 1, \lambda_j > 0 \right\} \tag{4}$$

The empirical estimate for technical productivity is then given by the insertion of (4) and (3) in (2):

$$\widehat{\mathrm{TP}}_i = \mathrm{Min}_{\theta,\lambda}\left\{\theta \middle| (\theta \mathbf{x}_i, \mathbf{y}_i) \in \lambda \hat{\hat{T}}\right\} = \mathrm{Min}_{\theta,\lambda}\left\{\theta \middle| (\theta \mathbf{x}_i, \mathbf{y}_i) \in \lambda \bigcup_t \bigcup_c \hat{T}^{tc}\right\} \qquad (5)$$

Bootstrapping DEA estimates

To calculate confidence intervals for the DEA estimates of technical productivity we use a bootstrapping technique developed by Simar and Wilson (1998). This also provides bias corrected estimates. The methods assume that we know the data generating process (DGP) and can replicate this in a pseudo world where our original reference set frontier estimate plays the role of the known pseudo-frontier.

The empirical distribution of the efficiency scores from the original DEA run is used to estimate a smoothed distribution by a kernel density estimate (KDE) using reflection to avoid the accumulation of efficiency score values of one (Silverman, 1986). This is necessary in order to have a consistent estimator of the efficiency score distribution at the efficient part of the distribution. The pseudo observations are then created by projecting all inefficient observations to the DEA frontier and drawing randomly an efficiency score for each unit from the KDE distribution. A new DEA frontier is then estimated on these pseudo observations, each generated by mimicking the original Data Generating Process (DGP), as if the original DEA estimated frontier were the true frontier. The new frontier must lie on the inside of the original DEA frontier. We then know the bias of the estimate in our pseudo world, and can use this as an estimate of the bias of our original estimator.

The estimated bias is used to calculate a corrected estimate of the original productivity measure for each observation. The replication is done in a large number of iterations (2000), and the resulting distribution is assumed to be an estimate of the real sampling distribution, and we are thus able to calculate mean bias-corrected estimates, standard errors of means and confidence intervals for the productivity estimates of each observation $\widetilde{TP}_i$ as well as for groups of observations such as countries and time periods.

Second stage regression methods

To test the statistical association of the productivity estimates with variables that are not inputs and outputs, including the Norwegian hospital reform, a second stage regression analysis is conducted. The bias-corrected estimates $\widetilde{TP}_{itc}$ from the bootstrapped DEA analysis is regressed on the explanatory variables:

$$\widetilde{TP}_{itc} = \alpha_0 + \alpha_r R_{itc} + \sum_j \alpha_j z_{jitc} + \alpha_t T_t + \alpha_i I_i + \varepsilon_{itc} \tag{6}$$

which is the fixed effect (FE) specification with a reform dummy *R*, a vector of other explanatory variables *z*, annual time dummies *T* and hospital dummies I_i. This FE model is tested against an ordinary least square (OLS) model with an F-test, and against a random effects (RE) model with a Hausman test. Both the OLS and the RE models include country dummies, but these are superfluous in the FE model.

Note that the distribution of the bias-corrected productivity estimates that result from the bootstrap analysis does not have a mass point at 1.0, and these are therefore not censored. The TOBIT analysis widely used in the literature for second stage analysis of DEA estimates is therefore inappropriate (Simar and Wilson, 2007)[6].

Stochastic frontier analysis (SFA) method

[6] Simar and Wilson (2007) further argue that a separability assumption is needed for the two-stage approach to be consistent when estimating technical input efficiency, in essence that the influence of the second stage variables be on the radial efficiency only and not on the mix of inputs and the level of outputs. Since we are concerned with technical *productivity* rather than efficiency, i.e. with differences in productivity derived from a common set of output weights rather than distance behind the frontier of the production possibility set, we do not feel it necessary to make such assumptions here.

As a robustness analysis, an SFA model is estimated, with assumptions as close as possible as those used in the DEA model. When there is only one input, the technology set can be expressed by the equivalent input requirement function:

$$F(\mathbf{y}) = \text{Min}_x \left\{ x \middle| (x, \mathbf{y}) \in T \right\} \tag{7}$$

which is also equivalent to a cost function when the input price is normalised to 1 as in our application. Technical efficiency is here the ratio of necessary to actual inputs

$$TE_i = \text{Min}_\theta \left\{ \theta \middle| (\theta x_i, \mathbf{y}_i) \in T \right\} = \text{Min}_x \left\{ \frac{x}{x_i} \middle| (\frac{x}{x_i} x_i, \mathbf{y}_i) \in T \right\} = F(\mathbf{y}_i) / x_i \tag{8}$$

which is equivalent to the cost efficiency defined as the ratio of necessary to actual costs. The Battese & Coelli (1995) SFA model in its cost function form can be estimated using the Frontier 4.1 computer program (Coelli, 1996), with F(**y**) parameterised as a Cobb-Douglas function $F(\mathbf{y}) = B_0 \prod_j y_j^{\beta_j}$. Since there are many zeroes in the DRG data, only two outputs can be used, outpatients and DRGs (j=1,2). To interpret the results as productivity measures, we need to impose constant returns to scale, $\sum_j \beta_j = 1$, and achieve this by dividing by the DRG output[7]. The estimated model is then

$$\ln\left(\frac{x_{itc}}{y_{itc}^2}\right) = \beta_0 + \beta_1 \ln\left(\frac{y_{itc}^1}{y_{itc}^2}\right) + U_{itc} + v_{itc} \tag{9}$$

$$V_{itc} \sim N(0, \sigma_V^2),\ U_{itc} \sim N(m_{itc}, \sigma_U^2) > 0 \tag{10}$$

where the V are normal error terms, but the U are one-sided inefficiency terms truncated at zero and with expected value as a function of the explanatory variables of the same form as (6), i.e. for the FE case,

FE: $$m_{itc} = \beta_0 + \beta_r R_{itc} + \sum_j \beta_j z_{jitc} + \beta_t T_t + \beta_i I_i \tag{11}$$

[7] To interpret the estimated productivity reference frontier as a consistent envelopment of the production possibility frontiers for country- and year-specific technologies, certain regularity conditions might be needed. Since we only use SFA as a robustness exercise in this analysis, we do not pursue these conditions further.

With constant returns to scale and all countries and time periods in the estimation, $e^{-U_{itc}}$ will have the interpretation of estimated technical productivity rather than technical productivity. To compare the SFA results with the second stage DEA results above, we report the estimated marginal effect of an explanatory variable z_j on productivity at the mean calculated as $-\beta_j e^{U_{itc}}$ rather than the estimated coefficients themselves.

5. Results and discussion

In Kittelsen et al. (2007) the productivity measures for Finland and Denmark are estimated as higher than for Norway, and these are again generally higher than the measures from Sweden. As they stand, these are pure productivity numbers, and may or may not reflect underlying differences in possibilities or country-specific factors that are not accounted for, rather than differences in efficiency. We pursue the differences between countries elsewhere (Linna et al., 2008), as in this paper the focus is on the effects of the Norwegian hospital reform. In this context the purpose of including the other three countries is to provide a control group.

Figure 1 shows the mean bias-corrected productivity measures for the three countries that have more than one year of observations for the period 1999-2004, normalised to the 2001 level for each country[8]. Also shown are the confidence intervals for these mean estimates for each country and year. The estimates are from the DEA bootstrap method, and do not reflect the second stage regressions or the parametric SFA estimates.

[8] 2001 is the earliest year with observations for three countries. The Danish observations are only for the year 2002, and have no influence on the control group productivity development, but are still influencing the estimated common reference frontier.

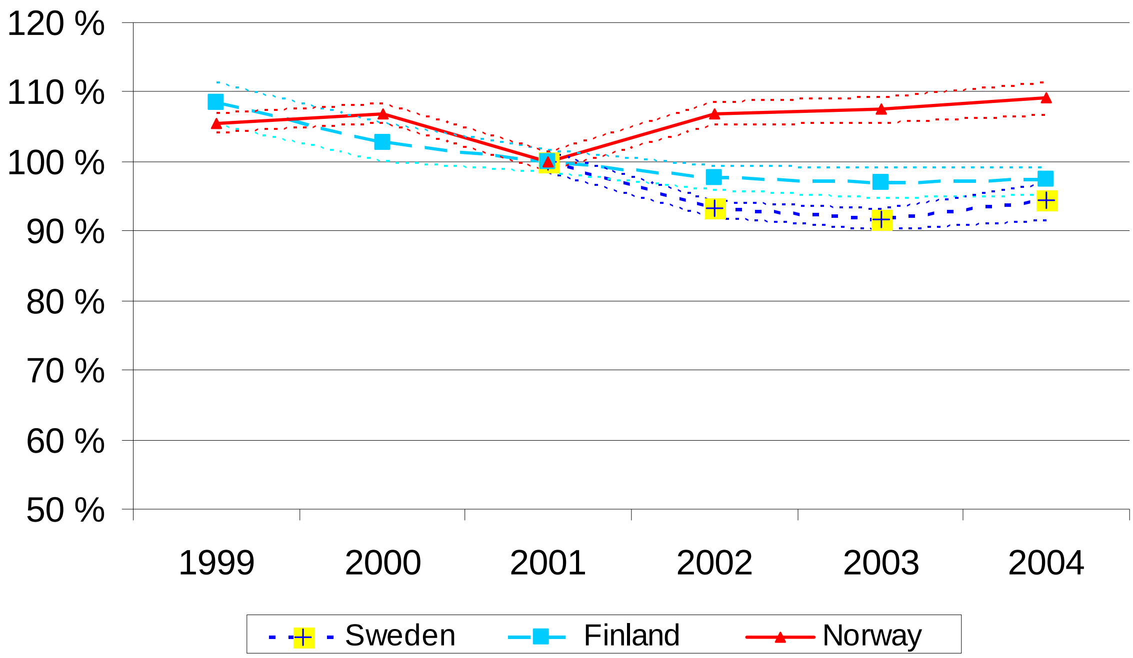

Figure 1: Average bias-corrected productivity levels and 95% confidence intervals by country and year, normalised to 2001 = 100% for each country

The figure clearly shows how the productivity development for the three last years almost coincide, and that the Norwegian development was characterised by a slight productivity increase from the years before the reform of 2002 to the years after the reform, while Sweden and Finland showed a decline over the same periods. The figure points to a specific Norwegian productivity development between the periods that could be associated with the reform. The underlying numbers shows that both the output of hospital services and of hospital costs have increased substantially in Norway over the period, but that former have increased somewhat more than the latter.

The purpose of the second stage analysis is to see whether the productivity development of the individual hospitals varies sufficiently systematically between countries to be associated with the reform, and to separate the effect of the reform from the effect of other factors. Such factors will partly be structural changes not related to the reform, changes in treatment practice and treatment technology and changes in incentive structures not related to the reform. Thus we focus on:

- Changes in payment system, operationalised through the changes and variations in the share of activity based financing (ABF). The share of ABF has varied between years in Norway, but for the most remained constant in the other countries, although it varies between Swedish counties.
- Major structural and technological changes, operationalised through annual time dummies to capture time-varying effects not captured by other variables.
- Hospital specific heterogeneity, through use of hospital fixed effects or random effects models.

In addition to this we include variables that may correct potential measurement errors. The variables included in the analysis are:

- Case-mix index (CMI) is included because we are not fully satisfied that the DRG based case mix adjustment full captures the variation in our material. Increased CMI implies a more resource demanding case-mix
- Length of stay deviation (LOS_D) is calculated as the difference between actual length of stay and the length of stay we would have expected had all patients in the hospital had average DRG-specific length of stay. This variable will in part correct for measurement error in the output variable and in part reflect true changes in productivity.
- Country dummies to capture country effects that may reflect institutional and geographic differences that are not captured by other variables. These dummies will be redundant in the hospital fixed effects model as each hospital belongs to the same country in all periods.

Descriptive statistics are given in table 4.

Table 4: Average values over individual hospitals (standard deviations) of independent variables in second stage regression, by year.

	1999	2000	2001	2002	2003	2004
Case-mix index (CMI)	0.79	0.79	0.83	0.85	0.85	0.82
	(0.09)	(0.08)	(0.09)	(0.10)	(0.10)	(0.09)
Length of stay deviation (LOS-D)	0.93	0.93	0.96	0.98	0.97	0.97
	(0.11)	(0.11)	(0.12)	(0.12)	(0.12)	(0.13)
Activity based financing percentage (ABF)	0.27	0.26	0.29	0.22	0.33	0.27
	(0.25)	(0.25)	(0.27)	(0.31)	(0.34)	(0.30)

CMI is hospital DRG points per patient relative to total DRG points per patients in sample as a whole. LOS-D is DRG-weighted average hospital length of stay in each DRG relative to total length of stay in the DRG for the whole sample

In this analysis the effect of the hospital reform will manifest itself as a shift in the productivity variable relative to trend. It is therefore of importance how we operationalise the reform. Two approaches are chosen. In model A we capture the reform through a dummy variable for the Norwegian hospitals for the years 2002, 2003 and 2004, which is the three years following the reform. Thus we expect to see a shift in productivity from the first year. In model B we capture the reform through dummy variables for the years 2003 and 2004 only. This specification has two advantages. First we open up for a reform effect that is not necessarily immediate, second we can account for some data inaccuracies related to the distribution of costs between 2001 and 2002[9] by including a separate dummy variable for these years.

Our data consists of an unbalanced panel for the period 1999 to 2004. The regression results are based on the bias corrected productivity measures from the bootstrap analyses. The analyses are done using i) no hospital effects (OLS), ii) hospital random-effects (RE) and iii) hospital fixed-effects (FE). Fixed effects imply time invariant hospital specific dummies and

[9] The transfer of hospitals from counties to state implied a change of accounting system that broadly led to measured costs in 2001 being too low and measured costs in 2002 being too high.

only utilise within hospital variation over the time period. All the factors assumed to be constant over time (i.e. relative differences between countries) will be captured by these dummies. The random effects model will utilise between variations as well and therefore generally be more efficient. If there are omitted variables that are correlated with the explanatory variables, the estimates may, however, be biased. A Hausman test rejects random-effect for both models, and an F-test rejects the no-hospital effects in favour of the Fixed-effects specification. The results of the second stage FE regression analysis are given in table 5, while the OLS and RE models are not included in the table since these specifications were rejected.

Table 5: DEA and SFA regression results. Marginal effect on productivity at sample mean and t-values.. N=728, * implies $p<0.05$, ** $p<0.01$

Model	DEA.A	DEA.B	SFA.A	SFA.B
Productivity estimates	Bootstrapped DEA	Bootstrapped DEA	SFA	SFA
Reform dummy	Reform_A	Reform_B	Reform_A	Reform_B
Hospital effects	Fixed (FE)	Fixed (FE)	Fixed (FE)	Fixed (FE)
Reform_A (Norway and 2002-2004)	**0.049**** (5.88)		**0.063**** (8.77)	
Reform_B (Norway and 2003-2004)		**0.043**** (4.05)		**0.148*** (2.01)
Norway 2001		-0.018 (1.57)		0.067 (0.88)
Norway 2002		**0.034**** (2.81)		**0.178*** (2.24)
Case-mix index (CMI)	0.093 (1.52)	0.090 (1.47)	-0.038 (1.24)	-0.334 (1.15)
Length of stay deviation (LOS-D)	**-0.285**** (6.37)	**-0.298**** (6.57)	**-0.298**** (13.34)	-0.043 (0.18)
Activity based financing percentage (ABF)	-0.0041 (0.14)	0.0053 (0.17)	**-0.099**** (3.11)	0.146 (0.80)
+ Hospital fixed effects	yes	yes	yes	yes
+ Annual time dummies	yes	yes	yes	yes
R^2	0.9056	0.9062		
F-test hospital fixed effect	14.35**	14.37**		
Hausman test random effect	31.47**	29.89**		
Log likelihood			1065.65	753.17

DEA OLS second stage regression with individual hospital bootstrap bias-corrected productivity estimates as dependent variable. SFA regressions with Battese & Coelli (1995) model specification transformed to marginal effects as -(regression coefficient)* (mean efficiency).

The effect of the hospital reform is positive and in the magnitude of more then 4 percentage points. We have performed several sensitivity analyses using alternative output specifications and the DEA method, and the results are robust to these changes (Kittelsen et al. 2007). Performing the analysis on the Norwegian hospitals only, gives similar results, although the reform effect loses its statistical significance, probably due to reduced number of degrees of freedom or multicollinearity between the reform dummy and time variables. Our interpretation is that the development in productivity in Norway parallels that in the other Nordic countries with the notable exception of the shift resulting from the reform.

A change in case-mix index does not have any effect on our estimates of productivity, while increased positive deviation from expected LOS is associated with reduced productivity (as expected). We also note that we fail to detect any effects of the changes in activity based financing in the preferred fixed effect models. The random effect specification associates higher productivity from higher ABF, with a 5 per cent confidence level, a result that is more in line with earlier studies (e.g. Biørn et al. 2003). This may be due to the fact that the FE model eliminates any variation in productivity that is associated with differences in ABF between hospitals, including those between Swedish countries, while the RE models includes these.

In addition to the second stage regression analysis of the DEA estimated productivity measures, we have examined robustness by estimating an SFA model with the same set of explanatory variables, including time and hospital fixed effects. The inefficiency term is estimated as a function of these explanatory variables, applying the single-stage approach of Battese and Coelli (1995). The marginal effects are shown in the rightmost two columns of table 5. For the model A with the reform effect as the years 2002-4 in Norway, the SFA results are very similar, with a highly significant reform effect of 6.3 per cent. While the effects of CMI and LOS are similar as well, there is a counter-expected negative productivity effect from ABF, which may be due the limited variability of this variable. For model B, the reform effect is quite a lot larger, but much less precisely estimated, as are several of the other coefficients. For this specification at least, the functional form imposed on the technology by SFA may limit the overall goodness of fit of the model.

6. Concluding comments

Our analysis indicates that the centralization of hospital ownership in the Norwegian ownership reform has had a positive effect on the productivity level of hospitals in the order of 4 per cent. The analysis does not show which aspects of the reform that have contributed to this improvement, but by controlling for some key variables the results indicate that the reform effect does not work through changes in case-mix (CMI), length of stay (LOS) or changes in activity based financing (ABF). Use of the Nordic countries as a control group has been useful by providing enough degrees of freedom to give the results statistical significance, and further shows that the reform effect is not due to changes in technology or other circumstances that are common to the Nordic countries. The reform effect is also robust to various model specifications.

Two themes emerge as areas for future research. First we need to acknowledge that “health reform”, even in the form of centralized ownership, is a somewhat fuzzy type of intervention. Thus more knowledge is needed before we can say *why* a recentralization of ownership would lead to higher levels of productivity. Possible explanations are structural changes – e.g. that the number of health enterprises has decreased substantially – better management or a tighter fiscal control. Previous analysis of the reform (Magnussen et al, 2007) suggests that the latter has not been the case, thus we are left with structural changes and better management. Both these factors, however, need closer investigation. Second, the results also indicate that there are clear differences in the productivity levels of the Nordic countries, and explanations of these differences will be the object of further research.

References:

Battese GE, Coelli TJ. A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data. Empirical Economics 1995; 20; 325-32

Berg SA, Førsund FR, Jansen ES. Malmquist indices of productivity growth during the deregulation of Norwegian banking 1980-1989. Scandinavian Journal of Economics, Supplement 1992; 94; 211-88

Biørn E, Hagen, TP, Iversen T, Magnussen J. The effect of activity-based financing on hospital efficiency: A panel data analysis of dea efficiency scores 1992-2000. Health Care Management Science 2003; 6; 271-83

Caves DW, Christensen L R, Diewert WE. The economic theory of index numbers and the measurement of input, output, and productivity. Econometrica 1982; 50; 1393-414

Charnes A, Cooper W, Rhodes E. Measuring the efficiency of decision-making units. European Journal of Operational Research 1978; 2; 429-444

Coelli T J. A Guide to Frontier Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation, Working paper 7 1996. Centre for efficiency and productivity analysis (CEPA), University of New England

Derveaux B, Ferrier GD, Leleu H, Valdmanis V. Comparing French and US hospital technologies: a directional input distance function approach. Applied Economics 2004; 36; 1065

Färe R, Grosskopf S, Lindgren B, Roos P. Productivity developments in Swedish hospitals; a Malmquist output index approach. In Charnes A, Cooper W, Lewin AY, Seiford LM. Data envelopment analysis: Theory, methodology and applications. Massachusets: Kluwer Academic Publishers; 1994, 253-72

Farrell MJ. The measurement of productive efficiency. Journal of the Royal Statistical Society; 1957; 120; 253-81

Grifell-Tatjé E, Lovell CAK . A Note on the Malmquist Productivity Index, Economics Letters; 1995; 47; 169-75

Kittelsen SAC, Magnussen J, Anthun K: Sykehusproduktivitet etter statlig overtakelse: En nordisk komparativ analyse HERO Working paper, 1 2007. Health Economics Research Programme, University of Oslo

Linna M, Häkkinen U, Magnussen J. Comparing hospital cost efficiency between Norway and Finland. Health Policy; 2006; 77; 268-278

Linna M, Häkkinen U, Peltola M, Magnussen J, Anthun K, Kittelsen S, Roed A, Medin E, Rehnberg C. Comparing efficiency between Nordic hospitals. Manuscript 2008, STAKES, Helsinki.

Magnussen J. Efficiency Measurement and the Operationalization of Hospital Production. Health Services Research 1996; 31; 21-37

Magnussen J, Hagen T P, Kaarboe O. Centralized or decentralized? A case study of Norwegian hospital reform. Social Science and Medicine; 2007 (in press)

Malmquist, S. Index numbers and indifference surfaces. Trabajos de Estadistica; 1953; 4; 209-24

Mobley L and J Magnussen: An International Comparison of Hospital Efficiency. Does Institutional Environment Matter? Applied Economics; 1998; 30; 1089-1100

Saltman R, Bankauskaite V, Vrangbæk K (eds): Decentralization in health care. 2007: Open University Press

Saltman R, Busse R, Figueras J (eds): Social Health Insurance systems in Europe, 2005 Open University Press.

Simar L, Wilson PW. Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. Management Science; 1998; 44; 49-61

Simar L, Wilson PW. Estimation and inference in two-stage, semi-parametric models of production processes. Journal of Econometrics; 2007; 136; 31-64

Steinmann L, Dittrich G, Karmann A, Zweifel P. Measuring and comparing the (in)effciency of German and Swiss Hospitals. The European Journal of Health Economics; 2004; 5; 216-226