Productivity growth, case mix and optimal size of hospitals. A 16-year study of the Norwegian hospital sector

Anthun, Kjartan Sarheim, Sverre Andreas Campbell Kittelsen, Jon Magnussen

Title: Productivity growth, case mix and optimal size of hospitals. A 16-year study of the Norwegian hospital sector

Authors: Kjartan Sarheim Anthun, Sverre A.C. Kittelsen, Jon Magnussen

PII: S0168-8510(17)30024-6
DOI: http://dx.doi.org/doi:10.1016/j.healthpol.2017.01.006
Reference: HEAP 3683

To appear in: *Health Policy*

Received date: 9-11-2016
Revised date: 17-1-2017
Accepted date: 19-1-2017

Please cite this article as: Anthun Kjartan Sarheim, Kittelsen Sverre AC, Magnussen Jon.Productivity growth, case mix and optimal size of hospitals.A 16-year study of the Norwegian hospital sector.*Health Policy* http://dx.doi.org/10.1016/j.healthpol.2017.01.006

Productivity growth, case mix and optimal size of hospitals. A 16-year study of the Norwegian hospital sector

Authors

Anthun, Kjartan Sarheim [a, b] kjartan.s.anthun@sintef.no

Kittelsen, Sverre A C [c] sverre.kittelsen@frisch.uio.no

Magnussen, Jon [a] jon.magnussen@ntnu.no

Affiliations

[a]

Norwegian University of Science and Technology

Department of Public Health and Nursing

PO Box 8905

7491 Trondheim

Norway

[b]

SINTEF Technology and Society

Department of Health Research

PO Box 4760 Sluppen

7465 Trondheim

Norway

[c]

The Ragnar Frisch Centre for Economic Research

Gaustadalléen 21

0349 Oslo

Norway

Corresponding author

Kjartan Sarheim Anthun

kjartan.anthun@ntnu.no

+47 934 43 132

Highlights:

- Norwegian hospital productivity increased by 24.6 percentage points from 1999 to 2014
- Most of the productivity growth was related to hospital reform of 2002
- Estimated optimal hospital size was smaller than actual size of most hospitals

ABSTRACT

Background and objectives: This paper analyses productivity growth in the Norwegian hospital sector over a period of 16 years, 1999 to 2014. This period was characterized by a large ownership reform with subsequent hospital reorganizations and mergers. We describe how technological change, technical productivity, scale efficiency and the estimated optimal size of hospitals have evolved during this period.

Material and methods: Hospital admissions were grouped into diagnosis-related groups using a fixed-grouper logic. Four composite outputs were defined and inputs were measured as operating costs. Productivity and efficiency were estimated with bootstrapped data envelopment analyses.

Results: Mean productivity increased by 24.6 percentage points from 1999 to 2014, an average annual change of 1.5%. There was a substantial growth in productivity and hospital size following the ownership reform. After the reform (2003–2014), average annual growth was <0.5%. There was no evidence of technical change. Estimated optimal size was smaller than the actual size of most hospitals, yet scale efficiency was high even after hospital mergers. However, the later hospital mergers have not been followed by similar productivity growth as around time of the reform.

Conclusions: This study addresses the issues of both cross-sectional and longitudinal comparability of case mix between hospitals, and thus provides a framework for future studies. The study adds to the discussion on optimal hospital size.

Keywords: Health care reform; Health services research; Organizational efficiency; Healthcare Financing; Diagnosis-related groups; Hospital Economics

1 Introduction

In the past 20 years, many countries have undergone large changes in the way health care is organized, financed and delivered. Under the umbrella of new public management, there has been an increase in quasi-markets, choice and competition, and increased use of activity- and results-based financing. In traditionally public tax-based systems, such as the UK and Norway, public hospitals have been reorganized into trusts with a large degree of autonomy [1, 2]. At the same time, several countries have pursued a policy of centralization, both in terms of exploiting perceived scale efficiency in the provision of services and by shifting power from local to central authorities [2, 3]. The recession in 2009 spurred a policy of fiscal austerity that has put health care, together with other publicly funded welfare services, under pressure.

Health care reforms, as well as increased fiscal pressure, infer an increased focus on how resources allocated to health care are used. Efficient use of available resources is an important policy goal in all health care systems. Regulators and policy makers will typically be interested in the level of productivity, whether and at what rate productivity increases or decrease over time, and the relationships between productivity and different regulatory, structural and financial policy measures. Hospitals constitute a major part of the health care sector; therefore, policy makers are particularly interested in assessing their performance. However, comparisons of productivity across hospitals are inherently difficult because of differences in case mix. Differences in case mix are often controlled for by using patient classification systems such as the diagnosis-related groups (DRGs) when describing hospital activity [4].

In this paper, we describe and discuss how hospital productivity has evolved in Norway from 1999 to 2014. Our analysis of this 16-year period enabled us to look at productivity from the long-term perspective of a period that included one major health care reform. Before 2002, all hospitals were owned and operated by the counties, and the hospitals had long waiting times and large deficits. The counties could not levy taxes themselves, so there was gaming of the budgeting and consequently soft budgeting as additional funding was provided by the central government [3, 5, 6]. In 2002, hospital ownership was transferred from 19 counties to the central government, and the responsibility for the provision of services was given to five (currently four) regional health authorities. The regional health authorities organized hospitals through hospital trusts.

Following the reform, there was major structural changes as the number of hospital trusts has decreased through mergers and reorganizations. Some minor hospitals that were located near larger hospitals were closed after mergers and reorganization of services. As a result, several hospital trusts are now multi-sited hospitals, and some even administer several multi-sited hospitals. According to Jacobs et al. [7], entities used in productivity analyses must have discretion about the conversion from inputs to outputs, must capture the entire production process and must be comparable. This applies to hospitals, hospital trusts and multi-site hospital trusts. Throughout this paper, we denote the organizational units as “hospitals” while acknowledging that these units often encompass several locations or physical hospitals.

This paper comprises three parts. First, we propose a way of describing hospital activity that captures both longitudinal and cross-sectional differences in case mix. This is crucial for capturing the effects of changes in treatment procedures on hospital productivity and enables us to relate the observed changes in productivity to the institutional and structural changes that have taken place during this period. In addition to the hospital reform in 2002, there has also been a substantial transition from inpatient to day care and outpatient treatment. To determine the long-term effects of reforms and policy changes, it is important to use data over a long time span. This is not commonly done, and most hospital efficiency analyses are either cross-sectional or span 1 year before and 1 year after a reform [8, 9]. In this analysis, we used data envelopment analysis (DEA) with a long time series and case mix-adjusted output measures. A long term approach was also presented by Halsteinli et al. [10], who used data for 9 years in their analysis of child and adolescent mental health services, and that of Biørn et al. [11], who used a 10-year span when evaluating a hospital financing reform. However, these two Norwegian studies did not adequately adjust for potential longitudinal changes in case mix.

Second, we estimate Malmquist indices [12, 13] to analyse to what extent the observed changes in productivity resulted from technical change (the best becoming better) or from changes in relative efficiency (“catching up”). Technical change is based on the performance of the best practice hospitals and it is as such it is often the result of a general development rather than the institutional environment or local policy initiatives. Thus, the relative share of the catching up and technical change elements provide an indication of the relationships between institutional environment, policy measures and provider performance.

Third, following the reform in 2002, the average hospital size has increased substantially, through reorganizations, mergers and hospital closures. There are arguments both for economies and diseconomies of scale in the literature [13-15], and we measured scale efficiency and estimated optimal scale and tracked the changes in these variables.

It is difficult to hypothesize the effects of the reform on catching up or technical change because the reformed implied both centralization and decentralization. If the governance of hospitals is strengthened, we might expect increased homogeneity in the results and thus reduced variance behind the frontier. However, technical change may not necessarily coincide with the reform if efficient hospitals already have exhausted their potential for improvement.

2 Materials and methods

2.1 Measuring hospital inputs and outputs

There are two issues that must be dealt with in an analysis of productivity growth. First, differences in case mix *between hospitals* must be adjusted for. Relative cost-weights (DRG prices) can be used to aggregate individual episodes into larger groups of hospital activity. However, such aggregation requires the assumption that relative treatment costs are independent of hospital case mix and size. Moreover, the results are usually sensitive to the type of case-mix adjustment that is chosen [8, 16-18]. Too many output categories can artificially inflate the number of efficient hospitals because rarer combinations of outputs determine the estimated best-practice front. Using all DRG groups as output dimensions would give no degrees of freedom because the number of DRGs would surpass the number of hospitals. A different approach would be to aggregate all hospital activity into one group, but that would underestimate differences between hospitals. Our point of departure is that the aim of any aggregation of groups should be to capture essential structural differences between hospitals. The distinguishing trends in hospital care during this period were the decreased length of stay and increased use of same-day surgery and outpatient treatments. The composite types listed below are well suited for capturing these trends, while at the same time allowing hospitals with different strategies related to the types of treatment to be considered efficient. We distinguish here between four composite types of hospital activity to summarize the activity and the technological profile of each hospital:

- Emergency inpatient discharges
- Elective inpatient discharges
- Day care treatments
- Outpatient visits and treatments

The second issue is related to case-mix comparability *between years.* Over the period covered in this analysis, there have been changes in the documentation of diagnoses and procedures, changes in DRG rules (grouper logic) and technological changes (which may shift patients between different types of hospital activities). The funding scheme was set up to accommodate this by annual adjustments of relative prices and, when needed, by changing the DRG logic (i.e., which procedures and treatments belong to which group). Consequently, technological changes may be captured by the annual changes in either DRG logic or the cost weights, and we would be unable to capture their effect on productivity. To avoid this, we regrouped all hospital episodes using a common grouper. We also applied fixed cost-weights. Technological change was thus captured as productivity growth and was not “hidden” in changes in relative prices or changes in grouper logic. For more details, we refer to the online supplementary material which describes in detail the output and input data.

Inputs were measured as deflated total hospital operating costs. These adjusted costs only included production in the four composite outputs; i.e. only DRG activity [19-21]. These costs are routinely used in productivity analyses and have been shown to have good comparability across time and units [19-21]. Capital costs and personnel were not used as inputs because of the lack of longitudinal comparability as data were available only for shorter time spans. The source of the hospital costs was Statistics Norway.

Information about hospital episodes in all Norwegian somatic hospitals in the period 1999–2014 were provided by the Norwegian Patient Register. Hospitals providing only elective care were excluded.

We aggregated the hospital activity data to the level of the cost data. The level of analysis was the operational ownership level; i.e. hospitals before the hospital ownership reform in 2002 and hospital trusts after the reform. There were 506 observations after aggregation.

2.2 Methods: estimating productivity and efficiency

Productivity is the relationship between inputs and outputs, whereas efficiency is the relationship between the observed productivity and the best possible productivity. A production frontier is the boundary of the production possibility set; the frontier thus describes the optimal possibility of conversion of inputs to outputs. Two related but distinctively different methodologies are commonly used for estimating frontiers and subsequently efficiency [7, 22-24]. Stochastic frontier analysis (SFA) is a parametric approach in which the efficiency frontier has a specific functional form, and the method incorporates random errors in the estimation of the production function. DEA is the other main approach. First suggested in 1957 [25], DEA was developed in 1978 [26]. The method has been used at different levels for comparing nations [27] and regions [28] but more commonly hospitals [23]. In Norway and the Nordic countries, various studies have used DEA for estimating hospital productivity [11, 13, 16, 21, 29-33]. This method is favoured in environments in which substantial measurement errors are unlikely and with multiple inputs and/or outputs [9, 24, 34]. This is assumed to apply to the public Norwegian hospital sector with the data described above, and we have thus chosen the DEA approach.

We estimated a production frontier and compared each hospital’s annual production to that frontier. A hospital on the frontier is, by definition, efficient and has a DEA score of 1.0. A hospital behind the frontier is considered inefficient and has a score <1.0, which represents the inefficiency as calculated as the relative distance to the frontier. As an example, a DEA score of 0.9 is interpreted as 10% inefficiency.

The Malmquist index of productivity growth compares the same decision-making unit between two years t and t+1. This index is then decomposed into an efficiency change and technical change. In the analysis of change, we compared two years so that a result >1.0 indicates increased productivity and <1.0 indicates decreased productivity. We assumed sequential (accumulated) frontiers. The productivity of a unit may be compared with best-practice hospitals at the frontier in the same year and in earlier years but never with observations in the future. It is likely that the productivity of a hospital in 1999 may be attainable by hospitals in 2014, but the opposite will generally not be the case.

In economic terms, it is common to distinguish between constant returns to scale (CRS) and variable returns to scale (VRS). CRS imply a linear conversion from inputs to outputs, whereas VRS technology allows a differentiated conversion ratio depending on the size of the hospital. For instance, fixed costs often imply increasing economies of scale where hospitals can increase their productivity by producing more as average costs decrease with increased volume. The term VRS is the general case, and CRS is a testable special case. In our analysis, we tested and failed to reject VRS, and therefore measured *scale efficiency* as the distance between the VRS and CRS frontier.

We estimated the optimal size of a hospital by multiplying each hospital's observed productivity by its cost, and then dividing by the relative distance to the optimal scale-efficient hospital. A confidence interval was created based on the distribution of the bootstrapped estimate of the relative distance to the optimal scale-efficient hospital.

When summarizing results from the analyses we calculate the annual mean level of productivity weighted by hospital operating cost. However, the annual scale efficiency of the average hospital is based on the sample average unit (SAU, which provides an implicit weighting). To calculate confidence intervals, the sample results are bootstrapped [35, 36].

3 Results

Table 1 shows the shares of each output relative to total case-mix adjusted number of episodes, and how average hospital size has changed during the period of analysis. The average hospital has more than tripled in size since 1999, primarily through reorganizations and mergers as the number of hospitals have more than halved in this period from 55 in 1999 to 22 in 2014. Outpatient treatment has increased in importance as the relative volume has increased by 45% since 1999. The share of day care of the total production gained 28%, while the share of inpatient treatment has reduced by 11% for emergencies and 5% for elective treatments.

----TABLE 1 ABOUT HERE----

In Figure 1 the development in productivity is shown with 1999 normalized to 1. We find an overall increase in annual mean productivity level by 24.6 percentage points from 1999 to 2014. There is a wide confidence interval, thus not all annual changes are significantly different from the previous year. In the years prior to the hospital reform, there are no significant changes in productivity. However, we find a substantial shift from 2002 to 2003, and also a positive trend from 2004 until 2014.

----FIGURE 1 ABOUT HERE----

Measures of total productivity growth, technical change and catching up are presented in Table 2. We found a large improvement in productivity around the time of the reform, with annual changes of 5–6% from 2001 to 2002 and from 2002 to 2003. However, in the years after 2003, annual changes were generally small, and in the whole period three of the years had a negative productivity growth (2000/2001, 2008/2009 and 2011/2012). The estimates for annual front shifts were significant only for 1999/2000 and 2003/2004. The estimates (and confidence intervals) for efficiency change for five of the years was >1.0, indicating that the mean hospital moved closer to the frontier. In these years, the hospitals were "catching up". In two years, the efficiency change was <1.0, which indicated that the mean hospital was not catching up but was "falling behind".

----TABLE 2 ABOUT HERE----

Although the mean size of hospitals more than tripled during this period, the average scale efficiency remained stable. There was a slightly higher scale efficiency in the period 2008 to 2014 than in 1999 to 2007. In 1999, 10 of the 55 hospitals were smaller than the estimated optimal hospital size, but in 2003 only three of the 30 hospitals were smaller than the estimated optimal hospital size. In 2012–2014, all of the hospitals were larger than the estimated optimal hospital size. Figure 2 shows the actual observed hospital sizes and the estimated optimal size (including 95% confidence intervals for the estimated optimal size). More detailed numbers relating to the scale efficiency and optimal size are available as online supplementary material.

----FIGURE 2 ABOUT HERE----

4 Discussion

The main policy reform relevant to this analysis was the hospital reform in 2002. The effects of the reform on hospital productivity were not obvious *a priori*, but one of the stated goals was to improve efficiency. It has been argued that the reform included both elements of centralization (state ownership) and decentralization (regional health authorities) [3]. Through centralization of the purchaser and provider role, productivity was expected to increase [37]. Another goal of the reform was to replace soft budgeting with hard budgeting. Early empirical work indicated that productivity increased by about 5% after the reform [37]. This present study confirms that the major shift in productivity in the 16-year period covered here occurred around the time of the reform. Overall, we found a productivity increase of 24.6 percentage points from 1999 to 2014. In the years after the reform (2003–2014), we found an average annual growth of <0.5%.

Separating productivity growth into “catching up” (the less efficient hospitals improving) and “technical change” (the production frontier shifting outwards) may give important information for policy makers. However, the results presented here fail to provide a clear picture. Mostly the estimates were not statistically significant, and for those that were, it was difficult to see a clear pattern and to link this to plausible policy explanations.

Following the hospital reform in 2002, the number of hospitals decreased substantially. This reduction was mainly the result of organizational mergers, although these were also accompanied by internal restructuring. The number of hospitals more than halved as the mean hospital size more than tripled. A major motivation for the transfer of hospital ownership to the state was to avoid unnecessary duplication of services across hospitals. Thus, we would expect scale efficiency to have increased and consequently that actual hospital size would have moved closer to the optimal hospital size. Although somewhat higher in the last years of our study, scale efficiency did not change significantly during the period. This suggests that merging hospitals is not a recipe for achieving efficiency. In our study, the aggregated productivity growth around the reform coincided with mergers, but later mergers did not coincide with similar growth.

There was larger activity growth in 2001-2003 than what was planned [5] due to budget gaming and soft budgeting constraint. Although this is also true for the costs (and subsequent deficits), hospitals seemed to increase size and productivity around the time of the reform. This could indicate that hospitals exploited economies of scale.

In this paper, the estimated optimal hospital size was quite small and, for the years 2012–2014, all of the hospitals were larger than the estimated optimal hospital size. However, the average optimal size did increase as the number of hospitals decreased from 2002 to 2005. This reflects that some of the merged hospitals performed well enough to define the optimal size from that point in time. However, these estimates rely on good case-mix adjustments, particularly for the largest hospitals. If cases were more severe or quality higher *within each DRG* in large than in small hospitals, the data do not fully capture the production of the large hospitals, and consequently, our estimate of optimal scale would have been downwards biased. However, a recent study reported no association between hospital size and differences in productivity, and no clear cost-quality trade-off pattern [30].

We have measured productivity at the operational level to compare hospitals before the 2002 reform with hospital trusts after the reform. Depending on the distribution behind the frontier, the level of analysis could cause some differences in the level of productivity; however, the development will be the same. The frontier depends only on the best units and is not as sensitive to marginal changes in the number of observations as are other methods. The scale efficiency measures should not depend on the level of analysis. Because the productivity change (as measured by the Malmquist index) is a product of scale change, technical change and pure efficiency change, it is possible to increase the overall productivity without an increase in all of the three components, and thus the productivity seemed to increase despite the fact that hospitals were larger than the estimated optimal size. Overall, these results for scale and size are consistent with the report by Kittelsen et al. [13], who estimated scale efficiency in Norway as comparatively high and found that hospitals were larger than the estimated optimal size.

Changes in the diagnostic registrations might be one explanation for the observed productivity growth. Norwegian hospitals are funded by a combination of fixed budgets and activity-based financing, and thus income relies partly on the recording of diagnoses and procedures. Recording of more diagnoses and procedures may have led to an increase in outputs through more expensive DRGs and subsequently higher estimates of technical productivity. However, this effect would only be short term as the prices are updated after 2 years to reflect average costs for each treatment in each group. Any large-scale wrongful upcoding would thus not yield long-term effects. There is currently more use of secondary diagnoses than in the early years, but this shift is linked only to a small extent to DRG-level prices [38]. In the present study, we used a common grouper and fixed weight for all years to avoid some of the issues related to upcoding. However, even after regrouping the data, we cannot exclude the possibility that some changes occurred because of upcoding rather than real changes in activity.

Some recent studies have incorporated quality indices as an output measure or have otherwise controlled for quality [30, 39, 40]. In this paper, we have measured only the hospital production by volume based on average costs, not the contents of that production. Both quality and actual health improvements are very important and may have changed over this long time span, and we suggest that further research is needed to answer this question.

In this paper, we have improved the comparability of the output measurement by using a fixed DRG logic and weights for all years. This enabled us to use a dataset of 16 years for hospital-level analyses which, to our knowledge, is an unprecedented long time span for hospital-level productivity analyses. We believe that it is important in studies of reform to apply a longitudinal perspective instead of analysing only 1 year before and 1 year after the reform.

5 Conclusion

The present study shows that mean productivity increased by 24.6 percentage points from 1999 to 2014, an average annual change of 1.5%. There was a substantial growth following the ownership reform in 2002. After the reform (2003–2014), the average annual growth was <0.5%. There was no evidence of technical change. The estimated optimal size is smaller than the actual size of most hospitals.

Conflict of interest statement

The authors report no conflicts of interest

Acknowledgements

This work was supported by The Research Council of Norway, grant number 214338. The authors wishes to thank Lise Rochaix and Youngho Oh who commented valuably on earlier drafts of this paper. Valuable comments were also provided by two anonymous reviewers of this journal.

6 References

[1] Martinussen P, Magnussen J. Health care reforms: the Nordic experience. Nordic health care systems Recent reforms and current policy challenges European Observatory on Health systems and Policies Series Maidenshead, Berkshire: Open University Press—McGraw-Hill Education 2009:21-52.

[2] Saltman RB, Bankauskaite V, Vrangbæk Ke. Decentralization in health care - Strategies and outcomes. Mc Graw Hill Open University Press, 2007.

[3] Magnussen J, Hagen TP, Kaarboe OM. Centralized or decentralized? A case study of Norwegian hospital reform. Social science & medicine 2007; 64:2129-37.

[4] Geissler A, Quentin W, Scheller-Kreinsen D, Busse R. Introduction to DRGs in Europe: Common objectives across different hospital systems. In: Busse R, Geissler A, Quentin W, Wiley M, editors. Diagnosis Related Groups in Europe: Moving Towards Transparency, Efficiency and Quality in Hospitals, 2011:9-21.

[5] Tjerbo T, Hagen TP. Deficits, soft budget constraints and bailouts: Budgeting after the Norwegian hospital reform. Scandinavian Political Studies 2009; 32:337-58.

[6] Hagen TP, Kaarbøe OM. The Norwegian hospital reform of 2002: central government takes over ownership of public hospitals. Health Policy 2006; 76:320-33.

[7] Jacobs R, Smith PC, Street A. Measuring efficiency in health care: analytic techniques and health policy. Cambridge University Press, 2006.

[8] Chowdhury H, Zelenyuk V, Laporte A, Wodchis WP. Analysis of productivity, efficiency and technological changes in hospital services in Ontario: How does case-mix matter? International Journal of Production Economics 2014; 150:74-82.

[9] O'Neill L, Rauner M, Heidenberger K, Kraus M. A cross-national comparison and taxonomy of DEA-based hospital efficiency studies. Socio-Economic Planning Sciences 2008; 42:158-89.

[10] Halsteinli V, Kittelsen SA, Magnussen J. Productivity growth in outpatient child and adolescent mental health services: the impact of case-mix adjustment. Social Science & Medicine 2010; 70:439-46.

[11] 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.

[12] Färe R, Grosskopf S, Lindgren B, Roos P. Productivity changes in Swedish pharamacies 1980–1989: A non-parametric Malmquist approach. International Applications of Productivity and Efficiency Analysis: Springer, 1992:81-97.

[13] Kittelsen SA, Winsnes BA, Anthun KS, Goude F, Hope Ø, Häkkinen U, Kalseth B, Kilsmark J, Medin E, Rehnberg C, Rättö H. Decomposing the productivity differences between hospitals in the Nordic countries. Journal of Productivity Analysis 2015; 43:281-93.

[14] Banker RD. Estimating most productive scale size using data envelopment analysis. European Journal of Operational Research 1984; 17:35-44.

[15] Dranove D. Economies of scale in non-revenue producing cost centers: implications for hospital mergers. Journal of Health Economics 1998; 17:69-83.

[16] Magnussen J. Efficiency measurement and the operationalization of hospital production. Health services research 1996; 31:21.

[17] Rosen AK, Loveland SA, Rakovski CC, Christiansen CL, Berlowitz DR. Do Different Case-Mix Measures Affect Assessments of Provider Efficiency?: Lessons from the Department of Veterans Affairs. The Journal of ambulatory care management 2003; 26:229-42.

[18] Vitikainen K, Street A, Linna M. Estimation of hospital efficiency—Do different definitions and casemix measures for hospital output affect the results? Health Policy 2009; 89:149-59.
[19] Samdata. SAMDATA Spesialisthelsetjenesten [Norwegian annual statistics on specialized hospital care]. Norwegian Directorate of Health, IS-2348, 2015.
[20] Anthun KS, Goude F, Häkkinen U, Kittelsen S, Kruse M, Medin E, Rehnberg C, Rättö H. Eurohope hospital level analysis: material, methods and indicators. Eurohope discussion papers Helsinki: THL, 2013.
[21] Medin E, Häkkinen U, Linna M, Anthun KS, Kittelsen SA, Rehnberg C, EuroHOPE-StudyGroup. International hospital productivity comparison: experiences from the Nordic countries. Health policy 2013; 112:80-7.
[22] Fried HO; Lovell Cak; Schmidt SSe. The measurement of productive efficiency and productivity growth. Oxford University Press, 2008.
[23] Hollingsworth B. The measurement of efficiency and productivity of health care delivery. Health economics 2008; 17:1107-28.
[24] Mortimer D. Competing methods for efficiency measurement: a systematic review of direct DEA vs SFA/DFA comparisons. 2002.
[25] Farrell MJ. The measurement of productive efficiency. Journal of the Royal Statistical Society Series A (General) 1957; 120:253-90.
[26] Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. European journal of operational research 1978; 2:429-44.
[27] Varabyova Y, Schreyögg J. International comparisons of the technical efficiency of the hospital sector: panel data analysis of OECD countries using parametric and non-parametric approaches. Health Policy 2013; 112:70-9.
[28] Halkos GE, Tzeremes NG. A conditional nonparametric analysis for measuring the efficiency of regional public healthcare delivery: An application to Greek prefectures. Health policy 2011; 103:73-82.
[29] Biørn E, Hagen TP, Iversen T, Magnussen J. How different are hospitals' responses to a financial reform? The impact on efficiency of activity-based financing. Health care management science 2010; 13:1-16.
[30] Kittelsen SA, Anthun KS, Goude F, Huitfeldt I, Häkkinen U, Kruse M, Medin E, Rehnberg C, Rättö H. Costs and quality at the hospital level in the Nordic countries. Health Economics 2015; 24:140-63.
[31] Linna M. Measuring hospital cost efficiency with panel data models. Health economics 1998; 7:415-27.
[32] Linna M, Häkkinen U, Peltola M, Magnussen J, Anthun KS, Kittelsen S, Roed A, Olsen K, Medin E, Rehnberg C. Measuring cost efficiency in the Nordic Hospitals—a cross-sectional comparison of public hospitals in 2002. Health care management science 2010; 13:346-57.
[33] Medin E, Anthun KS, Häkkinen U, Kittelsen SA, Linna M, Magnussen J, Olsen K, Rehnberg C. Cost efficiency of university hospitals in the Nordic countries: a cross-country analysis. The European journal of health economics 2011; 12:509-19.
[34] Hollingsworth B, Dawson P, Maniadakis N. Efficiency measurement of health care: a review of non-parametric methods and applications. Health care management science 1999; 2:161-72.
[35] Simar L, Wilson PW. Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. Management science 1998; 44:49-61.
[36] Simar L, Wilson PW. Statistical inference in nonparametric frontier models: The state of the art. Journal of productivity analysis 2000; 13:49-78.
[37] Kittelsen SA, Magnussen J, Anthun KS, Häkkinen U, Linna M, Medin E, Olsen KR, Rehnberg C. Hospital productivity and the Norwegian ownership reform: A Nordic comparative study. Stakes, 2008.
[38] Anthun KS, Bjørngaard JH, Magnussen J. Economic incentives and diagnostic coding in a public health care system. International Journal of Health Economics and Management 2016:1-19.

[39] Du J, Wang J, Chen Y, Chou S-Y, Zhu J. Incorporating health outcomes in Pennsylvania hospital efficiency: an additive super-efficiency DEA approach. Annals of Operations Research 2014; 221:161-72.
[40] Ferrier GD, Trivitt JS. Incorporating quality into the measurement of hospital efficiency: a double DEA approach. Journal of Productivity Analysis 2013; 40:337-55.

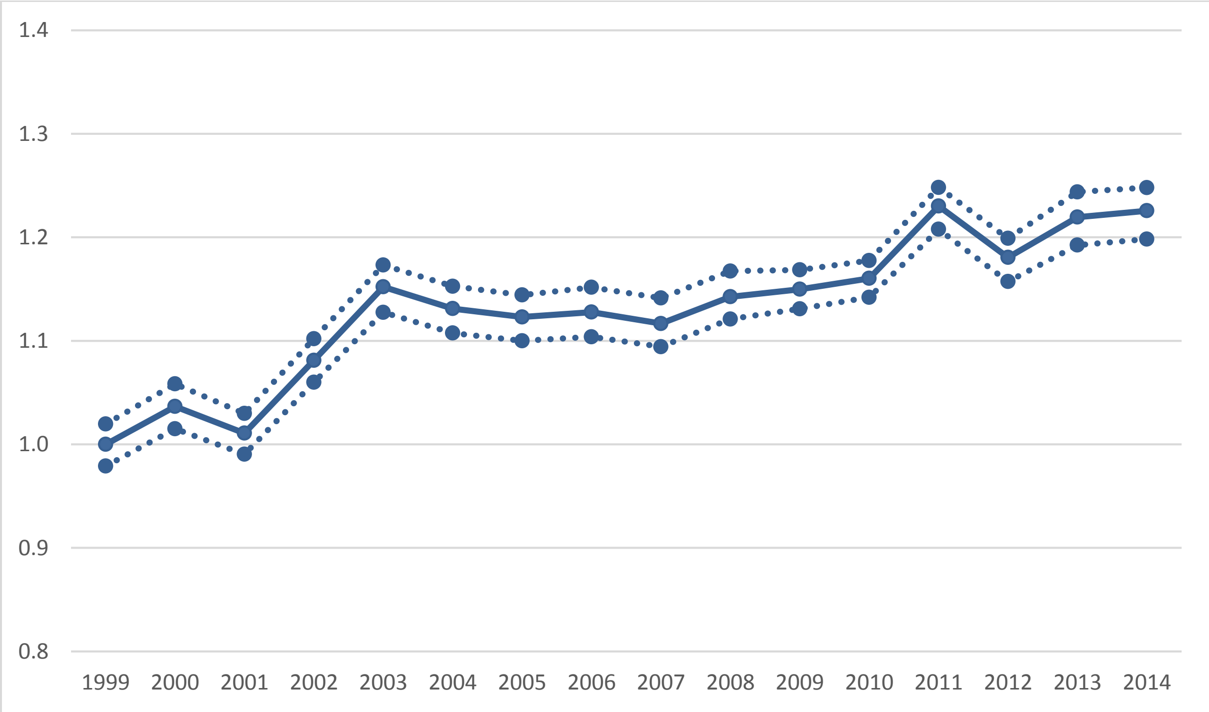

Figure 1 Bootstrapped productivity estimates, weighted mean unit, pooled base, and 95% confidence intervals, relative development from 1999

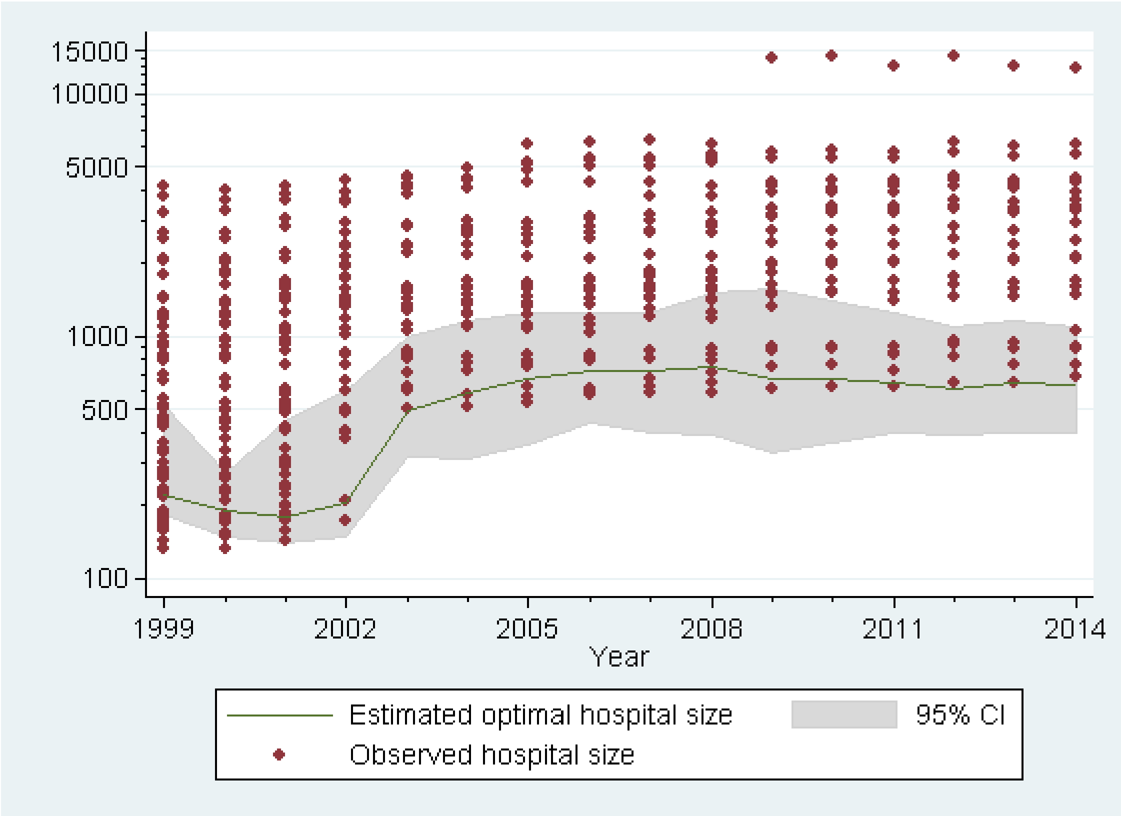

Figure 2 Observed hospital size and estimated optimal hospital size with 95% confidence intervals. The data are expressed on a logarithmic scale for hospital size, as measured by real operating costs (million NOK).

Table 1 Output shares, relative hospital size and number of hospitals, N = 506

Year	Emergency	Elective	Day care	Outpatient	Size (1999 = 1)	# Hospitals
1999	0.51	0.32	0.05	0.12	1.0	55
2000	0.50	0.32	0.05	0.13	1.0	54
2001	0.49	0.33	0.05	0.13	1.1	54
2002	0.49	0.33	0.05	0.12	1.7	36
2003	0.47	0.34	0.06	0.12	2.1	30
2004	0.48	0.34	0.06	0.13	2.1	30
2005	0.48	0.33	0.06	0.13	2.2	29
2006	0.48	0.33	0.06	0.13	2.3	29
2007	0.48	0.32	0.06	0.14	2.4	28
2008	0.48	0.32	0.06	0.14	2.4	28
2009	0.48	0.30	0.06	0.15	3.1	23
2010	0.49	0.29	0.07	0.15	3.3	22
2011	0.47	0.30	0.07	0.17	3.3	22
2012	0.46	0.30	0.07	0.17	3.5	22
2013	0.46	0.30	0.07	0.17	3.3	22
2014	0.45	0.30	0.07	0.18	3.4	22

Table note: Hospital size was measured as mean real operating costs relative to the 1999 mean.

Table 2 Decomposition of productivity growth (M) into catching up (MC) and front shift (MF), with 95% confidence intervals and weighted mean

Year	**N**	**Malmquist index of productivity growth**	**Efficiency change / catching up**	**Technical change / front shift**
1999/2000	54	1.046 (1.034–1.059)	0.963 (0.920–1.014)	1.086 (1.026–1.133)
2000/2001	53	0.971 (0.960–0.982)	0.966 (0.944–1.004)	1.005 (0.962–1.031)
2001/2002	19	1.056 (1.045–1.067)	1.069 (1.048–1.109)	0.987 (0.947–1.011)
2002/2003	30	1.062 (1.047–1.077)	1.024 (0.992–1.066)	1.038 (0.992–1.074)
2003/2004	30	1.000 (0.988–1.011)	0.936 (0.896–0.985)	1.069 (1.013–1.113)
2004/2005	28	1.007 (0.996–1.017)	1.018 (0.990–1.062)	0.989 (0.943–1.016)
2005/2006	29	1.008 (0.993–1.021)	1.003 (0.976–1.048)	1.005 (0.961–1.028)
2006/2007	28	1.005 (0.989–1.021)	1.009 (0.985–1.049)	0.996 (0.959–1.017)
2007/2008	28	1.021 (1.011–1.030)	1.028 (1.011–1.063)	0.993 (0.960–1.007)
2008/2009	21	0.981 (0.966–0.995)	0.984 (0.966–1.007)	0.996 (0.974–1.015)
2009/2010	21	1.001 (0.990–1.012)	0.988 (0.972–1.016)	1.013 (0.987–1.030)
2010/2011	22	1.054 (1.036–1.069)	1.032 (1.004–1.071)	1.021 (0.989–1.047)
2011/2012	22	0.958 (0.948–0.968)	0.963 (0.947–0.994)	0.995 (0.966–1.008)
2012/2013	22	1.037 (1.025–1.048)	1.042 (1.027–1.074)	0.995 (0.967–1.007)
2013/2014	22	1.006 (0.997–1.016)	0.996 (0.982–1.027)	1.010 (0.984–1.016)

Table note: The numbers represent a comparison of two consecutive years for each hospital, >1.0 indicates productivity growth and <1.0 indicates productivity decline.