STOCKHOLM SCHOOL
OF ECONOMICS
THE ECONOMIC RESEARCH INSTITUTE

Internal Markets and Performance in Swedish Health Care

Magnus Tambour and Clas Rehnberg

Working Paper No. 161
March 1997

Working Paper Series
in Economics and Finance

Internal Markets and Performance in Swedish Health Care

Magnus Tambour and Clas Rehnberg

Centre for Health Economics
Department of Economics
Stockholm School of Economics
PO Box 6501, S-113 83 Stockholm, SWEDEN

Fax: +46 8 30 21 15.
e-mail: hemt@hhs.se and hecr@hhs.se

March 1997

Abstract: A separation of provider and purchaser roles - together with performance-based reimbursement - were recently introduced in several county councils, which are the local governments that provide and fund almost all health care in Sweden. These changes have been motivated by efficiency and consumer choice arguments. In this paper we analyze the outcome of the former by studying differences in productive performance for county councils with and without new internal organizations. Performance is measured in terms of technical efficiency and productivity growth. The performance measures are estimated using data envelopment analysis. Using data for the 26 county councils between 1989 and 1994 we find differences in mean efficiency between the two groups, indicating higher technical efficiency in the group with performance based reimbursement (internal markets). Differences in productivity growth are also found, although the differences are significant in one period only.

Keywords: Internal Markets, Health Care, Productivity, Data Envelopment Analysis.

JEL Classification: D23, D24, I11.

1. Introduction

The Swedish health care system is an integrated system where the local governments control both the funding and the provision of health services. The system has been exposed for several reforms in recent years. One major reform is the purchaser/provider Split within the county councils. These changes involved the introduction of performance based reimbursement and competition among providers to increase incentives to use available resources more efficiently. The purpose of this paper is to analyze the outcome of these reforms by measuring performance at the county council level.

Frontier-based studies of the aggregate productivity in the Swedish health care are infrequent. Färe *et al*. (1995) and Färe, Grosskopf and Roos (1994) used data for samples of hospitals to estimate Malmquist productivity indices. The results showed a decline in productivity during the pre-reform period 1980-1991. Productivity studies of the reforms and organizational changes taken place in the Swedish health care system are also few. One study is Jonsson (1994) in which performance measures for one county council with purchaser/provider split were compared with measures for 14 county councils without new organizational systems. Jonsson used simple, non-frontier based, productivity estimates (ratio benchmark measures). Further, no hypotheses tests were made on the estimated performance measures. The results anyway indicated that productivity had improved more in the county council with performance-based reimbursement compared to the county councils in the control group. In this paper we perform a similar analysis, but with methods that allow representations of multiple outputs (and multiple inputs) on a richer set of data in terms of number of county councils.

The organizational changes in the Swedish health care system are described in Section 2 followed by a description of some models of hospital behavior in Section 3. The performance measures and the data envelopment analysis (DEA) estimation is described in Section 4. DEA is applied to production data in 1989 to 1994 for the 26 county councils to obtain estimates of technical efficiency and productivity growth. The input- and output variables as well as the performance results are presented in Section 5. Finally, Section 6 ends the paper with some concluding remarks.

2. Organizational change in Swedish health care

The Swedish health care system is basically an integrated system where the local governments (the county councils) control both funding and provision of health services. The internal allocation of resources has traditionally been based on budgets where providers receive an annual grant to cover all their services. This system of funding gave the providers a high degree of freedom to use resources for different purposes. The development in the health care system during the 1990s shows some radical changes in the principles for allocating resources in the system. The system has been exposed for several reforms introduced by the central government or the county councils themselves. One of the major reforms is the purchaser/provider split within the county councils. A common principle in the reform has been to induce politicians to concentrate on the interest of citizens by separating the consumer/purchaser and provider roles within the county councils. The providers in these counties remain under public ownership, but politicians have decided to not be represented on the boards of hospitals and health centers. They therefore have less decision-making power at the operational level. These changes aimed at increasing the incentives to use available resources more efficiently by introducing performance based reimbursement and competition among providers.

The movement toward the use of market mechanisms within public health care can be summarized as encompassing the following attributes:

- Contracts and performance-based reimbursement
- Establishment of collective purchasing units
- Freedom of choice for consumers
- Introduction of provider competition
- Provider autonomy

In contrast to earlier changes affecting the organizational and financial structure of the county councils, individual county councils now develop their own management control systems. Changes in such systems during the 1960s and 1970s were designed and developed centrally by the Federation of County Councils. The implementation of collective purchasing units has been structured differently. Some county councils have decentralized the purchasing function to local units corresponding to health districts and others have established central agencies acting as a collective purchaser of health care for all citizens. The models separating consumer and provider interests have been implemented successively. The collective purchasing units usually receive their resources based on population characteristics (number of inhabitants, age, etc.), which are used for purchasing health services from providers within the county council, but purchases from external providers are also allowed.

Through 1995 twelve of the 26 county councils had implemented some type of model featuring a purchaser/provider split. Most of those have only implemented internal markets in segments of their health service. However, five of those twelve county councils have implemented an organization based on a comprehensive purchaser/provider split. These county councils represent about one third of the total population in Sweden. Still, most county councils are still allocating resources by the traditional budget process. purchase of health services was first implemented in the surgery specialties. Specialties with a higher degree of uncertainty about costs and outcome and where the output of the service is more difficult to measure (psychiatry, geriatrics) were included later. Table 1 shows the county councils and specialties where a purchaser/provider split has been implemented.

Table 1: Implementation of purchaser/provider split.

Comprehensive system		
Year	*County Council*	*Specialty*
1991	Dalarna	Somatic care
1992	Stockholm	Part of surgery
	Örebro	Somatic care
1993	Bohuslän	All health care
	Sörmland	All health care

Source: Anell & Svarvar 1993, Bergman & Dahlbäck1995, Jonsson 1996

The implementation of internal markets has also changed the way providers are reimbursed. The budget process has been replaced by various performance-related reimbursement mechanisms (fee-for-service or fee-per-diagnosis). Such arrangements have been adopted slowly and applied where most appropriate.

3. Theoretical considerations of hospital behavior

In this study we limit the analysis to somatic illnesses, that is, mainly hospital services. There is no unified theory of hospital behavior but different models which differ according to the role of environmental and internal factors. Several models highlight the importance of internal factors (e.g., Harris 1977). Berki (1972) states that the internal structure of the hospital is of overriding importance as compared with environmental factors when it comes to explaining hospital behavior. The importance of internal factors draws attention to the principles for allocating resources and delegating responsibilities within the hospital. In health care organizations, as in any organization above a minimum size and complexity, there is a separation of ownership and management. This separation of ownership from control gives rise to problem of discretionary behavior by managers. The problem arises when one party (the principal) engages another party (the agent) to take action on behalf of the former. In principle, this framework has been used for a variety of agency relationships such as patient-physician, politician-bureaucrat, and shareholder-manager etc. The transfer (or change) of property rights between parties also has strong effects on the organizational structure. In public health care there are different complex patterns of transfers of property rights. If we consider the politicians as the principals there are several agents supposed to act in the interest of the former (managers, head of clinics etc.). Of major importance is of course the role of the physicians, who act as ”double”-agents on behalf of both politicians and the patients (Blomquist, 1991). Here, we are interested in the politician-physician relationship and how this is handled by payment systems as a tool of controlling the behavior of the agent.

One type of model focuses on different actors, where each actor is assumed to maximize his utility inside the hospital organizational framework. The outcome for the hospital is a function of these internal relationships. The most frequently studied actor is the physician. In a model developed by Pauly and Redisch (1973) the physicians dominate the principal-agent relationship through their professional position and informational advantages and they have access to any surplus of revenue over expenditures generated in the hospital. This model can also be used to analyze how physicians use their position to achieve other objectives than net incomes. In public health care systems the physicians have limited possibilities to maximize income, but can instead use their clinical freedom to achieve other objectives (e.g., medical research).

Another set of models consider the hospitals as an entity with a common objective. An objective frequently attributed to the hospital has been quantity-maximization of the output (e.g., Feldstein 1968). Newhouse (1970) combines these objectives in an equilibrium model where the reimbursement is based on some kind of fee-for-service-system, i.e., the hospital can influence its revenues.

Here we take the view that hospitals in the pre-reform system were reimbursed by global budgets (or block grants) not related to the output (or any other performance measure). In fact inputs were often used as proxies for outputs. This reimbursement system gave space for discretionary behavior. However, we do not know how this managerial discretion was used. Still, we assume that the lack of incentives for efficiency resulted in productivity losses. With the introduction of internal markets and performance-related reimbursement the degree of discretionary behavior has become more limited. The limitation of discretionary behavior can have had positive productivity effects, e.g., reducing the degree of X-inefficiency and - or - higher growth in productivity. At least, these were the intentions of the organizational reforms.

4. Productivity and DEA

There are a number of methods available for measuring (technical) efficiency and productivity growth. One way of representing the production technology is to use distance functions (Shephard, 1970). The distance function - which is a multiple-output generalization of the production function - can be conveniently estimated by data envelopment analysis (DEA), even without observing input- or output-price data. Since the distance function is the inverse of the efficiency measure proposed by Farrell (1957) it can also be used to measure (technical) efficiency as well as productivity growth by computing Malmquist productivity indices (Färe *et al.*, 1989). Moreover, the DEA approach is appropriate for our problem since it has the virtue that no behavioral assumptions (as cost minimizing behavior) have to be imposed.

Consider then a sample of K units (county councils) using $x \in R_+^N$ inputs in the production of $y \in R_+^M$ outputs. A multiple-input, multiple-output production technology can be represented by the output set, defined as

$$P(x) = \{y\text{: } y \text{ can be produced by } x\}. \qquad (1)$$

In this paper an output-based approach is used since it is reasonable that the county councils take the inputs as given, and hence ”choose” the level and mix of outputs. In an output-based approach, the production technology is completely characterized by the (output) distance function defined as

$$D_o(y,x) = min\left\{\theta\text{: } \frac{y}{\theta} \in P(x)\right\}. \qquad (2)$$

The distance function is less than, or equal to one if and only if the output *y* belongs to the output set $P(x)$. A unit is considered technically efficient if the distance function equals one and consequently, values less than one indicate inefficiency.

Estimating Technical Efficiency

Using a piecewise linear reference technology the distance function for unit *k* can be estimated as the solution to the following LP problem

$$\left[\hat{D}_o\left(y_k, x_k\right)\right]^{-1} = \max_z \left\{\theta : \theta y_k \leq Yz,\ x_k \geq Xz, 1_K' z \leq 1, z \in R_+^K\right\}, \qquad (3)$$

where y_k is an *M*-vector of outputs, x_k is an *N*-vector of inputs, *Y* is a (*M*×*K*) matrix of outputs, *X* is a (*N*×*K*) matrix of inputs and *z* is a *K*-vector of non-negative intensity variables. The estimated distance function satisfies $\hat{D}_{o,k} \leq 1$ (since unit *k* is represented in the reference technology) and returns to scale is assumed to be non-increasing. Other assumptions regarding returns to scale are also possible. Constant returns to scale (CRS) can be imposed by dropping the intensity sum restriction and variable returns to scale (VRS) can be modeled by restricting the intensity variables to sum to one.

Productivity change

Distance functions can also be used to analyze changes in productivity. Färe *et al.* (1989) showed how distance functions, estimated by DEA, can be used to compute Malmquist indices of productivity change. Following, Färe *et al.* (1989) the Malmquist productivity index between time period *t* and *t*+1 can be defined as

$$M_o^{t,t+1}\left(y^t, y^{t+1}, x^t, x^{t+1}\right) = \left[\frac{D_o^t\left(y^{t+1}, x^{t+1}\right)}{D_o^t\left(y^t, x^t\right)} \frac{D_o^{t+1}\left(y^{t+1}, x^{t+1}\right)}{D_o^{t+1}\left(y^t, x^t\right)}\right]^{\frac{1}{2}}. \qquad (4)$$

Values larger than one indicate increased productivity and values less than one indicate decreased productivity. The index can be interpreted as a measure of total factor productivity growth. For example, in a single-input single-output technology the Malmquist index (under constant returns to scale) simplifies to $M^{t,t+1} = \frac{y^{t+1}/x^{t+1}}{y^t/x^t}$. The Malmquist index can also be decomposed into two components as

$$M_o^{t,t+1} = \frac{D_o^{t+1}\left(y^{t+1}, x^{t+1}\right)}{D_o^{t}\left(y^{t}, x^{t}\right)} \left[\frac{D_o^{t}\left(y^{t}, x^{t}\right)}{D_o^{t+1}\left(y^{t}, x^{t}\right)} \frac{D_o^{t}\left(y^{t+1}, x^{t+1}\right)}{D_o^{t+1}\left(y^{t+1}, x^{t+1}\right)} \right]^{\frac{1}{2}} . \qquad (5)$$

The term outside the brackets measures change in efficiency or a ”catching up” to the frontier effect between period *t* and *t*+1. The term within brackets measures the shift in the production frontier as geometric mean of two ratios of distance functions. Values greater than one indicate improved efficiency and - or - progressive technology change and values less than one indicate the opposite.

Four separate LP-problems have to be solved to estimate the Malmquist index for each unit. The first two single period distance functions are solved using the formulation in (3). The other two distance functions need observations from two periods. The distance function $\hat{D}_o^t\left(y^{t+1}, x^{t+1}\right)$, where input and output observations from period *t*+1 are evaluated relative to the technology in period *t*, is estimated by (assuming CRS)

$$\left[\hat{D}_o^t\left(y_k^{t+1}, x_k^{t+1}\right)\right]^{-1} = \max_z \left\{\theta : \theta y_k^{t+1} \le Y^t z,\ x_k^{t+1} \ge X^t z,\ z \in R_+^K \right\}. \qquad (6)$$

Note that in the case of progressive technical change, y^{t+1} does not necessarily belong to $P^t\left(x^t\right)$. Hence, the cross-period distance function can take on values both greater or less than one. The second cross-period distance function is obtained by solving a similar LP-problem as in (6), although the time superscripts are switched.

It has been argued that constant returns to scale has to be imposed for the Malmquist index to be a total factor productivity index (Färe, Grosskopf and Roos, 1996). In the presentation of the Malmquist index results we report the estimates from the CRS model. For the (single period) distance function results from both the CRS and the non-increasing returns to scale models are presented.

5. Data and results

Data

As stated above most of the county councils have not introduced an internal market system (a purchaser/provider split) or performance based reimbursement. There is no official classification of the county councils as having implemented an internal market and those relying on the budget process. There are also differences regarding how performance-related reimbursement have been used in different specialties within each county council. In this study we have classified the county councils based on the sources presented in Table 1. According to those studies five county councils had implemented comprehensive systems of internal markets. Seven others have partially implemented systems based on market mechanisms and performance-related payments for short-term surgery and internal medicine. In the presented study we focus on the group with comprehensive internal market systems.

All production data are taken from various official records. It should be noted that the empirical application is concentrated to performance in short term (hospital) care. This is motivated by the fact that the internal markets have been focused (at least initially) on this type of health care delivery. We then use data for all hospital services (exc. psychiatry) within each county council in the comparison. Two variables are used as proxies for input use. The first variable is a cost measure for short-term care (*COST* in Table 2). The cost figures are taken from yearly published financial records (Landstingsförbundet, 1994a) and are deflated by the county council price index. Cost is used since data on labor are not easily accessible for the short term care. It is, however, recorded on an aggregated level, i.e., when somatic, primary and psychiatric health care is aggregated. The cost measure does however not reflect expenses for capital services. We therefore include the number of hospital beds (*BEDS*) as a proxy for capital cost. The figures for this input are taken from Landstingsförbundet (1994b)

As usual the conceptual output in health care delivery - change in health status - is difficult to measure which implies that a number of intermediate outputs will be used instead. Five types of output measures are used. The first is number of operations in short-term care (*OPERA*). These data are taken from the National Board of Health and Welfare-inpatient care register (Socialstyrelsen, 1996). The next output is number of admissions. This variable is divided in two variables: admissions in surgery- (*ADSUR*) and internal medicine (*ADMED*). The admission data are taken from the compilations made by the Federation of County Councils (Landstingsförbundet, 1994b). Finally, the last output category is number of physician visits in surgical care (*VISUR*) and internal medicine (*VIMED*), respectively. The data source is the same as for the admission data. Descriptive statistics are reported in Table 2.

Table 2: Descriptive statistics, inputs and outputs, 1989-1994. All variables except COST are measured in quantities (number of beds, operations etc.). COST in Million SEK, 1990 prices.

		1989	1990	1991	1992	1993	1994
	Max	7 714	8 863	9 868	8 607	9 080	9 621
COST	*Mean*	1 747	1 769	1 820	1 801	1 799	1 844
	Min	289	286	289	290	349	350
	Max	5 893	5 672	5 129	4 736	4 248	3 687
BEDS	*Mean*	1 328	1 289	1 227	1 134	1 050	975
	Min	238	238	238	238	199	186
	Max	113 090	113 367	114 402	127 913	145 536	146 785
OPERA	*Mean*	25 538	25 452	26 248	27 108	28 560	28 378
	Min	4 369	3 879	3 324	4 163	4 103	3 630
	Max	130 569	130 900	134 984	142 991	137 911	124 239
ADSUR	*Mean*	29 648	29 748	30 190	30 159	29 325	27 980
	Min	5 146	5 189	4 937	5 070	4 750	4 846
	Max	94 645	94 999	99 891	106 760	110 105	98 843
ADMED	*Mean*	24 018	24 178	24 856	25 729	26 036	25 333
	Min	4 738	4 909	4 868	4 675	5 055	5 021
	Max	1 013 500	934 100	899 600	986 000	1 011 900	964 100
VISUR	*Mean*	215 196	209 242	204 365	210 138	211 800	204 846
	Min	41 100	39 700	41 100	40 900	39 500	41 300
	Max	810 600	794 800	753 100	841 900	877 700	772 600
VIMED	*Mean*	138 177	137 377	134 915	141 181	142 023	138 600
	Min	22 000	21 900	22 600	22 000	26 200	24 300

The cost for short term care have been relatively stable over the period 1989 - 1994, but the number of beds on the other hand, have decreased steadily. For the output variables both (average) increases (OPERA, ADMED and VIMED) and decreases (ADSUR, VISUR) are found.

Performance results and tests

We estimated the distance functions for the 26 county councils in 1989 to 1994. Average results are given in Table 3. Using DEA-estimates of the distance functions we also computed

Malmquist productivity indices, for which average results are reported in Table 4. County council specific estimates are presented in Appendix 2. We also analyzed the effect of changing the reference technology by estimating the distance functions and Malmquist indices for the budget county councils without internal market county councils in the reference technology. These results are reported in Table 5. In order to test if there are significant differences between the two groups, a set of mean-difference tests was carried out. Those results are given in Table 6. It should be noted that the comparisons concern county councils with performance based allocation and the remaining county councils with budget based allocation.

Table 3: Sample means of the output distance functions. Non-increasing returns to scale (NIRS) and constant returns to scale (CRS).

		1989	**1990**	**1991**	**1992**	**1993**	**1994**
Nirs	Overall	0.953	0.957	0.962	0.961	0.952	0.939
	IM	0.950	0.954	0.983	1.000	0.999	0.995
	B	0.952	0.957	0.958	0.954	0.943	0.928
Crs	Overall	0.945	0.950	0.956	0.959	0.951	0.938
	IM	0.942	0.952	0.982	1.000	0.999	0.996
	B	0.945	0.950	0.951	0.952	0.943	0.927

The first row (*Overall*) shows the mean of the distance functions for the whole sample. The next row (*IM*) contains means for the county councils that have implemented internal markets in a full scale with performance based reimbursement. The mean for the remaining, budget, county councils are shown in the last row (*B*). From the result in Table 3 we observe wider variation between the two groups over time, where county councils using internal markets have higher efficiency on average. At the beginning of the period the potential output increase were about five percent for both groups. The situation is different in the end of the period. For example, in 1994 the internal market county councils were close to full efficiency whereas the budget group had an estimated potential output increase of about seven percent. From the county council specific results we note that for most of the county councils using internal markets efficiency were increased some year prior to the reform. The mean estimates thus show that county councils with internal markets did not differ from budget based ones in the beginning of the period. The qualitative results in this respect are similar for both the non-increasing and constant returns to scale technologies. Hence, this indicates that the results are not sensitive to the assumption of returns to scale.

Table 4: Sample means of the Malmquist index, efficiency change and technical change. Constant returns to scale.

	89/90	**90/91**	**91/92**	**92/93**	**93/94**
Malmquist index					
Overall	1.025	1.039	1.077	1.045	1.022
IM	1.021	1.059	1.129	1.062	1.033
B	1.028	1.036	1.068	1.040	1.019
Efficiency change					
Overall	1.006	1.006	1.004	0.991	0.985
IM	1.010	1.033	1.018	0.999	0.997
B	1.007	1.001	1.002	0.990	0.982
Technical change					
Overall	1.019	1.033	1.073	1.054	1.038
IM	1.011	1.025	1.109	1.064	1.036
B	1.021	1.035	1.066	1.051	1.038

In Table 4 the productivity change estimates indicate improved overall productivity in all periods. The increase is greater than two percent per year over the entire period. The positive productivity growth result is the opposite of what was found in the study by Färe, Grosskopf and Roos (1994) for the pre-reform period, where overall negative productivity growth is reported. Although a comparison with this study is not without problems the results here indicates that there can have been an overall shift from a negative to a positive productivity growth in health services. The rate of change is higher for county councils with internal markets for all periods except the first, although the mean difference is small in this case. The greatest improvement in productivity occurred between 1991 and 1992 followed by lower growth at the end of the period. This could be explained by some sort of once-for-all effect that occurred when the new payment mechanism was introduced. Another explanation is that several county councils introduced ‘ceilings’ in order to control total expenditures for health services. The latter seems reasonable as both groups experienced a large increase in productivity between 1991 and 1992.

The decomposition of the productivity index shows that progressive technical change is found for both groups of county councils over the whole period. Budget county councils have somewhat higher technical change on average in the beginning of the period. Change in efficiency was also positive in the beginning of the period for both groups. In the last two periods we, however, obtain worsened technical efficiency, although the changes were small for the internal market group.

If we, as in Table 3, observe that internal market county councils are less inefficient than budget county councils this can be due to two effects. First, the internal markets can have moved closer to the production frontier. This can be interpreted as a reduced organizational slack-effect. In this case we can think of the observations as narrowing a ”fixed” production frontier that has not changed due to the organizational changes. Another effect is the change in the production frontier. It is possible that the introduction of new organizational systems expands the output set in the sense that the frontier is ”farther away” than initially. The county councils with budget allocation are then estimated as more inefficient unless they narrow the frontier in the same degree. The empirical problem is of course that it is hard to disentangle the effect of new organizational systems from the effect of new medical technology and other ongoing changes in the health care system. The results reported in Table 4 show that the change in frontier effect was similar for the two groups. One exception is the technical change between 1991 and 1992. In this period the technical change effect was greater for the internal market county councils.

The effect of the reference technology

We also estimated the distance functions for the county councils with and without internal markets in the reference technology using a ”nested” procedure. That is, first only ”budget” county councils are used in the reference technology and these county councils are compared only with themselves. Then we included the remaining internal-market county councils in the reference technology as well. If the mean estimate is lower when internal markets are included in the reference technology this would indicate that the output set is ”expanded”. This in turn leads to a greater inefficiency compared to the case when only budget county councils are in the reference technology. We perform a similar analysis for the Malmquist productivity index. It should be noted that the effect of excluding observations in the reference technology is perhaps less clear when estimating the Malmquist index compared to the single period distance function. The results are shown in the Table 5 below.

Table 5: Distance function (1) and Malmquist productivity means (2) for budget county councils, without (a) and with (b) internal market county councils in reference technology.

Distance function (Nirs)						
	1989	**1990**	**1991**	**1992**	**1993**	**1994**
(1 a)	0.967	0.960	0.962	0.976	0.951	0.950
(1 b)	0.952	0.956	0.957	0.951	0.939	0.923
Diff (b-a)	-0.015	-0.004	-0.006	-0.025	-0.012	-0.027
Malmquist index (Crs)						
	89/90	**90/91**	**91/92**	**92/93**	**93/94**	
(2 a)	1.018	1.038	1.067	1.052	1.031	
(2 b)	1.026	1.036	1.067	1.043	1.021	
Diff (b-a)	-0.008	0.002	0	0.009	0.01	

The first part of the table show means of the distance functions for the ”budget” county councils. In all years between 1990 and 1994 the mean is higher when internal market county councils are not included in the reference technology. This can be interpreted as if the budget county councils are more inefficient when the internal market county councils are included in the reference technology. Hence, this indicates that internal market county councils expand the output set (isoquant) and the (budget) county councils are therefore, on average, estimated as more inefficient when they are compared with all county councils. The reference technology effect seems to be more strong at the end of the period, when the new organizational forms had been in effect. For the Malmquist index estimates the results are less clear. In the periods 90/91, 92/93 and 93/94 the productivity growth is higher when internal market county councils are not included in the reference technology. The opposite result is obtained for the first period, 1989 - 90 and between 1991 and -92.

Statistical analysis of mean differences - a bootstrap approach

Our main hypothesis concerns the difference in performance between county councils that have introduced an internal market organization (performance based reimbursement) and those that have kept the traditional integrated organization with budget based allocation. As a statistical test we could therefore compare group means of performance scores, e.g., the distance functions. The group-mean comparison approach has been used extensively in efficiency studies of health care provision (see, e.g., Grosskopf and Valdmanis, 1987; Borden, 1988; Burgess and Wilson, 1996). Examples of such tests are *t*-test, non-parametric tests as the Mann-Whitney *U*-test (Wilcoxon *W*-rank sum test) and bootstrap based tests (Atkinson and Wilson, 1995). One drawback with the *t*-test is that it relies on an assumption of normality, which can be questioned when comparing efficiency or productivity estimates from DEA-models (Grosskopf, 1996). This is not the case for the non-parametric or bootstrap based tests. Here we choose to employ a bootstrap procedure as it avoids unnecessary restrictive assumptions when differences in sample means are analyzed.

To compare the two groups we applied the sample mean bootstrap procedure proposed by Atkinson and Wilson (1995). The procedure is outlined in Appendix 1. The bootstrap procedure was applied both to the distance functions and the Malmquist productivity index for the whole period, 1989 to 1994. The county councils were divided into two groups with the same classification in all periods. The results are given in Table 6 below.

Table 6: Differences in group means, internal market and budget based county councils. Original estimated difference and 95% bias corrected bootstrap confidence intervals. Distance function (NIRS) and Malmquist index (CRS).

	1989	**1990**	**1991**	**1992**	**1993**	**1994**
Distance function						
lower	-0.050	-0.063	-0.009	0.025	0.032	0.037
original	-0.004	-0.003	0.025	0.047	0.056	0.067
upper	0.043	0.041	0.053	0.068	0.083	0.101
	89/90	**90/91**	**91/92**	**92/93**	**93/94**	
Malmquist index						
lower	-0.034	-0.025	0.012	-0.025	-0.017	
original	-0.005	0.025	0.064	0.020	0.013	
upper	0.030	0.068	0.113	0.064	0.048	

The first three rows show differences in means of the distance functions, or equivalently in technical efficiency. Note that a positive sign means that internal market county councils have higher efficiency and / or higher productivity change and vice versa. The original estimates show that internal market county councils have higher technical efficiency in all periods except in 1989 and 1990. The differences in 1992, 1993 and 1994 are significant implying that internal market county councils were significantly more efficient in those periods. We also used the same procedure on the CRS estimates of the distance functions. We do not report these results as they were qualitatively the same with the difference that a significant difference was obtained for 1991 as well. A Mann-Whitney test was also applied which gave significant differences in the years 1992 to 1994 on 10% and 5% level (see Tables in Appendix 2). We note that the differences are low in the beginning of the period and increases for each year. In 1989 and 1990 the means difference is negative. This means that budget county councils were more efficient in 1989 and 1990. These differences were however not significant.

For the Malmquist index the county councils with internal markets have higher scores on average in all periods except the first between 1989 and 1990. This implies that the productivity improvement was higher on average for these county councils. On the other hand, the differences are not statistically significant except for the productivity growth between 1991-92. It seems that the county councils with internal market experienced a major increase in productivity growth in this period. With the exception of that period the results do not give a clear indication of a different development in productivity for the two groups.

6. Concluding remarks

The high degree of vertical integration in Swedish health care has been exposed for several reforms in recent years. One major reform is the purchaser/provider Split within the county councils that provide and fund almost all health care services in Sweden. These changes aimed at introducing incentives for the public owned health care providers to use available resources more efficiently. The purpose of this paper was to analyze the outcome of these changes by estimating performance measures as technical efficiency and productivity growth.

The DEA-method was used to study the productive performance in the short-term somatic care for the Swedish county councils. The results are in accordance with other studies using non-frontier analysis. It shows an overall improvement in productivity for most county councils during the period 1990-1994. However, the results indicate that county councils which changed their internal resource allocation system into a performance based system have higher improvements in productivity, although the differences were not significant except for one period. It also seems that there was a clear improvement in productivity for those county councils between 1991 and 1992. This result is valid for both internal market and budget county councils, although the rate of change was higher for the former. The technical efficiency estimates also suggest that internal market county councils were more efficient compared to the county councils with budget based allocation from the middle to the end of the studied period. The results also shows that the two groups were about equally efficient in the beginning of the period. This can be interpreted as if the group with new organizational systems have reduced the degree of X-inefficiency. Although we found significant efficiency differences in all years from 1992 to 1994, the time period is to short to draw any conclusion regarding persistency of the organizational effects on performance. As more data will be available the question of once-and-for-all effect versus persistent effects of new organizational systems can be analyzed more formally.

Finally, it is worth to emphasize that the improvement in productivity cannot only be attributed to the implementation of market mechanisms. The group of county councils relying on budget allocation mechanisms also shows improvements in productivity. It has to be considered that there has been other initiatives and other reforms introduced during this period. The maximum waiting time guarantee and the care of the elderly reform are most likely to have an impact on the productivity in short-term somatic care. Still, these reforms where implemented all over the country simultaneously and can hardly explain the differences between the county councils.

References

Anell, A. and Svarvar, P., (1993), ”Reformerad landstingsmodell - En kartläggning och analys av pågående förnyelse”. Bilaga till Hälso- och sjukvården i framtiden - tre modeller, SOU 1993:38.

Atkinson, S. E. and Wilson, P. W. (1995), ”Comparing Mean Efficiency and Productivity Scores From Small Samples: A Bootstrap Methodology”, *Journal of Productivity Analysis*, 6 (2): 137-152.

Bergman, S-E. and Dahlbäck, U., (1995), ”Att beställa hälso- och sjukvård”, Landstingsförbundet. Stockholm.

Berki, S.E., (1972), ”Hospital Economics”, Lexington Books. Lexington, Mass.

Blomquist, Å., (1991), ”The Doctor as a Double Agent: Information Asymmetry, Health Insurance and Medical Care”, *Journal of Health Economics*, Vol. 10, 411-422.

Borden, J. P., (1988), ”An Assessment of the Impact of Diagnosis-Related Group (DRG)-Based Reimbursement on the Technical Efficiency of New Jersey Hospitals Using Data Envelopment Analysis”, *Journal of Accounting and Public Policy*, Vol. 7, 77-96.

Burgess, J. F. and Wilson, P. W., (1996), ”Hospital Ownership and Technical Inefficiency”, *Management Science*, 42, 111-123.

Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. S. (eds.), (1995), ”*Data Envelopment Analysis: Theory, Methodology and Applications*”, Kluwer Academic Publishers, Amsterdam.

Feldstein, P., (1968), ”Applying economic concepts to hospital care”. *Hospital Administration*, 13, 68-69.

Färe, R., Grosskopf, S., Lindgren, B and Roos, P., (1989), ”Productivity Developments in Swedish Hospitals: A Malmquist Output Index Approach”, Department of Economics Discussion Paper 89-3, Southern Illinois University, Carbondale. Published in Charnes *et al*. (1995).

Färe, R., Grosskopf, S. and Roos, P., (1996), ”Remarks on Malmquist Productivity Indexes”, IHE Working Paper, 1996:6, IHE, Lund.

Grosskopf, S. and Valdmanis, V., (1987), ”Measuring Hospital Performance - A Non-parametric Approach”, *Journal of Health Economics*, 6, 89-107.

Grosskopf, S., (1996), ”Statistical Inference and Nonparametric Efficiency: A Selective Survey”, *Journal of Productivity Analysis*, 7, 161-176.

Harris, J. E., (1977), ”The Internal Organization of Hospitals: Some Economic Implications”, *Bell Journal of Economics*, 8, 467-482.

Jonsson, E., (1994), ”Har den s.k. Stockholmsmodellen genererat mer vård för pengarna? - en jämförande utvärdering”, Institutet för Kommunal Ekonomi 1994:48, Stockholms Universitet.

Jonsson, E., (1996), Har betalning per prestation givit mer vård för pengarna? Spri. Stockholm.

Landstingsförbundet (Federation of County Councils), (1994a), ”Ekonomisk Basstatistik. Boksslutsstatistik 1994 Ekonomi”, Landstingsförbundet, Stockholm. (and equivalent reports 1990 - 1993)

Landstingsförbundet (Federation of County Councils), (1994b), ”Basårsstatistik 1994 Verksamhet. Sammanställning av vårddagar, vårdplatser, läkarbesök mm”, Landstingsförbundet, Stockholm. (and equivalent reports 1990 - 1993)

Newhouse, J. P., (1970), ”Toward a Theory of Nonprofit Institutions: An Economic Model of a Hospital, *American Economic Review*, 60, 65-74.

Pauly, M. and Redisch, M., (1973), ”The Not-For-Profit Hospital as a Physicians Cooperative, *American Economic Review*, 63, 87-99.

Shephard, R., (1970), *”Theory of Cost and Production Functions*”, Princeton University Press, Princeton, New Jersey.

Socialstyrelsen (National Board of Health and Welfare), (1996), Inpatient Register.

Appendix 1

Confidence intervals for mean efficiency and productivity estimates

The procedure is adapted from Atkinson and Wilson (1995) and briefly restated in this appendix. For a more complete discussion we refer to Atkinson and Wilson (1995). Let x_k denote an estimate of efficiency or productivity growth (e.g., ”firm” specific distance function and Malmquist index estimate). To obtain a bootstrap confidence interval for mean efficiency (and productivity) Atkinson and Wilson suggest the following algorithm:

(1) Estimate sample mean: $\bar{x} = \sum_k x_k \; / \; K$.

(2) Make a small sample correction: $\tilde{x}_k \; = x_k \sqrt{K / (K-1)} + \bar{x}\left(1 - \sqrt{K / (K-1)}\right)$

(3) Draw, with replacement, *K* elements from $\left\{\tilde{x}_k\right\}_{k=1}^{K}$.

(4) Compute a bootstrap estimate of the sample mean: $\bar{x}^{*b} \; = \sum_k \tilde{x}_k^{*b} \; / \; K$.

(5) Repeat steps (3) - (4) *B* times to obtain a set of bootstrap estimates of the sample mean.

By dividing the original sample in two samples we use the algorithm in step (1) to (5) to compute a bootstrap estimate of the mean differences. Note that step (2) and (3) are made independently for the internal market group and the budget county council group.

Different methods to obtain bootstrap confidence intervals for the sample mean differences can then be employed (see, e.g. Efron and Tibshirani, 1993). The simplest method of calculating bootstrap confidence intervals is to arrange the *B* bootstrap statistics in increasing order and then to cut off, say, five percent of the observations in the each tail. This procedure gives a 90 percent, double-sided, *percentile* confidence interval. Since this method does not take the original estimate into consideration, the issue of bias arises. One method that takes this into consideration is the bias-corrected (BC) confidence intervals (Efron and Tibshirani, 1993). A double sided $(1-2\alpha)\times 100$ percent BC confidence interval for the statistic of interest is bounded by the percentiles

$$\left(\hat{\theta}^{*b}\left(\alpha_L\right),\hat{\theta}^{*b}\left(\alpha_U\right)\right), \qquad \text{(A1)}$$

where $\hat{\theta}^{*b}$ denotes the vector of bootstrap sample mean difference estimates. The percentiles α_L and α_U are given by

$$\alpha_L = \Phi\left(2\hat{v} + \Phi^{-1}(\alpha)\right)$$

$$\alpha_U = \Phi\left(2\hat{v} + \Phi^{-1}(1-\alpha)\right). \qquad \text{(A2)}$$

$\Phi(\cdot)$ denotes the standard normal cumulative distribution function. The bias-adjustment term $\hat{v}$ is given by

$$\hat{v} = \Phi^{-1}\left(\frac{\#\left\{\hat{\theta}^{*b} < \hat{\theta}\right\}}{B}\right). \qquad \text{(A3)}$$

If $\hat{v} = 0$, i.e., *B*/2 of the bootstrap replications are less (larger) than the original estimate, the confidence interval is the *percentile* interval. Thus, the percentile interval is a special case of the BC interval.

Appendix 2

County council specific estimates of the distance function (CRS and NIRS) and Malmquist index with efficiency change and technical change decomposition. Bottom rows contain mean difference between internal market (performance based reimbursement) and budget county councils. A negative sign implies better performance for budget county councils and vice versa. The last row contain p-values obtained when performing a Mann-Whitney test for the differences. County councils with performance based reimbursement are: nos. 1, 3, 12, 16 and 18.

Table A2.1: Distance function estimates. A score = 1 indicates technical efficiency and a score less than one indicates inefficiency. Mean differences, internal market vs. budget and *p*-values Mann-Whitney test. Constant returns to scale.

County Council	1989	1990	1991	1992	1993	1994
1	**0.981**	**1**	**1**	**1**	**1**	**1**
2	0.870	0.853	0.884	0.877	0.872	0.833
3	**0.905**	**0.871**	**0.943**	**1**	**1**	**0.995**
4	1	1	1	0.994	0.971	0.938
5	0.958	0.939	0.963	1	1	1
6	0.923	0.952	0.911	0.973	0.866	0.867
7	1	1	1	1	1	1
8	0.960	0.958	0.926	0.903	0.835	0.874
9	0.851	0.858	0.919	0.902	0.897	0.945
10	0.959	1	0.967	1	1	1
11	1	1	1	1	1	1
12	**0.942**	**0.955**	**1**	**1**	**1**	**1**
13	0.854	0.864	0.886	0.916	0.922	0.933
14	1	1	1	1	1	1
15	0.916	0.914	0.873	0.908	0.902	0.913
16	**1**	**1**	**1**	**1**	**0.993**	**1**
17	0.872	1.000	1.000	0.995	1	1
18	**0.881**	**0.932**	**0.969**	**1**	**1**	**0.983**
19	1	1	1	1	1	1
20	0.829	0.861	0.948	0.978	0.922	0.817
21	0.961	0.929	0.978	0.895	0.896	0.782
22	1	0.981	0.865	0.958	0.947	0.912
23	0.997	0.952	0.913	0.879	0.818	0.797
24	0.929	0.902	0.909	0.855	0.889	0.927
25	1	1	1	1	1	0.864
26	0.977	0.986	0.993	0.906	1	1
Mean diff	-0.004	0.002	0.033	0.050	0.059	0.072
p-value	0.792	0.867	0.160	0.17	0.076	0.097

Table A2.2: Distance function estimates. A score = 1 indicates technical efficiency and a score less than one indicates inefficiency. Mean differences, internal market vs. budget and *p*-values Mann-Whitney test. Non-increasing returns to scale.

County Council	1989	1990	1991	1992	1993	1994
1	**1**	**1**	**1**	**1**	**1**	**1**
2	0.917	0.897	0.906	0.920	0.919	0.893
3	**0.905**	**0.871**	**0.943**	**1**	**1**	**0.994**
4	1	1	1	1	0.971	0.968
5	0.975	0.972	0.989	1	1	1
6	0.923	0.952	0.911	0.973	0.866	0.867
7	1	1	1	1	1	1
8	0.960	0.959	0.926	0.903	0.835	0.874
9	0.878	0.878	0.929	0.905	0.899	0.945
10	1	1	1	1	1	1
11	1	1	1	1	1	1
12	**0.942**	**0.955**	**1**	**1**	**1**	**1**
13	0.883	0.897	0.907	0.902	0.922	0.930
14	1	1	1	1	1	1
15	0.916	0.915	0.879	0.921	0.902	0.913
16	**1**	**1**	**1**	**1**	**0.995**	**1**
17	0.873	0.992	1	0.995	1	0.999
18	**0.909**	**0.950**	**0.973**	**1**	**1**	**0.983**
19	1	1	1	1	1	1
20	0.853	0.879	0.982	0.982	0.922	0.817
21	0.961	0.929	0.978	0.895	0.896	0.782
22	1	1	0.903	0.980	0.947	0.918
23	0.998	0.952	0.913	0.882	0.828	0.802
24	0.929	0.902	0.909	0.855	0.889	0.927
25	1	1	1	1	1	0.864
26	0.999	1	1	0.906	1	1
Mean diff	-0.004	-0.003	0.025	0.047	0.056	0.067
p-value	0.867	0.813	0.304	0.024	0.076	0.082

Table A2.3: Malmquist productivity index. Individual results and mean differences, internal market vs. budget and *p*-values Mann-Whitney test. Constant returns to scale.

County Council	89/90	90/91	91/92	92/93	93/94
1	**1.017**	**1.044**	**1.187**	**1.119**	**1.018**
2	0.987	1.108	1.016	1.005	0.999
3	**0.993**	**1.097**	**1.178**	**1.049**	**1.029**
4	1.054	1.052	0.976	1.076	1.031
5	1.018	1.019	1.107	1.030	1.018
6	1.057	1.002	1.163	0.945	1.026
7	1.036	1.015	1.048	1.028	1.034
8	1.027	1.060	1.044	1.011	1.099
9	1.035	1.086	1.053	1.048	1.035
10	1.046	1.009	1.106	1.093	1.017
11	1.027	1.009	1.047	1.026	1.019
12	**1.033**	**1.121**	**1.113**	**1.092**	**1.086**
13	1.035	1.068	1.123	1.116	1.058
14	1.050	1.056	1.014	0.904	1.025
15	1.031	0.987	1.135	1.031	1.049
16	**0.982**	**0.974**	**1.048**	**1.018**	**1.024**
17	1.063	0.953	1.058	1.087	0.988
18	**1.082**	**1.066**	**1.127**	**1.037**	**1.012**
19	1.034	1.087	1.066	1.087	1.054
20	1.065	1.089	1.105	0.991	0.927
21	0.985	1.101	0.996	0.961	0.924
22	1.004	0.975	1.181	1.068	1.023
23	0.977	1.020	1.033	0.967	1.045
24	1.003	1.019	0.989	1.063	1.076
25	0.982	0.988	1.101	1.107	0.892
26	1.036	1.047	1.036	1.260	1.099
Mean diff	-0.005	0.025	0.064	0.020	0.013
p-value	0.527	0.374	0.028	0.374	0.9

Table A2.4: Efficiency change. Individual results and mean differences, internal market vs. budget and *p*-values Mann-Whitney test. Constant returns to scale.

County Council	89/90	90/91	91/92	92/93	93/94
1	**1.020**	**1**	**1**	**1**	**1**
2	0.981	1.036	0.992	0.994	0.956
3	**0.962**	**1.082**	**1.061**	**1**	**0.995**
4	1	1	0.994	0.977	0.966
5	0.980	1.025	1.039	1	1
6	1.032	0.956	1.069	0.890	1.001
7	1	1	1	1	1
8	0.999	0.967	0.975	0.924	1.046
9	1.008	1.071	0.982	0.994	1.053
10	1.042	0.967	1.034	1	1
11	1	1	1	1	1
12	**1.014**	**1.047**	**1**	**1**	**1**
13	1.012	1.026	1.034	1.006	1.012
14	1	1	1	1	1
15	0.998	0.955	1.041	0.993	1.012
16	**1**	**1**	**1**	**0.993**	**1.007**
17	1.146	1.000	0.995	1.005	1
18	**1.058**	**1.040**	**1.032**	**1**	**0.983**
19	1	1	1	1	1
20	1.040	1.100	1.032	0.943	0.886
21	0.967	1.052	0.916	1.001	0.873
22	0.981	0.881	1.108	0.988	0.964
23	0.955	0.959	0.963	0.930	0.974
24	0.972	1.007	0.941	1.040	1.042
25	1	1	1	1	0.864
26	1.009	1.007	0.912	1.104	1
Mean diff	0.005	0.033	0.017	0.009	0.014
p-value	0.527	0.374	0.028	0.374	0.9

Table A2.5: Technical change. Individual results and mean differences, internal market vs. budget and *p*-values Mann-Whitney test.

Constant returns to scale.

County Council	89/90	90/91	91/92	92/93	93/94
1	**0.998**	**1.044**	**1.187**	**1.119**	**1.018**
2	1.006	1.069	1.024	1.011	1.046
3	**1.032**	**1.014**	**1.110**	**1.049**	**1.034**
4	1.054	1.052	0.983	1.101	1.067
5	1.038	0.994	1.066	1.030	1.018
6	1.024	1.047	1.088	1.062	1.026
7	1.036	1.015	1.048	1.028	1.034
8	1.029	1.096	1.070	1.094	1.050
9	1.027	1.014	1.073	1.055	0.982
10	1.003	1.044	1.070	1.093	1.017
11	1.027	1.009	1.047	1.026	1.019
12	**1.019**	**1.071**	**1.113**	**1.092**	**1.086**
13	1.023	1.041	1.087	1.109	1.046
14	1.050	1.056	1.014	0.904	1.025
15	1.033	1.034	1.090	1.039	1.037
16	**0.982**	**0.974**	**1.048**	**1.025**	**1.017**
17	0.927	0.953	1.063	1.081	0.988
18	**1.023**	**1.025**	**1.092**	**1.037**	**1.029**
19	1.034	1.087	1.066	1.087	1.054
20	1.025	0.990	1.071	1.051	1.047
21	1.019	1.047	1.088	0.960	1.059
22	1.023	1.107	1.066	1.081	1.062
23	1.023	1.063	1.073	1.040	1.073
24	1.032	1.012	1.050	1.022	1.032
25	0.982	0.988	1.101	1.107	1.033
26	1.026	1.039	1.135	1.141	1.099
Mean diff	-0.010	-0.010	0.045	0.011	-0.002
p-value	0.121	0.659	0.034	0.9	0.527