VATT-TUTKIMUKSIA
118
VATT RESEARCH REPORTS

Tarmo Räty - Kalevi Luoma
Juho Aaltonen - Maija-Liisa Järviö

PRODUCTIVITY AND ITS DRIVERS IN FINNISH PRIMARY CARE 1988–2003

Valtion taloudellinen tutkimuskeskus
Government Institute for Economic Research
Helsinki 2005

ISBN 951-561-587-9 (nid.)
ISBN 951-561-588-7 (PDF)

ISSN 0788-5008 (nid)
ISSN 1795-3340 (PDF)

Valtion taloudellinen tutkimuskeskus

Government Institute for Economic Research

Arkadiankatu 7, 00101 Helsinki, Finland

Email: etunimi.sukunimi@vatt.fi

Oy Nord Print Ab

Helsinki, Joulukuu 2005

RÄTY TARMO, LUOMA KALEVI, AALTONEN, JUHO JA JÄRVIÖ MAIJA-LIISA: PRODUCTIVITY AND ITS DRIVERS IN FINNISH PRIMARY CARE 1988–2003. Helsinki, VATT, Valtion taloudellinen tutkimuskeskus, Government Institute for Economic Research, 2005, (B, ISSN 0788-5008 (nid), ISSN 1795-3340 (PDF), No 118). ISBN 951-561-587-9 (nid.), ISBN 951-561-588-7 (PDF).

Abstract: We measure productivity changes of primary care in Finland between 1988 and 2003 as a ratio of key services produced and real operating costs. In the second stage we estimate a truncated regression model that quantifies the contribution of certain internal and exogenous factors to productivity. We use newly developed techniques to correct asymptotic bias in non-parametric efficiency scores and bootstrap the confidence intervals for the explanatory model parameter estimates. The bias accounts on average 2.8 percent decrease in efficiency level. From 1997 to 2003 the average productivity declined 13.7 percent; the result is insensitive to estimated bias. Even if standard parametric confidence intervals do not generally apply when efficiency scores are regressed, our bootstrapped intervals are almost equal to parametric ones. Of the correlates used the increased income subject to municipal taxation accounted for three percentage points of the productivity decrease. The correlates, that are expected to decrease the need of primary care services, had a negative impact on productivity, implying that health centres have not been able to adjust their resource usage correspondingly. Organisational changes that have taken place within primary care have not resulted in desired productivity improvements.

Key words: Primary care, health centres, productivity, bootstrap

Tiivistelmä: Tässä tutkimuksessa mittaamme Suomen perusterveydenhuollossa tuotettujen palvelujen ja kustannusten suhdetta, eli tuottavuuden kehitystä vuodesta 1988 vuoteen 2003. Tuottavuusmuutosten arvioinnissa käytämme vastikään kehitettyjä menetelmiä, joilla voidaan korjata ei-parametristen tehokkuusestimaattien kiinteiden otosten harha ja lasketaan tehokuutta selittävien mallien parametrien konsistenssit luottamusvälit. Harhaestimaattimme, eli määrä jolla perinteiset tehokkuusluvut yliarvioivat terveyskeskusten tehokkuuden on keskimäärin 2,8 prosenttia. Tuottavuudet vuosina 1988 ja 2003 ovat likimain samalla tasolla, mutta tuottavuus on vaihdellut lähinnä taloudellisten suhdanteiden mukaisesti. Vuodesta 1997 vuoteen 2003 tuottavuus on laskenut 13,7 prosenttia. Tehokkuusmuutoksia selittävän mallin luottamusvälien korjauksella ei tässä aineistossa ole suurta merkitystä. Mallin mukaan verotettavien ansiotulojen kasvulla on ollut noin kolmen prosentin vaikutus perusterveydenhuollon tuottavuuden laskuun. Terveyskeskuspalvelujen kysyntää vähentäneillä muutoksilla on ollut tuottavuutta alentava vaikutus, eli terveyskeskukset eivät ole sopeuttaneet resurssiensa käyttöä vastaavasti. Toteutetuilla organisaatiomuutoksilla ei ole ollut merkittäviä tuottavuusvaikutuksia käytetyssä tutkimusaineistossa.

Asiasanat: Perusterveydenhuolto, terveyskeskukset, tuottavuus, bootstrap

Contents

1. Introduction **1**

2. Data Description **4**

2.1 Primary Care Service Outputs and Inputs 4
2.2 Explanatory Variables for Efficiency 6

3. Methods **8**

3.1 Measures of Productive Efficiency and Productivity 8
3.2 Explanatory Models for Efficiency 11
3.3 Simulation of Bias and Confidence Intervals 11
3.4 Operationalisation of the Models 14

4. Results **16**

4.1 Productivity Trend in the Long Run 16
4.2 Productivity Changes 1997-2003 17
4.2.1 Productivity drivers 19

5. Discussion and Conclusions **23**

Notes **25**

References **26**

1. Introduction

During the last fifteen years the development of the Finnish health care system has in many ways been exceptional whether one compares it to the earlier periods or to other OECD countries. In the 1980s the GDP share of health expenditure rose rapidly from 6.4 percent in 1980 to 9.4 percent in 1992. By 2000 it had decreased to almost the same level (6.6 %) as in 1980 [1]. In the 1990s real per capita growth in health expenditure was very modest compared to other OECD countries. Between 1990 and 2002 the average annual rate of expenditure growth was just 0.9 percent while in other European countries it varied between 2.2 and 8.2 percent [2].

To a large extent the exceptional health expenditure development in the 1990s can be explained by the economic crisis which Finland experienced in the early 1990s. The development of Finnish economy during that period has been aptly described by the phrase: down from the heavens, up from the ashes [3]. The boom of the late 1980s turned into a bust with real GDP dropping by 14 percent from the peak in 1990 to the through in 1993 [4]. After the recession Finland experienced a strong recovery with average economic growth reaching 4.7 percent from the end of the recession to 2000. The exceptional volatility in economic development had also profound impacts on the development of the Finnish health care system. The deep recession led to severe public finance problems. The central government and local authorities (municipalities) responded to widening deficits by cutting expenditures and raising taxes. The largest single item in the central government's budget savings was its transfers to the local government sector. Municipalities, which are responsible for providing health care services, responded to cuts in these transfers by reducing health care resources.

During the 1990s significant reforms were also carried out in organising and funding public social and health care. The most important of these was the grant reform which came into force from the beginning of 1993. Under the reformed system central government grants to municipalities are non-earmarked lump-sump grants, which are calculated prospectively by using a specific need-based capitation formula. Prior to the reform municipal providers of social and health services got a certain fraction of their costs covered by matching grants, which varied between 34 and 66 percent depending on the economic capacity of the municipality. Under the new system municipalities have to bear the full marginal costs of health service inputs.

In the Finnish two-tier system of government, the responsibility for arranging and funding primary care as well as other health services lies with the municipalities. With the grant reform central government control was substantially reduced so that the municipal authorities reached almost full discretion regarding the scope, content and organisation of social and health services. Given the changes in financial incentives, budgetary stringency and increased local freedom in providing services brought about by the reforms and volatility in economic development, it is interesting to see how local decision makers reacted to these changes in the management of health care resources. We address this question by investigating the productive performance of primary care provided by public health centres over the period 1988-2003.

In the long run quality of care given in health centres has most unlikely stayed constant. With quality we mean a multidimensional concept covering everything between queuing times, good care indicators and how demanding the given care is. Whether its changes should be taken into account in performance measurement depends on the goal of the study. If we are interested in comparing the amount of output created to the resources used, it is usually desirable to adjust output measures with case-mix indices or other quality indicators (see e.g. [5,6,]). In this case quality is a deflator of the outputs, which makes comparisons over time, and between the units, more accurate.

Our main objective is in financial performance of primary care, for that purpose the need for quality adjustments is not so obvious. The productivity measure is calculated as a ratio of a number of different service outputs to operating costs. Deflating service outputs by (increased) quality would hide important information, namely the increased unit costs of operations. This is important information especially in primary care, which will first face the rising demand of services of the rapidly ageing population. If productivity declines, total expenditure will increase at higher rate than volume of services. Thus, output quality is an important factor in performance measurement, but if finance of services is considered it should be discussed separately. In this paper we evaluate the development of the financial pressure that primary care has created against local government budgets. The observed variation in productivity should be explained with quality changes or other factors, but is beyond our objective. In the first stage we will estimate the productivity changes over the whole period. In the second stage, we turn our interest on possible drivers that have contributed to the observed productivity changes.

Productivity is a powerful measure to indicate economic performance, but unfortunately unambiguous just in single input and single output production. In multidimensional case some kind of weights are needed to give value for each output or input. For the inputs market prices normally suffice. If outputs are sold on markets, and the market prices reflect the desired values, also outputs could be aggregated using their prices. The main problem in productivity measurement of public health services is how to value different service outputs. Usable market prices rarely exist. One could also try to value the different service outputs on the basis of their impact on health and welfare outcomes. For primary care the available evidence on this is so scant, that at present this is simply not a feasible option. A way to circumvent these problems is to use the tools of production economics and operation research to define a reasonable weighting scheme.

The basic tool is a production function, or in multiple outputs and inputs, distance functions. The estimation methods are roughly divided into two approaches, stochastic and non-parametric estimators. Stochastic frontier models are a common and flexible tool for single output production but for the cases of multiple outputs multicollinearity problems restrict their usefulness. The nonparametric tradition originates from the work of Farrell [7] and was made practical by Charnes et al. [8]. The models were developed to measure efficiency (available productivity improvements) within a group of similar production units by optimising the weights of each unit, so that its efficiency is maximized. As a result no price or target weight information is necessary to calculate productivity figures in multiple dimensions, but they are replaced by the most favourable set of weights for the unit. There are no guarantees that the weight structure is comparable with the valuations of service outputs, but at least productivity is not underestimated. Non-parametric methods have been subject of serious criticism in health economics [9], and they should not be used if better information is available. For Finnish primary care such information does not exist. If the long run extensive productivity information is needed, except for subjective evaluations non-parametric methods are the only possibility.

In this paper our first concern is restricting the variation of weights over time and between the units. We estimate long run productivity of Finnish primary care units using non-parametric Assurance Region Data Envelopment Analysis (AR-DEA). At first we measure productivity changes between 1988 and 2003. After that we take a closer look into the period 1997-2003 and present the bias estimates of efficiency scores as suggested in [10]. Next we discuss the so called two-stage DEA approach and prerequisites for using it. The main concern is that the traditional approaches used in the 2nd stage are inconsistent. The Tobit estimator is poorly justified and parametric statistical inference generally does not apply when dependent variable is serially correlated. We estimate a truncated regression model using panel data that quantifies the contribution of certain internal and exogenous factors to productivity development in primary care. On the basis of the specified data generation process, we bootstrap confidence intervals for the parameters.

In section 2 we introduce the data set of Finnish primary care over 1988-2003. Section 3 presents the models used to estimate productivity changes and covariates, discusses possible problems in methodology and interpretation of the results. Sections 4 and 5 present the actual result from the analysis and section 6 concludes.

2. Data description

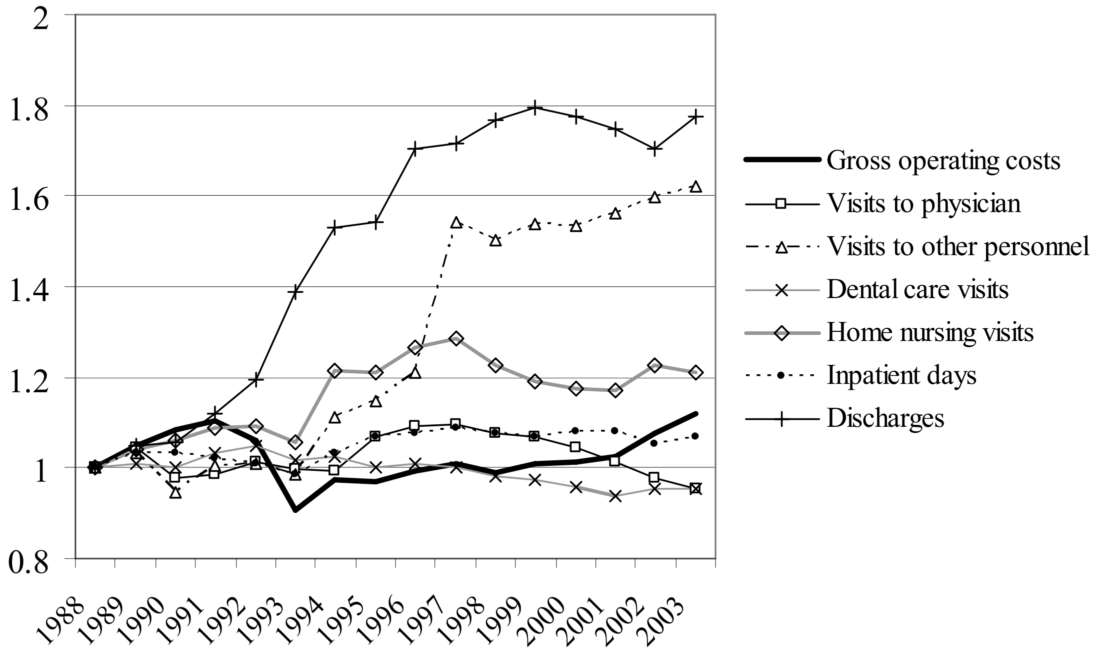
2.1 Primary care service outputs and inputs

As usual with this kind of study, available output data for assessing productivity development of health centres are far from ideal. Statistics and registers on health centres contain information only on what can be termed intermediate or service outputs e.g. the number of contacts or amount of inpatient days. It is clear that they should not be used, at least without quality or case mix corrections, when benchmarking of units is desired. Such extensive effectiveness data do not exist. However, service outputs carry important information on financial performance of units. In Finland primary care is almost completely financed by taxes, meaning that a certain amount of money is annually budgeted for operations by municipalities. The budget needed depends on the number of operations and their costs, thus from fiscal point of view it is important to monitor how unit costs of these operations are developing. When expenditures are the only input measure used, the inverse of productivity equals to unit cost.

The observation unit in this study is a public health centre, which provides both inpatient and outpatient care. The data are an unbalanced panel from 1988-2003. Due the mergers and decompositions and the lack of reliable data we have annually 204–254 health centres, which cover 67-96 percent of the Finnish population. The data on outputs and operating costs were carefully checked for, and the units which had extremely deviant information about their activity were excluded from the study. Some doubt concerning especially the early data remains. Therefore, we give a rough estimate on productivity improvement over whole period just for units led by GPs and execute a more detailed two-stage analysis for the sub period 1997-2003.

The aggregate view of outputs and expenditures of health centres led by GPs is plotted in Figure 1. We used six different output variables in this study. Visits to physician, visits to other health care personnel, dental care visits and home nursing visits represent outputs for outpatient care. The data between years 1988-2001 are derived from the financial statistics of health centres and the report information register (KETI) [a]. The data of the years 2002 and 2003 are instead derived from Stakes and Statistics Finland. Outputs for inpatient care are divided into short term acute care and long term chronic care. Short term care is measured by the number of admissions into inpatient wards. Long term care is measured by the number of bed days. The data are derived from the Care Registers for Social Welfare and Health Care (HILMO) [b].

Figure 1. *Primary care operating costs and services from 1988 to 2003, quantity 1988=1*



Visits to other health care personnel increased approximately by 28 percent from 1996 to 1997. Until 1996 the variable covered only nurse visits, but from 1997 onwards also visits to other personnel, for example visits to psychologists, is included. Visits to other personnel have increased since 1997 by five percent. Also admissions into inpatient wards have increased considerably, but the number of bed days has stayed quite stable in this period. The visits to physician were at the highest in year 1997, but since then they have decreased approximately 13 percent. Also the dental care visits have decreased about four percent between the years 1997 and 2003.

The input variable used in this study is operating costs of health centres measured in 2003 prices. Operating costs include personnel costs, material costs, costs of purchased services, rents and other costs. The average share of the personnel costs was approximately 70 percent. The operating costs increased from 1988 to 1991 but in 1992 and 1993 they decreased rapidly. This was due to the severe economic recession and the grant reform which took place in 1993. One of the chief purposes of the grant reform was to give municipalities stronger incentives to provide services economically and efficiently. Between the years 1993-1999 the real operating costs stayed almost unchanged and were approximately at the same level as in 1988. After that operating costs in Finnish primary care have increased sharply.

2.2 Explanatory variables for efficiency

In the second step of our analysis we estimate the impact of fixed set of variables on efficiency taking into account special features of the data generation process. The goal is to quantify how these variables, also called as the drivers, have contributed to productivity changes over the estimation period. The potential productivity drivers can roughly be divided into two groups: variables which describe the changes in economic and demographic environment of health centres and variables which describe how the service production of health centres is organised. We have considered seven variables relating to economic and demographic characteristics: the size of a health centre, the population share of elderly 85 years of age or older, per capita expenditure for specialized health care, private physician visits per capita, recipients of disability pension per working age population, the proportion of people living in old people's homes or service houses of the 65+ population and per capita income subject to local government taxation. The changes in the size of the health centre population did not have any significant impact on productivity so this variable was dropped from the model. Also, expenditures for specialized care appeared multicollinear with taxable per capita income, leaving five economic or demographic drivers in the model.

We had originally five variables purported to describe the organization of primary care production: outsourcing measured as the expenditure share of purchased client services, the system of local population responsibility indicated as a dummy variable, the share of emergency visits from all physician visits, the share of licensed doctors and share of nurses of health centre workforce. The preliminary estimates showed that emergency visits are not likely a significant driver. Comparable data on the shares of doctors and nurses was unfortunately not available from 2003, thus we ended up using only two organizational variables in estimations, outsourcing and personal GPs. These two variables have a special interest in our study. The municipalities have searched actively ways to improve the efficiency. We measure the extent to which purchaser–provider types of strategies are implemented as the amount of outsourcing. The advantages of system of local population responsibility of GPs are widely discussed in Finnish primary care. The proponents of the system claim, that population responsibility reduces unnecessary tests and visits of a client and consequently reduces the costs. Our data set and models will show if the savings have realised.

The discussed shift in definition of visits to other personnel in 1997 makes a more detailed analysis over the whole period somewhat questionable. Also, the data on health centres led by specialists is not comparable before and after 1993. Therefore, we execute a more detailed analysis from 1997 to 2003. For this period we can also include health centres led by specialists in the sample. Unfortunately as the full set of observations is required for both the DEA and the

2nd step model, the final number of health centres in this analysis is smaller. Of 247 different health centres we have 223-232 units each year in the data set; however we have been able to keep the coverage of primary care in over 90 percent. Table 1 presents the descriptive statistics of the variables from the first and last years of the data set. The figures are unweighted averages over health centre areas.

Table 1. The descriptive statistics of data from 1997 and 2003

	1997		2003	
	Mean*	Std. Dev	Mean*	Std. Dev
Input:				
Gross operating costs in 2003 prices	10244	20771	11300	24884
Outputs:				
Physician visits in public health centres	45924	71413	39019	61763
Visits to nurses and other personnel	61732	109102	63069	114864
Dental care visits	21957	33692	20500	35659
Home nursing visits	15559	33706	14020	35075
Care days, long term care	26637	66801	25812	63620
Discharges, number	1061	1382	1098	2011
Explanatory variables:				
Private physician visits, per capita	0.458	0.210	0.507	0.219
Taxable income, per capita euros	9299	1934	10319	2053
Share of 85+ years old population	0.015	0.005	0.018	0.006
Recipients of disability pension per 1000 people of working age	98.408	25.864	91.403	25.636
Share of 65+ population in alternative long term care units	0.048	0.019	0.051	0.018
Local population responsibility of GPs, dummy	0.209	0.407	0.407	0.492
Share of outsourcing services of operating costs	0.031	0.029	0.055	0.072

* unweighted averages over health centres

3. Methods

3.1 Measures of productive efficiency and productivity

Productivity is a ratio of outputs to inputs and easy to calculate only in the one output one input case. In the case of multiple outputs or multiple inputs one has to aggregate them by using some kind of weighting scheme. For a profit seeking firm it is natural to use market prices of outputs and inputs as weights. For public services where market prices are usually not available it has been a common procedure to use unit costs of different service outputs as weights. Obviously, this is not always satisfactory in health care, as unit costs do not necessarily carry information on social or medical value of these outputs. Also, health centres are independent municipal units, and seldom produce comparable price data for their outputs. We have only rough estimates on typical unit costs of health centre outputs available. Therefore we have ended up using output weights which optimize the measured performance of the units, given that they roughly follow typical cost structure. This method allows individual health centres to adjust they own output mix according the local needs, without their multi output productivity levels getting under estimated.

Instead of calculating absolute productivity levels, Data Envelopment Analysis (DEA, see e.g. [11, 12]) calculates efficiency scores (shortfall from optimal productivity) for each unit. Calculations can be done easily by using so called multiplier formulation (1). Let $S = 1594$ denote the number of observations. Also, denote $M\,(=6)$ vector of output multipliers as $\boldsymbol{\mu}^{\mathsf{T}}$ and $N\,(=1)$ vector (a scalar) of input multiplier as $\mathbf{v}^{\mathsf{T}}$. $\mathbf{Y}$ and $\mathbf{X}$ are the $M \times S$ and $N \times S$ matrices of outputs and inputs respectively with columns $\mathbf{y}_s$ and $\mathbf{x}_s$ for a health centre s respectively. The output oriented constant returns to scale DEA program for a health centre s is,

$$
\begin{array}{lrcll}
\min\limits_{\boldsymbol{\mu},\mathbf{v}} & q_s = \mathbf{v}^{\mathsf{T}}\mathbf{x}_s & & & \\
& \boldsymbol{\mu}^{\mathsf{T}}\mathbf{y}_s & = & 1 & \\
& -\boldsymbol{\mu}^{\mathsf{T}}\mathbf{Y} + \mathbf{v}^{\mathsf{T}}\mathbf{X} & \geq & \mathbf{0} & \\
& \mu_m & \geq & \varepsilon & \forall m = 1,\ldots,M \\
& \nu_n & \geq & \varepsilon & \forall n = 1,\ldots,N
\end{array}
\tag{1}
$$

where ε is an infinitesimal that forces multipliers to have a small positive value to ensure unique solution. Let's denote the optimal multipliers as $\boldsymbol{\mu}^{*\mathsf{T}}, \mathbf{v}^{*\mathsf{T}}$. Thus the problem is to find minimum aggregate input ($q_s^* = \mathbf{v}^{*\mathsf{T}}\mathbf{x}_s$) for each unit in turn. Inputs are aggregated such that unit s's aggregate input level is one. The second set of constraints implies that in the minimum, the ratio of aggregate output

$\boldsymbol{\mu}^{*\mathrm{T}}\mathbf{y}_i$ to aggregate input $\mathbf{v}^{*\mathrm{T}}\mathbf{x}_i$ for all the units $i = 1,\ldots,S$ is less or equal to unity and $q_s^* \geq 1$. If the constraint for unit s in the second set is not binding in the optimum, we could multiply its outputs by q_s^* and still maintain feasibility. Thus, the DEA optimums q_s^* from (1) are efficiency scores, which tell how much the units could increase outputs. Each unit will have its own set of optimal multipliers for aggregating outputs and inputs. For a cost minimizing units the solutions are shadow prices, and their ratios marginal rates of substitutions, which express the trade-offs between different outputs. For a rational decision making unit, these values should be very close to relative market prices or, in case of non-marketed outputs, relative valuations of different outputs.

Especially in higher dimensional problems some of the optimal multiplier values in (1) are infinitesimals, indicating that the variable in question does not contribute to efficiency. To avoid such cases and include some information on plausible valuations of each output we augment (1) by so called assurance region constraints. The idea is to set lower and upper bounds for the ratios of (μ_m, μ_i) $m,i \in \{1,\ldots,M\}$ pairs, i.e. to the marginal rates of substitution between selected outputs,

$$l_{mi} \leq \mu_m / \mu_i \leq u_{mi} \Leftrightarrow \begin{cases} \mu_m - l_{mi}\mu_i \geq 0 \\ -\mu_m + u_{mi}\mu_i \geq 0 \end{cases} \quad m,i \in \{1,\ldots,M\}. \tag{2}$$

(1) and (2) yield so called assurance region DEA (AR-DEA, see e.g. [13,14]) model. A desired number of AR-type constraints could be set for both inputs and outputs.

The original DEA problems are in nature nonparametric and non-stochastic. They do not assume any parametric form for the production function and solutions are searched without any stochastic errors in efficiency scores or without considering how binding the constraints are[c]. However, it is clear from (1) that the solution is sensitive to the number of units in data set. Additional units create a new constraints, thus keeping efficiency score intact or pushing it up. Banker [17] and Korostelev et al. [18] have shown that asymptotically DEA is a consistent estimator for the efficiency frontier, but in a fixed sample always biased downwards. Also, Kneip et al. [19] have shown that convergence rate with respect to the sample size is very slow and slows as the dimensions of the problem, $M+N$, increase. Simar and Wilson [20] and Kneip et al. [21] have presented bootstrap algorithms to simulate confidence intervals (CIs) for efficiency scores. In this paper we use a simpler approach due to Simar and Wilson [22] to estimate the bias correction. We present the algorithm in the next chapter.

Efficiency scores as such include important information on potential production gains if unused resources exist (output orientation), or inversely on possible savings if inputs are disposed (input orientation). Usually efficiency differences are analysed within a period. Often it is more interesting to investigate productivity changes. Generally, productivity index is a ratio of two productivity levels. DEA does not reveal productivity levels, but we can calculate productivity index from efficiency scores.

In accordance with (1) output oriented efficiency score q_s^* is a ratio of a reference (or optimal) productivity level to the unit's own productivity. Furthermore, productivity index PI, is a ratio of any two productivity levels. Let superscripts 0 and 1 denote the two periods of interest and upper bar denote the reference level of productivity. If p is productivity level, and ignoring the subscripts referring to units and superscript referring to optimum, we get

$$\begin{aligned} p &= \frac{\omega(\mathbf{y})}{\xi(\mathbf{x})}, \quad PI = \frac{p^1}{p^0} \\ q^0 &= \frac{\overline{p}^0}{p^0}, \quad q^1 = \frac{\overline{p}^1}{p^1} \end{aligned}, \tag{3}$$

where $\omega(\mathbf{y})$ and $\xi(\mathbf{x})$ are, not necessarily independent, scalar valued aggregator function of outputs and inputs respectively. DEA is just one way to define them. Using (3), we can calculate PI by comparing p^0 and p^1 to $\overline{p}^0$ or $\overline{p}^1$, i.e.

$$\begin{aligned} PI^0 &= \frac{\overline{p}^0 / p^0}{\overline{p}^0 / p^1} = \frac{q^0}{q^{1,0}} \\ PI^1 &= \frac{\overline{p}^1 / p^0}{\overline{p}^1 / p^1} = \frac{q^{0,1}}{q^1} \end{aligned}. \tag{4}$$

The cross efficiencies $q^{0,1}$ and $q^{1,0}$ compare the unit's productivity to another period's reference technology. In the most cases PI^0 and PI^1 differ. Therefore, productivity index is usually calculated as the geometric mean of them,

$$PI = \left(PI^0 \times PI^1\right)^{½}. \tag{5}$$

This way to calculate productivity indices is originated from Malmquist [23] where distance functions (inverses to efficiency scores) were first used. The method was introduced in production context in Caves et al. [24]. We must calculate all the four efficiencies to get *PI*. The cross efficiency $q_s^{1,0}$ can be calculated by replacing $\mathbf{x}_s^0$ by $\mathbf{x}_s^1$ in $\mathbf{X}^0$ and the objective, as well as $\mathbf{y}_s^0$ by $\mathbf{y}_s^1$ in $\mathbf{Y}^0$ and the first constraint in (1), and $q_s^{0,1}$ with analogous replacements. An alternative is to create a common reference technology for both periods by pooling the data sets from both periods. In that case substituting $\overline{p}^0 = \overline{p}^1$ to (4) and (5) yields

$$PI\Big|_{\overline{p}^0=\overline{p}^1} = \frac{q^0}{q^1}. \tag{6}$$

Thus, in the pooled data it is enough to compare two efficiency scores to calculate a productivity index.

3.2 Explanatory models for efficiency

Even if productivity is a measure that simply relates outputs to inputs used, it does not mean that the origin of efficiency differences is in the production process itself. It has been a common practice to estimate efficiencies and build an explanatory regression models for efficiencies or even try to isolate the impact of certain factors on efficiency (see e.g. [25] and for applications to Finnish primary care [26], [27], [28],[29]). Let $\mathbf{Z}$ be the $R \times S$ matrix of explanatory variables with for efficiencies with columns $\mathbf{z}_s$ and $\boldsymbol{\beta}$ the corresponding parameter vector. The second step model is,

$$q_s = g\left(\mathbf{Z};\boldsymbol{\beta}\right) + \varepsilon_s \tag{7}$$

The derivatives of the model imply marginal effects on efficiency. If pooled data is used, one can calculate relative productivity changes conditional on explanatory variables. Totally differentiating (7) and dividing by efficiency, using (6) yields,

$$TI\Big|_{\overline{p}^0=\overline{p}^1} - 1 = \Delta p\Big|_{\overline{p}^0=\overline{p}^1} = \frac{q^0 - q^1}{q^1} = \frac{-dq}{q^1} = \frac{-\sum_{r=1}^{R} g_r^{'}\left(\mathbf{Z};\boldsymbol{\beta}\right) dz_r}{q^1}. \tag{8}$$

Thus, the contribution of individual variable on productivity, Δp_r, is calculated as substituting $g_r^{'}(\cdot)$, dz_r and q^1 in $g_r^{'} dz_r \big/ q^1$.

3.3 Simulation of bias and confidence intervals

Simar and Wilson [22] have critically reviewed attempts to create two-step models for efficiency. They point out that no one has specified satisfactory data generation process (DGP), which tells what is estimated in two stage approach. A proper statistical specification should take into account a bias in efficiency scores and maintained serial correlation of DEA-based efficiency scores. The serial correlation arises because efficiencies are (by definition of efficiency) dependent on each other. DEA is a complicated discrete nonlinear function, making practically impossible to model this correlation parametrically. The serial correlation problem alone makes the traditional methods of stochastic inference to fail in the estimation of (7). Simar and Wilson [22] have developed a bootstrapping algorithm that overcomes both bias and correlation problem and animates DGP that gives a reasonable view on efficiency and its explanations.

For explanatory models like (7) it seems natural to think that a health centre $s, s \in \{1,\dots,S\}$ first faces environmental variables $\mathbf{z}_s$ drawn from density $f(\mathbf{z})$. Given $\mathbf{z}_s$, efficiency level q_s is drawn from $f\left(q|z_s\right)$, justifying the explanatory regression. The step from efficiencies to inputs and outputs is somewhat more complicated. After having the (output) efficiencies we already know something about outputs. We can express a vector using its length and $M-1$ angles (or directions) $\boldsymbol{\eta}$ i.e. in its polar coordinates. The point is that q_s is clearly related to length of $\mathbf{y}_s$ and the angles define the mix of outputs. Thus, given that environmental variables define efficiency, we can draw $\mathbf{x}_s$ and $\boldsymbol{\eta}_s$ from $f\left(\mathbf{x},\boldsymbol{\eta}|q_s,\mathbf{z}_s\right)$ resulting in a realization $\left(\mathbf{x}_s,\mathbf{y}_s,\mathbf{z}_s\right)$ from the joint density $f\left(\mathbf{x},\mathbf{y},\mathbf{z}\right)$, after transforming the polar coordinates $\left(q_s,\boldsymbol{\eta}_s\right)$ to Cartesian coordinates $\mathbf{y}_s$. A focal assumption in DGP is that a realisation of inputs and output mix is conditional on realised efficiency which is in part conditional on environment a unit has. Thus a health centre has no efficiency gains available by choosing an output mix. Generally this doesn't sound reasonable, but actually health centres have to respond in a stable flow of statutory needs and can't optimise their output mix, reducing the impact of conditioning of q_s on $\boldsymbol{\eta}_s$

It has been a common practice to estimate (7) as a Tobit model (see e.g. [30]) with a left censored distribution for ε. The key idea in censoring is that the modelled latent variable actually has values below the censoring point. This is not the case in described DGP, where in a fixed sample absolute lower limit for efficiency score is 1. To avoid using information not present in DGP a better choice is a truncated regression model (see e.g. [30]) for (7). Densities in left truncation are scaled such that density function above the truncation point integrates to one. We loose any information concerning the truncated observations, but reciprocally do not add any information not belonging in the model as Tobit-model does.

The greatest advance in DGP above is providing a firm base for a bootstrap algorithm to simulate the bias in efficiency and CIs for the regression parameters. Especially the simulation of the bias has previously been computationally demanding (see e.g. [20], [21]). The explanatory model in this DGP has information that can be used in modelling the variation of q. As preliminary steps we solve (1) for each $s, s \in \{1,\ldots,S\}$ to yield estimates of efficiency scores $\hat{q}_s$. Assuming that $\varepsilon_s \sim N(0,\sigma^2)$ with left truncation at $1-g(\mathbf{z}_s;\boldsymbol{\beta})$ we next regress $\hat{q}_s$ on $\mathbf{z}_s$ to obtain estimates $\hat{\beta}$ and $\hat{\sigma}$ for (7).

The first problem is to estimate the bias in $\hat{q}_s$. We can create pseudo efficiencies q_s^* by drawing ε_s from the $N(0,\hat{\sigma}_\varepsilon^2)$ distribution with left truncation at $1-g(\mathbf{z}_s;\boldsymbol{\beta})$ and calculate $q_s^* = g(\mathbf{z}_s;\boldsymbol{\beta})+\varepsilon_s$. The corresponding output levels for pseudo efficiencies are $y_s^* = \hat{q}_s y_s / q_s^*$. The random variation in the unit i's, $i \in \{1,\ldots,S\}$ efficiency score may now be simulated by replacing $\mathbf{y}_s$ by $\mathbf{y}_s^*\ \forall s \neq i, s \in \{1,\ldots,S\}$ in (1), that results in revised pseudo efficiency estimate $\hat{q}_i^*$. Repeating random draws and ensuing calculations L_1 times for each unit we get L_1 sample of efficiency scores $\hat{q}_{s,l}^*$ for each $s, s \in \{1,\ldots,S\}$. The estimate of the bias is the difference between the mean of revised pseudo efficiencies and original DEA efficiency estimate. The bias-corrected estimate of efficiency, $\hat{\hat{q}}_s$, is,

$$
\begin{aligned}
&BI\hat{A}S(\hat{q}_s) = \frac{1}{L_1}\sum_{l=1}^{L_1}\hat{q}_{s,l}^* - \hat{q}_s \\
&\hat{\hat{q}}_s = \hat{q}_s - BI\hat{A}S(\hat{q}_s) = 2\hat{q}_s - \frac{1}{L_1}\sum_{l=1}^{L_1}\hat{q}_{s,l}^*
\end{aligned}
\tag{9}
$$

Each round in L_1 includes solutions of S independent problems. The time needed to create a model and solve it increases sharply as S increases. To speed up calculations see e.g. [31, 32].

To simulate proper CIs for parameter estimates of (7), a regular parametric bootstrap suffices. The natural starting point is the set of bias-corrected efficiencies. Running a truncated regression of $\hat{\hat{q}}_s$ on $\mathbf{z}_s$ yields parameter estimates $\left(\hat{\hat{\beta}},\hat{\hat{\sigma}}_\varepsilon\right)$. Next we draw ε_s for each $s=\{1,\ldots,S\}$ from the $N\left(0,\hat{\hat{\sigma}}_\varepsilon^2\right)$ with the left truncation at $1-g\left(\mathbf{z}_s;\hat{\hat{\boldsymbol{\beta}}}\right)$ and create new pseudo efficiencies $q_s^{**} = g\left(\mathbf{z}_s;\hat{\hat{\boldsymbol{\beta}}}\right)+\varepsilon_s$. Finally we regress q_s^{**} on $\mathbf{z}_s$ using truncated regression, yielding a set of parameter estimates $\left(\hat{\hat{\beta}}^*,\hat{\hat{\sigma}}_\varepsilon^*\right)$. Repeating random draws L_2 times we obtain a sample of bootstrap estimates. The centiles of this bootstrap distribution of estimated parameters yield CIs with a desired probability. To draw random numbers from left truncated distribution using modified transformation method see [22].

The bootstrap algorithm follows the sequential nature of DGP. Apart from the limited capacity to optimize output mix the algorithm uses iid. draws from the $N\left(0,\sigma_{\varepsilon}^{2}\right)$. In simulating the bias iid sounds harmless, as basically draws are used to create pseudo outputs. Thus, one need not to do explicit assumption on the distribution of efficiency scores. Efficiency distributions of units can still depend on other units and they need not to be identical. This is also likely the case. iid assumption is more problematic for the CIs of (7). Function $g(\mathbf{Z};\boldsymbol{\beta})$ models the dependency of efficiency levels in the second stage, but the variation in efficiency scores is considered independent (but non-identical) between the units.

3.4 Operationalisation of the models

The DEA estimator of efficiency (1) used here assumes that optimal efficiency is not dependent on the scale of production, which equals to assuming constant returns to scale (CRS). Our productivity measure equals to the inverse of the unit cost of the aggregate output $\omega(\mathbf{y})$. Therefore, it does not sound sensible to let larger units to have higher unit costs and still be classified as efficient as variable returns to scale would entail. Also, according to earlier studies ([26], [27], [28], [29]) efficient health centres can be of very different sizes and so there seems to be no significant economies of scale in health centre production.

Our panel data makes it possible to measure efficiency variation and productivity changes in two different ways. We can construct the efficiency frontier for each year separately or form a common efficiency frontier for the whole period. The differences in measured productivity change with the two approaches are quite small; therefore we report productivity changes according the latter approach. Also, after deflating operating costs by the municipal health care price index, cost changes should reflect the resource usage. Therefore, it sounds sensible to assume that health centres ability to transform resources into service outputs has remained unchanged.

The model of six outputs and one input with over 3000 observations could be expected to suffer from large variation of marginal rates of substitution (MRS) and some output weights are likely to be zeros. We have ended up using AR-DEA constraints, (2), to mitigate these problems. They can be used efficiently to rule out any undesired zero weights, but some knowledge on the structure of MRS between outputs is needed. As the weights imply the ratios in which outputs may replace each other they are crucial to the interpretation of the results. From pure fiscal point of view, we should use relative unit prices. Alternatively from pure health promotion point of view relative impacts on health of each output should be used. We do not have complete information on either one. Actually the only information available are the estimates on typical unit prices of primary care service outputs in Finland from [33]. The estimates are rough and do not reflect optimal behaviour. We also had to adjust figures to fit into our classification of outputs. Therefore, to allow health centres to adjust to local prices and their targets in health promotion we let the price ratios vary within the range [-50%, 100%]. The unit prices used in this study, price ratios (in parentheses) and allowed variations of MRS around price ratios are reported in Table 2. We let the MRS between visits to other personnel and home nursing visits to adjust without restrictions. The shaded cells in Table 2 are implicitly restricted through other cells.

Table 2. Unit price ratios and allowed ranges of marginal rates of substitution for AR-DEA model

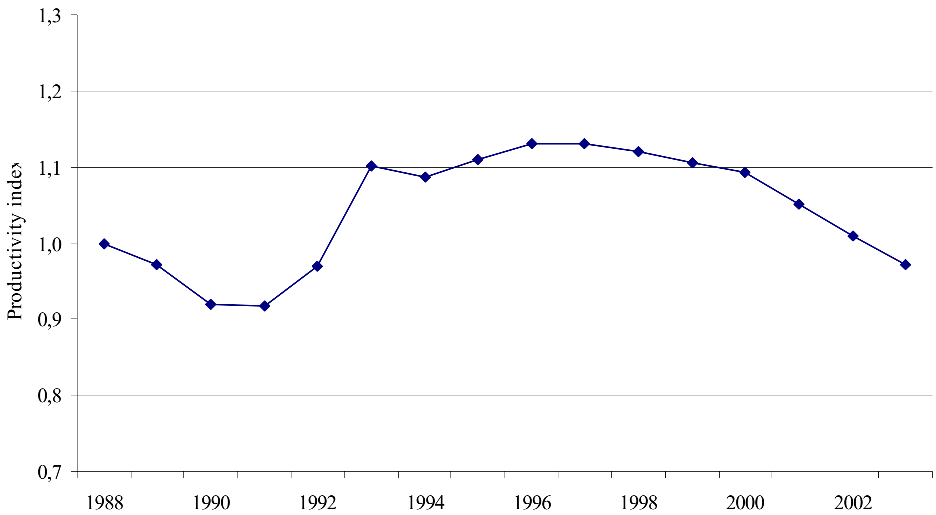
	Dischages	Bed Days	Visits to physician	Dental care visits	Unit price
Bed Days	2–20 (11)				95€
Visits to physician	2–35 (15)	1–3 (1.5)			65€
Dental care visits	2–35 (15)	1–3 (1.5)	0.5–2 (1.0)		65€
Visits to other personnel	4–65 (33)	2–6 (3.2)	1–3.5 (2.2)	1–3.5 (2.2)	30€
Home nursing visits	3–50 (25)	1.5–4.5 (2.3)	0.9–3 (1.6)	0.9–3 (1.6)	40€
Unit price	1000€	95€	65€	65€	

4. Results

4.1 Productivity trend in the long run

The two major changes in the data set complicate the productivity measurement over 1988-2003. First, the data on health centres led by a specialist is not comparable before and after 1993. Thus we have to restrict analysis to health centres led by GPs. Another problem is in visits to other health care personnel than doctors. In 1997 the level of this variable increased due to the incorporation of all visits to the nursing staff by 28 percent. In our productivity estimates before 1997 we have multiplied the reported number of visits in this variable by 1.28.

Figure 2. The average productivity of the health centres 1988-2003. Productivity has increased if the index is > 1



In figure 2 we have pooled the whole data and calculated a joint efficiency frontier over the whole period. The resulting frontier consists mainly of the most efficient health centres from 1993-1998. The productivity index is calculated as (6), but we have averaged annual efficiencies using operating costs as weights. The base year of all the figures is 1988. The productivity levels of 1988 and 2003 are almost equal, but there is a large variation in productivity within the period. The turnaround occurred in 1991, the year when Finland was hit by a very severe economic recession and when it became obvious that the grant system would be reformed from matching grant system to block grant system. From 1991 to 1993 productivity of health centres improved roughly 20 percent and it continued to improve steadily also after that until 1997. Thereafter productivity have had a falling trend and decreased altogether 14 percent between the years 1997 and 2003.

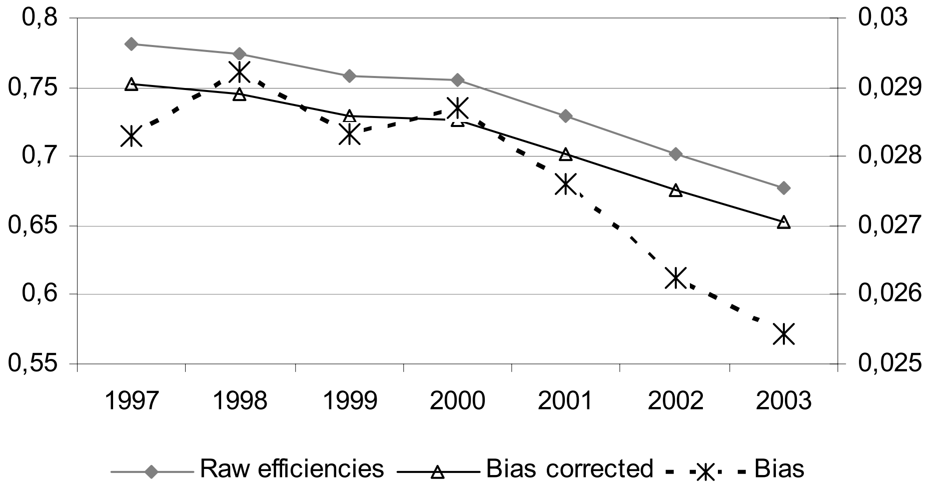
The severe economic crisis and large changes in financial environment appeared to have dramatic impact on primary care productivity. The impact of early 90's recession and the grant reform are interesting as such, but not helpful if we are interested in drivers that have a permanent impact on financial performance in primary care. Also the data problems discussed earlier makes it difficult to analyse pre 1997 development in detail. Thus, in next section we turn our interest in period 1997-2003, when Finnish economy experienced rather smooth economic development.

4.2 Productivity changes 1997-2003

Due the mergers of municipalities and decomposition of some health centre federations and discarding some inconsistent observations our unbalanced panel data set has altogether 247 health centres, with 223-232 units each year. Again, we pool the whole data set of 1594 observations to calculate efficiencies. The first step is to produce bias corrected estimates of efficiency (9). We estimate the AR-DEA model ((1) and (2)) and the 2nd step regression model (7) to yield an estimate of standard deviation, $\hat{\sigma}_{\varepsilon}$. We leave the discussion on regression models to the section 4.2.1. On each simulation round we need to solve 1594 individual linear programming models each having 1623 constraints. Without optimizing and pre evaluation of the units the Mathematica® [34] code to solve each AR-DEA round would take 3 hours on a standard desk top PC. Starting with previously found efficient units and iteratively increasing the set of constraints if necessary, the solution time was cut to less than a minute. Thus we set the number of simulation rounds, L_1, to 1000.

We have plotted the unweighted annual average efficiencies and the estimated bias in figure 3. The original output oriented efficiency scores as well as the bias is converted to show shortfall in productivity i.e. input oriented efficiency. The bias corrected average efficiency decreases from 75.2% to 65.2%. The decrease in average efficiency implies decreased productivity over time. The estimated bias varies between 2.5 and 2.9 percentage units, the amount by which raw DEA scores overestimate efficiencies. It accounts for uncertainty on output levels due to environmental and organisational factors. The level of bias seems to be important, and it is positively correlated with efficiency levels or time. As the range of bias in efficiencies remains relatively stable over time, ratios of efficiencies, i.e. productivity indices, are almost insensitive to estimated bias. This is illustrated in Figure 4.

Figure 3. Annual averages of raw and bias corrected input efficiency scores and estimated average bias, 1000 simulation rounds

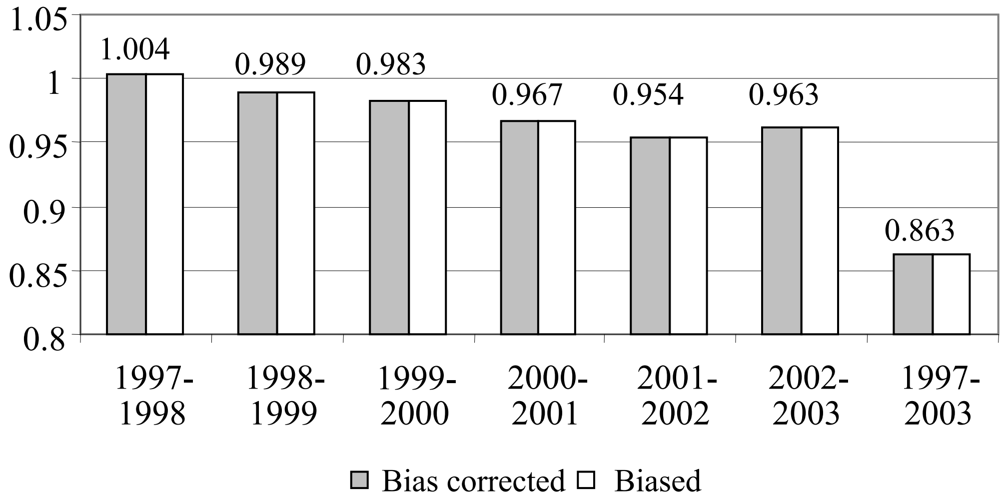


In Figure 4, we have calculated productivity indices as in (6) for consecutive years and first and last years. The indices are operating cost weighted productivity averages over units. The biased and bias corrected consecutive year indices do not differ remarkably from each other. Even if the data set is extended and the base year of the index is different the pattern is almost equal to figure 3. The aggregate productivity decline is 13.7 percent; the major part of it has taken place from 2000 to 2003. The selection of base year for the index and annual variation in units in the data set has a minor importance. Summing up annual productivity indices results in 14.1 percent decline.

The bias in efficiency scores appears to vanish when efficiency scores are turned into productivity index. This is good news if the target is not in benchmarking the units. In measuring aggregate productivity development of the units, laborious simulation of bias is not necessarily needed. However if the interests is in efficiencies alone or even in ranking of the units; these results emphasize the importance of bias.

The productivity figures are useful, but the usual question set is, what the drivers behind observed behaviour are. Instead of modelling the productivity, we use the explanatory models for efficiency scores to quantify the impact of some operation environment variables on productivity using formula (8).

Figure 4. Annual expenditures weighted productivity indices with and without estimated bias in efficiency scores



4.2.1 Productivity drivers

In the explanatory model we can exploit the panel features of the data. The data set covers almost all the public primary care in Finland, thus we consider the municipal, or health centre, effects being fixed. Even if it is technically possible to solve a truncated regression model with 247 health centre dummies, the solution times are relative long and numerical stability of solutions doubtful. Therefore we have converted all the variables to differences from unit means, this is equal to calculating within, or fixed effect estimator. Thus, the truncation point of the dependent variable is $1-\overline{q}_i$, where $\overline{q}_i$ is the average efficiency of the health centre *i*. We also estimate the model without a constant term, such that predicted (as well as pseudo) efficiencies are calculates as $\hat{q}_s = g\left(\tilde{\mathbf{z}}_s;\hat{\boldsymbol{\beta}}\right)+\overline{q}_i$, where $\tilde{\mathbf{z}}_s = \mathbf{z}_s - \overline{\mathbf{z}}_i$ and i is the health centre the observation s comes from. The parameter estimates of biased and bias corrected models of (7), as well as parametric and bootstrapped CIs from $L_2 = 1000$ simulation rounds are collected in Table 3.

Table 3. Truncated regressions of output efficiency scores on environmental and operational variables

	Regression of biased score Number of obs = 1594 Truncated obs = 15 Wald chi^2(7)=282.02			Regression of bias corrected score Number of obs = 1594 Truncated obs = 15 Wald chi^2(7)=231.01				
		95% parametric			95% Confidence Intervals			
	Coeff.	Confidence Interval		Coeff.	Bootstrap		Parametric	
Private physician visits	0.353	0.205	0.501	0.346	0.179	0.509	0.182	0.511
Log of taxable income per capita	0.419	0.288	0.550	0.406	0.275	0.543	0.280	0.532
Share of 85+ years old	8.018	3.812	12.225	7.877	3.350	12.346	3.314	12.441
Disability prevalence of the working-age population	-0.003	-0.005	-0.002	-0.003	-0.005	-0.002	-0.005	-0.002
Personal GPs	0.026	-0.003	0.054	0.024	-0.009	0.053	-0.007	0.056
Alternative long term care	0.139	-0.666	0.945	0.131	-0.737	0.960	-0.759	1.021
Outsourcing	0.316	0.112	0.520	0.318	0.089	0.550	0.092	0.543
Sigma	0.113	0.109	0.117	0.126	0.110	0.119	0.121	0.130

Parameter estimates and CIs of explanatory variables do not differ remarkably if biased or bias corrected dependent efficiency is regressed or if parametric or bootstrapped CIs are used. Thus, the inference on parameters gives inherently the same result. Neither the system of personal GPs, nor the supply of alternative long term care has significant impact on efficiency. However the bias seems to be important for the estimate of standard deviation. The bias corrected estimate $\hat{\hat{\sigma}}_\varepsilon$ is somewhat higher and outside the parametric CI of $\hat{\sigma}_\varepsilon$. For bootstrapped CIs it is not necessary that original parameter estimate is within the range as the simulated CI corrects the bias. This is true for the fixed sample estimate $\hat{\hat{\sigma}}_\varepsilon$ that overestimates σ_ε.

Using both the biased and unbiased parameters we have calculated the productivity effects Δp_r of each explanatory variable r in turn as discussed in section 3.2. Our interest is in development of productivity over the estimation period, thus productivity effects are calculated between 1997 and 2003 values. Figures in the Table 4 are operating cost weighted average effects over the health centres present in the panel both 1997 and 2003.

Table 4. Regressors' contribution to the productivity changes over 1997-2003, the productivity effects Δp_r. Weighted averages over health centres

	Model on biased efficiency scores	Model on bias corrected efficiency scores	dz_r from 1997 to 2003*
Private physician visits	-0.97%	-0.91%	0.04 (visits a person)
Taxable income per capita	-3.15%	-2.94%	11.5% (revenues)
Share of 85+ years old	-1.10%	-1.04%	0.2% (share)
Disability of working age population	-1.80 %	-1.71%	-7.9‰ (share)
Outsourcing	-0.34%	-0.33%	1.6% (share of expenditures)

*For tax revenues dz_r / z_r

Bias corrected productivity changes are slightly smaller. The calculated productivity changes count 6.9 percent of the total 13.7 productivity decrease over the whole estimation period.

The real income subject to local government taxation has grown within the six year period 11.5 percent and has pushed productivity down about three percent. This is not surprising as our productivity measure accounts for money spent for a unit of aggregate output. With favourable development of the local economy municipalities increased primary care resources, this has not led to a respective growth in the service output.

The three drivers connected to demand or need of primary care services account for together about 3.7 percent of the productivity decrease. The reasons for this can be either on more demanding service outputs needed or expenditures that has not reflected on lower demand. The share of elderly above 85 years has increased about 0.2 percentage units over the period. This is typically the part of the population that uses the most of the inpatients ward. Thus, roughly 1 percentage unit decrease in productivity may reflect the increased resource requirements due to more demanding services. The share of elderly in alternative long term care units had no significant independent impact on efficiency, but the structural change taken place in Finnish long term care may have contributed indirectly to decreased productivity in health centres. As the number of beds in alternative units has increased, inpatient wards in health centres take care of more demanding patients.

The increased supply of services which are to some extent substitutes for health centre service like private medical services, service housing and old people's homes reduces the demand for primary care services. If health centre operating costs are slow to adjust to demand for services the health centre productivity will decrease. Slightly less than one percentage fall in productivity of public services is associated with the increased usage of private medical services. Also the improved health status of working-age population has a similar impact. It has likely contributed to the decreased need of primary care, without corresponding change in operating costs of public services.

The widely discussed two organisational changes that have taken place in some health centres do not show success in our results. The system of personal GPs is not statistically significant and outsourcing, that accounts for the purchase of services, have had a small negative impact on productivity. Negative impact may well follow from unintentional nature of outsourcing. In many cases health centres have been forced to buy services from private suppliers, after failing to fill GP vacancies.

5. Discussion and conclusions

In this study we have estimated efficiency variation and productivity changes of health centres by relating the key service outputs to real operating costs. The focus has been on the productive efficiency and the economical use of resources. As a method in our calculation we have used data envelopment analysis which is well suited method for assessing productive performance in cases where price information for outputs is missing or where it is otherwise difficult to assign value weights to different outputs. The use of DEA does not require weights to be assigned a priori, but they are calculated for each service producing unit as a part of the optimization procedure. We, however, constrained the free determination of output weighs by the assurance region method.

The efficiency and parameter estimators we have used here are developed as a response to criticism levelled at non-parametric DEA estimates of efficiency and second stage explanatory regressions. The bootstrap algorithm that animates given data generation process gives a reasonable way to simulate both bias in efficiency scores and consistent confidence intervals (CIs) of regression parameters. The correction of bias is important in two ways. First, if we have a fixed sample of units, it estimates the impact of out of sample units on efficiency. Second, bias correction may also be seen as a fix for the “luck”. Some units may have faced worse or better environmental conditions than others, resulting in variation in output levels. Our bias estimate showed on average 2.8 percentage units shortfall in productive efficiency. However, it appears unimportant if productivity changes are the concern. Thereby, the need of laborious simulations seems to be important for efficiency rankings alone. Technically, we could have used the same simulation process to calculate confidence intervals of efficiency scores, but they are useful only if benchmarking of units is required.

Parametric CIs of regression parameters have been problematic, as the dependent variable in regression, efficiency, is by structure serially correlated. Our simulation results show that inconsistency of parametric CIs is dependent on the data used. In [22] parametric and bootstrapped CIs differed a lot, in our data the only difference was in CI for the standard deviation of error term. Further research in this field is clearly needed, but the likely reason for good performance of parametric CIs in our data is the relatively low number of efficient units.

Our productivity measure equals to the inverse of unit cost of the aggregate service output of primary care. This measure has a crucial role for the finance of primary care. Municipalities financed 89 percent (year 2003) of the gross operation costs of health centres through taxes and government subsidies. Operating costs have a direct impact on municipal budgets, which allocate tax and grant revenues also to specialized hospital care, social services, education, infrastructure maintenance and construction. The fall in productivity means increasing unit prices for services, it gets more expensive to provide the same volume of services. This is what happened in the late 1990s and early 2000s. The only service that has increased in quantities is visits to other personnel than physicians. At the same time expenditures have increased considerably. The resulting decrease in productivity between 1997 and 2003 is 13.7 percent.

Productivity changes over the period 1988-2003 reveal that financial incentives have a strong impact on productivity. Shoestring budgets tend to improve productivity in primary care but savings have not appeared to be permanent. Our regression model predicts roughly a half of the measured productivity change, with taxable income change being the largest single driver. It accounts for three percentage points of the 13.7 percent productivity fall. We assume that private physician visits and disability pension prevalence have a direct impact on demand for public services. Changes in these variables have decreased demand for public primary care, but have had a negative impact on productivity. We consider this as a sign of inflexibility of health centre production and budgeting. Given that some resources have been made available through decreased demand or need of service outputs, savings or reuse of resources are not visible in our data.

Productive measures we have used in this study are based on rather rough and aggregate output data. Therefore, one can question whether the finding that productivity in primary care has declined substantially in recent years is at all reliable. Even if we argue, that economic evaluation is necessary, it can not stand alone as an evaluation criterion. Health centres may have used their resources either to improve quality factors of care, improve working environment or the given care has simply turned to be more demanding. The problem is that no one seems to know what has happened. More demanding care is the usually given answer, but does not give wide support from our analysis. Just one percentage point of productivity loss may be assigned to the share of the oldest olds. Thus, the explanations of productivity fall need a lot more research. We should go inside the unit to see what kinds of changes really have taken place.

Notes

a. The report information register, which was based on annual reports of health centres, was in up till the year 1993 maintained by The Ministry of Social affairs and health.

b. The HILMO is a civil register comprising comprehensive health-care records on inpatient care provided by all hospitals and health centres in Finland.

c. See discussion on stochastic frontiers e.g. in [15] or chance constrained DEA in [16].

References

1. Häkkinen U. The impact of changes in Finland's health care system. *Health Economics* 2005; 14: S101-118.
2. Organisation for Economic Co-operation and Development. *OECD Health Data 2005*. OECD: Paris, 2005.
3. Kalela J, Kiander J, Kivikuru U, Loikkanen HA, Simpura J (eds). *Down from the heavens, Up from the ashes. The Finnish economic crisis of the 1990s in the light of economic and social research*, VATT publications 27:6: Helsinki, 2001.
4. Honkapohja S, Koskela E. The economic crisis of the 1990s in Finland. In *Down from the heavens, Up from the ashes. The Finnish economic crisis of the 1990s in the light of economic and social research*, Kalela J, Kiander J, Kivikuru U, Loikkanen HA, Simpura J (eds). VATT publications 27:6: Helsinki, 2001; 52-101
5. Linna M. Health care financing reform and the productivity change in Finnish hospitals. *J Health Care Finance* 2000; 26: 83-100.
6. Laine J, Linna M, Häkkinen U, Noro A. Measuring the productive efficiency and clinical quality of institutional long-term care for the elderly. *Health Economics* 2005; 14: 245-256.
7. Farrell MJ. The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A*, 1957; 120: 253-290.
8. Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. *European Journal of Operational Research* 1978; 2: 429-444.
9. Newhouse JP. Frontier estimation: how useful a tool for health economics. *Journal of Health Economics* 1994; 13: 317-322.
10. Simar L, Wilson PW. Statistical inference in nonparametric frontier models: the state of the art. *Journal of Productivity Analysis* 2000; 13: 49-78.
11. Charnes A, Cooper WW, Lewin AY, Seiford LM (eds.). *Data Envelopment analysis: Theory, Methodology and Applications*. Kluwer Academic Publishers: Boston, 1994.
12. Coelli T, Rao PDS, Battese GE. *An Introduction to Efficiency and Productivity Analysis*. Kluwer Academic Publishers: Boston, 1994.

13. Pedraja-Chaparro, FJ, Salinas-Jimenez J, Smith P. On the role of weight restrictions in data envelopment analysis. *Journal of Productivity Analysis* 1997; 8: 215-230.

14. Cooper, W., Seiford, L, Tone K. *Data Envelopment Analysis. A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. Kluwer Academic Publishers: Boston (MA), 2000.

15. Kumbhakar SC, Lovell CAK. *Stochastic Frontier Analysis*. Cambridge University Press, New York, 2000.

16. Olesen OB, Petersen NC. Chance constrained efficiency evaluation. *Management Science* 1995; 41 March: 441-457.

17. Banker RD. Maximum likelihood, consistency and data envelopment analysis: A Statistical Foundation. *Management Science* 1993; 39:1261-1264.

18. Korostelev AP, Simar L, Tsubakov AB. Efficient estimation of monotone boundaries. *The Annals of Statistics* 1995; 23: 476-489.

19. Kneip A, Park BU, Simar L. A note on convergence of nonparametric DEA estimators for production efficiency scores. *Econometric Theory* 1998; 14: 783-793.

20. Simar L. and Wilson PW. (2000): A general methodology for bootstrapping in nonparametric frontier models. *Journal of Applied Statistics* 2000; 27: 779-802.

21. Kneip A, Simar L, Wilson PW. *Asymptotics in Non-Parametric Frontier Models*. Discussion paper 0317, Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, 2003.

22. Simar L, Wilson PW. Estimation and inference in two-stage, semiparametric models of production processes. *Journal of Econometrics*, article in press.

23. Malmquist S. Index numbers and indifference surfaces. *Trabajos de Estatistica 1953*; 4: 209-242.

24. Caves DW, Christensen LR, Diewert WE. The economic theory of index numbers and measurement of input, output and productivity. *Econometrica* 1982; 50: 1393-1414.

25. Fried HO, Schmidt SS, Yaisawarng S. Incorporating the operating environment into a nonparametric measure of technical efficiency. *Journal of Productivity Analysis* 1999; 12 November: 249-267.

26. Luoma K, Järviö M-L, Suoniemi I, Hjerppe R. Financial Incentives and Productive Efficiency in Finnish Health Centres. *Health Economics* 1996; 5: 435-445.

27. Räty T, Luoma K, Koskinen, V, Järviö M-L. *Terveyskeskusten tuottavuus vuosina 1997 ja 1998 sekä tuottavuuseroja selittävät tekijät*. Tutkimuksia 88, Government Institute for Economic Research: Helsinki, 2002.

28. Aaltonen J, Järviö M-L, Luoma K, Räty T. *Terveyskeskusten tuottavuuden ja tehokkuuserojen kehitys vuosina 1988–2002*. VATT Discussion Papers 354, Government Institute for Economic Research: Helsinki 354.

29. Aaltonen J, Räty T, Järviö M-L, Luoma K. *Perusterveydenhuollon kustannukset ja tuotetut palvelut – tuottavuudenkehitys 1997–2003* Discussion Papers 372, Government Institute for Economic Research: Helsinki, 2005.

30. Greene HW. *Econometric Analysis*, 5th ed., Prentice Hall, cop: Upper Saddle River (N.J.), 2003.

31. Ali AI. Computational Aspects of DEA. In *Data Envelopment Analysis: Theory, Methodology and Applications*, Charnes A, Cooper W, Lewin AY, Seiford LM (eds.). Kluwer Academic Publishers: Boston, 1994; 63–88.

32. Kalvelagen E. *Efficiently solving DEA models with GAMS*. GAMS Development Corporation, Website. http://www.gams.com/~erwin/dea/dea.pdf. [Updated November 2004], 2002.

33. Hujanen, T. *Terveydenhuollon yksikkökustannukset Suomessa vuonna 2001*. Aiheita 1, Stakes: Helsinki, 2003.

34. Wolfram S. *The Mathematica Book*, 4th ed. Cambridge Univ. Press, 1999.

VATT-TUTKIMUKSIA -SARJASSA ILMESTYNEITÄ

PUBLISHED VATT RESEARCH REPORTS

80. Riihelä Marja – Sullström Risto: Tuloerot ja eriarvoisuus suuralueilla pitkällä aikavälillä 1971-1998 ja erityisesti 1990-luvulla. Helsinki 2001.
81. Ruuskanen Petri: Sosiaalinen pääoma – käsitteet, suuntaukset ja mekanismit. Helsinki 2001.
82. Perrels Adriaan – Kemppi Heikki – Lehtilä Antti: Assessment of the Macro-economic Effects of Domestic Climate Policies for Finland. Helsinki 2001.
83. Venetoklis Takis: Business Subsidies and Bureaucratic Behaviour, A Revised Approach. Helsinki 2001.
84. Moisio Antti – Kangasharju Aki – Ahtonen Sanna-Mari: Menestyksen mitta? Vaihtoehtoisia mittareita aluetalouden kehitykselle. Helsinki 2001.
85. Tuomala Juha: Työvoimakoulutuksen vaikutus työttömien työllistymiseen. Helsinki 2002.
86. Ruotoistenmäki Riikka – Babygina Evgenia: The Actors and the Financial Affairs of the Northern Dimension. Helsinki 2002.
87. Kyyrä Tomi: Funktionaalinen tulonjako Suomessa. Helsinki 2002.
88. Räty Tarmo – Luoma Kalevi – Koskinen Ville – Järviö Maija-Liisa: Terveyskeskusten tuottavuus vuosina 1997 ja 1998 sekä tuottavuuseroja selittävät tekijät. Helsinki 2002.
89. Hakola Tuulia: Economic Incentives and Labour Market Transitions of the Aged Finnish Workforce. Helsinki 2002.
90. Venetoklis Takis: Public Policy Evaluation: Introduction to Quantitative Methodologies. Helsinki 2002.
91. Berghäll Elina – Heikkilä Tuomo – Hjerppe Reino – Kiander Jaakko – Kilponen Juha – Lavrac Vladimir – Stanovnik Peter: The Role of Science and Technology Policy in Small Economies. Helsinki 2002.
92. Räisänen Heikki (toim.): Rakenteellinen työttömyys. Tutkimusinventaari ja politiikkajohtopäätökset. Helsinki 2002.
93. Moisio Antti: Essays on Finnish Municipal Finance and Intergovernmental Grants. Helsinki 2002.
94. Parkkinen Pekka: Hoivapalvelut ja eläkemenot vuoteen 2050. Helsinki 2002.
95. Junka Teuvo: Maailman kilpailukykyisin maa? Tuottavuus ja investoinnit Suomessa 1975-2000. Helsinki 2003.
96. Cogan Josehp – McDevitt James: Science, Technology and Innovation Policies in Selected small European Countries. Helsinki 2003.
97. Perrels Adriaan – Kemppi Heikki: Liberalised Electricity Markets – Strengths and Weaknesses in Finland and Nordpool. Helsinki 2003.
98. Sarvimäki Matti: Euroopan Unionin itälaajentuminen ja maahanmuutto Suomeen. Helsinki 2003.

99. Räty Tarmo – Luoma Kalevi – Mäkinen Erkki – Vaarama Marja: The Factors Affecting the Use of Elderly Care and the Need for Resources by 2030 in Finland. Helsinki 2003.

100. van Beers Cees: The Role of Foreign Direct Investments on Small Countries' Competitive and Technological Position. Helsinki 2003.

101. Kangasharju Aki: Maksaako asumistuen saaja muita korkeampaa vuokraa? Helsinki 2003.

102. Honkatukia Juha – Forsström Juha – Tamminen Eero: Energiaverotuksen asema EU:n laajuisen päästökaupan yhteydessä. Loppuraportti. Helsinki 2003.

103. Simai Mihály (ed.): Practical Guide for Active National Policy Makers – what Science and Technology Policy Can and Cannot Do? Helsinki 2003.

104. Luoma Arto – Luoto Jani – Siivonen Erkki: Growth, Institutions and Productivity: An empirical analysis using the Bayesian approach. Helsinki 2003.

105. Montén Seppo – Tuomala Juha: Muuttoliike, työssäkäynti ja työvoimavarat Uudellamaalla. Helsinki 2003.

106. Venetoklis Takis: An Evaluation of Wage Subsidy Programs to SMEs Utilising Propensity Score Matching. Helsinki 2004.

107. Räisänen Heikki: Työvoiman hankinta julkisessa työnvälityksessä. Helsinki 2004.

108. Romppanen Antti: Vakaus- ja kasvusopimuksen ensimmäiset vuodet. Helsinki 2004.

109. Vaittinen Risto: Trade Policies and Integration – Evaluations with CGE Models. Helsinki 2004.

110. Hjerppe Reino – Kiander Jaakko (eds.): Technology Policy and Knowledge-Based Growth in small Countries. Helsinki 2004.

111. Sinko Pekka: Essays on Labour Taxation and Unemployment Insurance. Helsinki 2004.

112. Kiander Jaakko – Martikainen Minna – Voipio Iikko: Yrittäjyyden tila 2002-2004. Helsinki 2004.

113. Kilponen Juha – Santavirta Torsten: Competition and Innovation – Microeconomic Evidence Using Finnish Data. Helsinki 2004.

114. Kiander Jaakko – Venetoklis Takis: Spending Preferences of Public Sector Officials. Survey Evidence from the Finnish Central Government. Helsinki 2004.

115. Hämäläinen Kari – Ollikainen Virve: Differential Effects of Active Labour Market Programmes in the Early Stages of Young People's Unemployment. Helsinki 2004.

116. Räisänen Heikki: Recent Labour Market Developments in Europe. Helsinki 2005.

117. Ropponen Olli: Kokonaiskulutuksen kehitys Suomessa talouden ulkopuolisten tekijöiden suhteen vuosina 1985-2001. Helsinki 2005.