MEMORANDUM

No 11/99

Efficiency and Productivity of Norwegian Colleges

By
Finn R. Førsund and Kjell Ove Kalhagen

ISSN: 0801-1117

Department of Economics
University of Oslo

This series is published by the
University of Oslo
Department of Economics

P. O.Box 1095 Blindern
N-0317 OSLO Norway
Telephone: + 47 22855127
Fax: + 47 22855035
Internet: http://www.sv.uio.no/sosoek/
e-mail: econdep@econ.uio.no

In co-operation with
The Frisch Centre for Economic Research

Gaustadalleén 21
N-0371 OSLO Norway
Telephone: +47 22 95 88 20
Fax: +47 22 95 88 25
Internet: http://www.frisch.uio.no/
e-mail: frisch@frisch.uio.no

No 01	By B. Gabriela Mundaca and Jon Strand: Speculative attacks in the exchange market with a band policy: A sequential game analysis. 45 p.
No 02	By Pedro P.Barros and Tore Nilssen: Industrial Policy and Firm Heterogeneity. 18 p.
No 03	By Steinar Strøm: The Economics of Screening Programs. 8 p.
No 04	By Kai Leitemo and Øistein Røisland: Choosing a Monetary Policy Regime: Effects on the Traded and Non-Traded Sectors. 39 p.
No 05	By Eivind Bjøntegård: The Composite Mean Regression as a Tool in Production Studies. 9 p.
No 06	By Tone Ognedal: Should the Standard of Evidence be reduced for White Collar Crime? 27 p.
No 07	By Knut Røed and Tao Zhang: What Hides Behind the Rate of Unemployment? Micro Evidence from Norway. 39 p.
No 08	By Geir B. Asheim, Wolfgang Buchholz and Bertil Tungodden: Justifying Sustainability. 24 p.
No 09	By Sverre A.C. Kittelsen: Monte Carlo Simulations of DEA Efficiency Measures and Hypothesis Tests. 61 p.
No 10	By Asbjørn Rødseth and Ragnar Nymoen: Nordic wage formation and unemployment seven years later. 52 p.

Efficiency and Productivity of Norwegian Colleges†

by

Finn R. Førsund

Department of Economics University of Oslo and

The Frisch Centre

Kjell Ove Kalhagen

The Frisch Centre

Abstract: Regional colleges in Norway were reorganised in 1994 with the purpose of promoting efficiency and productivity. This is the first effort of checking what actually has happened afterwards with efficiency and productivity. DEA and Malmquist index approaches are used. Data for three years, 1994, 1995 and 1996 at a department level for about 100 units where collected by questionnaire and direct contacts. The three outputs where final exams distributed on two types; short- and long studies, and research publications. Inputs where number of academic and non-academic staff in full time equivalents, current expenses other than salaries, and building size in square metres. Typical cross section efficiency results show a large share of efficient departments, with a disproportionate number of efficient departments giving theoretical general education, and a large variation within the group of inefficient units. The difference between professional and arts and science departments may be explained by the nature of the teaching production function, but calculations for a sub-sample of professional departments (e.g. nurses, engineers, teachers) show almost the same variation within this group. The productivity change each year was mainly positive, with most departments experiencing a positive productivity effect from frontier shift, but a greater variation from positive to negative as regards the contribution from catching up.

Keywords: *Colleges, efficiency, DEA, Malmquist productivity index*

JEL classification: *C 61, D24, I21*

† This work is part of a project at the Frisch Centre about efficiency in the college sector. The first report (in Norwegian) is Erlandsen, Førsund and Kalhagen (1998), and the second (in Norwegian) Kalhagen (1998). The latter is Kjell Ove Kalhagen's Master Thesis, and Finn R. Førsund was the thesis adviser. Dag Fjeld Edvardsen has assisted during the finishing phase of the report. We are indebted to Tor Jakob Klette for comments.

1. Background

Pressure on public sector expenditures has generated interest in performance indicators the last decades. Higher education in Norway is almost exclusively state run. The sector consists of colleges and universities. Recent interest in overhauling the performance of the public sector of Norway resulted in the creation of a Parliamentary Commission looking into cost efficiency. Performance of the college sector was paid special attention, because with effect from October 1994, 98 colleges were merged into 26 new ones. One purpose of the reform was to obtain a more efficient use of the resources according to educational- and research policy objectives. The task of the Commission in 1997 was to find out if this potential has been realised.

The new state run colleges consist of totally 109 departments, varying from 1 to 8 with an average of 4,5 departments. The colleges offers a lot of studies; professional studies (health and social studies, teacher training, engineering, media, and degrees of Bachelor of Commerce and graduate engineer), university subjects (minor and major subjects), or arts and science in general. The colleges are fully financed by the Ministry of Education, Research and Church Affairs.

In contrast to universities, colleges are relatively more teaching intensive. Another difference is that the colleges, although required to carry out research, do not have a national responsibility for performing basic research.

As a part of the work of the Commission the Frisch Centre has undertaken to investigate the efficiency and productivity of colleges for the relevant time period. The department level turned out to be the most disaggregated level suitable for data collection. In our analysis we will regard each department in the colleges as comparable production units producing education and research. A more ideal level would have been each study organised under departments.

The initial plan was to collect data for a suitable number of years before the reform and up to the latest available year, 1996. But it turned out to be impossible to get data for the pre-reform period for enough departments, leaving us with data for the years 1994, 1995 and 1996. With such a limited number of years our intention with the productivity part of the study is more to explore the possible methods and result presentations rather than offer conclusive insights. In defence of the exercise it may be underlined that this is the first time such an exercise is performed with the applied methodology, and it may serve as a catalyst for improving the data production in the sector, or as Rhodes and Southwick (1993, p.146) expressed it: "..our intention in this exploratory exercise is to identify areas for more thorough investigation and to bring some light, however dim, on a question of relative performance that has received little previous exposure".

When studying inefficiency there are two methodological problems that should be separated: i) establishing a frame of reference for efficient operations, ii) defining the efficiency measures. As to the former we will use the non-parametric approach of DEA, as introduced by Charnes et al. (1978) based on an idea of Farrell (1957), assuming a piecewise linear frontier production structure, and as to the latter we will use the Farrell (1957) efficiency measures. The motivation for imposing a minimal structure on the production possibilities is that the technology for college production is rather unknown, and typically multi-output. Furthermore, there are no prices on outputs; they are not traded in markets.

Among previous studies using DEA for analysing efficiency in higher education related to our study we would like to mention Tomkins and Green (1988), Ahn et al. (1989), Beasley (1990), Rhodes and Southwick (1993), Johnes and Johnes (1993) and (1995), Doyle and Arthurs (1995), and Sarafoglou and Haynes (1996). Typically, all studies have used proxies for the ideal output variables (Flemming, 1991), such as number of students at different levels, exam points, number of research publications of various categories, and research grants. Inputs used have been number of employees of different categories, especially faculty- and administrative staff, wage bill, building and equipment investments, expenditure general and maintenance, equipment, support functions, and research grants. Only Ahn et al. (1989) use data for several years, but do not calculate productivity changes, but focus on changes in efficiency scores by use of "windows analysis". We will explicitly calculate productivity changes. The studies all show a significant dispersion of efficiency scores, and deal with sensitivity analyses in different ways to illustrate the impact of choice of model specifications. We may note that quality issues seldom have been dealt with, probably due to lack of data, but Rhodes and Southwick (1993) do a two-stage analysis with quality-related variables in the second stage of correlating efficiency scores with explanatory variables.

Conceptual issues in defining outputs and inputs are dealt with in Section 2. The DEA method and Malmquist index are presented in Section 3, and the data structure is shown in Section 4. The efficiency distributions are given in Section 5, and productivity results and a more detailed analysis of productivity determinants performed in Section 6. Some remarks on further research are offered in Section 7.

2. Measures of outputs and inputs

When studying productivity and efficiency the key to success is, first of all, to base the study on theoretically satisfactorily definitions of outputs and inputs, and then to operationalise these definitions without compromising too much. A fruitful approach to understand what the institutions in question are producing, is to inspect the objectives of the activities. In general terms a college produces educational services, research, and dissemination of knowledge in society at large. Ideal measures of outputs may be measures of the human capital added for students taking degrees as to education, addition to scientific knowledge as to research (person-specific knowledge and general knowledge, according to Beasley, 1990), and increase in enlightenment of society at large as to interactions college – society (and contribution to "national culture" according to Higgins, 1989). Operational measures of the first category may be number and type of exams. Research may be measures by number of research publications of different types; from prestigious international journals to national language local working papers (see e.g. Johnes and Johnes (1993) for a classification). Interacting activities may be measures by newspaper articles, media appearances, participation of scientific staff in public commissions, and consulting for public and private sector. Ideal and most commonly used measures are presented in Table 1.

The classification of inputs can in general be cast in the KLEM format, i.e. Capital[1], Labour, Energy and Materials. Ideal measures of inputs are hours of labour of different types, such as scientific faculty, administration and support staff, building space, various categories of equipment, and current inputs such as energy, cleaning, maintenance, postage, telephones, stationary. It is usually possible to operationalise Labour straightforwardly by hours worked by different categories. Areas of buildings

[1] K is used instead of C due to tradition.

Table 1. Ideal output measures and operationalisations

Variables	**Ideal measures**	**Operationalisations**
Education	Addition to human capital	Stock of students, Flow of exams, degrees
Research	Addition to scientific knowledge	Research publications External research funds PhD's
Interaction society	Increase in general knowledge, impact on decision-making	Newspaper articles, media exposure, participation in public commissions, consultancies

may be supplemented with year of building to indicate functionality. Equipment should include PCs, but these are difficult to operationalise because ideally we are interested in the potential productivity of the PC, and actual purchase or replacement value do not correspond well to the role of the equipment in research. May be capacity in Bytes and speed in Herz could serve. Usually one has to use purchase figures, and we have to cope with the distortions created.

The quality dimensions are of especial importance for college outputs. Number and types of exams do not tell us the full story of the addition to human capital. One way of capturing the quality dimension of exams would be to have a measure of the success of the candidates after graduation. In a society where wages are strongly influenced by productivity a measure of lifetime income would serve as a quality measure of education. But such information is very difficult to come by, and the egalitarian structure of Norwegian wages makes the quality signals very weak. A more limited measure would be the time it takes for students to get jobs after graduation, assuming that people from the most prestigious colleges get jobs first (see e.g. Johnes et al. ,1987). But such measures, which are possible to get from special labour market surveys, depend heavily on the state of the relevant labour market. With a low rates of unemployment, as in Norway in the relevant years, many candidates experience such low waiting times that a correlation breaks down, e.g.

Table 2 Quality dimensions

Variables	Quality measures
Education	Time before getting first job Income level Reputation of college
Research	Citations Peer recognition
Interactions society	Impact on decisions
Student material	Qualifications at entry Number of hours studying
Staff material	Degrees Seniority Participation in networks International conferences

because a need for a holiday before entering the labour market may be more influential than the expected quality of the education.

Quality of research could be captured by influence measures by citation indices (but see e.g. Flemming (1991) and Higgins (1989) on problems using these). The extent (national/international) and type of networks of faculty could represent quality, and also international conference participation. Where relevant the diffusion of research into practical adaptations in business could be a measure of quality.

It is very difficult to measure the quality of the interactions with society. Impacts through citations of media exposure could be one way.

The role of students should be paid particular attention. Students are the “carriers” of education outputs, but are also inputs. The personal qualities of the students determine how much human capital is actually absorbed during the education. The number of hours used by students studying will obviously also influence the build-up of human capital.

Quality of staff may also be of importance. Measures used have been years of experience, seniority, etc. (see e.g. Johnes and Johnes (1993) p. 343).

The use of proxies for the ideal variables, as portrayed in Table 1, makes necessary explicit measures of quality. Some suggestions are provided in Table 2.

3. The method

The DEA Approach

The technology set, S, can in general be written:

$$S = \{ (y, x) \mid y \text{ can be produced by } x \} \quad \mathbf{(1)}$$

where y is the vector of M outputs and x a vector of R inputs. It is assumed that the set is convex and exhibiting free disposability of outputs and inputs. Farrell (1957) technical efficiency measures can be defined with respect to this set, and they are identical to *distance functions* (introduced to economics in Shephard, 1953) or their inverse. The input-oriented technical efficiency measure, $E_{1,j}$ for unit j is:

$$E_{1,j} = E_{1,j}(y_j, x_j) = \text{Min}_\theta \left\{ \theta \mid (y_j, \theta x_j) \in S \right\} \quad (2)$$

i.e. we seek the maximal uniform proportional contraction of all observed inputs allowed by the feasible technology set.

Introducing a set of N observations the set, S, is estimated as a piecewise linear set by:

$$S = \left\{ (y,x) \mid \sum_{n \in N} \lambda_n y_{nm} \geq y_m \ (m \in M),\ x_r \geq \sum_{n \in N} \lambda_n x_{nr} \ (r \in R),\ \lambda_n \geq 0 \ (n \in N) \right\} \quad (3)$$

where λ_n is the weight for observation n when defining the reference point on the frontier, and N, M, R are also used as symbols for the index sets. It is assumed that the envelopment of the data is done as ”tight” as possible, i.e. minimum extrapolation and inclusion of all observations are assumed. Further, constant returns to scale (CRS) is specified. A special form of variable returns to scale (VRS) is obtained by restricting the sum of the weights to be 1:

$$\sum_{j=1}^{N} \lambda_j = 1 \qquad (4)$$

A piecewise linear production set with (4) included was first formulated in Afriat (1972) as the relevant set for efficiency analysis.

The estimator for the input-saving efficiency measure for observation j is then:

$$E_{1,j} = \min_{\lambda,\theta} \left\{ \theta \mid \sum_{n\in N} \lambda_n y_{nm} \geq y_j m \ (\forall m\in M),\ \theta x_{jr} \geq \sum_{n\in N} \lambda_n x_{nr} \ (\forall r\in R),\ \sum_{n\in N} \lambda_n = 1\ ,\ \lambda_n \geq 0 \ (\forall n\in N) \right\} \qquad (5)$$

This problem is a linear programming problem with M+R (CRS) (+1 if VRS) constraints, and can be solved in a standard way[2]. Following Charnes et al. (1978) this is called the DEA model. The VRS case was reintroduced by Banker et al.(1984), without reference to Afriat (1972).

The Farrell technical efficiency measures are radial, and measure the relative distance to the frontier from an observation. There are two natural directions: keeping output fixed and *input-orient* the measure, and keeping input fixed and *output-orient* the measure. The efficiency measures can be interpreted as total factor productivity measures in the standard meaning of an index of outputs on an index of inputs. The input–oriented (or input-saving) measure is the ratio of the productivity of the observation and the corresponding reference point on the frontier, keeping outputs constant, the output-oriented (or output-increasing) measure is the ratio of the productivity of the observation and the corresponding reference point on the frontier, keeping inputs constant. Since the numerators (denominators) of the productivity indices in the input-oriented (output-oriented) case are identical, we do not have to worry about how the output (input) index is constructed. The efficiency score is based on proportional change of all magnitudes. Assuming that the input (output) index is homogenous of degree 1 in the inputs (outputs) the unknown input (output) index for

[2] We are using an in-house program of the Frisch Centre.

the observation cancels out, and we are left with the efficiency score (see Førsund (1997) for further explorations).

For a VRS frontier technology the basic efficiency measures are extended to cover scale (see, Førsund and Hjalmarsson 1974, 1979). A sort of a scale measure, termed gross scale measure in Førsund and Hjalmarsson (1979), but here renamed more appropriately *technical productivity measure*, is defined as the ratio of the productivity of the observation and the productivity at the corresponding (i.e. keeping observed output ratios and input ratios) technically optimal scale point on the frontier. We know (see Frisch (1965) or e.g. Førsund, 1996a) that the latter productivity is maximal. The pure scale measures defined in Førsund and Hjalmarsson (1979), here simplified to *scale measures*, may also be interpreted as productivity measures by forming ratios of productivities with the input- and output corrected reference points respectively on the frontier and optimal scale point. To realise that also in these cases we do not have to know the productivity indices is a little more involved, and require the introduction of the enclosure of the VRS production function by the smallest cone, i.e. a CRS technology. We will return to this explanation after the graphical presentation of the DEA frontier and the efficiency measures provided in Figure 1.

Two inefficient units, P_1 and P_2 are shown in Figure 1, and the concepts used in DEA analysis are introduced. The efficiency measures for observations P_1 are:

Input - saving efficiency: $E_1 = x_F / x_{1.}$

Output - increasing efficiency: $E_2 = y_1 / y_{G.}$

Technical productivity: $E_3 = (y_1/x_1)/(y_B/x_1) = E_1(CRS) = x_I/x_1 = E_2(CRS) = y_1/y_M$.

Scale efficiency, input orientation: $E_4 = E_3 / E_1 = (y_1/x_F) / (y_B/x_B)$.

Scale efficiency, output orientation $E_5 = E_3 / E_2 = (y_G/x_1) / (y_B/x_B)$.

The way these measures are defined they are all between zero and one. The productivity- and scale measures can be expressed as ratios of productivity of the observation, P_1, and its two corresponding frontier points, F and G respectively, and the maximal productivity at the frontier at B. These measures can also be expressed as ratios of the slopes of the rays from the origin through these points and the slope of the ray to the point of maximal productivity, B. Returning to the productivity interpretation above for the E_3, E_4 and E_5 measures in general, note that the productivity measure is identical to the input- and output-oriented efficiency measures with the CRS support technology as the frontier reference technology, as stated above for Figure 1. But this is a general result because with more dimensions we require that

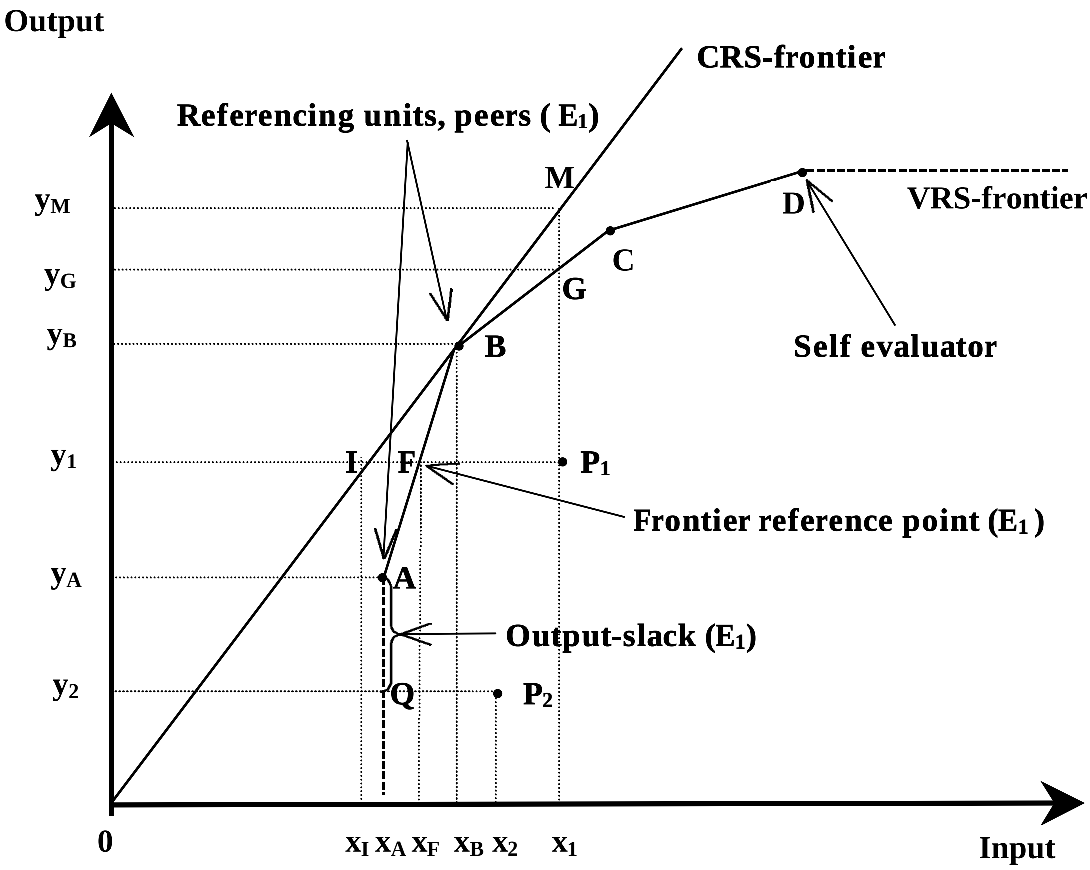


Figure 1. DEA frontier, concepts and efficiency measures

observed output ratios and input ratios are kept fixed. Therefore, the last two relations are also general. These can than be used to give E_4 and E_5 productivity interpretations.

The two main technologies, CRS and VRS are shown in the figure. We note the special feature of VRS in the DEA case: the technology does not include the origin. A non-increasing returns to scale technology (NIRS) could also be specified, in Figure 1 with OBCD as graph.

The terminology we will use is indicated in Figure 1. The efficient units when calculating the efficiency score for an inefficient unit are termed *referencing units,* or *peers* i.e. the efficient units with positive λ-weights in (5), and the point on the frontier is the *reference point*. Calculating, in the VRS case, E_1 for unit P_1, units A and B are referencing units (peers) and F is the reference point. Unit D is efficient, but is a *self-evaluator* calculating both input- and output- oriented measures.

We know slacks are an integral part of a LP problem. In Figure 1 we have an output-slack when calculating E_1 for unit P_2. With more dimensions we can also have input (output)- slacks when calculating input (output)-oriented efficiency, and we have a choice of presenting the radial efficiency measures, or non-radial ones including slacks (see e.g. Torgersen et al. (1996) for an overview).

Finally, the LP programme also calculates the duals and gives us all the shadow prices, which can be utilised to calculate marginal transformation rates and productivities.

The Farrell technical efficiency measure in the CRS case ($E_1 = E_2$) is the most used, but also the extended Farrell measures have been used in the literature under various names. However, the comprehensive scheme offered above, predating this literature, based on Førsund and Hjalmarsson (1974) and (1979), seems to have gone mainly unobserved[3]. Since the student enrolments of colleges are determined by the Government it is most relevant to calculate input-saving efficiency measures here.

The Malmquist productivity index

The productivity index is based on binary comparisons for a production unit between two time points (or between two different units at the same point in time). The time periods to be compared, are denoted 1 and 2 for short. Only quantities are involved, and at least one technology has to be known. As a convention we will compare a unit observed in period 2 with the same unit observed in period 1, i.e. expressions involving period 2 observations will be in the numerator and expressions involving period 1 observations will be in the denominator.

Introducing cross-section data sets for several years the technology set, S, has to be dated, e.g. S^t, $t \in T$, where T is the set of years. Caves et al. (1982) introduced productivity indices for discrete observations based on Malmquist (1953). The basic idea is to utilise Farrell efficiency measures, or distance functions, for the two

[3] For instance, Banker et al. (1984) call E_3 for "technical and scale efficiency", and E_4 for "(input) scale efficiency", while Färe and Lovell (1978), Färe et al. (1985), Färe et al. (1994a) do not recognise E_3 as a scale measure, but as a technical efficiency measure for CRS technology, probably due to $E_3 = E_1$ (CRS) = E_2(CRS), and call E_4 input scale efficiency measure and E_5 output scale efficiency measure.

observations against a common reference frontier. An efficiency measure can itself be interpreted as a ratio of the observed productivity and the productivity at the corresponding point on the reference technology. The Malmquist productivity index, $M_{j1}^{\ 1,2}$, for comparison between two time periods 1 and 2 for a unit *j* with frontier technology from period *1* as reference, based on input-oriented efficiency measures, is:

$$M_{j1}^{1,2} = M_{j1}(y_j^1, x_j^1, y_j^2, x_j^2) = \frac{E_{j1}^2}{E_{j1}^1} = \frac{\mathrm{Min}_{\theta^2}\left\{\theta^2 \mid (y_j^2, \theta^2 x_j^2) \in S^1\right\}}{\mathrm{Min}_{\theta^1}\left\{\theta^1 \mid (y_j^1, \theta^1 x_j^1) \in S^1\right\}} \quad , \ 1,2 \in T \ , \ j \in N \qquad (6)$$

The index notation system is that observation years are shown as superscripts, and technology year as subscript. We have picked out two years called 1 and 2 as observation years, and used one of them, 1, as technology reference. In general, any year in the set t can be used as technology reference. The numerator shows the proportional adjustment, by the scalar θ^2, of the observed input vector of the period 2 observation required to be on the frontier function of the reference period 1 with observed outputs. The denominator shows correspondingly the adjustment by θ^1 of the observed input vector of period 1 for the observation to be on the same period 1 frontier function. Note that the measure with different time reference for year of observation and reference technology now may be greater than one, if the observation is not feasible within the technology in question. In fact, the measure itself may be infeasible to calculate. If $M_{1j}^{\ 1,2} > (<) \ 1$, then the observation in period 2 is more (less) productive than the observation in period 1.

An output-oriented Malmquist index can be defined in a similar way. Under the CRS assumption it would be equal in value to the input-oriented index, and the efficiency measures will always be feasible in principle. If we want to interpret the index as a total factor productivity index we must base the efficiency measures on comparing the observations with the corresponding optimal scale points, i.e. we must use the measure we have termed E_3; the technical productivity measure. In practical applications this is *as if* we use a CRS reference technology enveloping the actual VRS technology (enveloping the VRS technology with the smallest cone) (see Førsund (1997) for further discussion).

In the presence of inefficient observations change in productivity is the combined effect of change in efficiency and shift in the frontier production function[4]. Färe et al. (1994b)[5] showed how the CCD index in the case of inefficient observations could be decomposed when there are two time periods and one of them is used as reference technology. The Malmquist productivity index, $M_{j1}^{1,2}$, can be multiplicatively decomposed into two parts showing the catching up, $MC_{j1,2}^{1,2}$, and the technology shift, $MF_{j1,2}^{2}$:

$$M_{j1}^{1,2} = \frac{E_{j1}^{2}}{E_{j1}^{1}} = \frac{E_{j2}^{2}}{E_{j1}^{1}} \cdot \frac{E_{j1}^{2}}{E_{j2}^{2}} = MC_{j1,2}^{1,2} \cdot MF_{j1,2}^{2} \;,\; 1,2 \in T \tag{7}$$

The catching-up effect, $MC_{j1,2}^{1,2}$, expresses the relative movement of the observed unit to the frontier, a higher (lower) "contemporary" efficiency score for the second period implying increased (decreased) efficiency. The frontier technology change is expressed by the ratio of the efficiency scores for the same second period observation relative to the two technologies. The numerator expresses the scaling of period 2 inputs in order to be on period 1 technology, while the denominator expresses the scaling of the same input vector in order to be on period 2 technology, in both cases subject to period 2 observed outputs. This then serves as a measure of technology shift, and is greater than one if period 2 technology is more efficient relative to period 1 technology for the input-output mix of the period 2 observation.

If another year than the two observation years is chosen as basis for reference technology, the expression for frontier shift is slightly complicated by imposing chaining (see Berg et al., 1992):

$$MF_{ji,2}^{1,2} = \frac{E_{ji}^{2}/E_{j2}^{2}}{E_{ji}^{1}/E_{j1}^{1}} \;,\; i,1,2 \in T \;,\; j \in N \tag{8}$$

The chained frontier technology change is a relative change between period i

---

[4] See e.g. Nishimizu and Page (1982) for such a decomposition in the parametric frontier case.

[5] Originally circulated as a working paper in 1989.

technology and period 2 technology in the numerator, and period i technology and period 1 technology in the denominator.

4. Data

We shall concern ourselves here with data for the 1994, 1995 and 1996 academic years (the latest year for which data are available at the time of writing). In addition to public data (NSD, 1997), the data used in the present paper was collected by the Foundation for Research in Economics and Business Administration (SNF) (now the Frisch Centre), at college department level. We sent out questionnaires to the 26 regional colleges comprising 109 departments and received data from 23 of them comprising 99 departments. Unfortunately the project had a very tight time schedule, so the quality of the data may be negatively influenced by this. In order to secure quality we followed up the questionnaire by telephone contact and gave all colleges the opportunity to see our first version of the data for themselves and communicate any corrections. Although there was some problems with interpretations of our variable definitions and the tight time schedule, in our opinion the data are of sufficient quality to express reliable structural features and trends in the regional college sector.

Output measures

As proxy-variables for research output (R&D) we asked for information according to the following typology

- (i) Papers in professional journals
- (ii) Papers in academic journals
- (iii) Authored books
- (iv) Contributions to edited books

Due to lack of information, some departments could not split up their research production into these four categories. We have therefore summed (no weighting) these four categories into one category called *R&D.*

As output-measure for person-specific increased knowledge we have used total number of exam credits[6] (product of candidates and exam credits). We have split this measure into two categories due to typical difference in resource usage between short and long education

(i) Short education: Studies that is stipulated from 6 months up to 2 years, plus one year extension course.
(ii) Long education: Studies that are stipulated for 3 years or more.

Input measures

Four input-measures are used in the analysis

(i) Faculty staff: Number of faculty staff man-labour year
(ii) Administrative staff: Number of administrative staff man-labour year
(iii) Net operating expenses: operating expenses minus wage costs
(iv) Building capital: Number of square meters

Two measures of staff man-labour year are used in the analysis. One for the staff with a solely research and teaching functions and one for staff with only administrative functions.

In the analysis we use net operating expenses (operating expenses subtracted by wage costs) for other inputs, such as cleaning, heating, stationary, telephones and postage, maintenance[7].

We have included size of building in the analysis. It is not obvious that this is an interesting input in our context. Of course, some minimum space is needed for the educational process and research, but above that it is difficult to argue that more space promotes the production of our outputs. Effects of space like it is more expensive to clean rooms in large buildings, and higher costs associated with central heating the more space, would be captured by operating expenses. Table 3 summarises the variables used in the analysis.

---

[6] By stipulation full time students will obtain 20 exam credits per year.

[7] Since Buildings are represented by area rent should have been taken out of the expenditure figures. It could also be argued that maintenance should be taken out, since it could be used as a proxy for buildings, see Ahn et al. (1989).

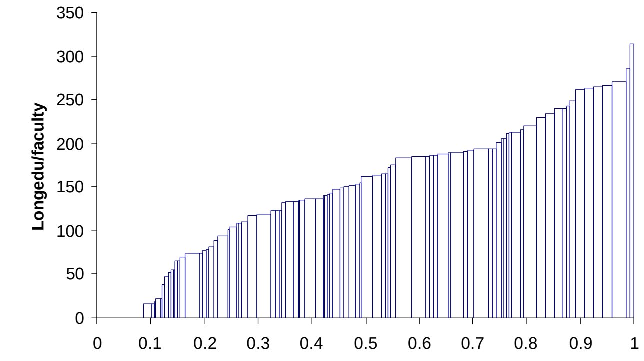
Table 3 Variables and definitions in DEA model

Variables	**Definitions**
Output measures:	
Shortedu	Studies that are stipulated from 6 months up to 2 years, plus one year extension courses
Longedu	Studies that are stipulated for 3 years or more
R&D	Research publications
Input measures:	
Faculty staff	Number of faculty staff man-labour year
Administrative staff	Number of administrative staff man-labour year
Net operating expenses	Operating expenses minus wage costs
Size of building:	Number of square meter building

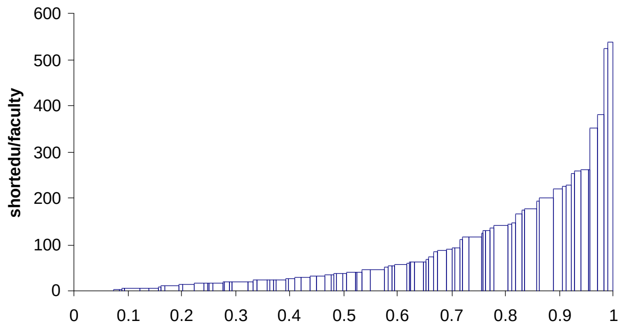
Structure of data

The structure of data can graphically be illustrated by joining variables in pairs as shown in Figure 2. Each histogram represents one department, and the width of the histogram represents the relative size of the department measured by the number of full time student equivalents. By ranking the departments in increasing order by the ratios we obtain information about the total variation in the distribution, the shape of the distribution, and the localisation of large and small units. The extent of the outlier problem will be revealed, and data to be double-checked are pinpointed. With totally seven variables, there are a lot of possible combinations to be shown. We have focused on six. We have done the calculations for all three years, but will only show the structure for 1996.

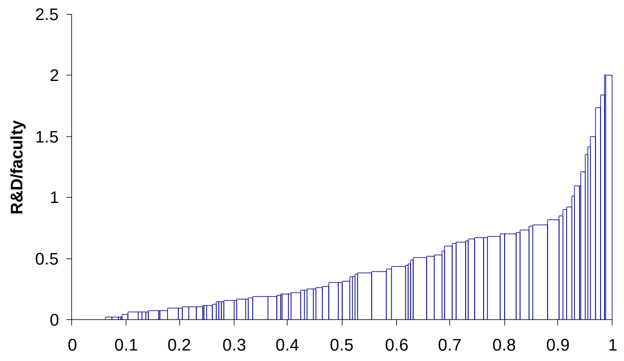
The three first distributions, *Panels a, b* and *c*, shows the ratio between the three products *short education*, *long education, R&D* and the input *faculty staff*. Exam



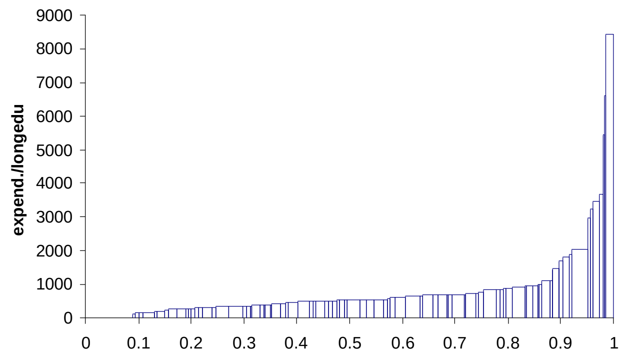
Panel a. Long education / faculty staff



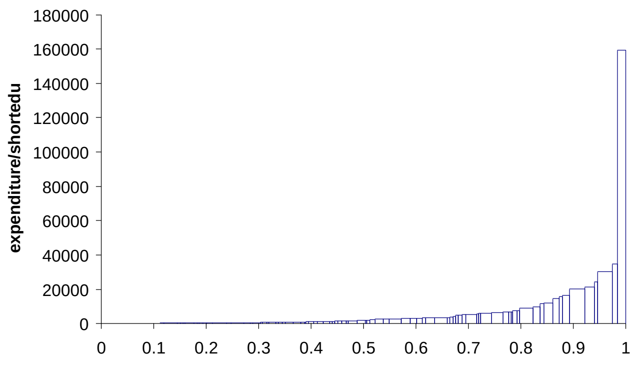
Panel b. Short education / faculty staff



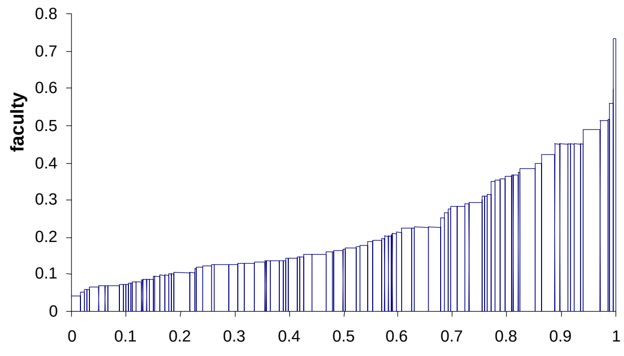
Panel c. R&D / faculty staff



Panel d. Expenditures / long education



Panel e. Expenditure / short education



Panel f. Administrative staff / faculty staff

Figure 2. Salter diagrams 1996. Relative size measured by full time student equivalents.

credits for the product *long education* per faculty staff varies gradually from 16 to 314. A little tail of departments representing about 9% of the population of students, has no *long education* at all. There is a tendency that medium-sized departments dominate the most ”productive” part of the distribution, but with exceptions.

Exam credits for the product *short education* per faculty staff (*Panel b*) varies from 4 to 538 with a median of 35. Also for *short education* we have departments representing about 9% of the population of students, with zero output. The distribution has a different shape with a large share of departments having modest productivity. Middle-sized departments dominate the most productive part of the distribution, which has a more marked “best practice” tail than for *long education*. Small and large departments dominate the part of the distribution with lowest productivity.

The distribution for *R&D* per faculty staff (*Panel c*) is somewhat skewed like the one for *Short education*. There are 12 departments that representing about 7% of the total number of students, with no *R&D* production at all. On the other side of the distribution a group of departments that represents 7% of the students has extreme high *R&D* production. These units are smaller than the average measured by relative student population. One of the two most extreme departments is very small and has over two *R&D* contributions per faculty staff. Generally we observe that small departments have *R&D* productions characterised by larger variance than larger departments.

In *Panel d* we have the ratio between operating expenses and exam credits for the product *long education*. The distribution shows large variation, from 99 to 8410 NOK per exam credits. The median is about 522. One department is extreme within a tail representing about 5% of the students.

In *Panel e* we have the ratio between operating expenses and exam credits for the product *short education*. The distribution shows large variation, from 63 to 159 275. The median is 1067.We recognise the same extreme department having almost 20 times as high ratio than the median. The distribution is visually dominated by this observation. Double-checking revealed that the department had had extremely low number of exams of both types that year.

In *Panel f* we look at the ratio between the inputs *administrative staff* and *faculty staff*. We find a smooth distribution with no extreme outliers, but the most extreme department has a somewhat higher ratio than the next one. We would expect to see a mix of economies of scale and professional departments needing more technical laboratory or equipment staff classified as administrative (not teaching). There is a relatively even mix of small and large departments in the distribution, but the lowest ratios are dominated by small departments, indicating *diseconomies* of scale, while around the median value medium-sized units dominate. Some large departments have relatively high ratios. These are professional departments and the technical staff effect could dominate. But it should be remembered that the distributions are all partial and that the simultaneous approach below is needed for a proper look into issues like economies of scale.

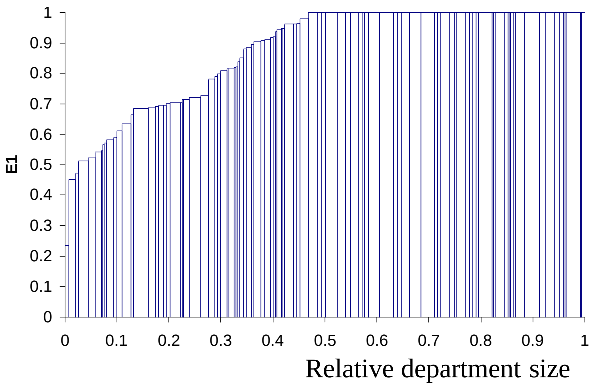
5. Efficiency results

Efficiency distributions

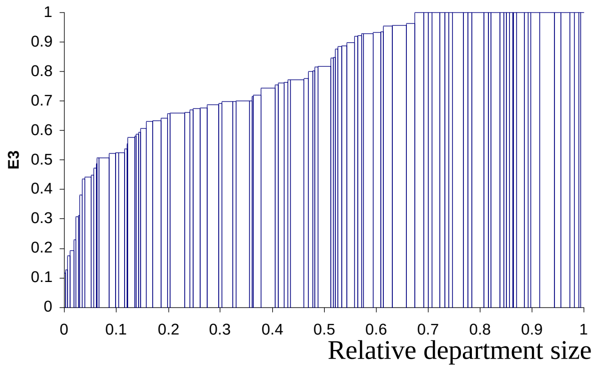
The technical efficiency of a college reflects the potential for increasing the college output without increasing the use of resources (output efficiency) or the potential for reducing the use of resources without reducing the school output (input efficiency). The analysis makes use of the input efficiency definition. This is due to the fact that student capacities are regarded as exogenous in the short run. We allow for variable returns to scale, which means we believe size of college is of importance calculating the efficiency scores.

The technical measure and the scale measure for 1996 are presented in Figure 3, *Panels a* and *b*. Along the horizontal axis we have all the 99 departments. Each histogram represents a department and the width is the ratio between student mass at a department related to the total student population in the college sector. Efficiency is measured along the vertical axis. The departments are ranked according to increasing efficiency score.

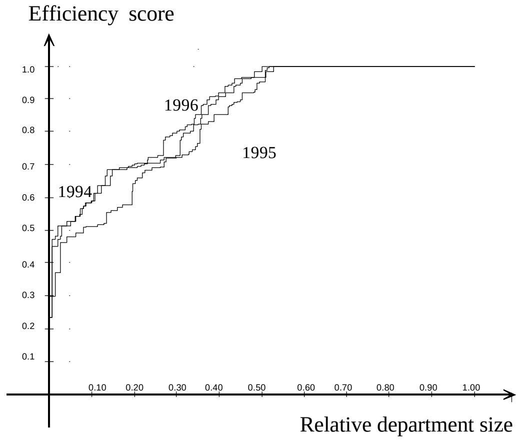
The distribution for the input saving technical efficiency measure shows that 47 departments of 99 are technically efficient (score equal to 1), and these best practice



Panel a. Input-saving efficiency 1996



Panel b. Technical productivity measure 1996



Panel c. Efficiency distributions.
Input-saving measure, 1994, 1995, 1996

Figure 3. Efficiency measure distributions.
Relative size measured by full time student equivalents 1996.

(BP) departments have a share of students at about 55% . Worst practice departments (WP) have a share of students at about 10% when WP is defined efficiency scores lower than 60% (or 0,6). From the figure we see that the BP units mainly consists of small and big departments, while WP mainly consists of medium-sized departments.

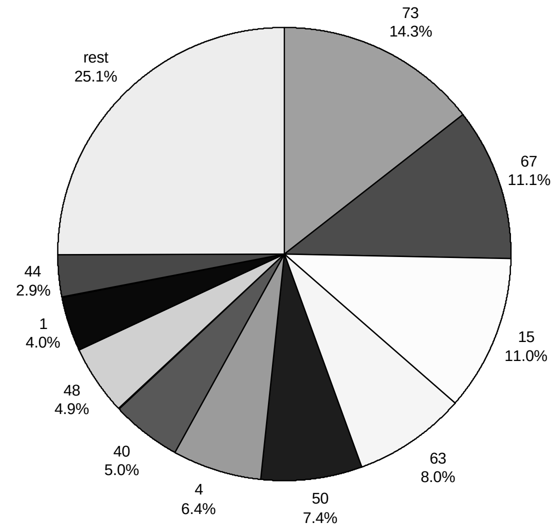
Panel b shows the distribution for the technical productivity measure. Of 99 departments 31 are scale efficient, and the optimal scale departments have a student mass at about 33%. The scale efficient units consist mainly of small and medium-sized departments. WP productivity departments have a student mass at about 15% efficiency, when WP is defined as efficiency lower than 60% (or 0.6). WP mainly consists of small and medium-sized departments, but the extreme worst tail consist of small ones.

Panel c shows the shift of the distribution for input-saving efficiency over the years 1994, 1995 and 1996. The tops of the histogram distributions like in Panel a are exhibited as step curves. We see that the shape and location of the distributions for 1994 and 1996 are quite similar (but note that movements of individual departments cannot be seen), and that the distribution for 1995 shows somewhat higher inefficiencies that year. The share of students at efficient departments is remarkably stable[8].

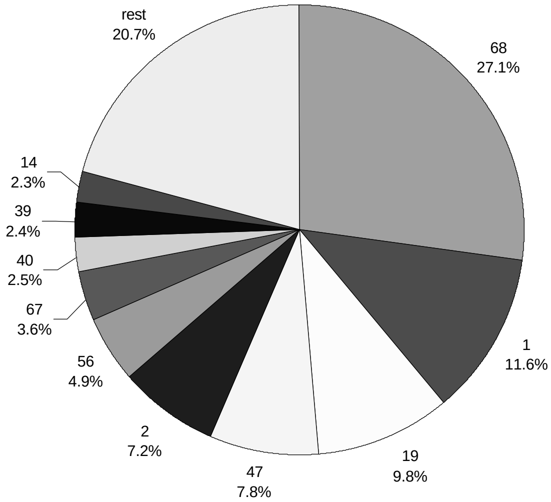
The Peer index

Panel c of Figure 3 shows us that the share of students at efficient units is relatively high at about a level of 50% for all years for input-saving efficiency. These units are the peers that inefficient units may study in order to improve their performance. The efficient units cannot be further ranked as to efficiency score. This has been pointed out as a problem in the literature, and ways of ranking them have been introduced (see Andersen and Petersen, 1993). We will here prefer to show an alternative ranking introduced in Torgersen et al. (1996). For each efficient unit we have in Figure 4 calculated the share of total potential input saving as to faculty staff that is represented by the inefficient units that have the efficient unit in question as a peer. We know that in general there may be several peers for an inefficient unit (in Figure 1 units A and B

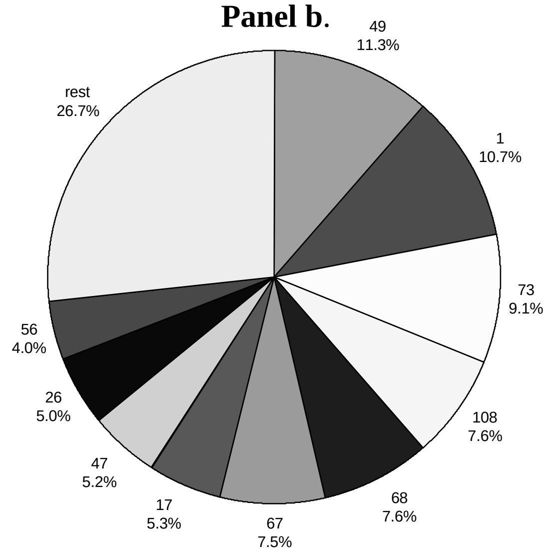
[8] Note that this approach is different from “window analysis” (Ahn et al. ,1989), where different cross section sets are created by dropping and adding years.



Panel a.



Panel b.



Panel c. 1996

Figure 4. Peer index for faculty staff (input saving efficiency).
Ten most important peers

are peers for unit P_1). The potential input saving is therefore weighted by the weight of the peer in the calculation of the frontier reference point (the λ_n in Eq. (5)). The peer index is input (or output) specific. We are only showing the index for faculty staff for the input-saving measure for the three years, identifying the ten most important peers.

Stability

A very important opportunity provided by times-series cross section data is to check on the stability of best practice units. If the turnover is very high then the yearly efficiency results are driven by time-specific conditions and it is difficult to learn from the exercise as to policy implications. If the set of best practice departments is fairly stable, then one has a much more reliable basis as to required policy actions in order to improve efficiency.

The VRS model yields a fairly high proportion of best practice departments for all years, 52% for 1994, 45% for 1995, and 43% for 1996. Such relations caution us to look for self-evaluators. There are 11 in 1994, 9 in 1995 and 10 in 1996, or a little in excess of 1/5 of the best practice departments each year. Of the best practice units in 1994, about 2/3 are also best practice ones in 1995, and of the remaining efficient ones a little less than 2/3 remains efficient also in 1996. Of the efficient units i 1995 above 2/3 remain efficient in 1996. The set of units remaining efficient in all years represents somewhat above ½ of the best practice units each year, or varying from 27 to 24% of the total number of departments. In this set no unit is a self-evaluator in all years, and only two are for two years, while the percentage of self-evaluators varies from 1/5 to less than 1/10 for each year.

Another way of looking at stability is to inspect the group of most influential best practice departments. The Peer index for each year in Figure 4 shows us the most influential peers. Choosing the faculty-oriented index, we have that of the 10 most influential peers each year, 6, 7, and 4 of the units in the years 1994, 1995, and 1996 respectively remaining efficient all the years belong to the 10 most influential. Of these, two units, no. 1 and 67, remain in the top-ten set all years, while five units are in the top-ten set two of the years. Although not based on any formal test, we conclude that there is enough stability in our results to claim that the study has revealed some structural features worth while pursuing for policy purposes.

6. The productivity development

The Malmquist productivity index

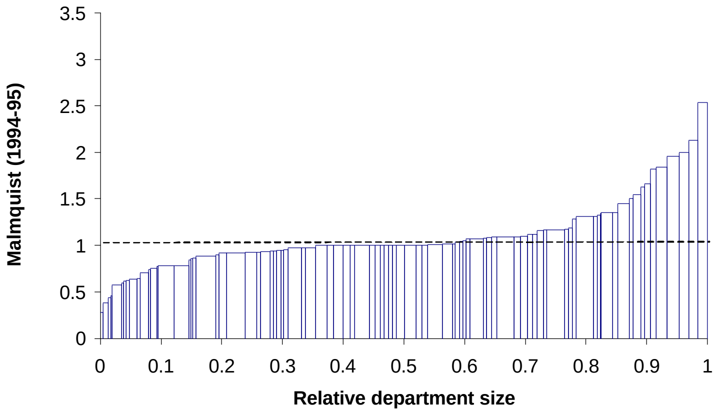
The strength of our approach to calculate productivity growth is that we get the development for each unit. As a background for a discussion of distributions of productivity change it may be useful to inspect the average changes of the variables, set out in Table 4.

Table 4. Percentage change in variables

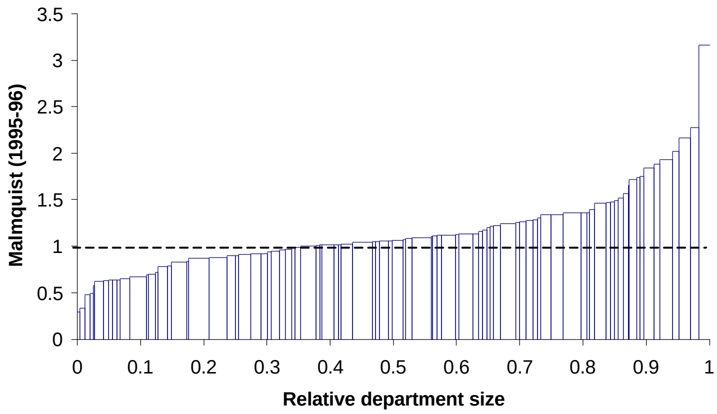
Variable	(95-94)/94%	(96-95)/95%	(96-94)/94%
Short edu	25.5	0.1	25.6
Long edu	3.4	15.9	23.0
R%D	16.0	10.8	28.6
Faculty	0.4	1.9	2.2
Adm. Staff	-0.5	2.9	0.0
Expenditure	-19.4	-26.2	-36.1
m2	-1.9	0.3	-1.6

Regarding the three outputs we see a strong average growth in short education in the first period and a moderate increase in long education, while short education is at a standstill in the second period while long education has strong growth. A strong substitution is indicated. Research and development has a high growth in both periods. As to the inputs all except expenditures (net of wages) are more or less at a standstill. The expenditures decrease quite strongly. This average development points to productivity increase on the average driven by output growth and expenditure decrease. The individual variability was demonstrated in Section 4. The variability in the outputs short education and research and development, and in the input expenditures, is much stronger than in the other variables.

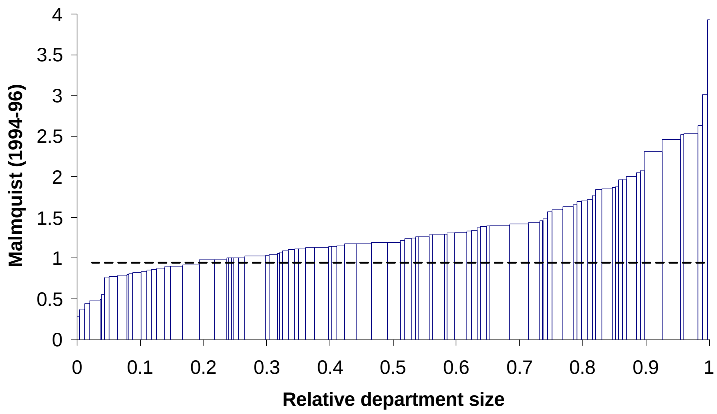
Figure 5 shows productivity distributions for pairs of years (1994-95, 1995-96 and 1994-96) in *Panels a-c*. The frontier for the starting year 1994 is used as reference technology. Since we are assuming VRS-technology, the Malmquist-index is based on the technical productivity measures. The productivity index is calibrated such that



Panel a. Productivity growth 1994-95



Panel b. Productivity growth 1995-96



Panel c. Productivity growth 1994-96

Figure 5. The Malmquist productivity index

productivity estimates lower than 1 are indicating decreased productivity, and larger than one increased productivity. If a unit obtains 1.10 this shall be interpreted as a 10% productivity growth. The width of the histogram is still proportional to the relative size measured by the number of full time student equivalents.

Panel a shows the productivity growth in 1994-95. Departments with decreased productivity represent about 35% of the student mass (in 1996), and departments with increased productivity growth represent 40%. These variations are large taking into account the short period, and as expected from the average changes set out in Table 4 and individual variability illustrated in Figure 2. There is a group of departments with almost no productivity growth covering about 25% of the students. We have a mix of medium-sized and small departments here. Small and medium sized departments also dominates the WP group with decreased productivity, and then some large departments, while medium-sized departments dominates the top group with productivity growth.

Panel b shows productivity growth distribution for the period 1995-96. Departments with positive productivity growth represent about 55% of the students (in 1996). In contrast to *Panel a*, there is no longer a group of departments with constant productivity. Large and small departments dominates the group with productivity decline, while small and medium-sized departments dominates the group with productivity growth, the latter again in the maximum growth group.

In *Panel c* we show the productivity growth for the whole period 1994-96. Since we are applying an index that is chained, productivity growth is simply the multiplication of the two corresponding numbers for a unit in Panels a and b. Therefore it is not surprising that we observe different trends regarding which type of departments having productivity growth. The share of departments with positive productivity growth increases further, with over 70% of the students at departments with productivity growth. For 1994-96 we observe no clear pattern indicating whether there are small or big departments dominating the group with productivity growth, but the positive productivity growth part of the distribution starts with large and medium-sized units dominating, then a part with small units, and lastly some large and small units at the top end. Note that the numbers are rather large for such a short period. The three large units in the top group in *Panel c* have an growth in productivity of about 150 % , while the small best practice outlier has a growth of almost 300%. But the significant changes in average values revealed in Table 4, and the large individual variation illustrated in Section 4 support the reliability of the results.

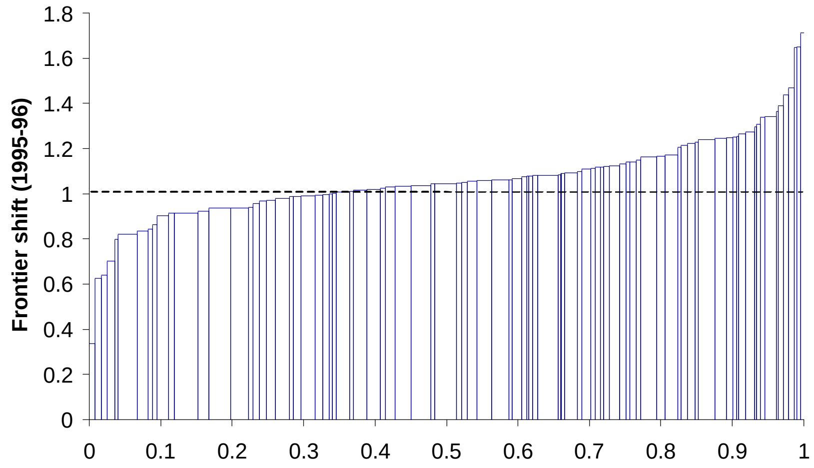
Decomposition of the Malmquist-index

In Figure 6 we have decomposed the productivity growth from 1995 until 1996 into a part called “frontier shift” (*Panel a*) and a part called “catching up” (*Panel b*), in accordance with Equations (7) and (8). From *Panel a* we can see that most of the departments have gained from a positive shift in the frontier transformation function. About 67% of the departments (relative size measured by full time student equivalents in 1996) have benefited from the frontier function shift. The large units have the most modest impact, while the top group consists of small departments. As to decline through frontier shift all the groups are represented, with medium-sized departments dominating the group with most modest impacts, and then large departments. There is a little tail (about 4% of the students) with a marked contribution in decrease in productivity from frontier shift.

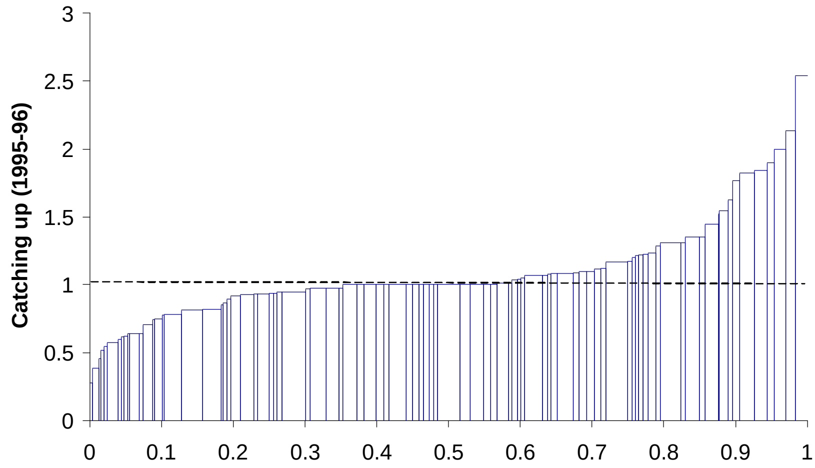
The “catching up” effect (*Panel b*) shows large variations, especially at the upper end of positive productivity growth contribution. The departments that are catching up the best practice departments represents about 45 % of the students. Large and medium-sized departments, and some small ones, dominate the group with productivity growth, with the latter group clearly dominating the top part. A share at about 20% shows now change. This group consists of the departments that are on the frontier both in 1995 and 1996. The units with reduced productivity growth represent about 35% of the students. Some large departments belong to this group with productivity decline contributed by catching-up, except from the worst practice group where small departments dominate (worst practice defined as productivity growth lower than 0.6, i.e. productivity decline of 40%). Summing up, it seems that positive frontier shift is most important for small units, negative shifts most important for large units, while positive catching-up is most important for medium-sized units, and negative catching-up for both small and large units.

Characterisation of productivity change

It is interesting to examine to what extent changes in the variables from 1995 to 1996 effects the estimated Malmquist index. The classic hypothesis of Verdoorn (see



Panel a. Malmquist decomposition, frontier shift



Panel b. Malmquist decomposition, catching up

Figure 6. Decomposition of the Malmquist productivity index for the period 1995- 1996. Relative size measured by full time student equivalents.

Førsund, 1996b) is that there must be growth in output in order to realise productivity growth. In the spirit of Verdoorn we want to investigate the average relationship between productivity and changes in *all* the variables. We have made a regression analysis were the regressors represent the percentage change in the variables from 1995 to 1996. The dependent variable is the estimated Malmquist index from in the relevant period.

Table 5 **D**rivers for the Malmquist productivity index 1995-96.
% change in the DEA variables as explanatory variables

Variable:	Estimate	St.dev.	t-value	p-value
Shortedu	0.002	0.002	0.96	0.34
Longedu	0.012	0.001	8.54	0.00
R&D	0.349	0.363	0.96	0.34
Faculty staff	-1.362	0.813	-1.67	0.10
Adm. staff	0.099	1.614	0.06	0.95
Net oper. exp.	0.000	0.000	-1.06	0.29
Building (m2)	0.031	0.052	0.60	0.55

Total number of observations: 89, R-squared: 0,449, F-value: 11,247

The results from the estimation process are presented in Table 5. In general one would expect positive signs for output growth, and negative for input growth. But we observe that there are only two variables having a significant effect on the estimated Malmquist index choosing a 10% rejection level. These are *long education* and *faculty staff*. As expected there is a positive correlation between growth in *long education* and the Malmquist index and a negative correlation for *faculty staff*. It is surprising that changes in operating expenditures are not significant, but this illustrates the great variance of this variable. The picture above is also relevant for the period 1994 – 1995, and also looking at the decomposition of the Malmquist index into “catching up”. For “frontier shift” it is interesting to note that there are *no* significant correlations. We therefore conclude that especially changes in the *long education* product and also the *faculty staff* input are the main drivers behind average productivity growth.

Anatomy of productivity change

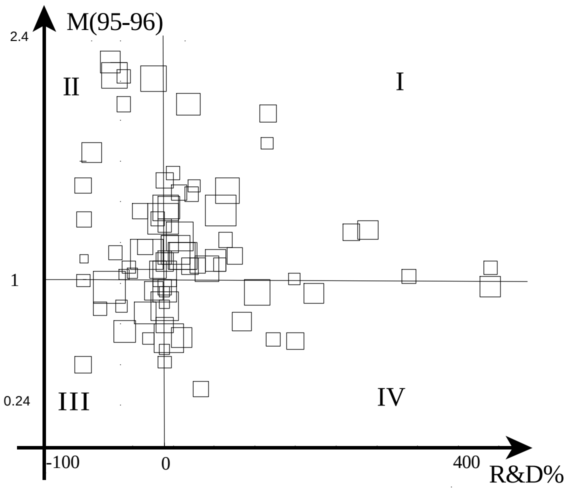
The development over time for each department that lies behind the average relations analysed in Table 5, can be illustrated graphically following the classification in Table 6. In Quadrant I we have departments that obtain both positive productivity growth and positive output growth. These departments have an *efficient expansion* because output is growing faster than inputs. In Quadrant II we have departments that

Table 6 Characterisation of change

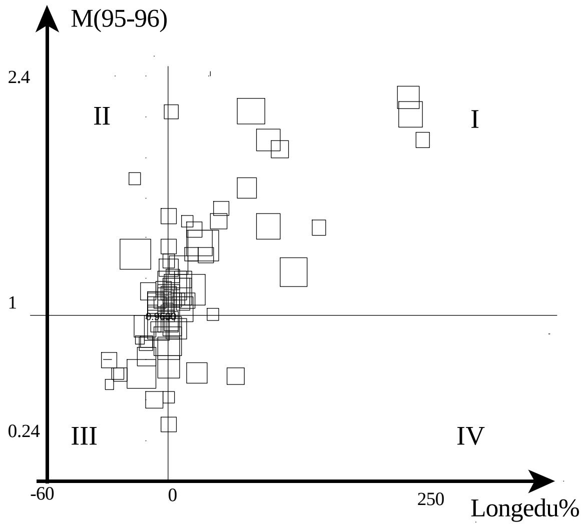
II **Positive adjustment capability** Positive productivity growth Negative output growth ($M > 1$, $\Delta y/y < 0$)	**I** **Efficient expansion** Positive productivity growth Positive output growth ($M > 1$, $\Delta y/y > 0$)
Negative adjustment capability Negative productivity growth Negative output growth ($M < 1$, $\Delta y/y < 0$) **III**	**Inefficient expansion** Negative productivity growth Positive output growth ($M < 1$, $\Delta y/y > 0$) **IV**

combines positive productivity growth with negative output growth. This is only possible if inputs are reduced more than outputs. These departments have *positive adjustment capability*. In Quadrant III we have departments that obtains a decrease in both productivity growth and output growth. These departments also have less reductions in inputs than in outputs, i.e. *negative adjustment capability*, because the reductions in inputs are not sufficient to obtain positive productivity growth. In quadrant IV we have departments that combines negative productivity growth with positive output growth. These have *inefficient expansion* because inputs are increasing more than outputs.

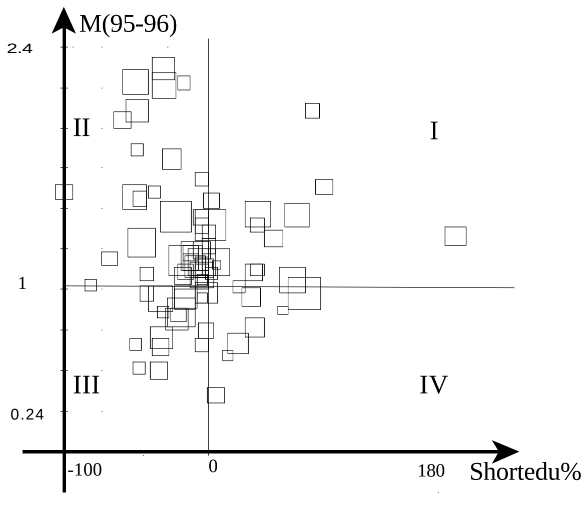
In Figure 7, *Panel a* we have shown the distribution on the four quadrants when productivity is linked with growth in *R&D*. Each square represents one department and the size of the square is proportional to the number of full time student equivalents in 1996. We can see that the departments are distributed on all quadrants. We observe units with both negative, zero and high *R&D* growth, the range is from 100% decline to 400% increase. (Units going from zero to a positive number have been excluded,



Panel a. Research



Panel b. Long education



Panel c. Short education

Figure 7. Productivity- and output growth

and units going from a positive number to zero have been given the figure 100). Some units have remarkable high productivity growth and reduction in *R&D* growth. This may indicate a substitution effect towards more teaching, meaning an increase in the number of grade points. But we should have in mind the possibility of lag effects between faculty input and *R&D*. One cannot expect a stable relationship year by year.

Panel b shows the distribution when we focus on the long education product. We observe a longitudinal pattern; growth in grade points is the main driver behind productivity growth. This is accordance to the average structure revealed in Table 5. There are relative few units in quadrants II and IV. The majority of departments experience an increase in long education, but there are also a number of departments with negative adjustment capability.

In *Panel c* we are comparing the productivity growth with growth in the short education product. We no longer find the longitudinal pattern as in Figure 6, in accordance with the insignificant coefficient in Table 5. Departments are spread over all quadrants. The average growth in short education is about zero, and it is noticeable that many departments show positive adjustment capability. There may be a substitution effect here: the departments with positive adjustment capability have managed to increase long education sufficiently to achieve positive productivity growth.

7. Conclusions and further research

In view of the variables we have had to use in the study and the ideal variables set out in Section 2, it is obvious that the study is far from perfect. However, in order to generate sufficient interest in engaging in the hard work at the institutional level of collecting new types of data we believe the study has been worth while. The proxies used for the three outputs where final exams distributed on two types; short- and long studies, and research publications. The four inputs where number of academic and non-academic staff in full time equivalents, current expenses other than salaries, and building size in square metres.

Typical cross section efficiency results show a large share of efficient departments, with a disproportionate number of efficient departments giving theoretical general education, and a large variation within the group of inefficient units. The difference between professional and arts and science departments may be explained by the nature of the teaching production function, but calculations for a sub-sample of professional departments (e.g. nurses, engineers, teachers) show almost the same variation within this group. The productivity change was mainly positive, with most departments experiencing a positive productivity effect from frontier shift, but a greater variation from positive to negative as regards the contribution from catching up. Positive frontier shift is most important for small units, negative shifts most important for large units, while positive catching-up is most important for medium-sized units, and negative catching-up for both small and large units.

Although some doubt has been voiced as to the legitimacy of the present study representing “true” efficiency, at least the structural differences between departments as to efficiency and productivity warrant further research.

There are several ways of improving upon the analysis:

Stage two- analysis

In order to address the question of why units differ in efficiency a second set of explanatory variables may be introduced (see e.g. Rhodes and Southwick, 1993). The stage two analysis tries to capture other variables that may affect the efficiency scores. In order for the procedure to be statistically sound, the new set of explanatory variables must be uncorrelated with the variables used in the first stage. It is usual to focus on non-discretionary variables outside the control of the units. We have tested the quality of staff by position, using as dependent variable the technical input-saving efficiency score obtained by DEA. It had a (weakly) significant effect on efficiency scores for two of the years. Number of individual studies offered by a department was not significant any year, but here we have a covariation problem with inputs used in the first stage. Other variables could be the location of the college (urban – rural, co-location with other institutions of higher learning), concentration or not of campus (spread out on different locations or in one location).

Separating professional and arts and science departments

It may be legitimate to question the assumption of the same technology for all types of departments. We have experimented with a subgroup of departments giving only professional education, since the lion's share of efficient departments are arts and science, and the underlying technology characterising professional education, like small student groups, need for laboratories, practice outside the college, etc. may well indicate different technologies. It turned out that the difference in efficiency scores and the shape of the distribution was very much alike the one for the total sample for 1996. Further investigations as to teaching technology is warranted. Are small teaching groups necessary, or just tradition, etc.

Quality variables

There is an obvious need for variables capturing quality aspects, as discussed in Section 2. There is also room for improvement of the variables used. The research output can be more elaborately designed by weighting, and research for departments like Music and Media must be introduced. Only written reports have been used in this study.

References

Afriat, S. (1972): "Efficiency estimation of production functions", *International Economic Review* 13(3), 568-598.

Ahn, T., V. Arnold, A. Charnes, and W. W. Cooper (1989): "DEA and ratio efficiency analyses for public institutions of higher learning in Texas", *Research in Governmental and Nonprofit Accounting*, 5, 165-185.

Andersen, P. and N. C. Petersen (1993): "A procedure for ranking efficient units in Data Envelopment Analysis". *Management Science*, 39, 1261-1264.

Banker, R. D., Charnes, A., and W. W. Cooper (1984): "Some models for estimating technical and scale inefficiencies", *Management Science* 39, 1261-1264.

Beasley, J.E. (1990): ”Comparing university departments”, *OMEGA Int. Journal of Management Sci.* 18(2), 171-183.

Berg S. A., F. R. Førsund, and E. Jansen (1992): "Malmquist indices of productivity growth during the deregulation of Norwegian banking, 1980 - 1989", *Scandinavian Journal of Economics*, 94, S211-S228,.

Busch, T., L. Fallan and A. Pettersen (1997): “Diciplinary differences in job satisfaction, self-efficacy, goal commitment, and organizational commitment among faculty employees in Norwegian colleges: an empirical assessment of indicators of performance”, *Rapport i TØH-serien 1997:2*, Avdeling for Økonomisk-Administrativ utdanning, Høgskolen i Sør-Trøndelag.

Caves, D.W., L.R. Christensen and W.E. Diewert (1982): "The economic theory of index numbers and the measurement of input, output and productivity", *Econometrica,* 50, 1393-1414.

Charnes, A., W. W. Cooper and E. Rhodes (1978): "Measuring the efficiency of decision-making units", *European Journal of Operational Research* 2, 429-444.

Doyle, J.R. and A.J. Arthurs (1995): “Judging the quality of research in business schools: the UK case study”, *Omega* 23, 257-270.

Erlandsen, E., F. R. Førsund og K. O. Kalhagen (1998): “Effektivitet og produktivitet i de statlige høyskoler “[Efficiency and productivity in the public colleges], *SNF – rapport 14/98,* Oslo .

Farrell, M. (1957): "The measurement of productive efficiency", *Journal of the Royal Statistical Society,* Series A (General), 120 (III), 253-281 (290)

Flemming, J. (1991): ”The use of assessments of British university teaching, and especially research, for the allocation of resources. A personal view”, *European Economic Review* 35, 612-618.

Frisch, R. (1965): *Theory of production,* Dordrecht: D. Reidel.

Färe, R. and C. A. K. Lovell (1978): "Measuring the technial efficiency of production", *Journal of Economic Theory* 19, 150-162.

Färe, R., S. Grosskopf and C. A. K. Lovell (1985): *The measurement of efficiency of production*, Boston: Kluwer - Nijhoff.

Färe, R., S. Grosskopf and C. A. K. Lovell (1994a): *Production frontiers*, Cambridge: Cambridge University Press.

Färe, R., S. Grosskopf, B. Lindgren and P. Roos (1994b): "Productivity developments in Swedish hospitals: a Malmquist output index approach", in Charnes, A., W. W. Cooper, A.Y. Lewin and L. M. Seiford (eds.), *Data Envelopment Analysis: theory, methodology, and applications*, Boston/Dordrecht/London: Kluwer Academic Publishers, 253-272.

Førsund, F.R. (1996a): "On the calculation of the scale elasticity in DEA models", *Journal of Productivity Analysis*, 7(2/3), 283-302, 1996.

Førsund, F.R. (1996b): "Productivity of Norwegian establishments: a Malmquist index approach", in D. G. Mayes (ed.) *: Sources of productivity growth*, Cambridge: Cambridge University Press, 315-330.

Førsund, F.R. (1997): "The Malmquist productivity index, TFP and scale", *Memorandum no. 233*, Dept. of Economics, School of Economics and Commercial Law, Göteborg University.

Førsund, F. R. and L. Hjalmarsson (1974): "On the measurement of productive efficiency", *Swedish Journal of Economics* 76 (2), 141-154.

Førsund, F. R. and L. Hjalmarsson (1979): "Generalised Farrell measures of efficiency: an application to milk processing in Swedish dairy plants", *Economic Journal* 89, 294-315.

Higgins, J. C. (1989): "Performance measurement in universities", *European Journal of Operational Research* 38, 358-368.

Johnes, G. (1990): "Measures of research output: university departments of economics in the UK, 1983-88", *The Economic Journal*, 100, 556-560.

Johnes, G. (1997): "Costs and industrial structure in contemporary British higher education", *The Economic Journal*, 107, 727-737.

Johnes, J. and G. Johnes (1993): "Measuring the research performance of UK economics departments: an application of Data Envelopment Analysis", *Oxford Economic Papers,* 45, 332-347.

Johnes, J. and G. Johnes (1995): "Research funding and performance in U.K. university departments of economics: a frontier analysis", *Economics of Education Review,* 14, 3, 301-314.

Johnes, J., J. Taylor, and G. Ferguson (1993): ”The employability of new graduates: a study of differences between UK universities”, *Applied Economics* 19, 695-710.

Johnes, J., J. Taylor, and B. Francis (1993): ”The research performance of UK universities: a statistical Aanalysis of the results of the 1989 Research Selectivity Exercise”, *J.R. Statistical Society*, 156, part 2, 271-286.

Kalhagen, K. O. (1998): “En analyse av teknisk effektivitet og produktivitet i høgskolesektoren basert på Data Envelopment Analysis” [An analysis of technical efficiency and productivity based on Data Envelopment Analysis*], SNF-arbeidnotat nr. 38/98*, Oslo.

Malmquist, S. (1953): "Index numbers and indifference surfaces", *Trabajos de Estadistica*, 4, 209-242.

Nishimizu, M. and J.M. Page (1982): "Total factor productivity growth, technological progress and technical efficiency change: dimensions of productivity change in Yugoslavia 1965-78", *Economic Journal* 92, 920-936.

Norsk Samfunnsvitenskaplige Datatjeneste (1997): Statistikk om høgre utdanning 1997. Økonomi, Studenter, Ansatte. NSD-publikasjon av desember 1997.

Rhodes, E.L. and L. Southwick (1993): ”Variations in public and private university efficiency", *Applications of Management Science, Public Policy Applications of Management Science* 7, 145-170.

Sarafoglou, N and Haynes, K.E. (1996): ”University production in Sweden: a demonstration and explanatory analysis for economics and business programs”, *The Annals of Regional Science,* 30, 285-304.

Shephard, R. W. (1953): *Cost and production functions,* Princeton: Princeton University Press.

Sinuany-Stern, Z., A. Mehrez, and A. Barboy (1994): ”Academic departments efficiency via DEA”, *Computers Ops Res*. 21, 543-556.

Tomkins, C. and Green, R. (1988): ”An experiment in the use of Data Envelopement Analysis for evaluating the efficiency of UK university departments of accounting”, *Financial Accountability & Management* 4, 147-164.

Torgersen, A.M., F.R. Førsund, and S.A.C. Kittelsen (1996): "Slack-adjusted efficiency measures and ranking of efficient units”, *Journal of Productivity Analysis,* 7, 379-398