University Efficiency: A Comparison of Results from Stochastic and Non-Stochastic Methods

by

Melville L McMillan
Department of Economics
University of Alberta
Edmonton, Alberta
Canada T6G 2H4
Phone: 780-492-7629
E-mail: melville.mcmillan@ualberta.ca

and

Wing Chan
Department of Economics
Wilfrid Laurier University
Waterloo, Ontario
Canada N2L 3C5
Phone: 519-742-0513
E-mail: wchan@.wlu.ca

Abstract

Efficiency scores are determined for Canadian universities using both Data Envelopment Analysis (DEA) and stochastic frontier (SF) methods for selected specifications. The scores are compared. Although there is some consistency, there is also considerable divergence in the efficiency scores and their rankings. Besides choice of the DEA or SF method, scores are sensitive to the definition of output, the inclusion of environmental (i.e., primarily firm-specific) factors, and the procedures for their inclusion (one-stage and two-stage methods). Despite the divergence among methods and specifications noted, the relative positions of individual universities across sets of several efficiency rankings (e.g., all the DEA and SF outcomes) demonstrate consistency. An analysis of rankings provides a range of potential rankings for each university. High and low efficiency groups are evidenced but the rank for most universities is not significantly different from that of many others. The results indicate that, while efficiency analysis can be helpful, decision makers need to be very cautious when employing efficiency scores for management and policy purposes and they recommend looking for confirmation of the usual efficiency analysis.

Keywords: universities, efficiency, DEA, stochastic frontier,

First Draft: July 2001
Revised: May, 2003

1. Introduction

Rising expectations for public sector performance are exposing publicly funded operations and organizations to greater scrutiny. Universities are no exception and these developments has spurred considerable scholarly investigation. Some researchers have used multi-product cost functions to explore (primarily) economies of scale and scope in institutions of higher education.[1] Of greater interest here, however, is the work of other researchers who have used frontier production methods to investigate the relative efficiencies of (usually) units within universities the results of which may be applied in efforts to enhance the utilization of resources without necessarily changing size or scope (e.g., Abbott and Doucouliagos, 2003, Ahn et al., 1988 and 1989, Ahn and Seiford, 1993, Athanassapoulos and Shale, 1997, Haksever and Muragishi, 1998, Izadi et al., 2003, Johnes and Johnes, 1995, Madden and Savage, 1997, McMillan and Datta, 1998, Sarafoglou and Haynes, 1996, Sinuany-Stern et al., 1994, Thursby, 2000, and Tomkins and Green, 1988). These studies have each used a single methodology to determine efficiency. However, as De Borger and Kerstens (1996) and others note, the efficiency rankings generated by alternative methodologies are not always uniform. If economic analysis is to be used for policy purposes, it is important that the outcomes are reliable. If policy implications from alternative methodologies are consistent, one can have greater confidence when making policy choices. Hence, methodology cross-checking is valuable. In addition, different approaches may afford different insights. Linking those together may enhance the quality of the final analysis (Mayston and Jesson, 1988).

In the pursuit of verification through comparative analysis, the focus of this paper is on the consistency of relative efficiency measures derived from popular alternative methods. Using one year's data from 45 Canadian universities, we compare the efficiency scores from data envelopment analysis (DEA), a non-stochastic linear programming based approach, to those derived from the stochastic method of Aigner et al. (1977) and Meeussen and van den Broeck (1977). Both are frontier methods. That is, both are based on identifying the best practice technology for producing outputs and on comparing individual decision-making unit performance to best practices. The two methods, however, have somewhat different strengths and limitations. Unfortunately, there are no established methods or criteria for choosing between them. Somewhat surprisingly, this work is also a rare instance of stochastic frontier analysis being applied to universities (also see Izadi et al., 2002) because previous studies have relied almost exclusively upon DEA.

The interest in production and cost efficiency has led to the widespread application of frontier methods in efforts to determine production technologies and relative efficiencies. Popular areas of analysis include farms, banks, utilities, and industrial plants.[2] Many studies have examined public or non-profit sector activities including, besides institutions of higher education, hospitals, nursing homes, schools, local governments, libraries and policing among others.[3] Although there are many studies of relative efficiency, few compare the results of applying alternative methods to the same data set.[4] In particular, we are unaware of any other research that compares the outcomes of different methods when applied to universities.

Besides the choice of frontier methodology, efficiency scores may be sensitive to choices among reasonable alternative specifications of the models. Definition and choice of inputs and outputs can vary. In addition, efficiency may be influenced by environmental factors; that is, determinants beyond the inputs and outputs that reflect the conditions under which the production process takes place. Hence, we consider alternative specifications and introduce environmental factors into both the DEA and stochastic analyses. The result is ten sets of efficiency scores. The efficiency rankings of the observations vary among them. Multiple outcomes are not usually presented and so the possibility of inconsistent rankings, which exists regardless of the frontier methodology used, is not often addressed. We extend the analysis to look for and assess consistency across the outcomes and, in an effort to benefit from the information provided by the alternative approaches, we utilize a procedure providing a relatively reliable range for the rank of each university.

To summarize, the paper, in part, provides a comparison of efficiency measures provided by DEA and stochastic methods. It also presents results from alternative but parallel specifications under each frontier approach. Finally, faced with a set of somewhat inconsistent efficiency rankings, we determine the range of potential rankings for each institution. The product offers insight into the choice of efficiency determination methods, alternative specifications, and use of the outcomes which can be of assistance to decision makers.

This paper is arranged in nine parts. Initially, the DEA method is outlined. That is followed by the description of the stochastic method which in this case is applied to a generalized cost function. Because efficiency can be influenced by firm-specific or environmental factors, models incorporating that extension are also employed. The data are discussed next. The data are similar to those employed by others and, as elsewhere, have several recognized limitations. The results of applying the DEA and the stochastic frontier (SF) procedures to the data follow. Those results are then compared and consistency in the efficiency rankings analysed. Conclusions and implications are discussed in the final section. To briefly summarize our overall conclusions, our analyses, not surprisingly, often yield considerable differences in the efficiency scores and the rankings of universities derived from alternative methods and specifications. Yet, when rankings are assessed across the set of outcomes, there is evidence of consistency in universities' relative positions yielding a sense of reliability in the low and high efficiency universities identified. Decision makers are advised to consider the results of efficiency analyses carefully and, especially in small numbers and without further analysis, regard them with caution.

2. Frontier Production Analysis: An Introduction[5]

The basic concepts of frontier productivity can be demonstrated with Figure 1 which illustrates the situation for decision making units, DMUs, (e.g., firms) producing one output using two inputs with a constant returns to scale technology. The axes show the inputs per unit of output. QQ' represents the input/unit output combinations possible when the available technology is efficiently utilized; i.e. the unit isoquant of the efficient producer. A DMU's operation can be represented by a point on or above QQ'. Point D represents a particular producer. D is not on QQ' so DMU D is not an efficient producer because input per unit output exceeds that technically possible. Technical efficiency could be realized by a proportional or radial reduction in inputs along OD to the point T on QQ'. Technical inefficiency is reflected in the distance DT and technical efficiency can be measured as OT/OD.

Figure 1

A DMU at T would be technically efficient but not necessarily economically efficient. Economic efficiency also requires allocative efficiency. Allocative efficiency necessitates using the appropriate combination of inputs on the isoquant QQ' -- that combination being determined by relative input prices. B_1B_1' represents the minimum cost or budget necessary to realize a unit of output for a given set of input prices. Given those prices, the least cost point is at E on QQ' rather than at T. Production at T requires a budget of B_2 which exceeds B_1. Allocative inefficiency is represented by the distance AT and allocative efficiency by OA/OT. Thus, the total inefficiency of D, AD, is a combination of technical inefficiency (DT) and allocative inefficiency (AT). The measure of economic efficiency is OA/OD. Due to the lack of input price variability or due to no or poor input price data, many efficiency studies, as this one, determine only technical efficiency.

Central to frontier productivity analysis is determining the efficient production technology; i.e., the QQ' in Figure 1. Various methods are used. The major distinction is between econometric and mathematical programming methods. Econometric methods are stochastic and parametric while the mathematical programming methods are non-stochastic and non-parametric.[6] The two methods have different relative merits. The econometric approach must estimate a (likely unknown) underlying input-output production relationship which characterizes the data. Choice of functional form and errors in the specification of the production function may result and constrain the estimates. However, the stochastic process allows statistical testing of alternative specifications and reveals information about the significance and roles of the included variables. It also allows for noise in the data. The mathematical programming approach is less restrictive about the character of the production relationship (i.e., it does not impose any functional form) but, instead, takes the bounding observations as defining the best practice efficient frontier. On the other hand, there are no tests in mathematical programming to help determine the appropriate selection or definition of inputs and outputs. Any and all variation between observed units and the frontier is attributed to inefficiency while, in the econometric method, that variation is (but also must be) separated into inefficiency and random error. The econometric approach may estimate the production function itself or, with adequate information on input or output prices, the associated dual cost or profit functions may be estimated. These latter specifications offer options that may be more consistent with the behavioural objectives of the organizations; e.g., cost minimization versus profit maximization. With price information, both the econometric and programming methods can be used to solve for allocative efficiency as well as technical efficiency.[7] Mathematical programming includes the free disposal hull (FDH) model and the more common data envelopment analysis (DEA) model. The efficiency scores from both econometric and programming approaches are often subject to second-stage regression analysis to help determine the impact of environmental factors beyond decision maker control upon efficiency.[8]

Both DEA and SF analysis are popular methods for assessing relative efficiency. Unfortunately, there is no definitive mechanism for selecting between them. The decision is a judgement call. Clearly,

where u_{rk} is the weight assigned each unit of output r from DMU_k and v_{ik} is the weight assigned each unit of input i used by DMU_k. That is, solutions are sought to maximize the ratio of weighted output to weighted input for each DMU (the ratio of virtual output to virtual input). By normalization, the efficiency scores range from zero to one. The same weights (virtual multipliers) that maximize h_k for DMU_k are applied to the inputs and outputs of all DMUs in the solution to the problem for DMU_k. This solution process is repeated for each DMU. Hence, because the weights can vary for each solution, the efficiency scores determined are those most favourable to each DMU.

The above is a fractional programming problem. Using a transformation process, Charnes, Cooper and Rhodes (1978) converted this non-linear problem into a standard linear programming problem. That transformation is indicated by the use of new variables μ and ν for u and v. Now the transformed programming problem becomes

maximize $_{\Upsilon \in \varphi}$

$$\sum_r \mu_{rj} y_{rj} + \mu_* \qquad \textbf{(2)}$$

subject to

$$\sum_r \mu_{rj} y_{rj} - \sum_i v_{ij} x_{ij} + \mu_* \le 0 \quad j = 1 \ldots, k, .. n$$

$$\sum_i v_{ij} x_{ij} = 1$$

$$\mu_r, v_i \ge 0$$

$$\mu_* \gtrless 0$$

for the input oriented, variable returns to scale DEA model.[9] In that this form solves for the multipliers μ and ν, this is the multiplier problem.

The corresponding dual problem solves for the surface which envelops the data. This envelopment problem is

minimize

$$\theta_k \qquad \textbf{(3)}$$

subject to

$$\sum_j y_{rj} \lambda_j \ge y_{rk}$$

$$\theta_k x_{ik} - \sum_j x_{ij} \lambda_j \geq 0$$

$$\sum_j \lambda_j = 1$$

$$\lambda_j \geq 0$$

Any inequality in the first two conditions implies a slack variable. Slacks represent excess inputs given the level of output or the potential to expand output given the inputs. Slacks only appear when DMUs are inefficient. Theta, θ, is the radial (Debreu-Farrell) measure of technical efficiency in that it reports the proportion of inputs needed to produce present output if the DMU operated efficiency. However, additional (now non-proportional) changes to inputs, and possibly even some output change, might further improve efficiency in some instances; that is, if the radially efficient point is on a portion of the isoquant parallelling an axis.[10] Also recognizing these possibilities conforms to the Koopman's concept of efficiency. While the Koopman's concept is more stringent and appealing, it is more difficult to measure (Lovell, 1993). The DEA method used for this study, IDEAS, calculates efficiency measures which proxy the Koopman's approach.

The mathematical programming problems presented above represent the variable returns to scale input oriented version of the DEA model. If constant returns to scale is assumed, the $\Sigma\lambda_j = 1$ constraint disappears from the envelopment problem and the $u_* = 0$ from the multiplier problem. Other scale restrictions can be imposed. DEA models may be input or output oriented. Input models measure efficiency in terms of the potential (proportional) reduction in input use while output models measure efficiency in term of the potential (proportional) output increase. While the efficient and inefficient units do not change, the efficiency scores can differ between the two orientations. Such differences do not arise in the constant returns to scale case. See, for example, Ali and Seiford (1993).[11]

This application of DEA adopts the input oriented variable returns to scale model. The input orientation is selected because we suspect that, at least from a longer term perspective, that outputs (e.g., student numbers) are less DMU choice variables than inputs for our universities so input choices are assumed to predominate. In undertaking previous work (McMillan and Datta, 1998), comparisons of input and output oriented DEA analyses suggested that the results were not sensitive to orientation. Although not unanimous, the input orientation appears to be the predominant choice in other DEA studies of university efficiency. The variable returns to scale model is more flexible and the previous work indicated that this was a legitimate choice.

A cost version of the DEA method is employed here. That is, the cost or expenditure made to produce the outputs replaces the several physical inputs in the analysis. In this case, c_k, the cost of DMU_k, replaces the x_{ik} in equations 1, 2 and 3. Unpriced multiple outputs remain as in the usual production form. Fare and Grosskopf (1985) and Fare, Grosskopf and Lovell (1988) show the correspondence between the primal production approach and the dual cost approach. This cost indirect production correspondence is appropriate for evaluating production efficiency when output is not priced but inputs are, so resource usage is reflected in costs. This approach evaluates output provided for given expenditure.

4. Stochastic Frontier Estimation [12]

The major development upon which most current stochastic estimation of production (and related) frontiers is based on is the work by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977) which advanced the composed error term

$$e_k = u_k + v_k$$

for observation k where v_k is the conventional symmetric random error term which is independently and identically distributed with zero mean, i.e. $N(0,\sigma_v^2)$, and u_k, representing inefficiency, which has a one-sided distribution (commonly half-normal or exponential) such that |u| is $N(0,\sigma_u^2)$. The u and v are unrelated. For the production function,

$$y_k = f(\boldsymbol{x_k} ; \boldsymbol{B}) \exp(v_k - u_k)$$

where y_k is firm output, $\boldsymbol{x_k}$ is the vector of inputs and $\boldsymbol{B}$ is the set of parameters to be estimated. If firm k is efficient, it has $u_k = 0$ and operates on the frontier $\ln y_k = \ln f(\boldsymbol{x_k} ; \boldsymbol{B}) + v_k$. If the firm is inefficient, $u_k > 0$ and the firm produces less than the output possible from the inputs used; i.e. $\ln y_k < \ln f(\boldsymbol{x_k} ; \boldsymbol{B}) + v_k$. The frontier itself is stochastic. The corresponding cost representation, the frontier cost function (see Schmidt and Lovell, 1979), is of the form

$$c_k = c(y_k , \boldsymbol{w_k} ; \boldsymbol{B}) \exp(u_k + v_k)$$

or

$$\ln c_k = \ln c(y_k , \boldsymbol{w_k} ; \boldsymbol{B}) + u_k + v_k$$

where c_k is cost, $\boldsymbol{w_k}$ is the set of input prices facing firm k while the other variables are as before.

Although the cost function above is written in terms of a single output, multiple outputs can be accommodated. Besides the cost and production approaches, profit and distance functions can also be defined and estimated. The model can be expanded to include input budget share equations but, because we work with the single equation cost function, we continue with the discussion of that approach.

Of further interest for estimation is that the Aigner, Lovell and Schmidt (1977) composed error model provides an estimate of average inefficiency, u , but not of the u_i for individual firms which are of interest for policy purposes. Jondrow, Lovell, Materov and Schmidt (1982) provide a means of estimating the u_i ; a piece of information important for policy and particularly for the comparison of efficiency estimates from DEA and stochastic alternatives.

We estimate stochastic frontier cost functions using maximum likelihood.[13] The cost model is selected because it accommodates well multiple outputs. Also, the cost model seems to suit reasonably well the university decision-making environment. That is, with input prices exogenous and university output (or the output expectations of public universities) given, the university is motivated to find input combinations minimizing cost which is the basis of the cost model.

5. Incorporating Environmental Factors

In both data envelopment and stochastic frontier analysis, it is common to explore the determinants of inefficiency in a two-stage process. That is, (in)efficiency is determined by the first-step procedures such as those already discussed and then, in a second step, the inefficiency scores are regressed against variables that would not be part of the production or cost function directly but are believed to affect firm performance.[14] To broaden our comparison of efficiency estimates, those implied after controlling for environmental factors are also examined in this analysis.

Kumbhakar et al. (1991) raise concerns about the two-stage process in the stochastic frontier context. Problems arise because the first stage estimation assumes the inefficiency effects are identically distributed while seeking an explanation of the inefficiency measured in the second stage implies that the errors are not identically distributed.[15] Kumbhakar et al. (1991) resolve this problem by incorporating the estimation of the determinants of inefficiency with estimation of the production frontier.[16] In their composed error term ($e_k = u_k + v_k$), they make the one-sided inefficiency component a function of z and w such that $u_k(z,w) = z\boldsymbol{\delta} + w_k$ where z is a vector of determinants of firm efficiency, $\boldsymbol{\delta}$ is the vector of parameters to be estimated, and w is the random error of the inefficiency component u. As before, v is the random error of the composed error term. To extend and enhance the treatment of the environmental variables, we also estimate this KGM (Kumbhakar, Ghosh and McGuckin) specification of the stochastic cost frontier model.

The two-step process is also not the only method of addressing environmental variables in DEA. Some users include all the variables in a one-step procedure in combination with computational extensions to calculated efficiency scores (e.g., Ruggerio, 1996). McCarty and Yaisawarng (1993) compared the two methods and found that they generated similar results. Wang and Schmidt (2002) compare one-step and two-step procedures for a stochastic model and found the two-step procedure led to biased results and so recommend against that approach.

Unfortunately, as is discussed below, there can be uncertainty as to just what are output measures and what are environmental factors. That situation occurs here. Consequently, results are presented for a narrow and a broad definition of output (with corresponding adjustment to what are environmental variables).

To review, three variations of the stochastic frontier approach are considered in this paper. First, we estimate a relatively simple function parallelling the specification of our DEA model. Second, we consider the consequences of introducing environmental factors using the common second-stage regression. Finally, we add the KGM variation as a preferred mechanism for incorporating the determinants of inefficiency into the specification. Each is done for the narrow and broad output definitions. We compare these SF results to those from the corresponding DEA on the same data within the narrow and broad output definition results. In particular, we compare the efficiency scores from the standard DEA to those from the standard SF, and those from the second-stage regression environment adjustments for both the DEA and SF outcomes, and these latter two to those scores from the KGM variation of the SF model. We consider the various specifications more closely below. Before doing that and looking at the results, we discuss the data.

6. The Data

The data used for our data envelopment and stochastic frontier analyses covers 45 Canadian universities for the academic year 1992-93.[17] The data come primarily from the Canadian Association of Business Officers (CAUBO) and from data collected by the Association of Universities and Colleges of Canada (AUCC). These data were supplemented by Statistics Canada data as required. The 1992-93 data were selected initially because of the relatively larger number of observations available for that year. The data and sources are listed in Table 1 and the means and ranges in Table 2. The data cover the full range of Canadian universities. Three major groups have been identified; primarily undergraduate, comprehensive without medical schools, and comprehensive with medical schools. Utilizing data from the full range of universities posed no difficulties and proved successful in McMillan and Datta (1998) and is continued for this analysis.[18]

Tables 1 and 2

Our data are quite similar to those used in most other studies of university cost and efficiency and, in part, contain the same weaknesses.[19] For example, university output is generally recognized as being teaching, research and service. Aggregated at the university level especially, neither teaching or research is particularly well measured and service is usually entirely unmeasured. To measure teaching, we use the full-time equivalent (fte) enrolment in programs. Among undergraduates, we distinguish between those in science areas and others and, at the graduate level, we have the advantage over many studies of being able to distinguish between master's and doctoral students. As have many others, we reluctantly (especially in regards to the fine arts, humanities and social sciences) resort to the amount of sponsored research funding as the measure of research output. These variables constitute the narrow definition of output. The broad definition is introduced to recognize some potentially important additional characteristics of university output (characteristics that might be considered a refinement of the output measurement). In an effort to recognize potential variation in research (and graduate education) quality, we include data on the relative success of eligible faculty in securing research funding from the national granting councils. In addition, to recognize better the division between science and other research (and possibly graduate education), we add the percentage of faculty in the medical, natural and engineering sciences as a broad definition output variable.

Measuring university inputs poses fewer difficulties than outputs but is not completely free of problems even when the only input is taken to be expenditure. Capital input is a problem and, as here, is often excluded from the models. Total expenditure here is total operating expenditure plus sponsored research expenditure. Excluded are ancillary enterprises, special purpose and trust funds and non-operating plant costs. Unfortunately for the cost function approach, satisfactory input price indices, even for major current inputs (e.g., academic and non-academic staff), are not common. We do engage one price variable; that is, average faculty salary as a proxy for the price of academic staff. Looking ahead, this variable did not prove important.[20] We suspect that, especially if quality adjusted, the prices of most important variable inputs are likely to be quite uniform to universities across the country. However, because price is a conceptually important variable, AVSALARY is retained in the equations estimated.

The same data are used for each of the frontier analyses reported. We turn to the efficiency analyses and those results next. Discussion of the environmental variables is deferred.

7. Efficiency Measures from Alternative Frontier Analyses

7.1 The Data Envelopment Analyses

7.1.a Standard DEA Scores

The results for the narrow and broad output DEA models are designated DEA-n and DEA-b respectively. The outputs in DEA-n are the four full time equivalent student enrolments; UG SCI, UG OTHER, MASTER'S and DOCTORAL; and sponsored research expenditures, RESEARCH$. Model DEA-b adds variables reflecting the faculty's success at obtaining research council grants, %SSHRC and %MRCNSE, in an effort to capture research and graduate education quality and, in addition, the variable % SCIFAC, the share of science faculty, to reflect the type (and perhaps the cost) of research and graduate education. Typically in DEA, variables measuring various physical inputs are also added. Here, however, to parallel to the frontier cost function, the only input variable is the total dollar value of operating expenditures plus sponsored research outlays; TOTALEXP. As seen from the results of the numerous specifications reported in McMillan and Datta (1998), input cost versions of the DEA were quite compatible with those using physical input measures. However, some universities found efficient when using physical measures were not cost efficient.[21]

The efficiency scores for the two models are reported in the first two columns of Table 3.

Efficiency scores range from 0.549 to 1.0 in the case of DEA-n. Eighteen of the 45 universities rate efficient and the average score is 0.914. A mix of efficient and inefficient universities exist in each of the three university classes. The inclusion of the grant success and the share of science faculty variables in DEA-b results in 17 more, 35 in total, scoring efficient and yielding a high average efficiency score of 0.974. The range of the scores narrowed to 0.732 to 1.0. Again, efficient and inefficient universities are found in each class. With the broad definition, the 18 universities deemed efficient under the narrow definition of output remained efficient. Relatively little change in efficiency is observed in the case of 13 universities, including five that go from relatively high scores (0.96 to 0.99) to 1.0. For 14 other universities, however, the scores change considerably; i.e., by more than 0.06. Some universities' scores are sensitive to the introduction of the grant success variables and some to the proportion of science faculty. While the changes in efficiency scores for eight universities are 0.15 or greater, the changes for two stand out. That for PEI increased from 0.55 to 1.0 and that for Lethbridge from 0.68 to 1.0. Lethbridge's score is sensitive to grant success while that for PEI is sensitive to the share of science faculty. One needs to regard the increase in efficiency scores with some caution because additional variables may attract a large weight relative to the previous variables (e.g., the sensitivity issue noted above) and also because additional variables increase the likelihood of efficient DEA scores whether or not the variables are important. These results also illustrate the potential sensitivity of DEA results to variable selection.

Table 3

7.1.b DEA Scores Adjusted for Environmental Factors

As a management tool, DEA focuses on the variables under decision makers' control. Efficiency may, however, be influenced by factors beyond the control of the decision making units' managers. Such factors are widely referred to as environmental variables. One should take into account the impacts of environmental variables in assessing performance. Therefore, analysts often seek to explain efficiency scores using environmental variables and regression analysis. Typically, Tobit regression is used to explain the *in*efficiency scores (i.e., one minus the efficiency score) and we follow that approach.

Before turning to our use of the second-stage regression method, a note on environmental variables is helpful. Unfortunately, the distinction between decision maker controlled and environmental variables is not always distinct. Generally, however, the actual inputs and outputs belong in the DEA while factors explaining the efficiency with which inputs produce outputs belong in the second-stage regression (Lovell, 1993). Eight environmental variables are defined in Table 1. Of those, the only variable that is unquestionably purely exogenous to university decision makers is ENROL200 (total enrolments in universities within 200 km). The others, it could be argued, if not choice variables at least might be influenced by management (especially in the longer term) although the environment of the public university may constrain some more than others. Production method factors like part-time enrolments, program diversity and class sizes are examples. Also, as already mentioned, factors characterizing outputs (e.g., science faculty share, grant success) represent another type of variable. Environmental variables here include purely exogenous variables as well as university-specific (more generally called firm-specific) variables representing production methods and output characteristics. We use the terms environmental variables and firm-specific variables interchangeably to refer to this broad set of factors that may influence university efficiency. We regard these potential regression variables as university-specific control variables.

Utilizing the second-stage regression approach, we adjust the efficiency scores to account for environmental variables. The methodology and selection of the environmental variables noted in Table 1 are discussed in McMillan and Datta and illustrative results from Tobit regressions are reported there. To explain the inefficiency scores from DEA-n, %SSHRC, %MRCNSE and %SCIFAC are treated as environmental (firm-specific) variables as well. In the second-stage Tobit regression, %SSHRC, %SCIFAC and FTE TOTAL had coefficients significant at the five percent level while the coefficient of CLASS SIZE was significant at the 10 percent level. To explain the inefficiency scores from DEA-b, only the eight environmental variables listed in Table 1 appeared in the regression because the three noted were treated as outputs. In that regression, only PART-TIME appeared with a significant coefficient and it was at the 10 percent level. Treating %SSHRC, %MRCNSE and %SCIFAC as outputs substantially reduced the explanatory power of the second stage regression. The results from the two Tobit regressions were used to adjust the efficiency scores from the DEA for differences in the control variables from their mean values. The resulting efficiency scores are reported in the final two columns of Table 3. That is, the reported scores are the predicted efficiency scores if all universities were exposed to the same university-specific conditions.

The environment adjusted efficiency scores differ from those of the standard DEA. Few universities achieve a perfect efficiency score (i.e., 1.0) but the bulk of the scores, particularly under the broad definition, are distributed between 0.949 and 0.999. The average efficiency scores change very little. A main effect of the control variables is a greater dispersion of efficiency scores at the upper end of the distribution. At the lower end, the number of universities do not change that much but the minimum scores increase; from 0.55 to 0.665 for PEI in the case of Adj DEA-n and from 0.732 (Memorial) to 0.846 (Chicoutimi) for Adj DEA-b. The individual scores are typically relatively close to those from the standard DEA but, for some, the change is notable. Twelve universities in the case of Adj DEA-n and eight universities in the case of Adj DEA-b have scores that change by at least 0.05. For Adj DEA-n, four have changes exceeding 0.10 (UNB, Hull and PEI) while that for Lethbridge changes from 0.682 to 0.956. For Adj DEA-b, Laurentian and Victoria see increases exceeding 0.10 while those for Memorial and Hull are about 0.15. If making an evaluation of university performance, the differences appear important enough that, if units are to be assessed on a comparable basis, environmental factors should be taken into account. Note, however, the adjustment for environmental factors has not eliminated differences arising between the narrow and broad definition of output. For five universities, the differences in the efficiency scores between Adj DEA-n and Adj DEA-b exceed 0.15 with that for PEI, which changes from 0.665 to 0.998, being the largest. For a number of universities, whether variables are treated as outputs or environmental variables is critical.

7.2 *The Stochastic Frontier Analyses*

7.2.a The Standard Model

Other than for the fact that it is the frontier cost function that is estimated, the cost function is similar to the conventional higher education multi-product cost functions appearing in the literature since the initial work of Cohn et al. (1989).[22] The primary measured outputs are teaching and research. Undergraduate instruction is captured by the fte enrolments in science and other programs; UG SCI and UG OTHER. Again, graduate instruction is measured by MASTER'S and DOCTORAL. A dummy variable equal to one if there is no PhD program is also included. Research output is indexed, albeit

to a broad definition of output causes the scores of 19 universities to change by at least 0.05 but those for four (Saskatchewan, Memorial, Bishop's and W Laurier) change by at least 0.10 and Memorial's score by 0.183. While efficiency scores improve for a number of universities, they decrease for eight of the 19.

Table 5

7.2.b The Standard Model with Environmental Adjustment

Environmental factors may affect performance. Hence here, as with the DEA analysis, the results of a second-stage regression to explain inefficiency is used to adjust the efficiency scores obtained from the standard stochastic frontier model.[27] The explanatory variables used in the Tobit regressions on the inefficiency scores from the Std SF-n specification are ENROL200, UG/DEG, PART-TIME, CLASS SIZE, H-INDEX, ENROL CHANGE, FUND CHANGE and also %SSHRC, %MRCNSE and %SCIFAC. In the case of the second-stage regression on the scores from Std SF-b, the last three variables listed are, of course, not included. These second-stage regressions perform poorly. The only variable to have a significant coefficient was PART-TIME in the narrow definition case and the squared correlations between predicted and actual inefficiency scores are only 0.339 and 0.084 for the narrow and broad definition models respectively.[28] Regardless, for comparison with the alternative KGM stochastic model, the adjustments implied are made to the efficiency scores. That is, the scores reported are for all units when experiencing the mean value of the environmental characteristics. The resulting scores appear in Table 5 as the 2nd Stage Adjusted scores for the two models; i.e., Adj SF-n and Adj SF-b. The ranges of the efficiency scores are from 0.765 to 0.949 for the narrow definition and 0.859 to 0.967 for the broad. The means, 0.894 and 0.917 respectively, are quite similar. The adjustment for environmental factors reduces the top scores in both cases but raises the bottom scores only in the case of Adj SF-b. For both the narrowly and broadly defined output cases, the environmental adjustment only changes the efficiency scores of three universities (different in each case) by more than 0.05.

7.3 *The KGM Stochastic Frontier Model*

To solve problems with inconsistencies in the underlying assumptions of the two-stage regression approach for accounting for environmental (or firm-specific) factors, Kumbhakar, Ghosh and McGukin (1991) proposed (see section 5) estimating an error term incorporating the environmental variables. That methodology is utilized here to estimate what we refer to as the KGM stochastic frontier model.[29] Estimation is done using the Frontier 4.1 program (Coelli, 1996) which is based on the model of Battese and Coelli (1995) for estimating this specification.

The estimation results are reported in Table 4. Both the narrow and broad definition versions of the model are estimated using the same non-environmental variables as before. These estimation results parallel closely those from the standard stochastic frontier models. The only changes to note are that LUG SCI now appears with a negative (but not significant) coefficient, the coefficient for LMASTER'S is larger and now significant, and, for KGM-n, the coefficient for LAVSALARY is significant (the only occasion in the set of estimates). Again, the pseudo R^2, exceeding 0.99, are high.

Of particular interest here are the results of introducing environmental variables in the error term. First, note that %MRCNSE, %SSHRC and %SCIFAC did not perform very well as environmental variables in KGM-n with none of the three having a coefficient significant at the five percent level. However, introducing these variables, instead, as outputs (i.e., the broad definition), resulted in each having a coefficient significant at the five percent level or better.[30] Because the contributions of the other environmental variables differed from the results in the two-stage model, we experimented with alternatives and finally selected those presented.[31] Three environmental variables are common to the narrowly and broadly defined KGM models; ENROL200, CLASS SIZE and H-INDEX. PART-TIME was included only for the narrow definition case because it was the only university-specific variable with a significant coefficient in that two-stage model, although its significance is only at the 10 percent level here. For KGM-n, the other three variables do not have significant coefficients. For KGM-b, however, CLASS SIZE has a significant coefficient at the five percent level, H-INDEX at 10 percent, but that for ENROL200 is not significant.[32] Because the KGM and standard models are not special cases of each other (i.e., they are non-nested), it is not possible to perform a likelihood ratio test in regards to the addition of the environmental terms in the error.

The KGM model is estimated here using a truncated normal distribution. Note also, efforts to estimate the standard stochastic frontier (with Frontier 4.1) did not converge for either the half-normal or the truncated normal specifications of the error. Recall that the standard stochastic frontier was estimated for an exponential distribution because it out performed the other two. Unfortunately, the Frontier package does not offer the exponential alternative. Hence, the KGM model estimated here with the truncated normal distribution is very likely a second-best alternative. To be noted, however, is that the addition of the KGM error term allowed estimation with the truncated normal which itself implies a superior specification. The gammas, γ, the estimated shares of the total variance of the error terms that are attributed to the inefficiency terms, are 0.31 and 0.63 for the narrow and broad cases respectively. Only the value for the broad case is significantly different from zero.[33]

The efficiency scores estimated by the KGM models are reported in Table 5. The ranges and the means for the narrow and broad models are close with ranges of 0.632 to 0.994 and 0.706 to 0.987 and means of 0.931 and 0.950 respectively.

8. Comparison of Efficiency Outcomes

Having obtained efficiency scores using various reasonable specifications of both the DEA and stochastic frontier methods, it is now possible to compare the results. Comparisons are made using correlations, scatter diagrams and by looking for consistency in the relative positions in the rankings.[34]

8.1 *Correlations*

Correlations between the efficiency scores obtained from the various approaches are provided in Table 6. The correlation coefficients of particular interest (e.g., those between parallel DEA and SF methods) are in bold. Note first the simple correlations in Part A of the Table. Looking initially at the correlations from the results when using the narrow definition of output, that for DEA-n vs Std SF-n, is 0.76 while the correlations for the environmental adjusted scores from Adj DEA-n and Adj SF-n and KGM-n are 0.81 and 0.88 respectively. These indicate relatively good correlations and they are not unlike what some other comparative analyses have reported. When one turns to the correlations when the broad definition of output is used, the correlations are much lower. The highest is that between DEA-b and Std SF-b at 0.16. Those between Adj DEA-b and Adj SF-b and KGM-b are even negative at -0.05 and -0.18. In the case of the broad output definition, there is essentially no correlation among the efficiency scores from parallel DEA and SF approaches.

Table 6

Clearly, the definition of output matters here. The correlations along the diagonal in the lower right hand corner of the matrix show the correlations between each method when only the definition of output differs. The highest correlation, at only 0.57, is between Std SF-n and Std SF-b. The scores from the two KGM estimates are the lowest at 0.13. Because adding variables permits greater distinctions to be made, it is not surprising that the efficiency scores vary between the two output definitions. What is surprising is the extent of the difference.

One can also note from this matrix that adjustments for environmental or firm-specific factors matters. Also, the correlations between the Adj SF and KGM outcomes, especially in the case of the broad definition, indicate that the method of incorporating the environmental adjustment affects the pattern of the efficiency scores that result.

The rank correlations appear in panel B of Table 6. Although the numbers vary somewhat, the broad pattern parallels that for the simple correlations.

8.2 *Scatter Diagrams*

Scatter diagrams lend a visual perspective to correlations. While, for reasons of space, no scatter diagrams are presented here, a few comments about those which plot the efficiency scores by DEA against the scores from the parallel SF method are helpful. If the methods yielded consistent results, the scores would plot a line passing through the (1.0, 1.0) point. In fact, one can see substantial scatter about any potential line if, indeed, any evidence of a line at all. The plots for Adj DEA-n and KGM-n (having a simple correlation of 0.88) are most like what is expected but, even there, it is not difficult to find observations differing by at least 0.1 (and so implying some considerable difference in rank) between the two approaches. The dispersion of the points is greater for the other results using the narrow definition of output. The problem is aggravated when the broad definition of output is employed. For example, when looking at the plots of DEA-b versus Std SF-b (simple correlation of 0.16), few observations are found in the interior as most lie scattered along the 1.0 boundaries. In the case of Adj DEA-n and KGM-n (simple correlation of -0.18, the points essentially form a right angle of points near the 1.0 boundaries with virtually no interior points.

(Figures 2-5 are illustrative and are provided only for the referees' information.)

8.3 *Finding Consistency Across the Results*

At this point, the evidence indicates that the efficiency scores and efficiency rankings for our universities are quite sensitive to the method of estimating efficiency (DEA versus SF), the definition of output used (narrow and broad), and both making an adjustment and the method of adjusting (one- or two-stage) for firm-specific factors. These results should be quite discouraging for those wanting to use efficiency estimates from (especially a single) frontier analysis for management and policy purposes because inappropriate conclusions could be drawn and efforts to enhance efficiency might easily be applied to the wrong institutions. In this section, we demonstrate that the results are not as dismal as they seem. In fact, there is an underlying consistency that may be helpful to policy analysts. The implications are, however, not as precise as many might like nor as precise as much empirical analysis appears to suggest.

8.3.a Kruskal-Wallis Tests

The Kruskal-Wallis non-parametric ANOVA test provides a means of testing statistically whether an individual unit maintains a relative position (i.e., its ordinal rank) in a group across different rankings.[35] That is, it is a test of whether there is some consistency in the rankings. The Kruskal-Wallis test statistic is

$$H = \frac{12}{nk(nk+1)} \left(\sum_{j=1}^{n} \frac{R_j^2}{k} \right) - 3(nk+1)$$

where:

n = the number of universities,

k = the number of rankings across which comparisons are being made, and

R_j = the sum of ranks for university j.[36]

The H statistic has a χ^2 distribution with n-1 degrees of freedom. Rejecting the null hypothesis, that all n universities have the same distribution of ordinal rankings, implies that there is some consistency in the relative positions of universities across alternative rankings.

We apply the Kruskal-Wallis test to seven sets of the ordinal rankings derived from our frontier analyses. Those sets are all ten results from both the DEA and the SF methods, the DEA and the SF outcomes separately, the results from the narrow and broad definitions of outputs separately, and, finally, the rankings for the efficiency adjusted and unadjusted results separately. The resulting H values are reported in Table 7 with the null hypothesis test results. For five of the seven sets of rankings, the null is rejected at the one percent level of significance and a sixth, all those outcomes without efficiency adjustment, is significant at the five, but not at the one, percent level. Only those rankings using the broad definition of output fail to show consistency in the rankings of the universities. Hence, despite the notable variation in the scores and rankings reported above, there is a statistically significant underlying consistency in the relative positions of the universities and this consistency holds across almost all of the methods and procedures utilized.

Table 7

8.3.b Rank of Sum of Ranks

The Kruskal-Wallis tests only reveals that some consistency in ranking exists. Table 8 indicates the consistency in the relative positions maintained by the different universities using the rank of the sum of ranks method. For each alternative set tested, the table reports the rank of the sum of ranks for each university, and the average and the range of those ranks across the seven sets. The universities are ordered from the lowest average rank (of sum of ranks), the most efficient universities, to the highest average, the least efficient universities. While there are occasional abnormalities, one typically observes a striking consistency in the rankings of the sum of ranks across the seven sets. For example, Toronto never ranks less than fifth, Guelph always between 20 and 27, Memorial never above 43. Primarily due to a broad range, Bishop's, Brandon, Chicoutimi, Laurentian, PEI, Sherbrooke, T Riviere and Western represent the more notable anomalies. Such variations should cause reflection, especially because wide differences are often observed in the rankings provided by DEA and SF.[37] The prevailing consistency of the rankings are reflected in the final two columns which show the number of times that the university ranked in the top 15 or the bottom 15. Universities ranked among the most efficient 15 (above Concordia) never have a rank of sum of ranks falling among the bottom 15 and universities among the 15 least efficient (below UNB) rarely appear among the top 15. Hence, the 'averaging' of the collective ranking (via the sum of ranks) for a university across a set of several efficiency score analyses suppresses the variations in the individual ranks that gave the discordant results noted earlier and, instead, reveals a consistency in ranking not previously apparent.

Table 8

While the consistency noted is encouraging, it has limitations. To demonstrate, we test between pairs for significance differences in the mean rankings of the sum of ranks distribution in Table 8.[38] The results appear in Table 9. The columns reporting those universities with a mean rank not significantly different from that of the specific university demonstrate that no university has a rank significantly different from those of every other university. In fact, the mean ranks of many universities are not significantly different from those of many others. Toronto, Winnipeg and Montreal each have mean ranks not significantly different from seven other universities. Eight universities could have a claim to first place. Memorial, ranked 45th, has a mean rank not significantly different from three other universities. That is, four universities could have a claim to last place. Many universities have means not significantly different from many others. For example, 22 universities have means not significantly different from 20 or more other universities. Nine universities' means are never significantly different from a university ranked 23rd or lower. Six have means not significantly different from universities ranked 14th or higher and only eight have means not significantly different from universities ranked 23rd or higher. Quite a number of universities have means not significantly different from those in the seventh or eighth position at the upper end but ranging to the upper thirties or even the low forties positions at the low end. Thus, while there is considerable scope for variation in the rankings, there appears to be a relative small group of highly efficient universities, a relatively small group of low efficiency universities, and a large group of those in between. Thus, even at this relatively aggregated level (ranks of sum of ranks), the individual mean efficiency scores do not reveal a university's ranking with much precision but probably does reveal the range of its relative position on the efficiency scale with some reliability. The extended analysis undertaken here demonstrates the potential ranges in the rankings and also reveals, for this data, groupings of low and high efficiency universities in which there can be some confidence.

Table 9

9. Summary and Conclusion

It is common to report efficiency scores determined from either the non-parametric DEA method or the stochastic frontier (SF) method. For efficiency analysis to be a reliable and valuable management tool, efficiency scores or rankings should be consistent and relatively insensitive to the efficiency determination method. Otherwise, reliance on the outcomes of one or another method may lead to inappropriate policies and decisions. This paper was undertaken in the spirit of methodological cross-checking and to link together, as advocated by Mayston and Jesson (1988), information from alternative models to identify problem areas.

The analyses in the paper focus on the consistency of efficiency measures determined by DEA and SF methods. Efficiency scores were determined for 45 Canadian universities based on 1992-93 data for relatively simple DEA and SF models and for some variations. The variants allowed for a narrow and a broad definition of outputs, made adjustments for environmental/firm-specific factors expected to influence performance, and utilized two approaches for adjusting for environmental factors (i.e., second-stage regression and a single-stage approach incorporated those variables directly into the SF estimation equation). Differences in the outcomes are to be expected. The methods are different, the assumptions are different, the variables included vary somewhat, the way variables interact (i.e., the particular specifications) can differ, etc. Still, DEA and SF analysis are both acceptable methods for determining relative efficiency (there is no definitive procedure for testing one against the other) and the particular variations employed are reasonable options. Hence, the need to compare outcomes and draw upon the results of both.

Comparison of efficiency outcomes from parallel DEA and SF models, one against the other, revealed considerable variation. Specifications using the broad definition of outputs showed little or no consistency between the DEA and SF determined efficiency scores and rankings. When the narrow definition was employed, there was much greater consistency with, for example, correlation coefficients similar to those reported in the few other studies comparing DEA and SF results. However, even there, substantial differences were found for some observations (more than enough to confound decision makers). Besides the method of determining efficiency and output definition, efficiency results were found to be sensitive (as expected) to adjustment for environmental factors and also (not necessarily anticipated) to the means by which that adjustment was introduced. Overall, reliance on efficiency outcomes from any one method could be problematic and, at least for this data, could be misleading and so result in, at least for some DMUs, inappropriate decisions.

From a more comprehensive perspective, however, the results were not as negative as thus far suggested. The analysis of efficiency rankings across alternative sets of efficiency outcomes (using the Kruskal-Wallis test based on the sum of ranks) revealed a significant consistency in universities' relative positions. Analysis of the means of the ranks of the sum of ranks demonstrates that most universities have mean ranks that are not significantly different from the means of many other universities. Despite that, a small number of universities can still be identified as belonging to a high efficiency group and a small number to a low efficiency group. We suspect that these results provide a more realistic assessment of relative efficiency than that suggested by many studies. Hence, efficiency analyses (with data such as ours) may yield information in which decision makers can place some confidence. Reliable information about efficiency ranking is not the product of a single efficiency analysis despite these methods' potentially beguiling sense of precision.[39] This work suggests that many reasonable alternatives, whether following one or more methods, may need to be estimated and their results analyzed.

Our analysis could be extended in various ways. One way is to extend the data to panel data. Besides the benefits of simply more data (such as permitting the estimation of more flexible functional forms of the stochastic models), panel data would allow searching for trends and more options for seeking consistency in the outcomes of individual observations. However, as the inter-year differences in efficiencies that are frequently reported in panel data studies imply, there may also be additional influences to explain. Also interesting would be to compare these results with those from determining confidence intervals for efficiency scores from individual evaluations.

To conclude, if our data are representative, this analysis suggests that, given the state of efficiency modelling, it is advisable to search for some consistency in efficiency outcomes from a variety of acceptable alternative efficiency analyses including both DEA and SF methods. The results of this extension beyond the usual efficiency analysis should cause policy analysts to appreciate better the limitations of frontier methods and the potential for better utilizing them.

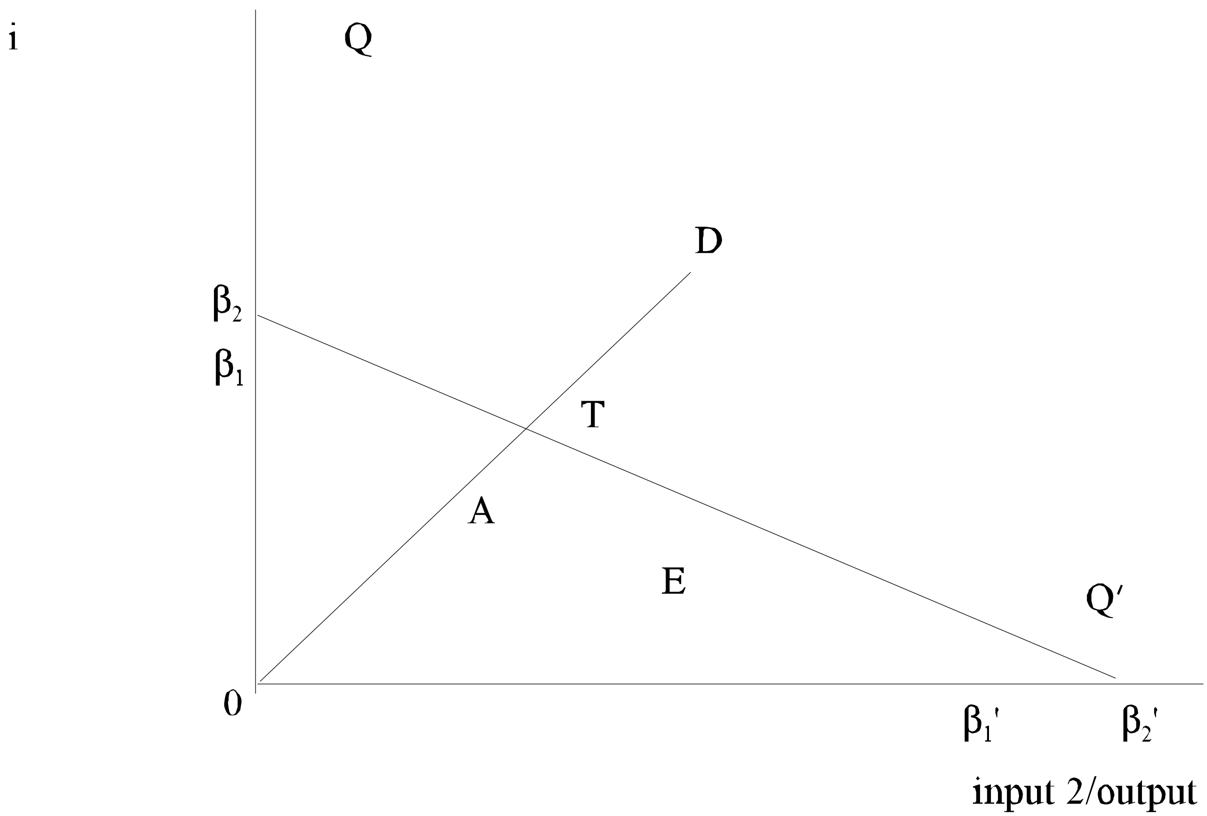
i
Q
D
β_2
β_1
T
A
E
Q′
0
β_1'
β_2'
input 2/output

Table 1. Variables used in the Analyses[a]	
Outputs in DEA / Independent variables in SF	
UG SCI	fte undergraduate enrolment in the sciences - approximated as enrolments in science programs, selected health sciences (medicine, dentistry, optometry, and veterinary sciences) and first-entry (i.e., directly from high school, or CEGEP in Quebec) professional programs (e.g., engineering). (A)[b]
UG OTHER	fte undergraduate enrolment in other than UG SCI programs - e.g., arts and social sciences, visual and performing arts, second-entry (i.e., university prerequisites required) professional programs (e.g., law) and unclassified students. (A)
MASTER'S	fte graduate enrolment in master's level programs (A)
DOCTORAL	fte graduate enrolment in doctoral stream programs (A)
RESEARCH$	total sponsored research expenditures (C)
AVSALARY	Average salary and benefit expenditure for faculty (SF only)
%SSHCC[f]	number of active SSHRC and Canada Council (for fine arts) grants as a percentage of eligible faculty (A)[c, d]
%MRCNSE[f]	number of active MRC and NSERC grants as a percentage of eligible faculty (A)[c,d]
%SCIFAC	proportion of full-time faculty eligible for MRC and/or NSERC grants (A)
DV	dummy variable equal to 1 if no PhD program (stochastic model only)
Input in DEA / Dependent variable for SF	
TOTAL EXP	total operating expenditure and sponsored research expenditure (C)[e]
Environmental Control (DEA and SF as required)	
ENROL200	total student enrolment in universities within 200 kilometers
UG/DEG	undergraduate fte enrolment per undergraduate degree awarded (A)
PART-TIME	part-time student enrolment divided by total student enrolment (A)
CLASS SIZE	the proportion of third and fourth year classes with less than 26 students (A)

H-INDEX	Herfindahl index denoting specialization among programs (smaller value, more diversity)
ENROL CHANGE	percentage change in total enrolment, 1990-91 to 1992-93 (SC)
FUND CHANGE	percentage change in total revenue, 1989-90 to 1992-93 (C)
FTE TOTAL	total full-time equivalent student enrolment; UG plus GRAD (A)

Notes:

a. Data for 1992-93 (see note c).

b. A and C indicate AUCC (Association of Universities and Colleges of Canada) and CAUBO (Canadian Association of University Business Officers) data respectively (or information derived from data there). SC indicates a Statistics Canada source.

c. Data averaged over 1992-93, 1993-94, and 1994-95 or as many years as data were available. Lower response rates for 1993-94 and 1994-95 resulted in no data for some universities in these years in which case the average is over fewer years.

d. SSHRC (Social Science and Humanities Research Council of Canada), MRC (Medical Research Council), and NSERC (Natural Sciences and Engineering Research Council).

e. Excluded from total expenditures are ancillary enterprises, special purpose and trust funds, and non-operating plant costs.

f. These variables; %SSHCC, %MRCNSE and %SCIFAC; are outputs in the broad definition of output but are environmental variables under the narrow definition of output.

Table 2. Descriptive Statistics			
Variable	Mean	Minimum	Maximum
Outputs in DEA / Independent variables in SF			
UG SCI	5471.5	276.3	18564.8
UG OTHER	6235.3	566.7	19170
MASTER’S	1151.4	5	5525
DOCTORAL	451.2	0	2877
RESEARCH[a]	34201.7	389	166735
AVSALARY[a]	86.67	63.2	107.51
%SSHCC[b]	16.2	2.1	45.3
%MRCNSE[b]	79.8	10.2	147.5
%SCIFAC[b]	0.406	0.16	0.694
DV	0.33	0	1
Inputs in DEA / Dependent variable in SF			
TOTALEXP[a]	180204.8	20522	698741
Environmental control (DEA and SF as required)			
ENROL200	68905	0	157450
UG/DEG	4.98	3.35	7.75
PART-TIME	0.311	0.026	0.651
CLASS SIZE	0.683	0.484	0.952
H-INDEX	0.31	0.19	0.57
ENROL CHANGE	8.4	-3.7	36.5
FUND CHANGE	21.6	-13.7	42.6
FTE TOTAL	13309	1981	44016

Note:

a. Thousands of dollars

b. Outputs under broad definition of output. Environmental variables under the narrow definition.

Table 3. Efficiency Scores for Canadian Universities from Data Envelopment Analyses

	DEA[a]		Environment Adjusted DEA[a]	
University	Narrow DEA-n	Broad DEA-b	Narrow Adj DEA-n	Broad Adj DEA-b
Comprehensive with Medical School				
1 Alberta	0.867	1	0.964	0.972
2 UBC	1	1	0.962	0.934
3 Calgary	0.849	1	0.859	0.962
4 Dalhousie	0.729	0.946	0.776	0.981
5 Laval	0.903	0.92	0.991	0.988
6 Manitoba	0.809	0.969	0.842	0.965
7 McGill	0.913	1	0.957	0.999
8 McMaster	1	1	0.97	0.995
9 Montreal	1	1	0.999	0.999
10 Ottawa	1	1	0.981	0.999
11 Queen's	0.968	1	0.948	0.991
12 Sask.	0.777	0.954	0.778	0.985
13 Sherbrooke	0.807	1	0.792	0.991

14 Toronto	1	1	1	1
15 Western	1	1	0.978	1
Average-All	0.908	0.986	0.92	0.984
Avg-Ineff obs	0.847	0.947	0.914	0.982
# Efficient	6	11	1	2
Comprehensive without Medical School				
16 Concordia	0.961	1	0.97	0.991
17 Guelph	0.97	1	0.925	0.999
18 Memorial	0.736	0.732	0.788	0.888
19 UNB	0.958	1	0.833	0.958
20 UQAM	1	1	1	0.995
21 S Fraser	0.974	1	0.992	0.967
22 T Riviere	0.892	0.95	0.891	0.923
23 Victoria	0.871	1	0.94	0.874
24 Waterloo	0.975	1	0.962	0.996
25 Windsor	1	1	0.922	0.999
26 York	1	1	1	0.9999
Average-All	0.94	0.971	0.929	0.963

Avg-Ineff obs	0.917	0.841	0.914	0.963
# Efficient	3	9	2	0
Primarily Undergraduate				
27 Acadia	0.924	1	0.889	0.985
28 Bishop’s	1	1	0.955	1
29 Brandon	1	1	0.959	0.973
30 Brock	1	1	0.969	0.992
31 Chicoutimi	0.848	0.892	0.827	0.846
32 Hull	0.788	0.793	0.914	0.953
33 Lakehead	0.919	1	0.937	0.937
34 Laurentian	0.809	0.818	0.9	0.957
35 Lethbridge	0.682	1	0.956	0.9999
36 Moncton	0.839	0.865	0.87	0.955
37 Mt Allison	1	1	0.946	0.999
38 Mt St Vin	1	1	0.962	0.989
39 PEI	0.549	1	0.665	0.998
40 Rimouski	1	1	0.974	0.984
41 St Mary’s	1	1	0.97	0.979

42 St F Xavier	0.898	1	0.82	0.972
43 Trent	0.935	1	0.993	1
44 W Laurier	1	1	0.986	0.999
45 Winnipeg	1	1	0.989	1
Average-All	0.905	0.967	0.92	0.975
Average-Ineff obs	0.819	0.842	0.92	0.97
# Efficient	9	15	0	3
All Universities				
Average-All	0.914	0.974	0.922	0.975
Average-Ineff obs	0.857	0.884	0.917	0.972
# Efficient	18	35	3	5
Range	0.549 - 1.000	0.732 - 1.000	0.665 - 1.000	0.846 - 1.000

Note: a. Narrow and broad refer to the definition of output. The narrow definition includes UG SCI, UG OTHER, MASTER'S, DOCTORAL and RESEARCH$. The broad definition expands output to include %SSHCC, %MRCNSE and %SCIFAC.

Table 4. Estimation Results for Stochastic Frontier Models

	Standard SF Model[a]				KGM SF Model[b]			
	Narrow Std SF-n		Broad Std SF-b		Narrow KGM SF-n		Broad KGM SF-b	
Variable	Estimated Coefficient	t-value	Estimated Coefficient	t-value	Estimated Coefficient	t-value	Estimated Coefficient	t-value
LUG SCI	0.22245	4.974	0.17709	3.62	-0.55656	-2.075	-0.07004	-1.68
LUG OTHER	0.33148	3.97	0.3149	6.552	0.45494	4.647	0.3619	4.668
LMASTER'S	0.02984	0.779	0.05888	1.9	0.07861	2.027	0.15299	3.891
LDOCTORAL	-0.16472	-2.498	-0.25541	-3.065	-0.13368	-1.209	-0.36634	-4.772
LRESEARCH$	0.23518	4.848	0.21586	5.183	0.2591	5.753	0.24722	6.803
LAVSALARY	0.21262	1.235	-0.0218	-0.134	0.40214	2.073	0.10729	0.701
L%SSHRC	—	—	-0.15376	-3.471	—	—	-0.12631	-3.063
L%MRCNSE	—	—	0.035858	0.499	—	—	0.17006	-2.624
L% SCIFAC	—	—	0.90692	2.823	—	—	1.53316	4.938
MD[c]	0.007163	0.812	-0.13653	-1.32	-0.00739	-0.559	-0.01912	-1.642
D2[c]	0.018909	2.085	0.046094	4.315	0.03513	2.792	0.05706	5.127
DI[c]	—	—	-0.08083	-3.699	—	—	-0.10809	-4.867
I2[c]	—	—	0.21878	1.797	—	—	0.45953	3.797
DV[c]	-0.25835	-1.511	-0.19947	-1.271	-0.05245	-0.219	-0.3489	-2.268

Constant	3.6376	4.974	5.5675	4.963	3.1888	3.066	6.36225	6.561
Variables and parameters in error term								
%SSHRC	—	—	—	—	-1.2054	-1.934	—	—
%MRCNSE	—	—	—	—	-0.07167	-0.39	—	—
%SCIFAC	—	—	—	—	0.32548	0.569	—	—
ENROL200	—	—	—	—	0	-1.474	0	0.885
CLASS SIZE	—	—	—	—	0.19577	0.315	2.34437	2.062
H-INDEX	—	—	—	—	0.11664	0.212	-1.29076	-1.763
PART-TIME	—	—	—	—	-0.6744	-1.894	—	—
Constant	—	—	—	—	-0.05065	-0.572	-1.41079	-1.956
θ	9.2107	3.507	10.319	5.676	—	—	—	—
σ_V	0.04192	1.335	0	0.001	—	—	—	—
γ	—	—	—	—	0.31161	0.284	0.63105	5.564
σ_v^2	—	—	—	—	0.0157	0.973	1.83106	15.133
Log likelihood	40.414	—	63.98	—	34.219	—	45.944	—
Pseudo $R^{2\ d}$	0.984	—	0.988	—	0.994	—	0.996	—

14 Toronto	0.972	0.998	0.937	0.931	0.984	0.975
15 Western	0.955	1	0.879	0.921	0.886	0.975
Average-All	0.907	0.932	0.894	0.922	0.906	0.957
Avg-Ineff obs	0.907	0.915	0.894	0.922	0.906	0.957
# Efficient[b]	0	3	0	0	0	0
Comprehensive without Medical School						
16 Concordia	0.927	0.882	0.934	0.938	0.987	0.977
17 Guelph	0.942	0.866	0.896	0.909	0.969	0.978
18 Memorial	0.635	0.818	0.83	0.931	0.785	0.974
19 UNB	0.927	0.975	0.827	0.903	0.88	0.97
20 UQAM	0.959	0.873	0.945	0.875	0.994	0.978
21 S Fraser	0.946	0.883	0.949	0.937	0.991	0.97
22 T Riviere	0.974	0.948	0.92	0.903	0.988	0.983
23 Victoria	0.913	0.871	0.942	0.937	0.991	0.974
24 Waterloo	0.977	0.995	0.915	0.917	0.978	0.983
25 Windsor	0.934	0.96	0.859	0.925	0.968	0.985
26 York	0.908	0.993	0.949	0.894	0.993	0.983
Average-All	0.913	0.915	0.906	0.915	0.957	0.978

42 St F Xavier	0.904	0.999	0.814	0.893	0.825	0.965
43 Trent	0.984	0.943	0.939	0.93	0.986	0.941
44 W Laurier	0.968	0.827	0.896	0.914	0.983	0.975
45 Winnipeg	0.957	0.992	0.948	0.967	0.992	0.957
Average-All	0.893	0.915	0.887	0.913	0.936	0.928
Average-Ineff obs	0.893	0.899	0.887	0.913	0.936	0.928
# Efficient[b]	0	4	0	0	0	0
All Universities						
Average-All	0.903	0.921	0.894	0.917	0.931	0.95
Average-Ineff obs	0.903	0.908	0.894	0.917	0.931	0.95
# Efficient[a]	0	7	0	0	0	0
Range	0.599 - 0.984	0.596 - 1.000	0.765 - 0.949	0.859 - 0.967	0.632 - 0.994	0.706 - 0.987

Notes:

a. Scores assume mean values of the environmental characteristics.

b. Efficient if score is 1.000.

Table 6. Correlations between Alternative Efficiency Measures

A. Simple Correlations

	Narrow[a]					Broad[a]				
	DEA-n	Adj DEA-n	Std SF-n	Adj SF-n	KGM-n	DEA-b	Adj DEA-b	Std SF-b	Adj SF-b	KGM-b
Narrow										
DEA-n	1									
Adj DEA-n	0.78	1								
Std SF-n	**0.76**[b]	0.53	1							
Adj SF-n	0.7	**0.81**	0.58	1						
KGM-n	0.73	**0.88**	0.62	**0.83**	1					
Broad										
DEA-b	**0.47**	0.35	0.37	0.27	0.2	1				
Adj DEA-b	0.29	**0.33**	0.07	0.13	0.04	0.52	1			
Std SF-b	0.27	-0.11	**0.57**	0.01	-0.06	**0.16**	0.00007	1		
Adj SF-b	0.13	0.13	0.19	**0.33**	0.12	-0.02	**-0.05**	0.29	1	
KGM-b	0.24	0.03	0.32	0.18	**0.13**	-0.11	**-0.18**	0.35	**0.12**	1

Table 6. (continued)

B. Rank Correlations

	Narrow[a]					Broad[a]				
	DEA-n	Adj DEA-n	Std SF-n	Adj SF-n	KGM-n	DEA-b	Adj DEA-b	Std SF-b	Adj SF-b	KGM-b
Narrow										
DEA-n	1									
Adj DEA-n	0.68	1								
Std SF-n	**0.47**[b]	0.48	1							
Adj SF-n	0.61	**0.77**	0.49	1						
KGM-n	0.6	**0.71**	0.53	**0.88**	1					
Broad										
DEA-b	**0.59**	0.73	0.57	0.9	0.97	1				
Adj DEA-b	0.5	**0.38**	0	0.26	0.1	0.09	1			
Std SF-b	0.47	0.54	**0.15**	0.28	0.18	**0.19**	0.37	1		
Adj SF-b	0.14	-0.12	0.43	**-0.09**	-0.14	-0.12	**0.04**	0.07	1	
KGM-b	0.17	0.17	0.22	0.34	**0.21**	0.17	**0.1**	0.01	**0.14**	1

Notes:

a. Narrow and broad definitions of output.

b. Correlations of particular interest (e.g., between parallel DEA and SF alternatives) are in bold.

Table 7. Kruskal-Wallis Test for Consistency in Rankings

Rankings Compared	Number of Rankings	Value of H	Reject Null[a]	
			5%	1%
All results	10	110.29	Yes	Yes
Narrow output	5	134.24	Yes	Yes
Broad output	5	36.221	No	No
All stochastic frontier	6	75.74	Yes	Yes
All DEA	4	87.195	Yes	Yes
All adjusted efficiency	6	76.778	Yes	Yes
All without adjustment	4	61.237	Yes	No

	Combinations(a)									Number of times in	
University	All results (10)	Narrow Output (5)	Broad Output (5)	All SF (6)	All DEA (4)	All adjusted efficiency	All without adjustment (4)	Average ranking	Range	Top 15	Bottom 15
Toronto	1	3	5	3	1	4	4	3.00	1-5	7	0
Winnipeg	2	4	4	2	5	1	6	3.43	1-6	7	0
Montreal	3	1	8	4	3	3	5	3.86	1-8	7	0
MtStVin	4	5	3	1	14	6	1	4.86	1-14	7	0
Ottawa	5	8	1	5	6	8	2	5.00	1-8	7	0
York	6	6	6	8	2	2	9	5.57	2-9	7	0
UQAM	8	2	25	17	4	5	14	10.71	2-25	5	0
Waterloo	7	13	9	6	20	14	8	11.00	6-20	6	0
Trent	10	11	12	9	17	7	17	11.86	7-17	5	0
Western	9	21	2	21	7	22	3	12.14	2-22	4	0
W Lauier	11	9	23	24	9	11	16	14.71	9-24	4	0
UBC	13	15	18	12	19	25	7	15.57	7-25	4	0
Windsor	14	23	10	19	16	19	11	16.00	10-23	3	0
S Fraser	12	7	27	13	21	9	26	16.43	7-27	4	0
St Mary's	15	12	26	23	13	17	15	17.29	12-26	4	0
Concordia	16	17	19	18	22	10	29	18.71	10-29	1	0
Queen's	19	22	14	16	24	18	18	18.71	14-24	1	0
McMaster	18	20	20	27	10	15	23	19.00	10-27	2	0
Brandon	17	10	35	22	18	24	10	19.43	10-35	2	1
Brock	21	16	29	28	11	12	28	20.71	11-29	2	0
Rimouski	20	14	32	29	12	16	24	21.00	12-32	2	1
Bishops	22	26	15	36	8	30	12	21.29	8-36	3	1
Mt Allison	23	18	31	31	15	32	13	23.29	13-32	2	3
Lakehead	26	28	16	14	30	31	19	23.43	14-31	1	1
Guelph	24	24	22	26	23	20	27	23.71	20-27	0	0
Laval	25	19	38	15	32	13	37	25.57	13-38	2	3
Sherbrooke	28	37	7	25	33	34	21	26.43	7-37	1	3
McGill	27	29	24	37	25	21	33	28.00	21-37	0	2
T Riviere	29	25	41	10	38	26	31	28.57	10-41	1	3
UNB	31	33	21	30	29	38	20	28.86	20-38	0	3
Victoria	30	27	39	20	35	23	34	29.71	20-39	0	3
Acadia	32	34	17	35	28	40	22	29.71	17-40	0	4

Laurentian	33	31	30	11	42	28	38	30.43	11-42	1	4
St F Xavier	35	39	13	32	31	41	25	30.86	13-41	1	5
Chicoutimi	34	30	37	7	44	35	30	31.00	7-44	1	4
Alberta	36	32	33	41	27	27	41	33.86	27-41	0	5
PEI	37	45	11	40	36	37	36	34.57	11-45	1	6
Manitoba	38	40	34	33	37	36	39	36.71	33-40	0	7
Lethbridge	39	36	40	44	26	33	42	37.14	26-44	0	6
Calgary	40	41	36	42	34	42	32	38.14	32-42	0	7
Saskatchewan	41	42	28	39	39	39	40	38.29	28-42	0	6
Hull	43	35	44	38	43	29	44	39.43	29-44	0	6
Moncton	42	38	42	34	41	44	35	39.43	34-44	0	7
Dalhousie	44	43	43	45	40	45	43	43.29	40-45	0	7
Memorial	45	44	45	43	45	43	45	44.29	43-45	0	7

Note: (a) Number of rankings in each combination analyzed.

Table 9. Test of Difference in the Means of the Sum of Ranks Distribution[a]

Rank	University	Number Not Different From	Not Significantly Different From Universities Ranked								
1	Toronto	7	2-7	10							
2	Winnipeg	7	1	3-7	10						
3	Montreal	7	1-2	4-7	10						
4	MtStVin	8	1-3	5-7	10-11						
5	Ottawa	8	1-4	6-7	10	14					
6	York	8	1-5	7-8	10						
7	UQAM	17	1-6	8-10	12-14	16-17	22	24	27		
8	Waterloo	15	6-7	9-16	18-22						
9	Trent	14	7-8	10-15	18-23						
10	Western	24	1-9	11-21	26	29	33	35			
11	W Lauier	19	4	8-10	12-19	22	24	26-27	29	33	35
12	UBC	19	7-11	13-23	25-27						
13	Windsor	19	7-12	14-23	26	29	35				
14	S Fraser	24	5	7-13	15-25	27	30	32	34	37	
15	St Mary's	22	8-14	16-22	24-27	29	33-35				
16	Concordia	23	7-8	10-15	17-24	26-27	29-30	32	34-35		
17	Queen's	18	7	10-16	18-24	26-27	29				
18	McMaster	26	8-17	19-27	29-35						
19	Brandon	28	8-18	20-35	37						
20	Brock	26	8-10	12-19	21-27	29-35	37				
21	Rimouski	27	8-10	12-20	22-35	37					
22	Bishops	26	7-9	11-21	23-31	33-35					
23	Mt Allison	25	9	12-14	16-22	24-37					
24	Lakehead	24	7	11	14-23	25-36					
25	Guelph	20	12	14-15	18-24	26-27	29-35	37			
26	Laval	30	10-13	15-25	27-41						
27	Sherbrooke	26	7	11-12	14-26	28-33	35-36	39	42		
28	McGill	15	19	21-24	26-27	29-35	37				
29	T Riviere	29	10-11	13	15-28	30-41					
30	UNB	23	14	16	18-29	31-37	39	42			
31	Victoria	23	18-30	32-41							
32	Acadia	23	14	16	18-21	23-31	33-39	42			
33	Laurentian	28	10-11	15	18-32	34-43					
34	St F Xavier	26	14-16	18-26	28-33	35-40	42-43				
35	Chicoutimi	31	10-11	13	15-16	18-34	36-44				
36	Alberta	17	23-24	26-27	29-35	37-38	40-43				
37	PEI	24	14	19-21	23	25-26	28-36	38-45			
38	Manitoba	14	26	29	31-37	39-43					
39	Lethbridge	15	26-27	29-35	37-38	40-43					
40	Calgary	13	26	29	31	33-39	41-43				
41	Saskatchewan	13	26	29	31	33	35-40	42-44			
42	Hull	15	27	30	32-41	43-45					
43	Moncton	11	33-42	44							
44	Dalhousie	6	35	37	41-43	45					
45	Memorial	3	37	42	44						

Notes: a. At the five percent level of significance.

Endnotes

1. Since Cohn et al. (1989) initially applied a multi-product cost function to institutions of higher education, various scholars have used this method to examine the production technology of universities (e.g., de Groot et al., 1991, Dunbar and Lewis, 1995, Glass and McKillop, 1995, Hashimoto and Cohn, 1997, Koshal and Koshal, 1999, and Nelson and Hevert, 1992).

2. For example, Coelli (1995), Bauer et al. (1993), Diewert and Nakamura (1999) and Forsund and Jansen (1977).

3. For example, Zuckerman et al. (1994), Vitaliano and Toren (1996), Grosskopf et al. (1997), Grossman et al. (1999), Hayes and Chang (1990), Vitaliano (1997), Worthington (1999), and Davis and Hayes (1993). Worthington (2001) surveys frontier studies of schooling and higher education.

4. See Bates (1997) on English local education authorities, Banker et al. (1986) on hospitals, De Borger and Kirstens (1996) and Vanden Eechaut et al. (1993) on Belgium local governments, Ferrier and Lovell (1990) on banks, Forsund (1992) on Norwegian ferries, and Hjalmarsson et al. (1996) on cement plants. Several comparative studies examine agricultural operations; Bravo-Ureta and Rieger (1990), Ekanayake and Jayasuriya (1987), Msafiri et al. (2000), Sharma et al. (1997) and Taylor and Shonkwiler (1986). Some have made comparisons using simulated or artificially generated data (e.g., Banker et al., 1988). Even among these, only five include a comparison of DEA and SF alternatives; De Borger and Kirstens (1996), Ferrier and Lovell (1990), Hjalmarsson et al. (1996), Msafiri et al. (2000) and Sharma et al. (1997). Mayston and Jesson (1988) draw their insights from an analysis of English local education authorities using DEA and ordinary regressions.

5. For reviews of the efficiency analysis and alternative frontier approaches see Coelli (1995), Greene (1999) and Lovell (1993). See Worthington (2001) for a review in the context of education.

6. There is also a deterministic approach which is parametric but non-stochastic.

7. In the absence of prices, programming methods (the cone ratio and the assurance region) have been introduced to derive allocative efficiency measures.

8. The success of both approaches relies on some common factors. These include that all inputs and outputs are homogenous across DMUs, are measurable, measured accurately and included, and that DMUs are relatively alike and employ a common technology.

9. See below for discussion of orientation and scale. The μ value determines returns to scale.

10. The isoquant of a DEA frontier is not the smooth function shown as QQ' in Figure 1 but a boundary made up of linear segments linking best practice observations the "tails" of which are parallel to the axes.

11. For further discussion of the DEA method, beyond Coelli (1995) and Lovell (1993), see Ali and Seiford (1993), Banker et al. (1989), the introductory chapters of Charnes, Cooper, Lewin and Seiford (1994), Coelli (1996b) and Seiford and Thrall (1990).

12. Surveys of the stochastic approach are provided by Bauer (1990) and Greene (1993, 1999).

13. Corrected ordinary least squares (COLS) is a less common alternative estimation method.

14. See Davis and Hayes (1993), De Borger and Kerstens (1996), McMillan and Datta (1998) and Parikh et al. (1995) for examples.

15. Also see Battese and Coelli (1995).

16. Also see Reifschneider and Stevenson (1991)

17. These universities, like almost all Canadian universities, are publicly funded in that they receive most of their operating revenue directly from the provincial governments. For an overview of university education in Canada, see Canadian Education Statistics Council (1996). Unlike some studies of university efficiency, the comparison of private and public funded institutions is not possible here.

18. Abbott and Doucouliagos (2003) also aggregated Australian universities. They reported DEA results for all 36 Australian universities together finding that separating them into distinct categories by various criteria generated similar outcomes. Also, although diverse, Izadi et al. (2002) use all 99 UK universities for their stochastic frontier analysis.

19. See McMillan and Datta (1998) for further discussion of data and alternatives.

20. Explorations with a price variable reflecting cost-of-living differences among cities offered no improvement.

21. This result may occur because the DEA cost models in McMillan and Datta (1998) failed to reflect potentially important cost considerations that %SCIFAC might capture (as suggested by their second-stage regression).

22. See, for example, de Groot et al. (1991), Glass and McKillop (1995), Koshal and Koshal (2000) and Koshal, Koshal and Gupta (2001). Izadi et al.(2002) and Johnes (1996) are the only other cases that we are aware of that have used the stochastic frontier method to estimate a higher education cost function.

23. Discussion of the implications of the cost functions (e.g., marginal costs, scale and scope economies) is deferred to another paper.

24. The model is estimated with LIMDEP 7.0 (Greene, 1996). The models were also estimated for the other common error specifications, the half normal and truncated normal forms. For the narrow output specification, both generated efficiency scores quite similar to those of the exponential model (e.g., correlation coefficients of 0.964 and 0.914 respectively) but for the broad definition, the scores for the half and truncated normal forms were high and relatively uniform and varied more from those of the exponential model (e.g., correlation coefficients of 0.709 and 0.761 respectively). In none of the half and truncated normal cases did the inefficiency term differ significantly from zero. There is no *a priori* basis for selecting among alterative specifications of the error term so, following Schmidt and Lovell (1979), that which performs well and consistently for both the narrow and broad definition cases is selected.

25. The actual to least cost ratio is $C(.)e^{u}/C(.) = e^{u}$ when $v = 0$. When $u = 0.1098$, for example, $e^{u} = 1.116$ which indicates costs on average 11.6 percent above those of efficient units.

26. The predicted values for the individual u are determined from $E(u|\, e) = z + \sigma_v \phi\,(z/\sigma_v) \,/\, \Phi\,(z/\sigma_v)$ where ϕ is the probability density function, Φ is the cumulative density function and $z = e$ -

$\theta\sigma_v^2$. See Greene (1996, ch. 29).

27. Examples of those using second stage regressions to explain stochastic frontier inefficiency scores are De Borger and Kerstens (1996), Davis and Hayes (1993), Parikh et al. (1995) and Vitaliano and Toren (1996).

28. Alternative specifications yielded no better results and had similar impacts on the scores.

29. Coelli et al. (1999) and Grossman et al. (1999) illustrate other applications of this method. Wang and Schmidt (2002) find the one-step method superior to the two-step approach but prefer another specification, that of Reifschneider and Stevenson (1991).

30. Because the two models are not nested, the likelihood ratio test cannot be used to determine which definition dominates.

31. Other environmental variables are available but taking into consideration that most did not have or rarely had significant coefficients, that we wanted to keep the set of these environmental variables included in the two KGM models the same or nearly so, and that, to some extent, the results can be sensitive to the variables selected, these were chosen.

32. ENROL200 is the environmental variable most beyond university management control. It is expected to reflect opportunities for provincial decision makers to realize economies through specialization and the potential for competition among universities.

33. In a production context, Coelli, Perelman and Romano (1999) compare placing environmental variables in the production component of the estimation equation to placing them in the error term as here. The former approach implies that those variables affect the production technology while the latter implies that they do not.

34. Recently, methods have been developed to estimate confidence intervals for efficiency scores obtained from both DEA and SF methods. See Simar and Wilson (1998 and 2000) for techniques related to DEA measures and Horrace and Schmidt (1996) for, and Jensen (2000) for an application of, those suitable for scores from SF methods. While confidence intervals are valuable additions to efficiency analysis, to introduce those here, with ten alternatives under study, would be a considerable exercise and could obscure some other comparisons. Hence, we opt for simpler, but still revealing, methods of comparison.

35. See Witte (1993) for details and Brockett, Golany and Li (1999) for a demonstration.

36. Ties are set at midrank.

37. Also, rankings resulting from use of the broad definition of output are often 'extreme' in that they are at one or the other end of the range.

38. Thursby (2000) applies this approach to the research rankings of economics departments.

39. The few studies providing confidence intervals for efficiency scores suggest that most of that apparent precision is lost in wide confidence intervals.

References

Abbott, M and C Doucouliagos (2003), "The Efficiency of Australian Universities: A Data Envelopment Analysis," Economics of Education Review, 22, 89-97.

Ahn, T, V Arnold, A Charnes and W W Cooper (1989), "DEA and Ratio Efficiency Analysis for Public Institutions of Higher Education," in Research in Governmental and NonProfit Accounting, 5, eds. J L Chan and J M Patton, Greenwich: JAI Press, 165-85.

Ahn, T, A Charnes and W W Cooper (1988), "Some Statistical and DEA Evaluation of Relative Efficiencies of Public and Private Institutions of Higher Education," Socio-Economic Planning Sciences, 6, 259-269.

Ahn, Taesik and Lawrence M Seiford (1993), "Sensitivity of DEA to Models and Variable Sets in a Hypothesis Test Setting: The Efficiency of University Operations," in Creative and Innovative Approaches to the Science of Management, ed. Yuji Ijiri, Westport, CT: Quorum Books, 191-208.

Aigner, D J, C A K Lovell and P J Schmidt (1977), "Formulation and Estimation of Stochastic Frontier Production Models," Journal of Econometrics, 6, #1, July, 21-37.

Ali, A I and L M Seiford (1993), "The Mathematical Programming Approach to Efficiency Analysis," in The Measurement of Productive Efficiency: Techniques and Applications, eds. Harold O Fried, C A Knox Lovell and Shelton S Schmidt, Oxford: Oxford University Press, 120-159.

Athanassapoulos, Antreas D and Estelle Shale (1997), "Assessing the Comparative Efficiency of Higher Education Institutions in the UK by Means of Data Envelopment Analysis," Education Economics, 5, #2, August, 119-134.

Banker, R D, A Charnes, W W Cooper and A Maindiratta (1988), "A Comparison of DEA and Translog Estimates of Production Function Frontiers Using Simulated Observations from a Known Technology," in Applications of Modern Production Theory: Efficiency and Productivity, eds. A Dogramaci and R Fare, Boston: Kluwer Academic Publishers, 33-55.

Banker, R D, A Charnes, W W Cooper, J Swarts and D A Thomas (1989), "An Introduction to Data Envelopment Analysis with Some of Its Models and Their Uses," in Research in Governmental and NonProfit Accounting, 5, eds. J L Chan and J M Patton, Greenwich: JAI Press, 125-163.

Banker, Rajiv D, Robert F Conrad and Robert P Strauss (1986), "A Comparative Application of Data Envelopment Analysis and Translog Methods: An Illustrative study of Hospital Production," Management Science, 32, #1, January, 30-44.

Bates, J M (1997), "Measuring Predetermined Socioeconomic 'Inputs' when Assessing the Efficiency of Education Outputs," Applied Economics, 29, 85-93.

Battese, G and T J Coelli (1995), "A Model for Technical Efficiency Effects in a Stochastic Frontier Production Function for Panel Data," Empirical Economics, 20, 325-332.

Battese, G and G Corra (1977), "Estimation of a Production Frontier Model with Application to the Pastoral Zone of Eastern Australia," Australian Journal of Agricultural Economics, 21, #3, December, 167-179.

Bauer, Paul W (1990), "Recent Developments in the Econometric Estimation of Frontiers," Journal of Econometrics, 46, 39-56.

Bauer, Paul W, Allen N Berger and David B Humphrey (1993), "Efficiency and Productivity Growth in US Banking," in The Measurement of Productive Efficiency: Techniques and Applications, eds. Harold O Fried, C A Knox Lovell and Shelton S Schmidt, (Oxford: Oxford University Press), 386-413.

Baumol, William J, John C Panzar and Robert D Willig (1982), Contestable Markets and the Theory of Industry Structure, (Orlando: Harcourt Brace Jovanovich, Inc).

Bjurek, Hans, Lennart Hjalmarsson and Finn R Forsund (1990), "Deterministic Parametric and Nonparametric Estimation of Efficiency in Service Production: A Comparison," Journal of Econometrics, 46, #1/2, October/November, 213-227.

Bravo-Ureta, B E and L Rieger (1990), "Alternative Production Frontier Methodologies and Dairy Farm Efficiency," Journal of Agricultural Economics, 41, #2, May, 215-226.

Brockett, P L, B Golany and Shan Li (1999), "Analysis of Intertemporal Efficiency Trends Using Rank Statistics with an Application Evaluating the Macro Economics Performance of OECD Nations," Journal of Productivity Analysis, 11, #2, April, 169-182.

Canadian Education Statistics Council (1996), A Statistical Portrait of Education at the University Level in Canada, Statistics Canada publication #81-579-XPB, Statistics Canada, Ottawa and Council of Ministers of Education, Toronto.

Charnes, Abraham, William W Cooper, Arie Y Lewin and Lawrence M Seiford (1994), Data Envelopment Analysis: Theory, Methodology and Applications, Boston: Kluwer Academic Publishers.

Charnes, Abraham, William W Cooper and E Rhodes (1978), "Measuring the Efficiency of Decision-Making Units," European Journal of Operational Research, 2, 429-444.

Coelli, T J (1995), "Recent Developments in Frontier Modelling and Efficiency Measurement," Australian Journal of Agricultural Economics, 39, #3, 219-245.

Coelli, T J (1996a), A Guide to FRONTIER Version 4.1, Centre for Efficiency and Productivity Analysis, Working Paper no. 7/96, University of New England, Armidale, Australia.

Coelli, T J (1996b), A Guide to DEAP Version 2.1: A Data Envelopment (Computer) Program, CEPA Working Paper no. 8/96, University of New England, Armidale, Australia.

Coelli, Tim, Sergio Perelman and Elliot Romano (1999), "Accounting for Environmental Influences in Stochastic Frontier Models: With Applications to International Airlines," Journal of Productivity Analysis, Vol. 11, June, 251-273.

Cohn, Elchanan, Sherrie L W Rhine and Maria C Santos (1989), "Institutions of Higher Education as Multi-Product Firms: Economies of Scale and Scope," Review of

Economics and Statistics, 71, May, 284-290.

Davis, Michael L and Kathy Hayes (1993), "The Demand for Good Government," Review of Economics and Statistics, 65, #1, February, 148-152.

De Borger, Bruno and Kristiaan Kerstens (1996), "Cost Efficiency of Belgian Local Governments: A Comparative Analysis of FDH, DEA and Econometric Approaches," Regional Science and Urban Economics, 26, 145-170.

de Groot, Has, Walter W McMahon, and J Fredericks Volkwein (1991), "The Cost Structure of American Research Universities," Review of Economics and Statistics, 73, August, 424-431.

Deller, S C and E Rudnicki (1992), "Managerial Efficiency in Local Government: Implications of Jurisdictional Consolidation," Public Choice, 74, #2, September, 221-231.

Diewert, W Erwin and Alice O Nakamura (1999), "Benchmarking and the Measurement of Best Practice Efficiency: An Electricity Generation Application," Canadian Journal of Economics, 32, #2, April, 570-588.

Dunbar, Halil and Darrell R Lewis (1995), "Departmental Productivity in American Universities: Economies of Scale and Scope," Economics of Education Review, 14, #2, 119-144.

Ekanayake, S A B and S K Jayasuriya (1987), "Measurement of Firm-Specific Technical Efficiency: A Comparison of Methods," Journal of Agricultural Economics, 38, #1, January, 115-122.

Fare, Rolf and Shawna Grosskopf (1985), "A Nonparametric Cost Approach to Scale Efficiency," Scandinavian Journal of Economics, 87, #4, 594-604.

Fare, R, S Grosskopf and CA K Lovell (1988), "An Indirect Approach to the Evaluation of Producer Performance," Journal of Public Economics, 37, #1, October, 71-89.

Ferrier, Gary D and C A Knox Lovell (1990). "Measuring Cost Efficiency in Banking: Econometric and Linear Programming Evidence," Journal of Econometrics, 46, 229-45.

Forsund, Finn R (1992), "A Comparison of Parametric and Non-Parametric Efficiency Measures: The Case of Norwegian Ferries," Journal of Productivity Analysis, 3, 25-43.

Forsund, Finn R and Eilev S Jansen (1977), "On Estimating Average and Best Practice Homothetic Production Functions," International Economic Review, 18, #2, June, 463-476.

Glass, J C and D G McKillop (1995), "The Achievement of Scale Efficiency in UK Universities: A Multiple-Input Multiple-Output Analysis," Education Economics, 3, #3, December, 249-263.

Greene, William H (1993), "The Econometric Approach to Efficiency Analysis," in The Measurement of Productive Efficiency: Techniques and Applications, eds. Harold O Fried, C A Knox Lovell and Shelton S Schmidt, (Oxford: Oxford University Press), 68- 119.

Greene, William H (1996), LIMDEP Version 7.0 Reference Guide, (New York: New York

University).

Greene, William H (1999), "Frontier Production Functions," in Handbook of Applied Econometric, Vol. II, Microeconomics, eds. M H Pesarsan and P Schmidt, (Oxford: Blackwell), 81-164.

Grosskopf, Shawna, Kathy J Hayes, Lori L Taylor and William L Weber (1997), "Budget-Constrained Frontier Measures of Fiscal Equality and Efficiency in Schooling," Review of Economics and Statistics, 74, #1, February, 116-124.

Grossman, Philip J, Panayiotis Mavros and Robert W Wassmer (1999), "Public Sector Technical Inefficiency in Large U.S. Cities," Journal of Urban Economics, 46, September, 278-299.

Hsksever, Cengiz and Yuri Muragihi (1998), "Measuring Value in MBA Programs," Education Economics, 6, #1, April, 11-25.

Hashimoto, K and E Cohn (1997), "Economies of Scale and Scope in Japanese Private Universities," Education Economics, 5, #2, August, 107-115.

Hayes, Kathy and Semoon Chang (1990), "The Relative Efficiency of City Manager and Mayor-Council Forms of Government," Southern Economic Journal, 57,#1, July, 167-177.

Hjalmarsson, Lennart, Subal C Kumbhakar and Almas Heshmati (1996), "DEA, DFA and SFA: A Comparison," Journal of Productivity Analysis, 7, 303-327.

Horrace, William C and Peter Schmidt (1996), "Confidence Statements for Efficiency Estimates from Stochastic Frontier Models," Journal of Productivity Analysis, 7, 257-282.

Izadi, H, G Johnes, R Oskrochi and R Crouchley (2002), "Stochastic Frontier Estimation of a CES Cost Function: The Case of Higher Education in Britain," Economics of Education Review, 21, 63-71.

Jensen, Uwe (2000), "Is It Efficient to Analyse Efficiency Rankings?," Empirical Economics, 25, 189-208.

Johnes, Geraint (1996), "Multi-Product Cost Functions and the Funding of Tuition in UK Universities," Applied Economics Letters, 3, 557-561.

Johnes, J and G Johnes (1995), "Research Funding and Performance in UK University Departments of Economics: A Frontier Analysis," Economics of Education Review, 14, 301-314.

Jondrow, James, C A Knox Lovell, Ivan S Materov and Peter Schmidt (1982), "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model," Journal of Econometrics, 19, August, 233-238.

Koshal, Rajindar K and Manjulika Koshal (1999), "Economies of Scale and Scope in Higher Education: A Case of Comprehensive Universities," Economics of Education Review

Koshal, Rajindar K, Manjulika Koshal and Ashok Gupta (2001), “Multi-Product Total Cost Function for Higher Education: A Case of Bible Colleges,” Economics of Education Review, 20, 297-303.

Kumbhakar, Subal C, Soumendra Ghosh and J Thomas McGuckin, (1991), “A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in U.S. Dairy Farms,” Journal of Business and Economic Statistics, 9, #3, July, 279-286.

Lovell, C A Knox (1993), “Production Frontiers and Productive Efficiency,” in The Measurement of Productive Efficiency: Techniques and Applications, eds. Harold O Fried, C A Knox Lovell and Shelton S Schmidt, Oxford: Oxford University Press, 3-67.

Madden, Gary and Scott Savage (1997), “Measuring Public Sector Efficiency: A Study of Economics Departments at Australian Universities,” Education Economics, 5, #2, 153- 168.

Mayston, David and David Jesson (1988), “Developing Models of Educational Accountability,” Oxford Review of Education, 14, #3, 321-339.

McCarty, TA and S Yaisawarng (1993), “Technical Efficiency in New Jersey School Districts,” in The Measurement of Productive Efficiency: Techniques and Applications, eds. Harold O Fried, C A Knox and Shelton S Schmidt, Oxford: Oxford University Press.

McMillan, Melville L and Debasish Datta (1998), “The Relative Efficiencies of Canadian Universities: A DEA Perspective,” Canadian Public Policy, 24, #4, 485-511.

Meeusen, W and J van den Broeck (1977), “Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error,” International Economic Review, 18, #2, June, 435-444.

Msafiri, Mbaga, Robert Romain, Bruno Larue and Luc Lebel (2000), “Measures of Technical Efficiency in Dairy Production: Comparison and Decomposition,” paper presented at the Meetings of the Canadian Agricultural Economics Society, Vancouver, June, 2000.

Nelson, Randy and Kathleen T Hevert, “Effect of Class Size on Economies of Scale and Marginal Costs in Higher Education,” Applied Economics, 24, May, 473-482.

Parikh, Ashok, Farman Ali and Mir Kalan Shah (1995), “Measurement of Economic Efficiency in Pakistani Agriculture,” American Journal of Agricultural Economics, 77, August, 675-685.

Reifschneider, David and Rodney Stevenson (1991), “Systematic Departures from the Frontier: A Framework for the Efficiency Analysis of Firm Inefficiency,” International Economic Review, 32, August, 715-723.

Ruggiero, John (1996), “On the Measurement of Technical Efficiency in the Public Sector,” European Journal of Operations Research, 90, 553-565.

Sarafoglou, N and K E Haynes (1996), “University Productivity in Sweden: A Demonstration and Explanatory Analysis for Economics and Business Programs,” Annals of Regional Science, 30, 285-304.

Schmidt, Peter and C A Knox Lovell (1979), “Estimating Technical and Allocative Inefficiency

Relative to Stochastic Production and Cost Frontiers," Journal of Econometrics, 9, February, 343-366.

Seiford, L M and R M Thall (1990), "Recent Developments in DEA: The Mathematical Programming Approach to Frontier Analysis," Journal of Econometrics, 46, 7-38.

Sharma, Khem, Pingsun Leung and Halina M Zaleski (1997), "Productive Efficiency of the Swine Industry in Hawaii: Stochastic Frontier vs. Data Envelopment Analysis," Journal of Productivity Analysis, 8, 447-459.

Simar, Leopold and Paul W Wilson (1998), "Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models," Management Science, 44, January, 49-61.

Simar, Leopold and Paul W Wilson (2000), "Statistical Inference in Nonparametric Frontier Models: The State of the Art," Journal of Productivity Analysis, 13, 49-78.

Sinuany-Stern, Z, A Mehrez and A Barboy (1994), "Academic Departments' Efficiency via DEA," Computers and Operations Research, 21, 543-556.

Taylor, Timothy G and J Scott Shonkwiler (1986), "Alternative Stochastic Specification of the Frontier Production Function in the Analysis of Agricultural Credit Programs and Technical Efficiency," Journal of Development Economics, 21, 149-160.

Thursby, Jerry G (2000), "What Do We Say About Ourselves and What Does It Mean? Yet Another Look at Economics Department Research," Journal of Economic Literature, 38, #2, June, 383-404.

Tomkins, C and R Green (1988), "An Experiment in the Use of Data Envelopment Analysis for Evaluating the Efficiency of UK University Departments of Accounting," Financial Accountability and Management, 4, 147-164.

Vanden Eechaut, Phillippe, Henry Tulkens and Marie-Astrid Jamar (1993), "Cost Efficiency in Belgium Municipalities," in Measurement of Productive Efficiency: Techniques and Applications, eds. Harold O Fried, C A Knox Lovell and Shelton S Schmidt, Oxford: Oxford University Press, 300-334.

Vitaliano, Donald F (1997), "X-inefficiency in the Public Sector: The Case of Libraries," Pubic Finance Review, 25, #6, November, 629-641.

Vitaliano, Donald F and Mark Toren (1996), "Hospital Cost and Efficiency in a Regime of Stringent Regulation," Eastern Economic Journal, 22, #2, Spring, 161- 175.

Witte, Robert S (1993), Statistics, fourth ed., New York: Holt, Rinehart and Winston.

Worthington, Andrew (2001), "An Empirical Survey of Frontier Efficiency Measurement Techniques in Education," Education Economics, 9, 3, 245-268.

Worthington, Andrew (1999), "Performance Indicators and Efficiency Measurement in Public Libraries," Australian Economic Review, 32, March, 31-42.

Zuckerman, Stephen, Jack Hadley and Lisa Iezzoni (1994), "Measuring Hospital Efficiency with Frontier Cost Functions," Journal of Health Economics, 13, 255-280.