Citation for published version

Sharp, John A. and Meng, Wei and Liu, Wenbin (2005) A Modified Slacks Based Measure Model for Data Environment Analysis with 'Natural' Negative Outputs and Inputs. Working paper. Kent Business School, University of Kent, Canterbury

DOI

Link to record in KAR

https://kar.kent.ac.uk/8504/

Document Version

UNSPECIFIED

Kent
Business
School

Working Paper Series

A Modified Slacks Based Measure Model for Data Envelopment Analysis with 'Natural' Negative Outputs and Inputs

J Sharp, W Meng, S Liu
Kent Business School

KBS
UNIVERSITY OF KENT

Working Paper No. 84
May 2005

A Modified Slacks Based Measure Model for Data Envelopment Analysis with "Natural" Negative Outputs and Inputs

J.A. Sharp, W.Meng and W.B. Liu
Kent Business School, University of Kent

Abstract

This paper is primarily concerned with Data Envelopment Analysis (DEA) of systems where negative outputs and negative inputs arise naturally. Examples of situations in which both negative inputs and negative outputs occur are given. More attention has been paid, in the literature, to the former type of problem. Most available DEA software does not solve this type of problem or copes with negative outputs and possibly negative inputs by assigning zero weights to them. A Modified Slacks Based Measure (MSBM) model is presented, in which both negative outputs and negative inputs occur. The MSBM model overcomes the lack of translation invariance in the Slacks-Based Measure model proposed by Tone (2001) by drawing on the *directional distance function approach* of Silva Portela et al (2004). The MSBM model takes into account individual input and output slacks, which provides more precise evaluation of inefficient DMUs. It therefore, generally leads to lower efficiencies for inefficient DMUs than the Range Directional Model proposed by Silva Portela et al (2004).

Keywords: Data Envelopment Analysis, Negative Inputs, Negative Output, Modified Slacks-Based Measure

Introduction

Discussions and applications of DEA have, in the main, assumed that all the inputs and outputs of DMUs are positive. There has, however, been significant consideration of the issue of *externalities,* which are likely to be particularly important in the public systems. These have been an important focus of DEA applications, e.g. Yu (2004), Korhonen and Luptacik (2004). Externalities are, of course, merely undesirable outputs of the system, i.e. the more we have of them the less attractive they are. This property is most easily reflected by assigning them a negative value.

Although negative outputs are a more obvious candidate for inclusion in DEA models there are situations, e.g. in disposing of pollutants or repairing defective items, where it is desirable to use as much as possible of the input. Such inputs can be reflected by assigning them negative values.

The focus of this paper is on DEA models for systems with at least one positive and one or more negative inputs as well as at least one positive and one or more negative outputs. Much DEA software does not permit negative outputs or inputs to be used directly, partly because no feasible solutions may exist in such cases. Although some software, for example, the DEA-Solver package (Cooper et al, 2000) does permit the use of negative outputs and inputs with some models, in most cases the weights corresponding to negative inputs and outputs are zero, which rather undermines the introduction of the negative outputs and inputs.

Systems with negative inputs and both positive and negative outputs have received rather less attention in the literature than systems with all positive inputs but both positive and negative outputs, though there are significant exceptions, such as Pastor (1996) and Seiford and Zhu (2002). Liu et al (2005) locate DEA models in a very general goal programming based framework, that is shown to apply to such systems. A major contribution to the computation of efficiency for systems with negative input and negative output data is that of Silva Portela et al (2004). They introduce a *Range Directional Model (RDM)* based on what we will henceforth refer to as the *SP Range (P)*. This is defined for the output variable y_r as:

$$P_{r0}^{+} = \max_{j}(y_{rj}) - y_{r0} \qquad r = 1,...,s \qquad (1)$$

and for the input variable x_i as:

$$P_{i0}^{-} = x_{i0} - \min_{j}(x_{ij}) \qquad i = 1,...,m \qquad (2)$$

The SP Range effectively measures the distance of the reference DMU from an “ideal” point represented by the maximum of any particular output over the data set or the minimum of any particular input over the data set. The SP Range has the important property that it is an upper bound on the slacks associated with the variables concerned.

The Range Directional Model of Silva Portela er al (2004) can be formulated as:

$$\begin{array}{ll} \text{Max} & \beta \\ \text{subject to :} & Y\lambda \geq y_0 + \beta P_0^+ \\ & X? \leq x_0 - \beta P_0^- \\ & e\lambda = 1 \\ & \lambda, \beta, P_0^+, P_0^- \geq 0 \end{array} \qquad (3)$$

where P_0^+ and P_0^- denote the vectors of output *SP Ranges* and input *SP Ranges* for DMU$_0$ respectively.

Silva Portela et al show that their model is both units invariant and translation invariant and that 1–ß can be considered a measure of efficiency. They note, however, that ß does not normally encapsulate all sources of inefficiency since at its optimum value some input and output variables may well have non-zero slacks. This suggests that a tighter measure of efficiency might be generated by means of the Slacks Based Measure approach of Tone (2001): a question that is explored more fully below.

Examples of Systems with Negative Outputs and Negative Inputs

Negative outputs and negative inputs can each be subdivided into two types: *Naturally Negative* and *Avoidably Negative. Naturally Negative Outputs/Inputs* are variables measured on a ratio scale, i.e. which have a natural zero. Physical quantities are typical of this type of negative variable but Silva Portela et al (2004) show how such variables may also arise through the use of growth rate variables.

Avoidably Negative Outputs/Inputs are those that are measured only on an interval or ordinal scale without a natural zero, e.g. canonical variates (Ueda and Hoshiai, 1997).

For *Avoidably Negative Outputs/Inputs* both units invariance and translation invariance are highly desirable (Seiford and Zhu, 2001). For *Naturally Negative Outputs/Inputs* only units invariance is needed since, given a natural zero, translation invariance is not relevant. Four examples of this type of system are given below.

Example 1 Plant for Pollutant Disposal (e.g. sewage sludge)

Positive Inputs	Labor Costs, Material Costs
Negative Inputs	Sewage Sludge
Positive Outputs	Compost produced, Electricity Produced
Negative Outputs	Heavy Metal Pollutants
Notional DMU	Week

Example 2 Pan European Logistics Operation

Positive Inputs	Personnel Costs, Vehicle Cost Costs, Own country delivery request
Negative Inputs	Consignments for delivery from each other European Country
Positive Outputs	Items Delivered (Ton Km)
Negative Outputs	Number of Incorrect Deliveries, Number of Complaints
Notional DMU	Country

Example 3 Evaluating Different Configurations of a Business System

Positive Inputs	Desirable Orders
Negative Inputs	Undesirable Orders
Positive Outputs	Desirable Orders Successfully Delivered, Undesirable Orders rejected
Negative Outputs	Number of Incorrect Deliveries, Number of Complaints, Number of Undesirable Orders Delivered
Notional DMU	Configuration = A unique combination of decisions with regard to sourcing of component business processes

Example 4 Evaluating Efficacy of Different Advertising Campaigns

Positive Inputs	Campaign Cost
Negative Inputs	Percentage of consumers with negative brand perception before campaign
Positive Outputs	Percentage of consumers with positive brand perception after campaign
Negative Outputs	Percentage of consumers with negative brand perception after campaign
Notional DMU	Campaign

A Modified Slacks Based Measure (MSBM) to solve DEA problems in which both negative inputs and negative outputs are allowed

This section presents the formulation of a model based on a generalisation of the Slacks Based Measure (Tone, 2001) in which both negative outputs and negative inputs occur. We assume that there are:

a) At least one positive output;

b) At least one positive input;

c) A number of negative inputs and outputs.

The starting point is Tone's (2001) elaboration of the Additive Model (Charnes et al, 1985). It is well known that this model applies the **Primal Form,** as redefined by Charnes et al (1994, p26), leads to a more straightforward interpretation of the notion of efficiency where there are negative outputs and negative inputs (Liu et al 2005), so we confine discussion to that. We choose the VRS form of the model since this not only guarantees translation invariance but also leads naturally to an efficiency measure in our MSBM model in the range [0, 1].

Tone's (2001) model for a system with m positive inputs and s positive outputs is:

$$\min \quad \rho = \frac{1-\frac{1}{m}\sum_{i=1}^{m} s_i^- / x_{i0}}{1+\frac{1}{s}\sum_{r=1}^{s} s_r^+ / y_{r0}} \qquad (4)$$

$$\begin{aligned} \text{subject to : } & Y\lambda - s^+ = y_0 \\ & X\lambda + s^- = x_0 \\ & e\lambda = 1 \\ & \lambda, s^+, s^- \geq 0 \end{aligned} \qquad (5)$$

Minor adjustments need to be made to (4) where either the inputs or the outputs are zero. The model is clearly units invariant but not translation invariant.

Amongst the attractions of the Slacks-Based Measure are (Tone, 2001):

a) It obviates the problems associated with zero weights being assigned to inputs or outputs;

b) The efficiency lies between 0 and 1. It attains the value 1 only if there are no inefficiencies of any type associated with DMUs, i.e. it accounts for sources of inefficiency other than technical inefficiency.

Property b) depends crucially on the fact that the numerator of (4) can be shown to lie between 0 and 1 and the denominator is greater than or equal to 1 (Tone. 2001). In both cases, they attain the value 1 only if all the slacks are zero.

The generalisation of the measure (4) to the case of both negative inputs and negative outputs is unfortunately not straightforward. Whereas in the case of positive inputs we have $s_i^- \le x_{i0}$, as $X > 0, \lambda \ge 0$, this is not necessarily the case for negative inputs and there is, therefore, the possibility that the measure (4) can become negative. Recalling the property of the Silva Portela et al RDM that the slacks for both inputs and outputs are less than or equal to the SP Range suggests, however, that it might be fruitful to modify (4) by using as divisors the relevant SP ranges instead of the reference DMU output and input values. However, we also take the opportunity to generalise (4) somewhat at the same time. We therefore consider the Modified Slacks Based Measure (MSBM) model:

$$
\begin{array}{ll}
\min & \rho = \dfrac{1-\sum_{i=1}^{m} w_i s_i^- / P_{i0}^-}{1+\sum_{r=1}^{s} v_r s_r^+ / P_{r0}^+} \\
\text{subject to :} & \sum_{r=1}^{s} y_{rj}\lambda_j - s^+ = y_{r0}, \quad r = 1,\dots,s \\
& \sum_{i=1}^{m} x_{ij}\lambda_j + s^- = x_{i0}, \quad i = 1,\dots,m \\
& \sum_{j=1}^{n} \lambda_j = 1, \quad j = 1,\dots,n \\
& \sum_{i=1}^{m} w_i = 1, \ \sum_{r=1}^{s} v_r = 1 \\
& v_r \ w_{\mathrm{i}}, \lambda_j, s^+, s^- \ge 0
\end{array}
\qquad (6)
$$

Where P_{i0}^- or P_{r0}^+ are zero we assume that the corresponding term is dropped from the numerator/denominator respectively.

Let us show that the above measure is in the range [0, 1]. Note

$$
\begin{aligned}
s^- &= x_0 - X\lambda \le \max_{\lambda}(x_0 - X\lambda) \\
&= x_0 + \max_{\lambda}(-X\lambda) = x_0 - \min_{\lambda} X\lambda
\end{aligned}
$$

as $\min_{\lambda} X\lambda \ge Z$, where $Z_i = \min_{j} (x_{ij})$. Since $P_0^- = x_0 - Z$, therefore $s^- \le P_0^-$. Also

$$s^+ = Y\lambda - y_0 \le \max_{\lambda}(Y\lambda - y_0) = \max_{\lambda}(X\lambda) - y_0$$

as $\max Y\lambda \le M$, where $M_r = \max_{j}(y_{rj})$. Note $P_0^+ = \max(Y) - M$, therefore $s^+ \le P_0^+$.

Therefore, the efficiency measure in Model (6) is in the range [0, 1].

Since the Additive model is translation invariant, the slacks in Model (6) are translation invariant. In addition the SP Range is translation invariant. It follows, then, that the MSBM model is translation invariant. It is easily verified that like the Tone (2001) Slacks Based Measure it is also units invariant.

Following Tone (2001) the fractional programming problem represented by Model (6) can be transformed to the following problem.

$$
\begin{aligned}
\min \quad & \tau = t - \sum_{i=1}^{m} w_i s_i^- / P_{i0}^- \\
\text{subject to :} \quad & Y\Lambda - S^+ = t y_0 \\
& X\Lambda + S^- = t x_0 \\
& \sum_{r=1}^{s} v_r s_r^+ / P_{r0}^+ + t = 1 \qquad\qquad (7) \\
& \sum_{i=1}^{m} w_i = 1, \quad \sum_{r=1}^{s} v_r = 1 \\
& t, \Lambda, S^+, S^- \ge 0
\end{aligned}
$$

Let an optimal solution to the Model (6) be $(\tau^*, t^*, \Lambda^*, S^{-*}, S^{+*})$. Then we have an optimal solution of model (6) as:

$$\rho^* = \tau^*,\ \lambda^* = \Lambda^* / t^*,\ s^{-*} = S^{-*} / t^*,\ s^{+*} = S^{+*} / t^*$$

Relationship of MSBM model to the Silva Portela et al efficiency measure

The Silva Portela et al efficiency measure $1-\beta^*$ appears intuitively likely to be greater than ρ^* because in the SP model both inputs and outputs are extended by the same proportion *ß*. This can be shown using similar arguments to those of Tone (2001). Let

$$Y\lambda^* - \beta P_0^+ - t^{+*} = y_0, \qquad X\lambda^* - \beta P_0^- - t^{-*} = x_0, \qquad e\lambda^* = 1$$

Where t^{+*} and t^{-*} are the slack values (= 0), being an optimal solution for the Silva Portela et al model.

If we set

$$s^+ = \beta^* P_0^+ + t^{+*}, \qquad s^- = \beta^* P_0^- + t^{-*}$$

The corresponding value of *?* given by:

$$\begin{aligned}\rho &= \frac{1-\sum_{i=1}^{m} w_i s_i^- / P_{i0}^-}{1+\sum_{r=1}^{s} v_r s_r^+ / P_{r0}^+} \\ &= \frac{1-\sum_{i=1}^{m} w_i(\beta^* P_{i0}^- + t^{-*}) / P_{i0}^-}{1+\sum_{r=1}^{s} v_r(\beta^* P_{r0}^+ + t^{+*}) / P_{r0}^+} \qquad (8) \\ &\le 1-\sum_{i=1}^{m} w_i \beta^* \\ &\le 1-\beta^*\end{aligned}$$

An Example

To illustrate the ideas above, consider the notional effluent processing system of Table 1, which has one positive input (cost), one negative input (effluent), one positive output (saleable output) and two negative outputs (methane and CO_2).

To facilitate comparison with the Silva Portela et al model, which weights all inputs equally and all outputs equally, the slack weights in the MSBM model are similarly set equal as in Table 2. Table 3 gives the ? weights, the slacks, and the values of *?*. For comparison purposes the values of *ß* and (*1-ß*) from the Silva Portela et al RDM solution are also shown.

Table 1

DMU	(I$_1$) Cost	(I$_2$) Effluent	(O$_1$) Saleable Output	(O$_2$) CO$_2$	(O$_3$) Methane
1	1.03	-0.05	0.56	-0.09	-0.44
2	1.75	-0.17	0.74	-0.24	-0.31
3	1.44	-0.56	1.37	-0.35	-0.21
4	10.80	-0.22	5.61	-0.98	-3.79
5	1.30	-0.07	0.49	-1.08	-0.34
6	1.98	-0.10	1.61	-0.44	-0.34
7	0.97	-0.17	0.82	-0.08	-0.43
8	9.82	-2.32	5.61	-1.42	-1.94
9	1.59	0.00	0.52	0.00	-0.37
10	5.96	-0.15	2.14	-0.52	-0.18
11	1.29	-0.11	0.57	0.00	-0.24
12	2.38	-0.25	0.57	-0.67	-0.43
13	10.30	-0.16	9.56	-0.58	0.00

Table 2

indicators	w_1 Cost	w_2 Effluent	v_1 Saleable	v_2 Methane	v_3 CO$_2$
weights	0.5	0.5	0.333	0.333	0.333

Table 3 Results for MSBM model and RDM model

Ref DMU	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
$?_1$													
$?_2$													
$?_3$		0.093	1	0.223		0.971				0.857		0.205	
$?_4$													
$?_5$													
$?_6$													
$?_7$	1	0.383			0.526		1					0.795	
$?_8$				0.537				1					
$?_9$													
$?_{10}$													
$?_{11}$		0.524			0.474				1		1		
$?_{12}$													
$?_{13}$				0.24		0.029				0.143			1
Score t	0.449	0.509	1.000	0.467	0.537	0.635	1	1	0.656	0.536	1	0.406	1
t	0.949	0.803	1	0.808	0.753	0.836	1	1	0.894	0.904	1	0.761	1
s_1	0.06	0.569		2.735	0.178	0.280			0.3	3.254		1.314	
s_2	0.12	0.005		1.190	0.716	0.448			0.11	0.353		0	
s_3	0.26	0		0	0.212	0			0.05	0.4		0.363	
s_4	0.01	0.177		0	1.038	0.083			0	0.137		0.535	
s_5	0.05	0		2.700	0	0.136			0.13	0		0.045	
ß	0.036	0.096	0	0.530	0.082	0.037	0	0	0.006	0.393	0	0.217	0
Efficiency 1-ß	0.964	0.904	1	0.470	0.918	0.963	1	1	0.994	0.607	1	0.783	1

As can be seen from Table 3, the MSBM model finds only 5 DMUs efficient (units 3, 7, 8, 11, 13). All the efficient units are same as those generated by the RDM. All the MSBM model efficiency scores of inefficient units are less than or equal to those generated by the RDM, as Figure 1 shows.

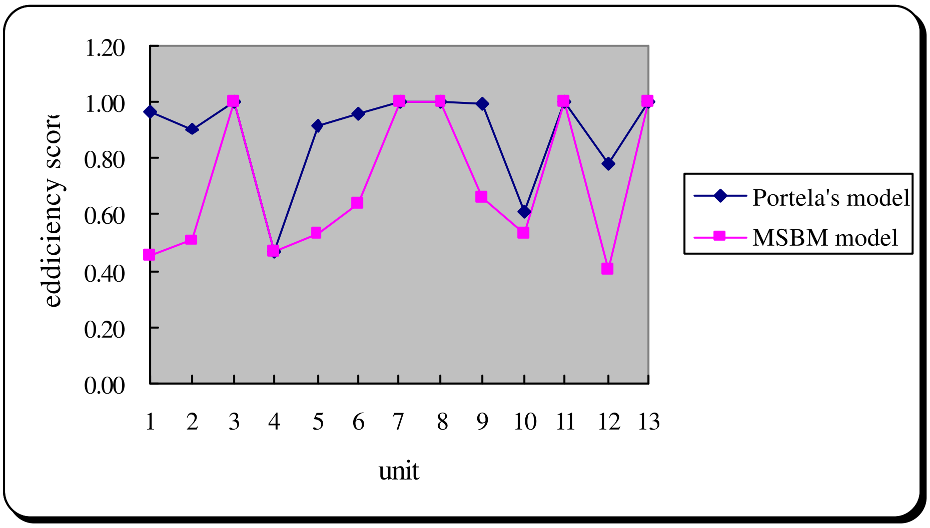

Fig 1 Comparison of efficiency Score from Silva Portela's model and MSBM model

As shown above, the MSBM efficiency score cannot be greater than the efficiency score generated by the Silva Portela RDM. For inefficient units the MSBM model efficiency is generally significantly below that of the Silva Portela et al model. Thus, efficiency scores vary from 0.406 to 1 under the MSBM model, whereas they vary from 0.47 to 1 under the Silva Portela model. Since the MSBM takes account of the individual variations in inputs and outputs rather than attempting a uniform shrinking across all inputs and outputs, the MSBM efficiency does not necessarily bear any direct relationship to that of the Silva Portela et al model. Thus under the MSBM model DMU 12 has the lowest efficiency, whereas for the Silva Portela et al model DMU 4 is the least efficient.

The MSBM model with unequal weights

The MSBM model allows us to alter the slack weights, as discussed above. Suppose, for instance, that in the example problem just discussed we decide that more weight should be given to the negative input and to the saleable output and the methane output. A suitable pattern of slack weights might then be that of Table 4. Now, the negative input weight (0.6) is higher than that of the positive input weight (0.4) whereas the positive output weight (0.5) is equal to the sum of the negative output weights (0.1 and 0.4).

Table 4

indicators	w_1 Cost	w_2 Effluent	v_1 Saleable	v_2 Methane	v_3 CO_2
weights	0.4	0.6	0.5	0.1	0.4

The corresponding lambda weights and efficiencies for the MSBM model are shown in Table 5 (we omit comparisons with the RDM, which are no longer directly relevant, though the MSBM model efficiencies are, of course, still lower than those computed under the RDM). Comparison with Table 3, shows that changes in the slack weights can affect the λ weights, e.g. for DMU2 and DMU12. The efficient DMUs remain, of course, efficient. However, the relative positions of the inefficient DMUs are changed. Thus, the least efficient DMU now becomes DMU4.

Table 5 The results of the MSBM model with unequal weights

Ref DMU	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
$?_1$													
$?_2$													
$?_3$		0.593	1	0.223	0.409	0.971				0.857		1	
$?_4$													
$?_5$													
$?_6$													
$?_7$	1	0.407			0.591		1						
$?_8$				0.537				1					
$?_9$													
$?_{10}$													
$?_{11}$									1		1		
$?_{12}$													
$?_{13}$				0.240		0.029				0.143			1
Score t	0.549	0.655	1	0.427	0.686	0.651	1	1	0.750	0.609	1	0.496	1
t	0.967	0.965	1	0.778	0.898	0.848	1	1	0.963	0.949	1	0.771	1
s_1	0.06	0.501		2.735	0.138	0.283			0.3	3.254		0.94	
s_2	0.12	0.231		1.19	2.26	0.448			0.11	0.353		0.31	
s_3	0.26	0.406			0.555				0.05	0.4		0.8	
s_4	0.01	0			0.89	0.83				0.137		0.32	
s_5	0.01	0.01		2.701		0.136			0.13			0.22	

Conclusions

Systems with "natural" negative inputs and outputs occur in a variety of situations. The Slacks-Based Measure (SBM) model proposed by Tone (2001) appears to be an attractive way to deal with cases where both negative outputs and negative inputs occur. However, it has the drawbacks that it is not translation invariant and that it can generate negative efficiencies. The MSBM model presented overcomes both these problems. While this could have been achieved in other ways, e.g. by using the ordinary Range (Max – Min) for scaling the input and output slacks, the use of the SP Range enables the MSBM model to be compared with that of Silva Portela et al (2004).

Arguably, the SP Range is a more appropriate measure than the input and output values of the reference DMU used for scaling the input and output slacks in Tone's Slacks Based Measure. In practical terms, since the SP Range is greater than zero or the corresponding term is dropped from the MSBM model, it obviates the necessity for modifications to the measure where some reference inputs or outputs are zero, as in the example discussed. Although the focus here is on "natural" negative input/output systems, the MSBM model applies to all situations where negative input and/or output values occur.

The MSBM model efficiencies take into account individual input and output slacks, which are ignored in Range Directional Model. This generally leads to lower efficiencies for inefficient DMUs than with the RDM model. The differences in the efficiencies between the two models are likely to be particularly great where the maximum value of ß in the RDM is small because particular outputs or inputs are used very efficiently by all DMUs. Such situations may well arise in comparing different engineering systems either as a result of industry-accepted design efficiency levels or as a result of responses to a design brief that specifies, say, a particular output/ input should be produced/consumed efficiently. The superior discriminative ability of the MSBM model is, therefore, likely to be useful in practice.

A further feature of the MBSM model is the ability to weight different slacks according to their importance to decision makers thus mirroring goal programming approaches to DEA as discussed in Liu et al (2005).

References

1. Charnes, A, Cooper, WW, Golany, B, Seiford, LM, Stutz, J (1985) 'Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions', *Journal of Econometrics,* **30** (1/2) 91-107
2. Charnes, A, Cooper WW, Lewin, AY and Seiford, L (1994) *Data Envelopment Analysis: Theory, Methodology and Applications*, Kluwer Academic Publishers, Boston MA
3. Cooper WW, Seiford, L and Tone, K (2000) *Data Envelopment Analysis: a comprehensive text with models, applications, references and DEA-Solver software*, Kluwer Academic Publishers, Boston MA
4. Korhonen, PJ and Luptacik, M (2004) 'Eco-efficiency of power plants: An extension of data envelopment analysis', *European Journal of Operational Research,* **154,** 437-446
5. Liu, WB, Wu, Z, Meng, W and Sharp, JA (2005) 'Preference, Production and Performance in Data Envelopment Analysis', *Annals of Operations Research*, to appear
6. Pastor, JT (1996) 'Translation invariance in data envelopment analysis: A generalization', *Annals of Operations Research'*, **66**, 93-102
7. Seiford. LM and Zhu, J (2002) 'Modeling undesirable factors in efficiency evaluation', *European Journal of Operational Research*, **142**, 16-20
8. Silva Portela, MCA, Thanassoulis, E and Simpson, G (2004) 'Negative Data in DEA: a directional distance approach applied to bank branches', Journal of the Operational Research Society, **55,** 1111-1121
9. Tone, K (2001) 'A slacks-based measure of efficiency in data envelopment analysis', *European Journal of Operational Research*, **130**, 498-509
10. Ueda, T and Hoshiai, Y (1997) 'Application of component analysis for parsimonious summarization of DEA inputs and/or outputs', *Journal of the Operations Research Society of Japan,* **40**, 466-478
11. Yu, MM (2004) 'Measuring physical efficiency of airports in Taiwan with undesirable outputs and environmental factors', *Journal of Air Transport Management*, **10**, 295-303

UNIVERSITY OF KENT

http://www.kent.ac.uk/kbs/research-information/index.htm