DEA Efficiency Analysis Involving Multiple Production Processes with an Application to Urban Mass Transit

Darold T. Barnum*, University of Illinois at Chicago
John M. Gleason, Creighton University

*Darold T. Barnum is the corresponding author
Great Cities Institute
College of Urban Planning and Public Affairs
University of Illinois at Chicago

GREAT *Cities*

UIC'S METROPOLITAN COMMITMENT

UIC

The Great Cities Institute
The Great Cities Institute is an interdisciplinary, applied urban research unit within the College of Urban Planning and Public Affairs at the University of Illinois at Chicago (UIC). Its mission is to create, disseminate, and apply interdisciplinary knowledge on urban areas. Faculty from UIC and elsewhere work collaboratively on urban issues through interdisciplinary research, outreach and education projects.
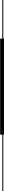
About the Author

Darold T. Barnum, University of Illinois at Chicago is Professor of Management and Professor of Information & Decision Sciences at the University of Illinois at Chicago. He formerly was an Associate Director at the Indiana University Institute for Urban Transportation, where he participated in the training of transit managers from across the nation. His research focuses on performance measurement.

John M. Gleason, Creighton University is a Professor in the Department of Information Systems and Technology within in the College of Business Administration at Creighton University in Omaha.

UIC Great Cities Institute

DEA Efficiency Analysis Involving Multiple Production Processes
With an Application to Urban Mass Transit

Abstract

This paper addresses Data Envelopment Analysis (DEA) efficiency analysis in organizations with multiple production processes. It shows how to measure the impact on an organization's overall efficiency of (a) inefficient and superefficient subunits, and (b) the efficiency with which input resources are allocated to the subunits. It introduces a simple model for efficiently allocating inputs among subunits, and applies the entire analytical process to a large urban mass transit agency.

INTRODUCTION

Organizations usually operate multiple production processes, each process supported by its own set of resource inputs. Multiple production processes can be nonjoint, or networked in various ways such as joint-interrelated and sequential, and inputs can be allocable or nonallocable to multiple outputs (Beattie and Taylor, 1985; Färe and Grosskopf, 2000; Barnum and Gleason, 2006a).

Many recent Data Envelopment Analysis (DEA) articles have addressed the efficiency of Decision Making Units (DMUs) with multiple production processes or subunits. One set of articles suggests methods for maximizing a DMU's reported efficiency when compared to its peer DMUs by assigning hypothetical shares of nonallocable input costs subunits (Yu, 2007). A second set of articles addresses the physical re-allocation of allocable inputs to those subunit outputs that will maximize the DMU's efficiency (Beasley, 2003; Lozano, Villa and Adenso-Diaz, 2004; Lozano and Villa, 2004; Lozano and Villa, 2005; Fang and Zhang, 2007; Nesterenko and Zelenyuk, 2007). Efficiency comparisons involve only subunits within one organization (intra-DMU comparisons), and assume that the DMU's subunit production technologies are identical (Nesterenko and Zelenyuk, 2007)

There are issues that that have not been addressed by either set of articles, but that are important for informing decision makers about how to maximize a target DMU's efficiency. It would be useful to know which of a target DMU's subunits are inefficient, efficient or superefficient compared to subunits with the same production technologies in other organizations, and the impact of each of these subunits' inter-DMU efficiency on the target DMU's efficiency. It also would be useful to know whether allocable input resources are currently being efficiently allocated among the target DMU's subunits when compared to other DMUs' allocations and the impact of inefficient allocation on the target DMU's efficiency.

This paper introduces a DEA procedure that addresses the preceding issues, thereby contributing to a series of articles in Applied Economics and Applied Economics Letters that deal with the impact of disaggregation on efficiency measurement (Tauer, 2001; Färe and Zelenyuk, 2002; Diez-Ticio and Mancebon, 2002; Färe, Grosskopf and Zelenyuk, 2004; Barnum and Gleason, 2005; 2006a; 2006b; 2007; Yu, 2007).

DISAGGREGATING DMU EFFICIENCIES INTO SUBUNIT EFFICIENCIES

Suppose we wish to analyze eight DMUs and two production technologies. Each DMU has two subunits with the first subunit producing with the first technology and the second subunit producing with the second technology. Each type of subunit (technology) produces one output using two inputs. The same kind of output is produced by both subunits, and they both use the same two kinds of inputs. (We have used a simple example to make our illustration intuitively clear, but similar results occur for any number of subunits, inputs, and outputs.) Data for all eight DMUs by technology/subunit type are presented in Table 1.

Table 1. Data for Eight DMUs by Subunit Type

DMU	Type 1 Subunits		Type 2 Subunits		Total DMU	
	Input 1	Input 2	Input 1	Input 2	Input 1	Input 2
A	16	80	16	80	32	160
B	18	60	18	60	36	120
C	40	35	40	35	80	70
D	62	10	62	10	124	20
E	100	6	100	6	200	12
F	17.5	65	16.5	75	34	140
G	16.5	75	81	8	97.5	83
H	28	16	56	32	84	48

Note: Output is 1 for each subunit and 2 for each DMU.

Graphical Introduction

Figure 1 reflects DMU efficiencies, and Figures 2 and 3 show subunit efficiencies. DMU C is inefficient overall; its Subunit 2 is efficient, but its Subunit 1 is not. DMU H appears to be efficient when looking at DMU efficiency, but its Subunit 1 is superefficient and its Subunit 2 is inefficient; because the input it saves in Subunit 1 exceeds the input it wastes in Subunit 2, it is on the overall frontier. Both of G's subunits are efficient, but the DMU is not. This occurs because the DMU's subunits employ different rates of technical substitution (RTS) between their inputs, as shown by the differing isoquant slopes at their points of production. Unless input RTSs are equal across all of a DMU's subunits, the DMU will have misallocated its inputs among subunits, and therefore will require more input than necessary for its levels of output (Barnum and Gleason, 2006a).
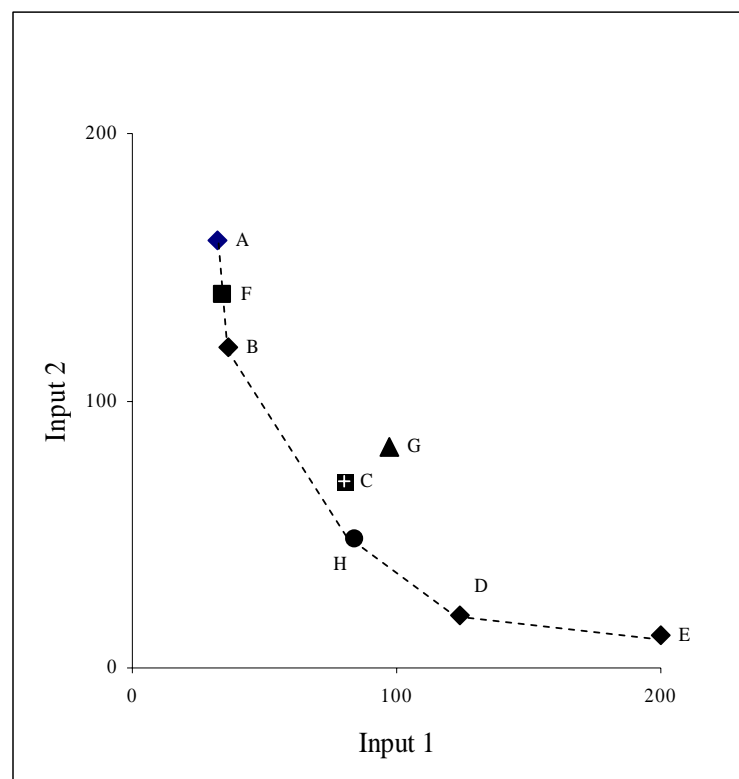
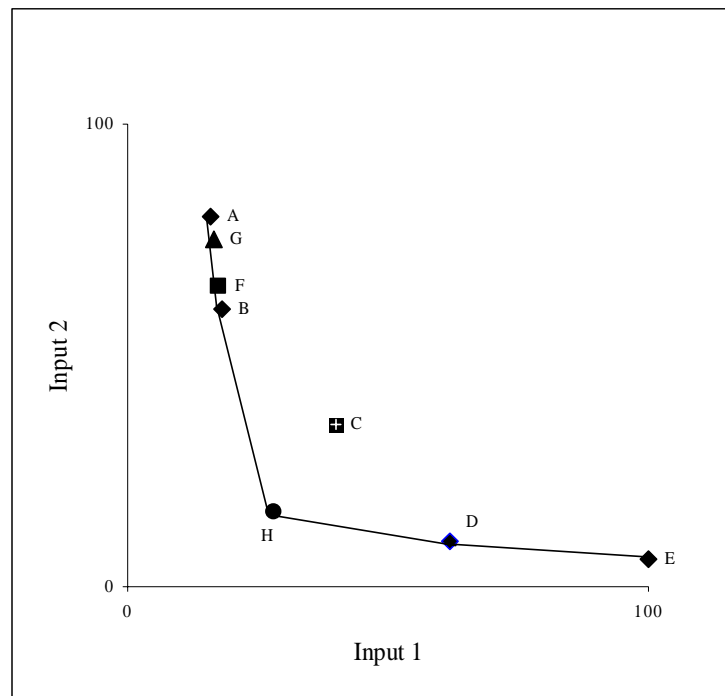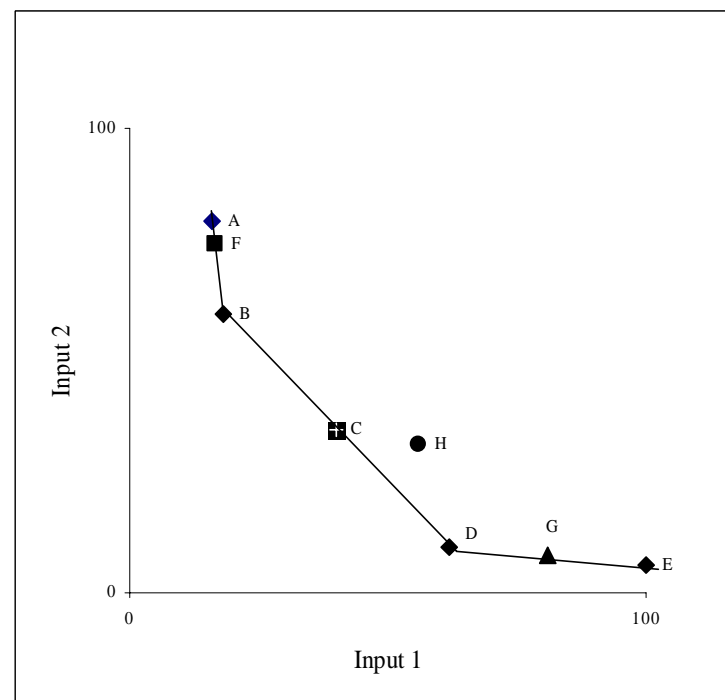Figure 1. DMU Efficiencies

Figure 2. Type 1 Subunit Efficiencies

Figure 3. Type 2 Subunit Efficiencies

Steps in DEA Analysis

In order to mathematically identify the causes of a target DMU's efficiency level and the effect of each cause on the DMU's overall efficiency, a number of steps are necessary. In all DEAs, the target DMU or target subunit is being compared with all of its peers; for DMU efficiency, all DMUs are being compared and for subunit efficiency, all subunits of the same type from all DMUs are being compared. We assume an input orientation, but the steps can be changed to reflect an output orientation.

1. Compute <u>DMU efficiency</u> (subunit inputs and outputs aggregated) for the target DMU. [Scores identify the target DMU's overall efficiency.]

2. Compute <u>Subunit efficiency</u> by subunit/technology type for each subunit of the target DMU, using each subunit's own inputs and outputs. [Scores identify each subunit as inefficient, efficient, or superefficient when compared to its subunit peers.]

3. For a target DMU with one or more inefficient subunits, do the following separately for each inefficient subunit.

   a. Proportionally lower the inputs of the inefficient subunit to the point that the subunit is efficient ($\theta = 1$).

   b. Using the lowered inputs for the single originally-inefficient subunit of the target DMU, and the original inputs for all other subunits of both the target and the remaining DMUs, compute the revised DMU efficiency score for the target DMU. [The difference between the revised DMU efficiency and the original DMU efficiency estimates the effect of the inefficient subunit on the target DMU.]

4. For a target DMU with one or more superefficient subunits, do the following separately for each superefficient subunit.

   a. Proportionally increase the inputs of the superefficient subunit to the point that the subunit is efficient ($\theta = 1$).

b. Using the higher inputs for the single originally-superefficient subunit of the target DMU, and the original inputs for all other subunits of both the target and the remaining DMUs, compute the revised DMU efficiency score for the target DMU. [The difference between the revised DMU efficiency and the original DMU efficiency will estimate the effect of the superefficient subunit on the target DMU.]

5. For all DMUs, not just the target DMU, do the following simultaneously.

a. For all inefficient subunits, proportionally lower their inputs to the point that each is efficient ($\theta = 1$).

b. Using the lowered inputs for all originally-inefficient subunits, and the original inputs for all originally efficient and superefficient subunits, compute DMU efficiency (subunit inputs and outputs aggregated) for the target DMU. [If a target DMU's score is one or greater ($\theta \geq 1$), the target DMU is allocatively efficient when compared to its peer DMUs, as Barnum and Gleason (2006a) demonstrate].

Application

We next apply these steps to our data. Efficiency scores for the first two steps are computed with a linear program (1-5), which is input oriented and assumes constant returns to scale. The DEAs are conducted with Scheel's EMS software (2000). For each observation $j = 1, ..., J$ there are data on $n = 1, ..., N$ inputs and on $m = 1, ..., M$ outputs, where $x^j = (x_{j1}, ..., x_{jN}) \in \Box^N_+$ and $y^j = (y_{j1}, ..., y_{jM}) \in \Box^M_+$. The DEA score θ estimates the technical superefficiency of the target DMU k (Andersen and Petersen, 1993).

	$\min_{\lambda} \theta$				(1)
	subject to	$\sum_{j=1}^{J} x_{jn}\lambda_j \leq \theta x_{kn}$	$n = 1, ..., N$		(2)

		$\displaystyle\sum_{j=1}^{J} y_{jm}\lambda_j \geq y_{km}$	$m = 1, ..., M$		(3)
		$\lambda_k = 0$			(4)
		$\lambda_j \geq 0$	$j = 1, ... J; \ j \neq k$		(5)

All DMUs but C and G are reported to be technically efficient in Step 1, with C's efficiency being 0.92 and G's 0.76. There are two types of subunits with eight observations in each, so one set of DEAs in Step 2 involves the subunits of the first type and the other set involves subunits of the second type. All subunits of DMUs A, B, D, E, F and G are efficient (Table 2).

Table 2. DMU and Subunit Efficiency Scores, Original Data

	Type 1 Subunits	Type 2 Subunits	DMU
A	1.03	1.03	1.06
B	1.02	1.14	1.09
C	0.66	1.00	0.92
D	1.07	1.21	1.27
E	1.67	1.33	1.67
F	1.00	1.00	1.00
G	1.00	1.00	0.76
H	1.68	0.84	1.12

Now, consider DMU C, which Step 1 reports to have an overall technical efficiency of 0.92. Step 2 tells us that C has one inefficient subunit, Subunit 1, which has a technical efficiency of 0.66. The effect on the DMU's overall efficiency is the difference between the DMU's score in step 3, which is 1.10, and its original efficiency score of 0.92. The 0.18 difference indicates how much the DMU's efficiency can be improved by improving its Subunit 1 efficiency.

DMU H's original aggregated superefficiency score was 1.12, with step 2 reporting that its Subunit 1 score is 1.68 and its Subunit 2 score is 0.84. Decreasing the inputs of Subunit 2

(Step 3) increases H's overall score from 1.12 to 1.25, a difference of 0.13. Increasing the inputs of Subunit 1 (Step 4) decreases H's overall score from 1.12 to 0.91, a difference of 0.23. Thus, the increase in efficiency of DMU H resulting from the superefficient Subunit 1 exceeds the decrease in efficiency resulting from the inefficient Subunit 2 (0.23 > 0.13). So, if a choice has to be made between maintaining Subunit 1's superefficiency at its current level, or making Subunit 2 efficient, the former would be the better choice.

Now, consider DMU G, whose subunits are both efficient (Step 2) although the DMU itself is inefficient (0.76) (Step 1). Steps 3 and 4 are inapplicable, so we proceed to Step 5. DMU G's Step-5 score is 0.69, so its inefficient allocation yields a DMU efficiency that is 0.31 lower than it should be, given both subunits are efficient. Thus, the only action necessary is to reallocate inputs between subunits of DMU G. Returning to Figures 2 and 3, G's two inputs can be reallocated to any point on the subunits' respective isoquants such that their RTSs are closer together than any efficient peer. (That is, the RTSs don't have to be exactly equal, but they must be closer to being equal than any benchmark competitors.) We now provide a methodology by which DMU G can reallocate its inputs between its two subunits, to decrease the total input amounts needed to produce its current level of outputs.

IMPROVING ALLOCATION EFFICIENCY

When a DMU has a high ratio of subunits to inputs-plus-outputs, and assigns each production process to its own subunit, then it may be possible to use "input allocation" models to measure and improve allocation efficiency (Beasley, 2003; Lozano, Villa and Adenso-Diaz, 2004; Lozano and Villa, 2004; Podinovski, 2004; Lozano and Villa, 2005; Podinovski, 2007; Fang and Zhang, 2007; Nesterenko and Zelenyuk, 2007). In the present case, however, as would be true in many industries, there would be too few subunits to make these input allocation models practical.

Accordingly, we provide a mathematical programming model that can be used to improve allocation efficiency in these cases. Our objective is to minimize inputs. We want to insure that each subunit will not produce less output, and that the DMU will not use more of any input than is currently available. Changes in input values cause changes in the output values, so it is not possible to reallocate the inputs without taking into account the effects on outputs. Unless we know the true functional relationships between inputs and outputs for each production process, it is necessary to estimate them using suitable econometric methods. If the estimated relationships are all linear, then linear programming can be used to minimize inputs. If the estimated relationships are not linear, then nonlinear programming must be used. In either case, mathematical program 6- 10 applies.

$\min\limits_{\hat{y}_m, x_{nm}} \sum\limits_{n=1}^{N}\sum\limits_{m=1}^{M} x_{nm}$			(6)
subject to	$\sum\limits_{m=1}^{M} x_{nm} \leq \iota_n$	$n = 1,...,N$	(7)
	$\hat{y}_m \geq \omega_m$	$m = 1,...,M$	(8)
	$x_{nm} \geq 0$	$n = 1,...,N; m = 1,...M$	(9)
	$\hat{y}_m = f(x_{nm})$	$m = 1,...,M$	(10)

The objective function minimizes the sum of all N inputs used to produce the M outputs, where x_{nm} is the amount of input n used to produce output in subunit m, and $\hat{y}_m$ is the revised value of output m. Constraint set 7 insures that the sum of the reallocated inputs will not exceed the current availability of the input, ι_n. Constraint set 8 requires that the revised value of each output, $\hat{y}_m$, is at least as great as the original value of that output, ω_m. Constraint set 9 requires that the amount of each input used by each of the outputs be non-negative. Constraint

set 10 estimates the revised value of each of the outputs based on the values of its inputs and the estimated relationship.

For DMU G, mathematical program 6-10 was solved using the mathematical programming module (Solver) in Excel 2002, which uses the generalized reduced gradient method to solve nonlinear problems (Fylstra *et al.,* 1998). Revised output values for subunits of DMU G were estimated based on equation 11.

$\ln(\hat{y}_m) = -3.4198 + 0.60302\ln(x_{1m}^*) + 0.40004\ln(x_{2m}^*)$	$m = 1, 2$	(11)

DMU G's use of Input 1 can be reduced from 97.5 to 70.6, and its use of Input 2 can be reduced from 83 to 46.8, while maintaining the original levels of outputs in both subunits, if it reallocates its inputs so each subunit receives 35.3 units of Input 1 and 23.4 units of Input 2. As a result, DMU G's Step 1 efficiency increases from 0.76 to 1.07, an increase of 0.31.

APPLICATION TO AN URBAN TRANSIT AGENCY

Large U.S. urban transit agencies can provide on-the-street service with up to four technologies/ subunits: self-operated demand-responsive service, outsourced demand-responsive service, self-operated fixed-schedule service, and outsourced fixed-schedule service. The Maryland Transit Administration (MTA) uses all four options. We conducted a DEA with the MTA as the target DMU, using 2006 data from 52 transit systems with 150 or more vehicles. Estimated seat-hours for each of the four subunits was used as the output, and operating expenses adjusted for price differences as the input (United States Federal Transit Administration, 2007). Thus, the MTA had four production technologies (or subunits), each utilizing one input to produce one output.

The MTA's Step 1 DMU technical efficiency was 0.67, and its Step 2 subunit technical efficiencies were 0.29 for self-operated demand-responsive service (lowering DMU efficiency by

0.02), 0.63 for outsourced demand-responsive service (lowering DMU efficiency by 0.03), 0.70

for self-operated fixed-schedule service (lowering DMU efficiency by 17.5), and 0.23 for

outsourced fixed-schedule service (lowering DMU efficiency by 0.06). Thus, it would appear

that the biggest increase in MTA efficiency could be made by improving the subunit that already

has the highest efficiency, self-operated fixed-schedule service. The MTA was almost

completely efficient in allocating its resources to its four subunits, as evidenced by the result of

Step 5, in which its DMU efficiency, when all subunits of all 52 agencies had been made

efficient, was 0.99. This value can be improved however with a reformulation of mathematical

program 6-10 (model 12-18), where $N = 1$ because there is a single input:

	$\displaystyle\min_{\hat{y}_m, x_m} \sum_{m=1}^{4} x_m$			(12)
	subject to	$\displaystyle\sum_{m=1}^{M} x_{1m} \leq \iota_1$		(13)
		$\displaystyle\sum_{m=1}^{4} \hat{y}_m \geq \sum_{m=1}^{4} \omega_m$		(14)
		$\hat{y}_m \geq 0.2\, y_m$	$m = 1,2,3,4$	(15)
		$\displaystyle\sum_{m=1}^{2} \hat{y}_m \geq 0.9 \sum_{m=1}^{2} \omega_m$		(16)
		$x_{1m}^{*} \geq 0$	$m = 1,2,3,4$	(17)
		$\hat{y}_m = f(x_{1m}^{*})$	$m = 1,2,3,4$	(18)

In this case, there is only one type of input $(N = 1)$, which is to be allocated among the

$M = 4$ production processes. That is, $m = 1$ is self-operated demand-responsive service, $m = 2$

is outsourced demand-responsive service, $m = 3$ is self-operated fixed-schedule service, and

$m = 4$ is outsourced fixed-schedule service. Constraint 13 requires that the total of the

reallocated single input not exceed the input's original value. Constraint set 14 requires that the

total output not decline. Constraint set 15 requires that each process retain at least 20 percent

of its original service (for political reasons) and constraint 16 requires that demand-responsive operations retain at least 90 percent of the service by that mode (under the assumption that 90 percent of the riders would be physically incapable of using other transit). Panel Data Analysis of the four modes for all 52 systems for 2002-2006 resulted in four estimation equations for the relationships between input-output pairs, each adjusted for the MTA subunits' levels of efficiency (constraint set 18).

Applying model 12-18 to the problem results in the following input reallocation recommendations. The agency should decrease self-operated demand-responsive service by 80 percent, increase outsourced demand-responsive service by 1 percent, decrease outsourced fixed-schedule transit by 80 percent, and increase self-operated fixed-schedule transit by 12 percent. This leaves the total number of seat-hours the same and decreases annual expenses by $11 million, which is 4.4 percent of the 2006 expenditures. As a result of these changes, DMU efficiency would increase from 0.67 to 0.70. Obviously, there may be practical, legal or political reasons why the recommended re-allocations cannot be fully implemented, but the suggested changes do inform management of the directions to be pursued for efficiency gains.

SUMMARY

This paper presents a new procedure for identifying and correcting internal causes of a multi-product DMU's inefficiency. The procedure makes it possible to evaluate the inter-DMU efficiency of the subunits responsible each of the input-to-output transformation processes, whether input resources are efficiently allocated among a target DMU's subunits, and the impact of subunit allocations, inefficiencies and superefficiencies on their DMU's overall efficiency. Further, it introduces a new model for efficiently allocating inputs among a DMU's subunits when the number of subunits is small, and demonstrates the process with an urban transit example.

References

1. Andersen, P. and Petersen, N. C. (1993) A procedure for ranking efficient units in data envelopment analysis, *Management Science*, *39* (10), 1261-1265.

2. Barnum, D. T. and Gleason, J. M. (2005) Technical efficiency bias caused by intra-input aggregation in data envelopment analysis, *Applied Economics Letters*, *12* (13), 785-788.

3. Barnum, D. T. and Gleason, J. M. (2006a) Biases in technical efficiency scores caused by intra-input aggregation: Mathematical analysis and a DEA application using simulated data, *Applied Economics*, *38* 1593-1603.

4. Barnum, D. T. and Gleason, J. M. (2006b) Measuring efficiency in allocating inputs among outputs with DEA, *Applied Economics Letters*, *13* (6), 333-336.

5. Barnum, D. T. and Gleason, J. M. (2007) Technical efficiency bias in data envelopment analysis caused by intra-output aggregation, *Applied Economics Letters*, *14* (9), 623-626.

6. Beasley, J. E. (2003) Allocating fixed costs and resources via data envelopment analysis, *European Journal of Operational Research*, *147* (1), 198-216.

7. Beattie, B. R. and Taylor, C. R. (1985) *The economics of production,* New York, Wiley.

8. Diez-Ticio, A. and Mancebon, M. J. (2002) The efficiency of the Spanish police service: An application of the multiactivity DEA model, *Applied Economics*, *34* (3), 351-362.

9. Fang, L. and Zhang, C.-Q. (2007) Resource allocation based on the DEA model. Forthcoming in the *Journal of the Operational Research Society*.

10. Färe, R., Grosskopf, S. and Zelenyuk, V. (2004) Aggregation bias and its bounds in measuring technical efficiency, *Applied Economics Letters*, *11* (10), 657-660.

11. Färe, R. and Grosskopf, S. (2000) Network DEA, *Socio-Economic Planning Sciences*, *34* (1), 35-49.

12. Färe, R. and Zelenyuk, V. (2002) Input aggregation and technical efficiency, *Applied Economics Letters*, *9* (10), 635-636.

13. Fylstra, D., Lasdon, L., Watson, J. and Waren, A. (1998) Design and use of the Microsoft Excel Solver, *Interfaces*, *28* (5), 29-55.

14. Lozano, S. and Villa, G. (2005) Centralized DEA models with the possibility of downsizing, *Journal of the Operational Research Society*, *56* (4), 357-364.

15. Lozano, S., Villa, G. and Adenso-Diaz, B. (2004) Centralized target setting for regional recycling operations using DEA, *Omega*, *32* (2), 101-110.

16. Lozano, S. and Villa, G. (2004) Centralized Resource Allocation Using Data Envelopment Analysis, *Journal of Productivity Analysis*, *22* (1), 143-161.

17. Nesterenko, V. and Zelenyuk, V. (2007) Measuring potential gains from reallocation of resources, *Journal of Productivity Analysis*, *28* (1), 107-116.

18. Podinovski, V. V. (2004) Production trade-offs and weight restrictions in data envelopment analysis, *Journal of the Operational Research Society*, *55* (12), 1311-1322.

19. Podinovski, V. V. (2007) Improving data envelopment analysis by the use of production trade-offs, *Journal of the Operational Research Society*, *58* (10), 1261-1270.

20. Scheel, H. (2000) EMS: Efficiency Measurement System, http://www.wiso.uni-dortmund.de/lsfg/or/scheel/ems/ .

21. Tauer, L. W. (2001) Input aggregation and computed technical efficiency, *Applied Economics Letters*, *8* (5), 295-297.

22. United States Federal Transit Administration (2007) National transit database, http://www.ntdprogram.gov/ntdprogram/.

23. Yu, M. M. (2007) Measuring the efficiency and return to scale status of multi-mode bus transit-evidence from Taiwan's bus system. Forthcoming in *Applied Economics Letters*.