Chapter 6
Scale Efficiency Measurement in Two-Stage Production Systems

Chiang Kao and Shiuh-Nan Hwang

Abstract One important objective in measuring efficiency is to find the factors that cause inefficiencies so that its performance can be improved. The conventional data envelopment analysis approach is able to decompose the overall efficiency of a system into the product of the technical and scale efficiencies when the internal structure is ignored. For two-stage systems, where the inputs are supplied to the first process to produce intermediate products for the second process to produce the final outputs, the system efficiency can be decomposed into process efficiencies. This paper further decomposes each process efficiency into the product of the technical and scale efficiencies via an input-oriented model for the first process and an output-oriented one for the second. The decomposition also reveals that the overall efficiency of the two-stage system, when the operations of the two processes are considered, is still the product of the technical and scale efficiencies. The concept is illustrated using an example of 24 non-life insurance companies in Taiwan.

Keywords Data envelopment analysis • Two-stage system • Scale efficiency • Technical efficiency

---

C. Kao (✉)
Department of Industrial and Information Management, National Cheng Kung University, Tainan, Taiwan
e-mail: ckao@mail.ncku.edu.tw

S.-N. Hwang
Department of Business Administration, Ming Chuan University, Taipei, Taiwan
e-mail: snhwang@mail.mcu.edu.tw

W.D. Cook and J. Zhu (eds.), *Data Envelopment Analysis*,
International Series in Operations Research & Management Science 208,
DOI 10.1007/978-1-4899-8068-7_6,

6.1 Introduction

Charnes et al. (1978) developed a data envelopment analysis (DEA) model, conventionally referred to as the CCR model, to measure the relative efficiency of a set of decision making units (DMUs) that use multiple inputs to produce multiple outputs under the assumption of constant returns to scale (CRS). This model was later extended by Banker et al. (1984), conventionally referred to as the BCC model, to measure efficiency under the assumption of variable returns to scale (VRS). The CCR model is able to detect inefficiency due to the aggregate effect of insufficient technology and improper scale, while the BCC model is only used to detect insufficient technology. By comparing the efficiency scores calculated from these two models, the effects due to improper scale can be identified. The identified sources of inefficiency can then enable decision makers to design suitable alternatives to improve the performance of a system.

Conventional DEA models treat the system as a whole unit, and thus only the inputs supplied to the system and the outputs produced from it are considered in measuring efficiency. However, in many situations a system is composed of several interrelated processes, with the outputs of one process being used by some others for production, and ignoring the operations of the internal processes will produce misleading results. In response to this, Färe and Grosskopf (2000) proposed a network DEA model to take the operations of the internal processes into account.

The simplest structure of network systems is a two-stage system, where all the inputs are supplied to the first process to produce intermediate products for the second process to produce the final outputs. Several models have been proposed to measure the efficiency of this type of system (see the review of Cook et al. 2010), and Kao (2009) classified these into independent, connected, and relational. The independent model is typified by that presented in Seiford and Zhu (1999), which treats the two stages as two independent DMUs, and their efficiencies are calculated separately. Therefore, the scale efficiency of each process can be calculated by applying the CCR and BCC models. However, it is still not possible to measure the scale efficiency of the system when considering the interrelation of the two processes.

Färe and Grosskopf (2000) remains the most representative work with regard to a connected model, in that the technologies of the two processes are considered in measuring the overall efficiency of the system. Although the efficiency can be measured under both CRS and VRS, how to calculate the scale efficiency is still a problem, because the relationship between the overall, technical, and scale efficiencies in the two-stage system is not known.

For relational models, the system and process efficiencies can be calculated at the same time; moreover, there exist mathematical relationships between them. For example, the model in Kao and Hwang (2008) shows a multiplicative relationship between the system and process efficiencies, while that in Chen et al. (2009) shows an additive one. The slacks-based measures model also exhibits an additive relationship (Tone and Tsutsui 2009; Kao 2013). For models with an additive relationship, the efficiencies of both the system and processes can be calculated under both CRS and VRS. However, similar to the case of the connected model, the scale efficiency cannot be obtained from the other two, because the relationship between the overall, technical, and scale efficiencies is not known. It thus seems that only the multiplicative form of the relational model can be applied to measure scale efficiencies, and Kao and Hwang (2011) proposed an approach to accomplish this task.

The major problem in measuring scale efficiency from the overall and technical efficiencies for the two-stage system is that the outputs of the first process are the inputs of the second. If one wishes to improve the efficiency of the first process by increasing its outputs, the efficiency of the second process will then be reduced due to the increased amount of inputs. Similarly, if one tries to raise the efficiency score of the second process by reducing the amount of its input, then the efficiency of the first process will be decreased due to producing output. To resolve this conflict, Kao and Hwang (2011) used an input-oriented model to measure the efficiency of the first process by fixing the amount of the output and an output-oriented model to measure the efficiency of the second process. In this way, the efficiency of the first process can be improved by reducing the amount of the input and the second process by increasing that of the output. The CCR and BCC models are then applied to measure scale efficiencies for the two processes, which then represent the scale efficiency of the system as a whole.

In the following sections, the input- and output-oriented DEA models are first briefly reviewed. The measurement of the scale efficiencies is then illustrated graphically using a simple example. After this, the models for measuring scale efficiencies for general cases are developed, and the technical and scale efficiencies of the system and two processes of non-life insurance companies in Taiwan are calculated. Finally, some conclusions are drawn from the discussion of these results.

6.2 Input- and Output-Oriented Models

Both the CCR and BCC models can be formulated from the input and output sides. The objective of the former is to examine how much of the input can be reduced while producing the same amount of output, while that of the latter is to examine how much of the output can be increased by using the same amount of input for production. In the following discussions, we use X_{ij}, $i = 1, \ldots, m$ and Y_{rj}, $r = 1, \ldots, s$ to denote the ith input and rth output of the jth DMU, $j = 1, \ldots, n$, respectively, with v_i and u_r being the virtual multipliers associated with X_{ij} and Y_{rj}.

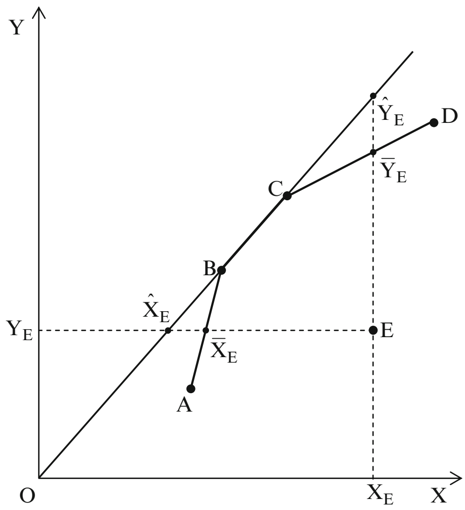


Fig. 6.1 Input- and output-oriented efficiency measures

6.2.1 Input-Oriented Model

Banker et al. (1984) developed the following model to measure the technical efficiency of the kth DMU from the input side under VRS:

$$
\begin{aligned}
E_k = \max.&\sum_{r=1}^{s} u_r Y_{rk} - u_0\\
\text{s.t.}\quad &\sum_{i=1}^{m} v_i X_{ik} = 1\\
&\left(\sum_{r=1}^{s} u_r Y_{rj} - u_0\right) - \sum_{i=1}^{m} v_i X_{ij} \le 0, \quad j = 1, \ldots, n\\
&u_r, v_i \ge \varepsilon, \quad r = 1, \ldots, s, \quad i = 1, \ldots, m\\
&u_0 \ \text{unrestricted in sign}
\end{aligned}
\tag{6.1}
$$

where ε is a small non-Archimedean number imposed to avoid ignoring any factor (Charnes et al. 1979; Charnes and Cooper 1984). When the term u_0 is omitted, Model (6.1) becomes the CCR model (Charnes et al. 1978), and the corresponding efficiency is the overall efficiency. The ratio of the overall efficiency to the technical efficiency is the (input) scale efficiency.

Consider a one-input one-output example with five DMUs, labeled as A, B, C, D, and E, as shown in Fig. 6.1. The straight line OBC and the connected line segments $ABCD$ are the production frontiers constructed under CRS and VRS, respectively. DMUs A, B, C, and D are technically efficient, among which B and C are also overall efficient. The input-oriented model calculates efficiencies based on the amount of input consumed for production. The (input) overall and technical efficiencies for DMU E are the ratios of $\hat{X}_E$ to X_E and $\overline{X}_E$ to X_E, respectively, and the (input) scale efficiency is the ratio of the overall efficiency to the technical efficiency, which is the ratio of $\hat{X}_E$ to $\overline{X}_E$.

6.2.2 Output-Oriented Model

Banker et al. (1984) also developed a model for measuring efficiencies from the output side under VRS, which can be formulated as:

$$
\begin{aligned}
\frac{1}{E_k} = \min. & \sum_{i=1}^{m} v_i X_{ik} + v_0 \\
\text{s.t.} \ & \sum_{r=1}^{s} u_r Y_{rk} = 1 \\
& \left(\sum_{i=1}^{m} v_i X_{ij} + v_0 \right) - \sum_{r=1}^{s} u_r Y_{rj} \geq 0, \quad j = 1, \ldots, n \\
& u_r, v_i \geq \varepsilon, \quad r = 1, \ldots, s, \quad i = 1, \ldots, m \\
& v_0 \text{ unrestricted in sign}
\end{aligned} \tag{6.2}
$$

where the efficiency is expressed in reciprocal form. The efficiency thus obtained is the output-oriented technical one. Similar to the previous case, when the term v_0 is omitted, Model (6.2) becomes the CCR model, and the resulting measure is the overall efficiency. The ratio of the overall efficiency to the (output) technical efficiency is the (output) scale efficiency.

The output-oriented model measures efficiencies based on the amount of output produced. For DMU E in Fig. 6.1, the output-oriented overall and technical efficiencies are the ratios of Y_E to $\hat{Y}_E$ and Y_E to $\overline{Y}_E$, respectively. The (output) scale efficiency, which is the ratio of the overall efficiency to the (output) technical efficiency, is the ratio of $\overline{Y}_E$ to $\hat{Y}_{\mathrm{E}}$.

Note that the overall efficiencies calculated from the input Model (6.1), with the term u_0 omitted, and output Model (6.2), with the term v_0 omitted, are the same. The technical and scale efficiencies calculated from these two models, however, may not be the same. In the example shown in Fig. 6.1, only the DMUs using the frontier facet BC to calculate efficiencies will result in the same measures.

6.3 Graphical Illustration

The two-stage system is a system composed of two processes connected in series, where all the inputs are supplied to the first process to produce intermediate products, and all of them in turn are used by the second process to produce the final outputs of the system. Figure 6.2 shows a typical two-stage system, where $Z_{fj}, f = 1, \ldots, g$ denote the intermediate products.

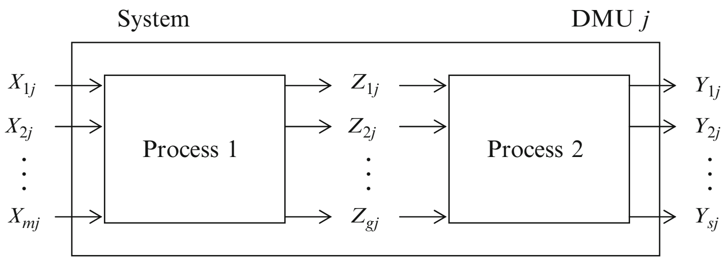


Fig. 6.2 Two-stage system with inputs *X*, outputs *Y*, and intermediate products *Z*

Table 6.1 Data and efficiencies measured from output-oriented Model (6.2) for five DMUs of a two-stage system

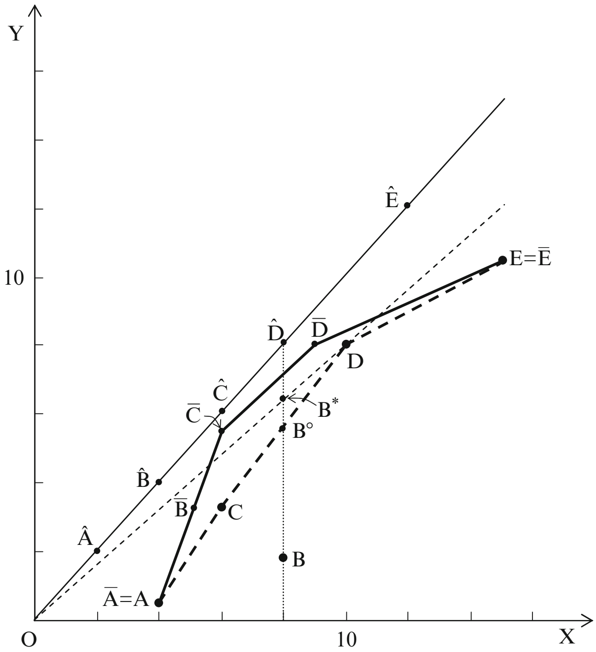
DMU	Input *X*	Intermediate product *Z*	Output *Y*	Overall efficiency	Technical efficiency	Scale efficiency
A	4	2	0.5	5/32	1	5/32
B	8	4	1.8	9/32	36/113	113/128
C	6	6	3.3	11/16	1	11/16
D	10	8	8	1	1	1
E	15	12	10.5	7/8	1	7/8

Consider a simple example of five DMUs, *A*~*E*, where each applies one input *X* to produce one intermediate product *Z* in the first process, and the intermediate product *Z* is then used in the second process to produce one output *Y*. The left part of Table 6.1 shows a set of hypothetical data for these five DMUs. The conventional DEA approach ignores the operations of the two processes, assuming that output *Y* is directly produced by input *X*. In this case, the production frontiers constructed under CRS and VRS are the dashed straight line *OD* and connected line segments *ACDE*, respectively, shown in Fig. 6.3. For DMU *B*, the overall and (output) technical efficiencies are B/B^* and B/B°, respectively, which produce an output scale efficiency of B°/B^*. The right part of Table 6.1 shows the overall, (output) technical, and (output) scale efficiencies of the five DMUs.

The network DEA approach, on the other hand, takes the operations of the two processes into consideration. Figure 6.4 depicts the production process in a counter-clockwise orientation, where the right side shows that Process 1 applies input *X* to produce intermediate product *Z*, and the left side shows that Process 2 applies intermediate product *Z* to produce output *Y*. The superscripts associated with the DMUs indicate the process. The straight lines OC^1 and OD^2 passing through the origin are the production frontiers under CRS for Processes 1 and 2, respectively. Note that here two production frontiers are constructed for the two processes, which is different from the idea of using one frontier for the two processes, as in Chen et al. (2010).

On the right side, the kinked line $A^1C^1E^1$ is the production frontier for Process 1 under VRS. The three DMUs on the frontier, *A*, *C*, and *E*, are thus technically efficient. DMU *C* is also overall efficient, because it lies on the frontier constructed

Fig. 6.3 Different types of production frontier on the *X*-*Y* plane



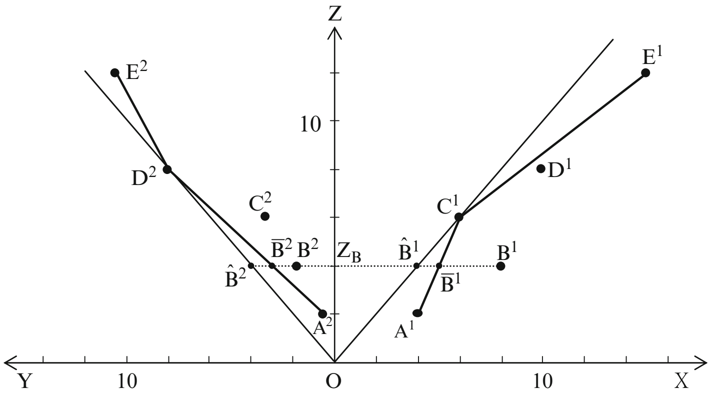


Fig. 6.4 Efficiency measurement for processes in a two-stage system

under CRS. The other two DMUs, *B* and *D*, are technically inefficient. Consider DMU *B*. Its overall and technical efficiencies measured from the input side are $\hat{B}^1Z_B/B^1Z_B$ (=1/2) and $\overline{B}^1Z_B/B^1Z_B$ (=5/8), respectively, which produce an (input) scale efficiency of $\hat{B}^1Z_B/\overline{B}^1Z_B$ (=4/5). These three types of efficiency for the other four DMUs can be calculated similarly, with the results shown in Table 6.2 under the heading of "Process 1."

The left side of Fig. 6.4 shows the production of Process 2. Note that the vertical axis represents the input of this process, the intermediate product *Z*, and the

Table 6.2 System and process efficiencies for the example, taking into account the operations of the processes

	System			Process 1			Process 2		
DMU	Overall	(Tech.	Scale)	Overall	(Tech.	Scale)	Overall	(Tech.	Scale)
A	1/8	(1	1/8)	1/2	(1	1/2)	1/4	(1	1/4)
B	9/40	(3/8	3/5)	1/2	(5/8	4/5)	9/20	(3/5	3/4)
C	11/20	(3/5	11/12)	1	(1	1)	11/20	(3/5	11/12)
D	4/5	(9/10	8/9)	4/5	(9/10	8/9)	1	(1	1)
E	7/10	(1	7/10)	4/5	(1	4/5)	7/8	(1	7/8)

horizontal axis shows the output Y of this process. The kinked line $A^2D^2E^2$ is the production frontier constructed under VRS, and DMUs A, D, and E are technically efficient. For the two technically inefficient DMUs, B and C, consider B. Its overall and technical efficiencies measured from the output side are $B^2Z_B/\hat{B}^2Z_B$ (=9/20) and $B^2Z_B/\overline{B}^2Z_B$ (=3/5), respectively, which result in an (output) scale efficiency of $\overline{B}^2Z_B/\hat{B}^2Z_B$ (=3/4). The efficiencies of the other four DMUs are measured similarly, with the results shown in Table 6.2, under the heading of "Process 2."

The results from the two processes can be aggregated to form that of the system. Consider DMU B again. This DMU uses 8 units of input X to produce 1.8 units of output Y (via 4 units of intermediate product Z), with a rate of 1.8/8. If it is technically efficient in both processes, then only 5 units of input X (corresponding to $\overline{B}^1$) are needed to produce 3 units of output Y (corresponding to $\overline{B}^2$), with a rate of 3/5. This corresponds to point $\overline{B}$ on the X-Y plane of Fig. 6.3. Comparing the actual rate of 1.8/8 to the technically efficient rate of 3/5, an overall technical efficiency of [(1.8/8)/(3/5)], or 3/8, is obtained for the system. This efficiency is clearly the product of the technical efficiency of the two processes, 5/8 and 3/5. By the same token, if both processes are overall efficient, then DMU B only requires 4 units of input X (corresponding to $\hat{B}^1$) to produce 4 units of output Y (corresponding to $\hat{B}^2$), with a rate of 4/4. This corresponds to point $\hat{B}$ on the X-Y plane of Fig. 6.3. Comparing the actual rate of 1.8/8 to this overall efficient rate of 4/4, an overall efficiency of [(1.8/8)/(4/4)], or 9/40, is obtained for the system. The ratio of the overall efficiency to the technical efficiency of the system, (9/40)/(3/8) = 3/5, is the scale efficiency of the system. From the graphical relationship shown in Fig. 6.4, this value is clearly the product of the scale efficiency of the two processes, 4/5 and 3/4. The efficiencies corresponding to the other four DMUs can be calculated similarly, with the results shown in Table 6.2 under the heading of "System."

From the discussion regarding the target points, we conclude that the straight line $O\hat{A}\,\hat{B}\,\hat{C}\,\hat{D}\,\hat{E}$ in Fig. 6.3 is the system frontier under CRS and the connected line segment $\overline{A}\,\overline{C}\,\overline{D}\,\overline{E}$ is that under VRS. These two frontiers lie above their counterparts OD and $ACDE$, respectively, constructed from the conventional DEA approach, indicating that ignoring the operations of the processes will overstate the measured efficiencies.

In this example, we find, first, the overall efficiency of the system is the product of those of the two processes, second, the overall efficiency of each process is the product of their technical and scale efficiencies, third, the technical efficiency of the system is the product of those of the two processes, and fourth, the scale efficiency of the system is the product of those of the two processes.

6.4 Measurement Models for General Cases

To measure the efficiency of the two-stage system for DMU k, Kao and Hwang (2008) proposed the following model:

$$
\begin{aligned}
E_k^S = \max. & \sum_{r=1}^{s} u_r Y_{rk} \\
\text{s.t.} \quad & \sum_{i=1}^{m} v_i X_{ik} = 1 \\
& \sum_{r=1}^{s} u_r Y_{rj} - \sum_{i=1}^{m} v_i X_{ij} \leq 0, \quad j = 1, \ldots, n \\
& \sum_{f=1}^{g} w_f Z_{fj} - \sum_{i=1}^{m} v_i X_{ij} \leq 0, \quad j = 1, \ldots, n \\
& \sum_{r=1}^{s} u_r Y_{rj} - \sum_{f=1}^{g} w_f Z_{fj} \leq 0, \quad j = 1, \ldots, n \\
& u_r, v_i, w_f \geq \varepsilon, \quad r = 1, \ldots, s, \quad i = 1, \ldots, m, \quad f = 1, \ldots, g
\end{aligned}
\tag{6.3}
$$

where the three sets of constraints correspond to the system, Process 1, and Process 2, respectively. Since the constraints corresponding to the system are the sum of those corresponding to the two processes, they are redundant, and can thus be deleted.

At optimality, the system and process efficiencies, based on Model (6.3), are calculated as:

$$
\begin{aligned}
E_k^S &= \sum_{r=1}^{s} u_r Y_{rk} \Big/ \sum_{i=1}^{m} v_i X_{ik} \\
E_k^{(1)} &= \sum_{f=1}^{g} w_f Z_{fk} \Big/ \sum_{i=1}^{m} v_i X_{ik} \\
E_k^{(2)} &= \sum_{r=1}^{s} u_r Y_{rk} \Big/ \sum_{f=1}^{g} w_f Z_{fk}
\end{aligned}
$$

Clearly, the system efficiency is the product of the two process efficiencies; that is, $E_k^S = E_k^{(1)} \times E_k^{(2)}$.

Model (6.3) may produce multiple solutions for the two process efficiencies. When this happens, the two process efficiencies do not have common bases for comparison. To make $E_j^{(1)}$ (and $E_j^{(2)}$) of different DMUs comparable, Kao and Hwang (2008) suggested using the maximum value of $E_k^{(1)}$ or $E_k^{(2)}$ for comparison, depending on which process is considered more important. Suppose the first process is of major concern, and the maximum value of $E_k^{(1)}$ is sought. The objective function of Model (6.3) is replaced by the formula of Process 1 efficiency, with the system efficiency maintained at the level of E_k^S obtained from Model (6.3). In symbols, it is:

$$\begin{aligned} E_k^{(1)} = \max. & \sum_{f=1}^{g} w_f Z_{fk} \\ \text{s.t.} \quad & \sum_{i=1}^{m} v_i X_{ik} = 1 \\ & \sum_{r=1}^{s} u_r Y_{rk} = E_k^S \sum_{i=1}^{m} v_i X_{ik} \\ & \sum_{f=1}^{g} w_f Z_{fj} - \sum_{i=1}^{m} v_i X_{ij} \leq 0, \quad j = 1, \ldots, n \\ & \sum_{r=1}^{s} u_r Y_{rj} - \sum_{f=1}^{g} w_f Z_{fj} \leq 0, \quad j = 1, \ldots, n \\ & u_r, v_i, w_f \geq \varepsilon, \quad r = 1, \ldots, s, \quad i = 1, \ldots, m, \quad f = 1, \ldots, g \end{aligned} \tag{6.4}$$

Note that in this formulation the constraints corresponding to the system have been deleted for simplicity.

To calculate the (input) technical efficiency of Process 1, one simply replaces the part related to the CCR model in Model (6.4) by that of BCC Model (6.1), and the model is thus:

$$\begin{aligned} T_k^{(1)} = \max. & \sum_{f=1}^{g} \widetilde{w}_f Z_{fk} - \widetilde{w}_0 \\ \text{s.t.} \quad & \sum_{i=1}^{m} \widetilde{v}_i X_{ik} = 1 \\ & \sum_{r=1}^{s} u_r Y_{rk} = E_k^S \sum_{i=1}^{m} v_i X_{ik} \\ & \sum_{f=1}^{g} w_f Z_{fj} - \sum_{i=1}^{m} v_i X_{ij} \leq 0, \quad j = 1, \ldots, n \\ & \sum_{r=1}^{s} u_r Y_{rj} - \sum_{f=1}^{g} w_f Z_{fj} \leq 0, \quad j = 1, \ldots, n \\ & \left(\sum_{f=1}^{g} \widetilde{w}_f Z_{fj} - \widetilde{w}_0 \right) - \sum_{i=1}^{m} \widetilde{v}_i X_{ij} \leq 0, \quad j = 1, \ldots, n \\ & u_r, v_i, w_f, \widetilde{w}_f \geq \varepsilon, \quad r = 1, \ldots, s, \quad i = 1, \ldots, m, \quad f = 1, \ldots, g \\ & \widetilde{w}_0 \text{ unrestricted in sign} \end{aligned} \tag{6.5}$$

Models (6.4) and (6.5) can be combined to calculate the overall and (input) technical efficiencies of Process 1 at the same time:

$$
\begin{aligned}
\text{max.}\ & \sum_{f=1}^{g} w_f Z_{fk} + \left(\sum_{f=1}^{g} \widetilde{w}_f Z_{fk} - \widetilde{w}_0 \right) \\
\text{s.t.}\ & \sum_{i=1}^{m} v_i X_{ik} = 1 \\
& \sum_{i=1}^{m} \widetilde{v}_i X_{ik} = 1 \\
& \sum_{r=1}^{s} u_r Y_{rk} = E_k^S \sum_{i=1}^{m} v_i X_{ik} \\
& \sum_{f=1}^{g} w_f Z_{fj} - \sum_{i=1}^{m} v_i X_{ij} \le 0, \quad j = 1, \ldots, n \\
& \sum_{r=1}^{s} u_r Y_{rj} - \sum_{f=1}^{g} w_f Z_{fj} \le 0, \quad j = 1, \ldots, n \\
& \left(\sum_{f=1}^{g} \widetilde{w}_f Z_{fj} - \widetilde{w}_0 \right) - \sum_{i=1}^{m} \widetilde{v}_i X_{ij} \le 0, \quad j = 1, \ldots, n \\
& u_r, v_i, w_f, \widetilde{w}_f \ge \varepsilon, \quad r = 1, \ldots, s, \quad i = 1, \ldots, m, \quad f = 1, \ldots, g \\
& \widetilde{w}_0 \text{ unrestricted in sign}
\end{aligned}
\tag{6.6}
$$

At optimality, the overall and (input) technical efficiencies of Profess 1 are calculated as:

$$
E_k^{(1)} = \sum_{f=1}^{g} w_f Z_{fk}
$$

$$
T_k^{(1)} = \sum_{f=1}^{g} \widetilde{w}_f Z_{fk} - \widetilde{w}_0
$$

Consequently, the input scale efficiency is calculated as their ratio:

$$
S_k^{(1)} = E_k^{(1)} / T_k^{(1)}
$$

Similarly, the (output) technical efficiency of Process 2 is calculated by replacing the part related to the CCR model in Model (6.4) by that of output

BCC Model (6.2), while maintaining the system efficiency at E_k^S and Process 1 efficiency at $E_k^{(1)}$:

$$
\begin{aligned}
\frac{1}{T_k^{(2)}} = \min. & \sum_{f=1}^{g} \hat{w}_f Z_{fk} + \hat{w}_0 \\
\text{s.t.} \ & \sum_{r=1}^{s} \hat{u}_r Y_{rk} = 1 \\
& \sum_{r=1}^{s} u_r Y_{rk} = E_k^S \sum_{i=1}^{m} v_i X_{ik} \\
& \sum_{f=1}^{g} w_f Z_{fk} = E_k^{(1)} \sum_{i=1}^{m} v_i X_{ik} \\
& \sum_{f=1}^{g} w_f Z_{fj} - \sum_{i=1}^{m} v_i X_{ij} \leq 0, \qquad j = 1, \ldots, n \\
& \sum_{r=1}^{s} u_r Y_{rj} - \sum_{f=1}^{g} w_f Z_{fj} \leq 0, \qquad j = 1, \ldots, n \\
& \left(\sum_{f=1}^{g} \hat{w}_f Z_{fj} + \hat{w}_0 \right) - \sum_{r=1}^{s} \hat{u}_r Y_{rj} \geq 0, \ j = 1, \ldots, n \\
& u_r, v_i, w_f, \hat{w}_f \geq \varepsilon, \quad r = 1, \ldots, s, \quad i = 1, \ldots, m, \quad f = 1, \ldots, g \\
& \hat{w}_0 \text{ unrestricted in sign}
\end{aligned}
\tag{6.7}
$$

At optimality, the (output) technical efficiency of Process 2 is:

$$T_k^{(2)} = 1 \Big/ \left(\sum_{f=1}^{g} \hat{w}_f Z_{fk} + \hat{w}_0 \right)$$

and the corresponding (output) scale efficiency is:

$$S_k^{(2)} = E_k^{(2)} / T_k^{(2)}$$

If Process 2 is considered more important, then $E_k^{(2)}$ and the associated $T_k^{(2)}$ are calculated first. The overall efficiency of Process 1 is calculated as $E_k^{(1)} = E_k^S / E_k^{(2)}$, and the technical and scale efficiencies are calculated by a similar procedure.

Combining the above discussions, one obtains the following properties:

$$
\begin{aligned}
E_k^S &= E_k^{(1)} \times E_k^{(2)} = \left(T_k^{(1)} \times S_k^{(1)} \right) \times \left(T_k^{(2)} \times S_k^{(2)} \right) \\
E_k^S &= T_k \times S_k = \left(T_k^{(1)} \times T_k^{(2)} \right) \times \left(S_k^{(1)} \times S_k^{(2)} \right)
\end{aligned}
$$

where T_k and S_k are the technical and scale efficiencies of the system, respectively.

6.5 Non-life Insurance Companies in Taiwan

Kao and Hwang (2008) developed a relational model to calculate the system and process efficiencies for two-stage systems, and used an example of 24 non-life insurance companies in Taiwan to illustrate it. In order to have a common basis for comparison, the same data set is used in this paper to calculate the technical and scale efficiencies of the system and processes.

The operations of a non-life insurance company can be separated into two processes, premium acquisition and profit generation. In the first process, clients are attracted to pay direct written premiums, and reinsurance premiums are received from other insurance companies. In the second process, premiums are loaned and invested to earn profit. The inputs are classified into two categories:

Operating expenses (X_1): salaries of the employees and various costs incurred in daily operations.
Insurance expenses (X_2): expenses paid to agencies, brokers, and solicitors, and other expenses associated with marketing insurance.

The intermediate products considered are:

Direct written premiums (Z_1): premiums received from insured clients.
Reinsurance premiums (Z_2): premiums received from ceding companies.

The outputs include two types of profit:

Underwriting profit (Y_1): profit earned from the insurance business.
Investment profit (Y_2): profit earned from the investment portfolio.

Table 6.3 shows the original data.

By applying Model (6.3), the system efficiency, E_k^S, is calculated for each company, as shown in the second column of Table 6.4. Model (6.6) is then applied to calculate the overall and (input) technical efficiencies, $E_k^{(1)}$ and $T_k^{(1)}$, for the first process. The ratio of $E_k^{(1)}$ to $T_k^{(1)}$ is the (input) scale efficiency, $S_k^{(1)}$, and that of E_k^S to $E_k^{(1)}$ is the overall efficiency of the second process, $E_k^{(2)}$. The results are shown in the central part of Table 6.4 under the heading "Process 1". Finally, Model (6.7) is used to calculate the (output) technical efficiencies, $T_k^{(2)}$, for the second process. Similar to Process 1, the ratio of $E_k^{(2)}$ to $T_k^{(2)}$ is the (output) scale efficiency, $S_k^{(2)}$, of Process 2. The results are shown on the right side of Table 6.4 under the heading "Process 2". The products of the two process technical efficiencies and two process scale efficiencies are the system technical and system scale efficiencies, respectively, as shown on the left side of Table 6.4 under the heading "System".

As indicated by the scores shown in the second column of Table 6.4, none of the 24 companies is efficient for the whole system. This is simply because none of the companies is efficient for both Processes 1 and 2, although four companies have an efficient Process 1 and two have an efficient Process 2. In order to get a general idea of the performance of both processes, and to preserve the relationship of

Table 6.3 Data of 24 non-life insurance companies in Taiwan

Company	Operating expenses (X_1)	Insurance expenses (X_2)	Direct written premiums (Z_1)	Reinsurance premiums (Z_2)	Underwriting profit (Y_1)	Investment profit (Y_2)
1. Taiwan Fire	1,178,744	673,512	7,451,757	856,735	984,143	681,687
2. Chung Kuo	1,381,822	1,352,755	10,020,274	1,812,894	1,228,502	834,754
3. Tai Ping	1,177,494	592,790	4,776,548	560,244	293,613	658,428
4. China Mariners	601,320	594,259	3,174,851	371,863	248,709	177,331
5. Fubon	6,699,063	3,531,614	37,392,862	1,753,794	7,851,229	3,925,272
6. Zurich	2,627,707	668,363	9,747,908	952,326	1,713,598	415,058
7. Taian	1,942,833	1,443,100	10,685,457	643,412	2,239,593	439,039
8. Ming Tai	3,789,001	1,873,530	17,267,266	1,134,600	3,899,530	622,868
9. Central	1,567,746	950,432	11,473,162	546,337	1,043,778	264,098
10. The First	1,303,249	1,298,470	8,210,389	504,528	1,697,941	554,806
11. Kuo Hua	1,962,448	672,414	7,222,378	643,178	1,486,014	18,259
12. Union	2,592,790	650,952	9,434,406	1,118,489	1,574,191	909,295
13. Shingkong	2,609,941	1,368,802	13,921,464	811,343	3,609,236	223,047
14. South China	1,396,002	988,888	7,396,396	465,509	1,401,200	332,283
15. Cathay Century	2,184,944	651,063	10,422,297	749,893	3,355,197	555,482
16. Allianz President	1,211,716	415,071	5,606,013	402,881	854,054	197,947
17. Newa	1,453,797	1,085,019	7,695,461	342,489	3,144,484	371,984
18. AIU	757,515	547,997	3,631,484	995,620	692,731	163,927
19. North America	159,422	182,338	1,141,950	483,291	519,121	46,857
20. Federal	145,442	53,518	316,829	131,920	355,624	26,537
21. Royal & Sunalliance	84,171	26,224	225,888	40,542	51,950	6,491
22. Asia	15,993	10,502	52,063	14,574	82,141	4,181
23. AXA	54,693	28,408	245,910	49,864	0.1	18,980
24. Mitsui Sumitomo	163,297	235,094	476,419	644,816	142,370	16,976
Mean	1,544,215	828,963	7,832,893	667,964	1,602,873	477,733

multiplicity, geometric means for the system efficiency, E_k^S, Process 1 efficiency, $E_k^{(1)}$, and Process 2 efficiency, $E_k^{(2)}$, are calculated, as shown in the last row of Table 6.4. As expected, the system mean, 0.3705, is the product of the two process means, 0.7831 and 0.4731. The values also indicate that Process 1 is more efficient than Process 2. This is reasonable, because the task of premium acquisition is relatively straightforward, and neither significant mistakes nor breakthroughs can be made. Profit generation, in contrast, involves high risk, which may produce large differences among companies. Therefore, the efficiencies are high for the former

Table 6.4 Various efficiencies measured from the two-stage model for 24 non-life insurance companies in Taiwan

	System			Process 1			Process 2		
Co.	Overall	(Tech.	Scale)	Overall	(Tech.	Scale)	Overall	(Tech.	Scale)
1	0.6992	(0.7845	0.8914)	0.9926	(1.0000	0.9926)	0.7045	(0.7845	0.8980)
2	0.6246	(0.7240	0.8626)	0.9982	(1.0000	0.9982)	0.6256	(0.7240	0.8642)
3	0.6900	(0.6903	0.9996)	0.6900	(0.6903	0.9996)	1.0000	(1.0000	1.0000)
4	0.3042	(0.3284	0.9264)	0.7242	(0.7258	0.9979)	0.4200	(0.4524	0.9283)
5	0.7670	(1.0000	0.7670)	0.8307	(1.0000	0.8307)	0.9233	(1.0000	0.9233)
6	0.3897	(0.5357	0.7273)	0.9606	(0.9636	0.9969)	0.4057	(0.5559	0.7297)
7	0.2766	(0.4654	0.5943)	0.6706	(0.7520	0.8918)	0.4124	(0.6189	0.6664)
8	0.2752	(0.6816	0.4037)	0.6630	(0.8156	0.8130)	0.4150	(0.8358	0.4966)
9	0.2233	(0.2955	0.7557)	1.0000	(1.0000	1.0000)	0.2233	(0.2955	0.7557)
10	0.4658	(0.6403	0.7275)	0.8611	(0.8612	0.9999)	0.5409	(0.7434	0.7276)
11	0.1637	(0.3710	0.4414)	0.6476	(0.7406	0.8744)	0.2528	(0.5009	0.5047)
12	0.7596	(0.8658	0.8773)	1.0000	(1.0000	1.0000)	0.7596	(0.8658	0.8773)
13	0.2078	(0.7552	0.2752)	0.6720	(0.8652	0.7767)	0.3093	(0.8729	0.3543)
14	0.2886	(0.4236	0.6813)	0.6699	(0.7248	0.9243)	0.4309	(0.5845	0.7371)
15	0.6138	(0.9377	0.6546)	1.0000	(1.0000	1.0000)	0.6138	(0.9377	0.6546)
16	0.3202	(0.4153	0.7709)	0.8856	(0.9107	0.9724)	0.3615	(0.4560	0.7928)
17	0.3600	(0.7239	0.4974)	0.6276	(0.7239	0.8670)	0.5736	(1.0000	0.5736)
18	0.2588	(0.5502	0.4705)	0.7935	(1.0000	0.7935)	0.3262	(0.5502	0.5929)
19	0.4112	(0.7775	0.5288)	1.0000	(1.0000	1.0000)	0.4112	(0.7775	0.5288)
20	0.5465	(0.9990	0.5471)	0.9331	(0.9990	0.9340)	0.5857	(1.0000	0.5857)
21	0.2008	(0.2955	0.6795)	0.7321	(0.9131	0.8018)	0.2743	(0.3236	0.8476)
22	0.5895	(1.0000	0.5895)	0.5895	(1.0000	0.5895)	1.0000	(1.0000	1.0000)
23	0.4203	(0.6042	0.6957)	0.8426	(0.9877	0.8530)	0.4989	(0.6117	0.8155)
24	0.1348	(0.3769	0.3577)	0.4288	(1.0000	0.4288)	0.3144	(0.3769	0.8343)
Mean	0.3705	(0.5921	0.6257)	0.7831	(0.8953	0.8747)	0.4731	(0.6614	0.7153)

and low for the latter. An effective way to increase the efficiency of a company is thus to improve the performance of Process 2.

The overall efficiency of both processes can be decomposed into the product of technical and scale efficiencies. For Process 1, ten companies are technically efficient and four have perfect scale efficiency. Their averages, as shown in the last row of Table 6.4, are 0.8953 and 0.8747. The product of the technical and scale efficiencies is equal to the overall efficiency of the process (0.7831). For Process 2, five companies have perfect technical efficiency and two have perfect scale efficiency. The geometric means are 0.6614 and 0.7153, respectively, whose product is exactly the overall efficiency, 0.4731.

The products of the technical efficiencies and scale efficiencies of the processes are the technical efficiency and scale efficiency of the system, respectively. The last row of Table 6.4, under the heading "System", shows that the average technical and scale efficiencies of the system are 0.5921 and 0.6257, respectively, which are exactly the products of those of the two processes, 0.8953×0.6614 and

0.8747 × 0.7153. If every company is operating efficiently from the technical point of view, then the system efficiency can be improved from 0.3705 to 0.6257. This improvement is accomplished by reducing the amount of input of Process 1 by 10.47 % (=1−0.8953), and increasing the output of Process 2 by 33.86 % (=1 − 0.6614). Each company can thus identify sources of inefficiency and make appropriate amendments to improve its overall efficiency.

6.6 Conclusion

The measurement of scale efficiency is quite straightforward in conventional DEA when only the aggregate operation of the system is considered. However, when the operations of the individual processes of the system are also considered, the measurement becomes a little complicated. This is primarily because the relationship between the efficiencies calculated under CRS and VRS are not known. In this paper we investigate the simplest case, the two-stage system.

The problem in decomposing the technical efficiency of the system into those of the two processes is that the outputs of the first process are the inputs of the second, such that to improve the efficiency of the first by increasing its outputs will affect that of the second, and to improve the efficiency of the second by reducing its inputs will affect that of the first. To resolve this conflict, this paper fixes the amounts of the intermediate products, which are the outputs of the first process and the inputs of the second, and uses an input-oriented model to measure the technical efficiency of the first process and an output-oriented one to measure that of the second. Based on the relational model of Kao and Hwang (2008), where the overall system and process efficiencies are calculated first and the technical efficiencies of the processes are calculated second, the scale efficiency of each process is calculated as the ratio of their respective overall efficiency to technical efficiency. The product of the technical efficiency of the two processes is that of the system. Similarly, the product of the scale efficiency of the two processes is that of the system. Moreover, the overall efficiency of the system is the product of its technical and scale efficiencies.

Decomposing the system efficiency into the product of the two process efficiencies, and each process efficiency into the product of their respective technical and scale efficiencies, enables decision makers to identify the sources of inefficiency and to find effective alternatives for making improvements to the system.

The efficiency measures used in this paper are radial, but other measures are also discussed in the literature. How to calculate scale efficiency based on these other measures, so that the performance of the system can be improved more effectively, is one direction for future research. Finally, the network system discussed in this paper is the simplest one; and calculating the scale efficiency for more complicated systems will be a more challenging task.

References

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale efficiencies in data envelopment analysis. *Management Science, 30*, 1078–1092.
Charnes, A., & Cooper, W. W. (1984). The non-Archimedean CCR ratio for efficiency analysis: A rejoinder to Boyd and Färe. *European Journal of Operational Research, 15*, 333–334.
Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*, 429–444.
Charnes, A., Cooper, W. W., & Rhodes, E. (1979). Short communication: Measuring the efficiency of decision making units. *European Journal of Operational Research, 3*, 339.
Chen, Y., Cook, W. D., Ning, L., & Zhu, J. (2009). Additive efficiency decomposition in two-stage DEA. *European Journal of Operational Research, 196*, 1170–1176.
Chen, Y., Cook, W. D., & Zhu, J. (2010). Deriving the DEA frontier for two-stage processes. *European Journal of Operational Research, 202*, 138–142.
Cook, W. D., Liang, L., & Zhu, J. (2010). Measuring performance of two-stage network structure by DEA: A review. *Omega, International Journal of Management Science, 38*, 423–430.
Färe, R., & Grosskopf, S. (2000). Network DEA. *Socio-Economic Planning Sciences, 34*, 35–49.
Kao, C. (2009). Efficiency decomposition in network data envelopment analysis: A relational model. *European Journal of Operational Research, 192*, 949–962.
Kao, C. (2013). Efficiency decomposition in network data envelopment analysis with slacks-based measures, Omega (2013), http://dx.doi.org/10.1016/j.omega.2013.12.002
Kao, C., & Hwang, S. N. (2008). Efficiency decomposition in two-stage data envelopment analysis: An application to non-life insurance companies in Taiwan. *European Journal of Operational Research, 185*, 418–429.
Kao, C., & Hwang, S. N. (2011). Decomposition of technical and scale efficiencies in two-stage production systems. *European Journal of Operational Research, 211*, 515–519.
Seiford, L. M., & Zhu, J. (1999). Profitability and marketability of the top 55 US commercial banks. *Management Science, 45*, 1270–1288.
Tone, K., & Tsutsui, M. (2009). Network DEA: A slacks-based measure approach. *European Journal of Operational Research, 197*, 243–252.