Scale efficiency in two-stage network DEA

Kun Chen & Joe Zhu

To cite this article: Kun Chen & Joe Zhu (2019) Scale efficiency in two-stage network DEA, Journal of the Operational Research Society, 70:1, 101-110

To link to this article: https://doi.org/10.1080/01605682.2017.1421850

Published online: 16 Feb 2018.

Submit your article to this journal

Article views: 80

View Crossmark data
Scale efficiency in two-stage network DEA

Kun Chen and Joe Zhu

ABSTRACT
Network data envelopment analysis (DEA) considers internal structures of decision-making units. Unlike the standard DEA, network DEA imposes hurdles for measuring scale efficiency due to the fact that (i) overall efficiency aggregated by the stage or divisional technical efficiencies is highly non-linear and only solvable in a heuristic manner, or (ii) the overall efficiency which concerns exclusively inputs and outputs of a system is difficult to be decomposed into divisional efficiencies. In this paper, we establish a mathematical transformation to convert the corresponding non-linear programming problem into second order cone programming programme. The transformation is shown to be versatile in dealing with both constant returns to scale and variable returns to scale models under the two-stage network DEA. Meanwhile, our numerical results reveal that overall scale efficiency in two-stage network DEA is consistent with scale efficiency in conventional DEA.

1. Introduction
Data envelopment analysis (DEA) is a data-oriented methodology for performance evaluation and benchmarking. One of the typical uses of DEA is measuring the technical efficiency under the assumption of variable returns to scale (VRS) (see, for example, Banker, Charnes, & Cooper, 1984). The original constant returns to scale (CRS) model of Charnes, Cooper, & Rhodes, 1978 measures the technical and scale efficiency. The ratio between the CRS and VRS scores is defined as scale efficiency. In other words, the overall CRS efficiency can be decomposed into technical efficiency and scale efficiency. Under the standard DEA, all these efficiency scores are not greater than one.

In recent years, researchers start to examine the internal structures of peer decision-making units (DMUs). The resulting DEA approach is referred to as network DEA (Cook, Tone, & Zhu, 2014). The internal structures of DMUs can be as simple as a two-stage production process or as complex as a network with many nodes and arcs (Kao, 2016). In a basic two-stage production process, the first stage of a DMU uses inputs to generate outputs, and all of the outputs become inputs to the second stage. These first stage outputs that link the second stage are called intermediate measures. The second stage then generates final outputs. For example, Chilingerian and Sherman (2004) describe a two-stage process in measuring physician care. Their first stage is a manager-controlled process with inputs including registered nurses, medical supplies, and capital and fixed costs. These inputs generate the outputs or intermediate measures, including patient days, quality of treatment, drugs dispensed, among others. The outputs of the second physician controlled stage include research grants, quality of patients, and quantity of individuals trained, by specialty. Note that a general two-stage network structure can have outputs that leave the system and do not become the inputs to the second stage, and/or (additional) inputs to the second stage.

There are two types of approaches to solving two-stage network DEA. One is called additive approach where the overall efficiency is defined as a weight average of the two stage efficiency scores (Chen, Cook, Li, & Zhu, 2009). The other is called multiplicative approach where the overall efficiency is aggregated or decomposed as a product of the two-stage efficiency scores Kao and Hwang (2008) and Liang, Cook, and Zhu (2008).

Measuring scale efficiency in two-stage network DEA cannot be extended directly from the standard DEA approaches due to the limitations inherent in the additive and multiplicative two-stage network approaches. In additive two-stage DEA, the efficiency scores of CRS and VRS models are contingent on predetermined weights.
(Guo, Shureshjani, Foroughi, & Zhu, 2017). Thus, in theory, there could be infinitely many scale efficiencies which are too tedious to compute.

While the multiplicative approach does not require a priori weights on the two stage efficiencies, measuring scale efficiency in the aggregation form of multiplicative two-stage network DEA is still a challenge even for the basic two-stage network structure studied in Kao and Hwang (2011). The difficulty lies in the fact that VRS model in multiplicative approach is a highly non-linear optimisation problem which cannot be transformed into linear problem using the Charnes–Cooper transformation (Charnes & Cooper, 1962).

To circumvent this difficulty, researchers switch to decomposition method to measure scale efficiency in two-stage network DEA using both multiplier and envelopment forms (Kao & Hwang, 2011; Sahoo, Zhu, Tone, & Klemen, 2014). In particular, in order to fix the level of intermediate measures whose changes are likely to affect both divisional efficiencies simultaneously, Kao and Hwang (2011) propose to use the input-oriented VRS model for first stage and the output-oriented VRS model for second stage. However, after carefully examining VRS models proposed by Kao and Hwang (2011), we find that the link between the first and second stages, which is the key characteristic of network DEA, is not well defined in measuring technical divisional efficiency. Kao and Hwang (2011) merely compute the technical efficiency of each stage separately, and combine the results as the overall technical efficiency. To improve upon Kao and Hwang (2011), one has to solve the two-stage network DEA under VRS which is highly non-linear.

Another limitation of a decomposition method (e.g., Kao & Hwang, 2011) is that it cannot be applied to general two-stage network structures where some outputs from stage 1 can leave the system and/or stage 2 may have additional inputs. This limitation of decomposition approach is due to its definition of overall efficiency which only considers inputs and outputs of a system, and therefore overlooks the relationship between stage efficiency and overall efficiency. In other words, although the defined overall efficiency could be derived, method for computing stage efficiency is usually not clear. Consequently, using decomposition approach to measure scale efficiency for individual stage in general network is also impractical.

Given the above issue with the decomposition approach, it appears that the aggregation approach to measuring scale efficiency in two-stage network DEA is a viable way because in the aggregation approach, the overall efficiency can be defined as a geometric mean of the individual stages’ scores. Although the proposed aggregation method is highly non-linear, we show that it can be solved by second order cone programming (SOCP) technique which is already a mature technique in the discipline of convex optimisation (Boyd & Vandenberghe, 2004; Lobo, Vandenberghe, Boyd, & Lebret, 1998). Note that Chen and Zhu (2017) applied SOCP to solve multiplicative aggregation model in two-stage network DEA, but scale efficiency for each stage has not been studied. Compared with the decomposition method proposed by Kao and Hwang (2011), our aggregation model for measuring scale efficiency addresses the link between the CRS and VRS models in multiplicative two-stage network DEA appropriately and can be applied in a general two-stage network structure.

Note that one can also use the heuristic method to solve for the overall efficiency which is defined as an aggregation of individual stages’ efficiency scores, see, for example, Li, Lei, Dai, and Liang (2015). However, heuristic approach does not guarantee global optimal solution to be found. In optimisation theory, heuristic method is usually the last choice. For heuristic search method, it is well known that its accuracy to global optimal solution depends on continuity property and unimodal property of optimal solutions as a function of varying parameters. In comparison, our approach is a sequential convex optimisation technique.

The reminder of the paper is organised as follows. In Section 2, we discuss the decomposition and aggregation methods in measuring scale efficiency of two-stage network DEA, and point out the major contribution of the current study. In Sections 3 and 4, we use SOCP to correctly model and measure scale efficiency in two-stage DEA with basic and general two-stage network structures. Section 5 concludes.

2. Measuring scale efficiency: decomposition vs. aggregation

We consider a general two-stage network structure shown in Figure 1. Each DMU, $j = 1, 2, \ldots, n$ has $m$ inputs $x_{ij}$ ($i = 1, 2, \ldots, m$) to the first stage and $P$ outputs $y_{pj}^1$ ($p = 1, 2, \ldots, P$) that leave the system. In addition to these $P$ outputs, stage 1 has $D$ outputs $y_{pj}^2$ ($d = 1, 2, \ldots, D$) called intermediate measures or links that become inputs to the second stage. The second stage has its own

**Figure 1.** General two-stage network structure.
inputs \( x_{i,j}^h \) \((h = 1, 2, \ldots, H)\). The outputs from the second stage are \( y_{r,j} \) \((r = 1, 2, \ldots, s)\).

When \( y_{r,j} \) and \( x_{i,j}^h \) are absent, Figure 1 becomes the standard or basic two-stage network structure studied in Kao and Hwang (2008) and Liang et al. (2008), as shown in Figure 2.

Under the CRS assumption when the basic two-stage network structure is considered, a relational model is proposed to compute overall efficiency (Kao, 2009).

\[
\begin{align*}
\text{max} & \quad \frac{\sum_{r=1}^{s} u_r y_{r,0}}{\sum_{j=1}^{m} v_j x_{j,0}} \\
\text{s.t.} & \quad \sum_{j=1}^{m} y_{j,0} \leq 1, \quad \forall j \\
& \quad \sum_{j=1}^{m} v_j x_{j,0} \leq 1, \quad \forall j \\
& \quad \eta_j, u, v \geq \varepsilon, \quad \forall d, r, i
\end{align*}
\]

where \( v_j, \eta_j, \) and \( u \) are weights which are assumed to be positive in the current study, by incorporating the small non-Archimedean \( \varepsilon \) in the DEA models. Note that \( \eta_j \) is weights on the links or intermediate measures between the two stages and are assumed to be equal.

Note that in the above model the overall efficiency for the standard two-stage network as shown in Figure 1 is defined as

\[
\sum_{r=1}^{s} u_r y_{r,0} + \sum_{p=1}^{p} w_j y_{j,0}^p \\
\sum_{j=1}^{m} v_j x_{j,0} + \sum_{h=1}^{H} Q_h x_{h,0}
\]

where the intermediate measures, \( z_{d,j} \), are ignored and only the inputs to the first stage (treated as the system inputs) and the outputs from the second stage (treated as the system outputs) are considered. In the same manner, if one only considers the system inputs and outputs, the overall efficiency of the general two-stage network shown in Figure 1 can be defined as

\[
\sum_{r=1}^{s} u_r y_{r,0} + \sum_{p=1}^{p} w_j y_{j,0} \\
\sum_{j=1}^{m} v_j x_{j,0} + \sum_{h=1}^{H} Q_h x_{h,0}
\]

In the discussion to follow, we will demonstrate that the overall efficiency defined in (3) is impossible to be decomposed into individual component efficiency scores.

However, one can apply the decomposition approach to (2) where one derives divisional efficiency scores, denoted as \( E_r^1 \) and \( E_r^2 \) for the first and second stages, respectively, from the obtained overall efficiency denoted as \( E_c \). Kao (2009) proposes the following decomposition model when the overall efficiency in (2) is fixed at the optimality.

\[
\begin{align*}
\text{max} & \quad \frac{\sum_{r=1}^{s} u_r y_{r,0}}{\sum_{j=1}^{m} v_j x_{j,0}} \\
\text{s.t.} & \quad \sum_{j=1}^{m} y_{j,0} \leq 1, \quad \forall j \\
& \quad \sum_{j=1}^{m} v_j x_{j,0} \leq 1, \quad \forall j \\
& \quad \eta_j, u, v \geq \varepsilon, \quad \forall d, r, i
\end{align*}
\]

Note that the maximal objective value of model (4) is \( E_c^1 \) and \( E_c^2 \) can be computed as \( E_c^1 = \frac{E_c}{E_c^2} \). Also, note that there is an alternative way of decomposing overall efficiency \( E_c \) by replacing the objective function in model (2) with \( \frac{\sum_{r=1}^{s} u_r y_{r,0}}{\sum_{j=1}^{m} v_j x_{j,0}} \). Obviously, these two types of decompositions can yield different divisional efficiencies.

In order to compute scale efficiency while maintaining the same levels of intermediate measures between the two stages, Kao and Hwang (2011) suggest to utilise the input-oriented VRS model to compute technical efficiency of the first stage and the output-oriented VRS model to measure technical efficiency of the second stage. These two models are summarised as follows.

\[
\begin{align*}
\text{max} & \quad \frac{\sum_{r=1}^{s} u_r y_{r,0} + \sum_{d=1}^{d} w_j y_{j,0}^d}{\sum_{j=1}^{m} v_j x_{j,0}} \\
\text{s.t.} & \quad \sum_{j=1}^{m} y_{j,0} \leq 1, \quad \forall j \\
& \quad \sum_{j=1}^{m} v_j x_{j,0} \leq 1, \quad \forall j \\
& \quad \eta_j, u, v \geq \varepsilon, \quad \forall d, r, i, \quad u^1 \text{ free in sign}
\end{align*}
\]

\[
\begin{align*}
\text{max} & \quad \frac{\sum_{r=1}^{s} u_r y_{r,0} + \sum_{d=1}^{d} w_j y_{j,0}^d}{\sum_{j=1}^{m} v_j x_{j,0}} \\
\text{s.t.} & \quad \sum_{j=1}^{m} y_{j,0} \leq 1, \quad \forall j \\
& \quad \sum_{j=1}^{m} v_j x_{j,0} \leq 1, \quad \forall j \\
& \quad \eta_j, u, v \geq \varepsilon, \quad \forall d, r, i, \quad u^2 \text{ free in sign}
\end{align*}
\]
The advantage of decomposition approach is that the objective function remains as one linear fractional form as in the standard DEA and the corresponding problem can be transformed into a linear model by the Charnes–Cooper transformation (Charnes & Cooper, 1962). Then classic issues associated with DEA such as scale efficiency and frontier projection can possibly be studied in accord with classic theories of conventional DEA (Chen, Cook, Kao, & Zhu, 2013).

However, there are two issues with the decomposition approach. First, the weights on the intermediate measures denoted as \( \eta_j \) in two-stage DEA is not correctly modelled in models (5) and (6). Note that models (5) and (6) are actually two independent optimisation problems and there is no linkage between them. Although the link may appear to be the same weight \( \eta_j \) used in models (5) and (6), \( \eta_j \) can have different optimal values in models (5) and (6) due to the fact that the two models are solved independently. Namely, there is no reason to assume that \( \eta_j \) yields the same value when both models achieve optimal solutions separately. As a result, the link in measuring the scale efficiency in the method proposed by Kao and Hwang (2011) can be a false one. In addition, since in models (5) and (6), different \( \eta_j \) values can be generated, the VRS models are found to be independent of CRS models which are inconsistent with conventional DEA Chen et al. (2013). While the proposed joint use of the input- and output-oriented models is the way to address the scale efficiency in two-stage network DEA, in order to establish a more accurate measure of scale efficiency, improvements are needed to correctly model the link between the two stages and extend to general network structure.

Second, decomposition approach is restricted to the basic two stage network structure depicted in Figure 2 and cannot be applied to general network structure shown in Figure 1. Note that the overall efficiency under model (1) can be expressed as the product of efficiency scores of the two stages \( \frac{\sum_{p=1}^{n} \eta_p x_{0p} \cdot \sum_{m=1}^{M} \eta_m y_{0m}}{\sum_{i=1}^{n} y_{ix} \cdot \sum_{m=1}^{M} \eta_m y_{0m}} \). Because the weighted intermediate measures can cancel out, the overall efficiency or the objective function in model (1) under Figure 1 becomes \( \frac{\sum_{m=1}^{M} \eta_m y_{0m}}{\sum_{m=1}^{M} \eta_m y_{0m}} \). This is the reason that the decomposition approach works. In fact, under the basic two-stage network structure, model (1) can also be regarded as an aggregation approach, namely both decomposition and aggregation approaches are identical in model (1).

However, in the case of general network structure, the first and second stage efficiency are expressed as \( \frac{\sum_{m=1}^{M} \eta_m y_{0m} \cdot \sum_{p=1}^{n} \eta_p x_{0p}}{\sum_{m=1}^{M} \eta_m y_{0m} \cdot \sum_{p=1}^{n} \eta_p x_{0p}} \) and \( \frac{\sum_{m=1}^{M} \eta_m y_{0m}}{\sum_{m=1}^{M} \eta_m y_{0m} + \sum_{m=1}^{M} Q_m y_{0m}} \), where \( v_j \), \( \eta_j \), \( \lambda_j \), \( u_j \), and \( Q_j \) are weights. It is extremely difficult or impossible to decompose the overall efficiency defined in (3) into the above two divisional efficiencies. That is the reason equation (3) is hardly used as a definition of the overall efficiency.

Therefore, the current study suggests an aggregation approach which defines the overall efficiency of the general two-stage network as

\[
\frac{\sum_{j=1}^{J} \eta_{j0} x_{d0} + \sum_{j=1}^{J} \eta_{j0} y_{r0}}{\sum_{j=1}^{J} v_{j0} x_{d0} \cdot \sum_{j=1}^{J} v_{j0} y_{r0} + \sum_{h=1}^{H} \eta_{h0} y_{r0}} \cdot \sum_{j=1}^{J} \eta_{j0} y_{r0} \cdot \sum_{h=1}^{H} Q_{h0} y_{r0} \quad (7)
\]

Note that (7) is extremely non-linear and beyond the application scope of the Charnes–Cooper transformation. Although (7) cannot be expressed as a single linear fractional term as in the objective function of model (1), Li et al. (2015) propose a heuristic method to solve it. In optimisation theory, it is well known that the accuracy of heuristic search method depends on continuity property and unimodal property of optimal solutions as a function of varying parameters. Since examining these properties usually is not an easy task, adopting heuristic search method could be a risky choice in finding global optimal solution.

The current study shows that the corresponding non-linear optimisation problem can be solved non-heuristically by the mature technique of SOCP. With this technical breakthrough, aggregation approach not only addresses the link in two-stage network DEA appropriately in formulas, but also is able to solve and be utilised to measure scale efficiency regardless of network structures.

The next section shows that since the link or multipliers for intermediate measures is incorporated or embedded in a single objective function as a single vector parameter, it has to be identical while the corresponding optimisation problem is being solved.

3. Scale efficiency under basic two-stage network structure

3.1. The VRS model

As shown in the previous section, the aggregation approach corresponds with decomposition method for CRS model under the basic two-stage network structure. Hence, we only study VRS model under basic two-stage network structure in this section. Following Kao’s method, we consider input-oriented model for the first stage and output-oriented model for the second stage.

\[
\max \left( \frac{\sum_{i=1}^{I} \eta_i x_{0i} + u^i}{\sum_{i=1}^{I} v_i x_{0i}} \cdot \frac{\sum_{j=1}^{J} \eta_j y_{0j}}{\sum_{j=1}^{J} \eta_j y_{0j} + \sum_{j=1}^{J} Q_j y_{0j}} \right) \\
\text{s.t. } \frac{\sum_{i=1}^{I} \eta_i x_{0i} + u^i}{\sum_{i=1}^{I} v_i x_{0i}} \leq 1, \quad \forall j \\
\frac{\sum_{j=1}^{J} \eta_j y_{0j}}{\sum_{j=1}^{J} \eta_j y_{0j} + \sum_{j=1}^{J} Q_j y_{0j}} \leq 1, \quad \forall j \\
\eta_j, v_i, y_{rj} \geq \epsilon, \quad \forall d, r, i \\
u^i, u^i \text{ free in sign}
\]
Note that unlike in Kao’s method, the link denoted as \( \eta_d \) has to take the same optimal value in (8). Subsequently, we show that problem (8) corresponds to a sequence of feasibility problem of SOCP. Model (8) is equivalent to the following.

\[
\begin{align*}
\min & \quad \sum_{i=1}^{D} \eta_d z_{d0} + u^2 \\
\text{s.t.} & \quad \frac{\sum_{i=1}^{D} u_i y_{i0}}{\sum_{i=1}^{D} u_i y_{i0}} \leq 1, \quad \forall j \\
& \quad \frac{\sum_{i=1}^{D} u_i y_{i0}}{\sum_{i=1}^{D} u_i y_{i0}} \leq 1, \quad \forall j \\
& \quad \eta_d u_i v_i \geq \epsilon, \quad \forall d, r, i \\
& \quad u^1, u^2 \text{ free in sign}
\end{align*}
\]

By an epigraph transformation, we have

\[
\begin{align*}
\min & \quad \theta \\
\text{s.t.} & \quad \frac{\sum_{i=1}^{D} u_i y_{i0} - \sum_{i=1}^{D} u_i y_{i0}^2}{\sum_{i=1}^{D} u_i y_{i0}^2} \leq \theta \\
& \quad \frac{\sum_{i=1}^{D} u_i y_{i0}}{\sum_{i=1}^{D} u_i y_{i0}} \leq 1, \quad \forall j \\
& \quad \frac{\sum_{i=1}^{D} u_i y_{i0}}{\sum_{i=1}^{D} u_i y_{i0}} \leq 1, \quad \forall j \\
& \quad \eta_d u_i v_i \geq \epsilon, \quad \forall d, r, i \\
& \quad u^1, u^2 \text{ free in sign}
\end{align*}
\]

Thus, the model (10) is transformed equivalently to the following model

\[
\begin{align*}
\min & \quad \bar{\theta} \\
\text{s.t.} & \quad \frac{1}{2} \left( \bar{\theta} \sum_{i=1}^{D} \eta_d z_{d0} + \theta u^1 - \sum_{i=1}^{D} u_i y_{i0} \right) \leq 1, \quad \forall j \\
& \quad \frac{\sum_{i=1}^{D} u_i y_{i0}}{\sum_{i=1}^{D} u_i y_{i0}} \leq 1, \quad \forall j \\
& \quad \eta_d u_i v_i \geq \epsilon, \quad \forall d, r, i \\
& \quad u^1, u^2 \text{ free in sign}
\end{align*}
\]

When \( \bar{\theta} \) is given, model (16) is a feasibility problem in the form of SOCP. Therefore, model (8) can be solved by searching the true value of \( \bar{\theta} \), denoted as \( \bar{\theta}^* \), in its domain. Specially, let the initial guess be \( \bar{\theta}^1 \). If problem (16) is feasible, we have \( \bar{\theta}^* \leq \bar{\theta}^1 \). Otherwise, we obtain \( \bar{\theta}^* > \bar{\theta}^1 \). Continuing this process, we can approximate \( \bar{\theta} \) on an accurate level.

It is well known that global optimal solution of convex optimisation which includes SOCP and linear programming is guaranteed by supporting hyperplane theorem (Boyd & Vandenberghe, 2004). Besides, SOCP is also known as a generalisation of linear programming. In general case, conic constraint can be written as \( \|Ax + b\|_2 \leq c^T x + d \) where \( x \in R^n \) are the decision variables, \( A \in R^{m \times n} \), \( b \in R^m \), \( c \in R^m \) and \( d \in R \). If \( A = 0 \), then the corresponding SOCP reduces to a (general) linear programming. Moreover, from the perspective of algorithm, both linear programming of large scale data-sets and SOCP adopt interior point method which is close to polynomial complexity and very efficient in practice. Thus, SOCP is considered not only as effective and accurate as the linear programming, but also capable of dealing with a larger class of non-linear but convex problems as shown in our current study. In that sense, SOCP is considered as effective and accurate as the linear programming. Therefore, in our current study, SOCP is a liable alternative to linear programming.

Suppose the optimal objective value of model (8) is denoted as \( T_v \). To ensure unique divisional technical efficiency, we propose the following model by fixing the value of \( T_v \). For simplicity, we only consider the case of maximising the efficiency of first stage.

\[
\begin{align*}
\max & \quad \sum_{i=1}^{D} u_i y_{i0} \sum_{i=1}^{D} \eta_d z_{d0} + u^2 \\
\text{s.t.} & \quad \frac{\sum_{i=1}^{D} u_i y_{i0}}{\sum_{i=1}^{D} u_i y_{i0}} \leq 1, \quad \forall j \\
& \quad \frac{\sum_{i=1}^{D} u_i y_{i0}}{\sum_{i=1}^{D} u_i y_{i0}} \leq 1, \quad \forall j \\
& \quad \eta_d u_i v_i \geq \epsilon, \quad \forall d, r, i \\
& \quad u^1, u^2 \text{ free in sign}
\end{align*}
\]
Since $\sum_{i=1}^{m} \eta_{i} y_{i} + u^{i} \leq T_{v}$ is equivalent to the intersection of the following two inequalities.

$$\sum_{i=1}^{D} \eta_{i} z_{d0} + u^{i} \leq T_{v}$$

(17)

Then, by similar transformations shown informal as (10)-(15), we find that formula (17) and formula (18) correspond to the following two conic constraints respectively.

$$\sum_{i=1}^{m} v_{i} x_{i0} + T_{v} \sum_{i=1}^{D} \eta_{i} z_{d0} + T_{v} u^{i} \leq T_{v}$$

(19)

Also maximising $\sum_{i=1}^{m} v_{i} x_{i0}$ is equivalent to minimising $\sum_{i=1}^{m} v_{i} x_{i0}$ and $\sum_{i=1}^{m} v_{i} x_{i0} \leq \theta$ can also be converted into a conic constraint. Finally, the corresponding SOCP form of problem (16) is as follows.

$$\min \theta$$

s.t. $$\left( \begin{array}{c} \sum_{i=1}^{m} v_{i} x_{i0} + 1 \\ \theta - \sum_{i=1}^{D} \eta_{i} z_{d0} + u^{i} \end{array} \right) \leq \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \forall j$$

(21)

$$\sum_{i=1}^{m} v_{i} x_{i0} \leq \sum_{i=1}^{m} \sum_{j=1}^{D} \eta_{j} z_{d0} + u^{i}$$

$$\sum_{i=1}^{m} v_{i} x_{i0} \leq 1, \forall j$$

$$\eta_{d} u_{r} v_{i} \geq 1, \forall d, r, i$$

$$u^{i}, u^{2} \text{ free in sign}$$

Compared with the decomposition method proposed by Kao and Hwang (2011) which derives (input) scale efficiency and (output) scale efficiency separately, our method computes (input) scale efficiency and (output) scale efficiency simultaneously in model (8). Specially, given models (8) and (16) are solved, we have unique (input) technical efficiency of the first stage as $T_{v}^{i} = \sum_{i=1}^{m} \frac{v_{i} x_{i0}}{v_{i}}$ and unique (output) technical efficiency of the second stage as $T_{v}^{o}$. Then, (input) scale efficiency and (output) scale efficiency can be computed as $E_{i}^{i} = E_{i}^{i} \times T_{v}^{i}$ and $E_{i}^{o} = E_{i}^{o} \times T_{v}^{o}$ respectively.

Note that if model (16) and its alternative model which is maximising technical efficiency of the second stage yield different decompositions of overall technical efficiency, there will be multiple divisional scale efficiency with respect to a single stage in network DEA. Thus, unlike the scale efficiency in the conventional DEA, network scale efficiency for each stage may not be unique.

3.2. Numerical example: Taiwanese non-life insurance companies

We consider the data set of 24 Taiwanese non-life insurance companies used in Kao and Hwang (2008). The two inputs to the first stage (premium acquisition) are operating expenses and insurance expenses. The intermediate measures (or the outputs from the first stage) are direct written premiums and reinsurance premiums. The outputs of the second stage (profit generation) are underwriting profit and investment profit.

Note that Kao and Hwang (2008) only compute CRS model with respect to this data set. Although Chen et al. (2009) develop an additive efficiency approach to process VRS model and demonstrate it using this data set, the proposed approach depends on a set of predetermined weights.

The results of our approach are summarised in Table 1. According to Chen et al. (2009), we know the CRS model of this data set has a unique decomposition. Coincidentally, model (16) and its alternative model of maximising $\sum_{i=1}^{m} \frac{v_{i} x_{i0}}{v_{i}}$ reveal that VRS model of this data set also has a unique decomposition. Subsequently, in this case, divisional scale efficiency is also unique. The relationship among these efficiencies are $E_{i} = E_{i}^{i} \times T_{v}^{i}$, $T_{v}^{i} = T_{v}^{i} \times T_{v}^{o}$, $E_{i} = E_{i}^{i} \times E_{i}^{o}$. $E_{i}^{o} = E_{i}^{o} \times T_{v}^{o}$, $E_{i}^{i} = E_{i}^{i} \times T_{v}^{i}$, and $E_{i} = E_{i} \times T_{v}^{i}$.
Table 1. Scale efficiency in basic two-stage network.

<table>
<thead>
<tr>
<th>DMU</th>
<th>CRS model</th>
<th>VRS model</th>
<th>Scale efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$E_1$</td>
<td>$E_2$</td>
<td>$E_3$</td>
</tr>
<tr>
<td>1 Taiwan Fire</td>
<td>0.993</td>
<td>0.704</td>
<td>0.699</td>
</tr>
<tr>
<td>2 Chung Kuo</td>
<td>0.998</td>
<td>0.626</td>
<td>0.625</td>
</tr>
<tr>
<td>3 Tai Ping</td>
<td>0.690</td>
<td>1</td>
<td>0.690</td>
</tr>
<tr>
<td>4 China Mariners</td>
<td>0.724</td>
<td>0.420</td>
<td>0.304</td>
</tr>
<tr>
<td>5 Fubon</td>
<td>0.831</td>
<td>0.923</td>
<td>0.767</td>
</tr>
<tr>
<td>6 Zurich</td>
<td>0.961</td>
<td>0.406</td>
<td>0.390</td>
</tr>
<tr>
<td>7 Tai'an</td>
<td>0.671</td>
<td>0.412</td>
<td>0.277</td>
</tr>
<tr>
<td>8 Ming Tai</td>
<td>0.663</td>
<td>0.415</td>
<td>0.275</td>
</tr>
<tr>
<td>9 Central</td>
<td>1</td>
<td>0.223</td>
<td>0.223</td>
</tr>
<tr>
<td>10 The First</td>
<td>0.862</td>
<td>0.541</td>
<td>0.466</td>
</tr>
<tr>
<td>11 Kuo Hua</td>
<td>0.647</td>
<td>0.253</td>
<td>0.164</td>
</tr>
<tr>
<td>12 Union</td>
<td>1</td>
<td>0.760</td>
<td>0.760</td>
</tr>
<tr>
<td>13 Shing Kong</td>
<td>0.672</td>
<td>0.309</td>
<td>0.208</td>
</tr>
<tr>
<td>14 South China</td>
<td>0.670</td>
<td>0.431</td>
<td>0.289</td>
</tr>
<tr>
<td>15 Cathay Century</td>
<td>1</td>
<td>0.614</td>
<td>0.614</td>
</tr>
<tr>
<td>16 Allianz President</td>
<td>0.886</td>
<td>0.362</td>
<td>0.320</td>
</tr>
<tr>
<td>17 Newa</td>
<td>0.628</td>
<td>0.574</td>
<td>0.360</td>
</tr>
<tr>
<td>18 AIU</td>
<td>0.794</td>
<td>0.326</td>
<td>0.259</td>
</tr>
<tr>
<td>19 North America</td>
<td>1</td>
<td>0.411</td>
<td>0.411</td>
</tr>
<tr>
<td>20 Federal</td>
<td>0.933</td>
<td>0.586</td>
<td>0.547</td>
</tr>
<tr>
<td>21 Royal &amp; &amp; Alliance</td>
<td>0.732</td>
<td>0.274</td>
<td>0.201</td>
</tr>
<tr>
<td>22 Asia</td>
<td>0.590</td>
<td>1</td>
<td>0.590</td>
</tr>
<tr>
<td>23 AXA</td>
<td>0.843</td>
<td>0.499</td>
<td>0.420</td>
</tr>
<tr>
<td>24 Mitsu Sumitomo</td>
<td>0.429</td>
<td>0.314</td>
<td>0.135</td>
</tr>
</tbody>
</table>

Notes: $E_j$, $T_j$ and $E_i$ stand for overall efficiency, technical efficiency and scale efficiency, respectively. superscripts 1 and 2 stand for different stage.

4. Scale efficiency in general two-stage network

4.1. The CRS model

In the general two-stage network structure, there are additional inputs in the second stage and some outputs leave the system. Therefore, aggregation method and decomposition method diverge from each other since the weighted intermediate measures cannot cancel out in multiplicative form. The corresponding non-linear programming problem of aggregation approach is as follows.

min $\theta$

s.t. $\sum_{j=1}^{m} \eta_j x_{j0} + \sum_{j=1}^{m} \lambda_j y_{j0} + \sum_{j=1}^{m} Q_{j0} x_{j0}^2 \leq \theta \sum_{j=1}^{m} \eta_j x_{j0} + \sum_{j=1}^{m} \lambda_j y_{j0} + \sum_{j=1}^{m} Q_{j0} x_{j0}^2 \leq 1$, $\forall j$

$\eta_j, \lambda_j, y_{j0}, Q_{j0} \geq \epsilon$, $\forall d, r, i, h$ (23)

Likewise, problem (23) may not ensure unique divisional efficiency and we need to solve the following optimisation problem to examine the existence of multiple divisional efficiencies by fixing the overall efficiency $E_t$ obtained from problem (23). Also, for simplicity, we only consider the case of maximising the efficiency of the first stage.

max $\sum_{j=1}^{m} \eta_j x_{j0} + \sum_{j=1}^{m} \lambda_j y_{j0}$

s.t. $\sum_{j=1}^{m} \eta_j x_{j0} + \sum_{j=1}^{m} \lambda_j y_{j0} + \sum_{j=1}^{m} Q_{j0} x_{j0}^2 \leq 1$, $\forall j$

$\eta_j, \lambda_j, y_{j0}, Q_{j0} \geq \epsilon$, $\forall d, r, i, h$ (24)

Similarly, problem (22) can first be converted into a minimising problem in reciprocal form and then transformed into the following feasibility problem of SOCP.

$\sum_{j=1}^{m} \eta_j x_{j0} + \sum_{j=1}^{m} \lambda_j y_{j0} \leq \theta$ is a conic constraint in disguise and $\sum_{j=1}^{m} \eta_j x_{j0} + \sum_{j=1}^{m} \lambda_j y_{j0} + \sum_{j=1}^{m} Q_{j0} x_{j0}^2 = E_t$ also corresponds to one conic constraint. Eventually, model (24) is equivalent to the following feasibility problem of SOCP form.
4.2. The VRS model

Based upon the VRS model in the basic two-stage network structure, we adopt the combination of input- and output-oriented models to compute overall technical efficiency in general two-stage network structure. The corresponding VRS model is as follows.

\[
\begin{align*}
\min \quad & \theta \\
\text{s.t.} \quad & \left\| 1 + \sum_{i=1}^{m} v_i x_{i0} + \sum_{p=1}^{\rho} \lambda_p y_{p0} \right\|_2 \leq \theta + \sum_{d=1}^{D} \eta_d z_{d0} \\
& \left\| \sum_{i=1}^{m} v_i x_{i0} + E_i \sum_{d=1}^{D} \eta_d z_{d0} + E_c \sum_{h=1}^{H} Q_h x_{h0} \right\|_2 \leq \sum_{d=1}^{D} \eta_d z_{d0} + \sum_{p=1}^{\rho} \lambda_p y_{p0} + \sum_{r=1}^{R} u_r y_{r0} \\
& \sum_{d=1}^{D} \eta_d z_{d0} + \sum_{p=1}^{\rho} \lambda_p y_{p0} + \sum_{r=1}^{R} u_r y_{r0} \leq \sum_{i=1}^{m} v_i x_{i0} + E_i \sum_{d=1}^{D} \eta_d z_{d0} + E_c \sum_{h=1}^{H} Q_h x_{h0} \\
& \sum_{i=1}^{m} v_i x_{i0} - E_i \sum_{d=1}^{D} \eta_d z_{d0} - E_c \sum_{h=1}^{H} Q_h x_{h0} \leq \sum_{i=1}^{m} v_i x_{i0} + E_i \sum_{d=1}^{D} \eta_d z_{d0} + E_c \sum_{h=1}^{H} Q_h x_{h0} \\
\end{align*}
\]

(25)

Similarly, model (28) can be converted equivalently into the following SOCP form.

\[
\begin{align*}
\min \quad & \theta \\
\text{s.t.} \quad & \left\| 1 + \sum_{i=1}^{m} v_i x_{i0} \right\|_2 \leq \theta + \sum_{d=1}^{D} \eta_d z_{d0} + \sum_{p=1}^{\rho} \lambda_p y_{p0} + u^1 \\
& \left\| \sum_{i=1}^{m} v_i x_{i0} + T_v \sum_{d=1}^{D} \eta_d z_{d0} + T_c \sum_{h=1}^{H} Q_h x_{h0} + T_{uv} \right\|_2 \leq \sum_{d=1}^{D} \eta_d z_{d0} + \sum_{p=1}^{\rho} \lambda_p y_{p0} + u^1 + \sum_{r=1}^{R} u_r y_{r0} \\
& \sum_{i=1}^{m} v_i x_{i0} - T_v \sum_{d=1}^{D} \eta_d z_{d0} - T_c \sum_{h=1}^{H} Q_h x_{h0} - T_{uv} \leq \sum_{i=1}^{m} v_i x_{i0} + T_v \sum_{d=1}^{D} \eta_d z_{d0} + T_c \sum_{h=1}^{H} Q_h x_{h0} + T_{uv} \\
\end{align*}
\]

(29)

4.3. Numerical example: Regional R&D process

As a demonstrative example of measuring scale efficiency in general network structure, we consider the data-set of regional R&D process of 30 Provincial level regions in China studied by Li, Chen, Jiang, and Xie (2012) and Guo et al. (2017).

In the first stage (the technology development process), the inputs are: R&D personnel, R&D expenditure and the proportion of regional science and technology funds in regional total financial expenditure. The outputs of the first stage are the number of patents and papers which are also inputs to the second stage (namely, these are intermediate measures). The second stage (the economic application) also has an input of contract value in technology market. The final outputs are GDP, total
exports, and urban per capita annual income and gross output of high-tech industry. Note that in this case, $y_{ij}^{\lambda}$ are not present.

The results based on our approach are summarised in Table 2. The divisional efficiency of each stage is known to be unique for the data-set (Li et al., 2012). The results of model (24) and model (28) indicate that divisional scale efficiency of each stage is also unique in this case. Further, similar to the results in the basic network structure, divisional scale efficiencies are found to be greater than one for the first and the second stages. For example, for DMU 2, the scale efficiency of the first stage is 1.145. For DMU 20, the scale efficiency of the second stage is 1.058. Moreover, consistent with results from basic network structure, overall efficiency is found to be less than one uniformly.

5. Conclusions

While computing scale efficiency in two-stage DEA using decomposition method (Kao & Hwang 2011), the link between individual stages in VRS models is not carefully modelled. Further, in order to keep the objective function as a single fractional form, decomposition method is limited to the basic two-stage network structure. In the current study, we discover that aggregation model for measuring scale efficiency correctly characterises the link between individual stages. In addition, aggregation approach allows general two-stage network structures.

We develop a SOCP-based transformation to solve the non-linear programming problem corresponding to the aggregation approach. The mathematical transformation we established in current study is flexible. It can be applied to both CRS and VRS models under the two-stage network DEA with general network structures as long as the objective function or the constraints remains the form of the product of two linear fractional terms.

The results of the numerical examples imply that overall scale efficiency in the two-stage network DEA is in accord with scale efficiency in conventional DEA with respect to the fact that VRS efficiency is always not less than the CRS efficiency. While the divisional VRS efficiency can be less than CRS efficiency. Since the combination of two input-oriented VRS models is found to yield overall scale efficiency being greater than one (Chen et al., 2009), the numerical finding in current study may serve as a validation of combined use of input- and output-oriented models for measuring technical and scale efficiencies.

Acknowledgements

The authors thank the helpful suggestions and comments made by an anonymous reviewer. The study is funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

Disclosure statement

No potential conflict of interest was reported by the authors.
References