Efficiency Measurement of Multistage Processes: Context Dependent Numbers of Stages

Wade D. Cook
Schulich School of Business, York University
4700 Keele Street, Toronto, Ontario M3J 1P3, Canada
wcook@schulich.yorku.ca

Chuanyin Guo
School of Economics and Management
Beihang University, Beijing 100191, P. R. China
cyguo@buaa.edu.cn

Wanghong Li*
Schulich School of Business, York University
4700 Keele Street, Toronto, Ontario M3J 1P3, Canada
whl@schulich.yorku.ca

Zhepeng Li
Schulich School of Business, York University
4700 Keele Street, Toronto, Ontario M3J 1P3, Canada
zli@schulich.yorku.ca

Liang Liang
He Fei University of Technology
He Fei, An Hui Province 230009, P. R. China
liang@ustc.edu.cn

Joe Zhu
School of Business
Worcester Polytechnic Institute, Worcester, USA

Received 20 January 2017
Revised 22 June 2017
Accepted 14 August 2017
Published 23 November 2017

An important area of research involving the benchmarking methodology data envelopment analysis (DEA), concerns the modeling of multistage situations. In the usual

*Corresponding author.
multistage settings, it is generally assumed that all decision-making units (DMUs) have the same number and configuration of stages. However, in many real-world examples, this assumption does not hold. Consider, for example, a supply chain setting where for some DMUs, products are shipped directly from a supplier to a retailer (single-stage), while for other DMUs, products can be transshipped through distribution centers (two or more stages). In the current paper, we investigate an efficiency measurement situation where the DMUs exhibit a mix of single and two-stage setups. The particular case examined involves a set of high technology firms that can be thought of as falling into two groups; those firms where the output of interest is the annual revenue generated, and those that only generate revenue, but as well invest a portion of that revenue in R&D. Firms in the first group can be viewed as being single-stage DMUs while those in the other group are of the two-stage type. The modeling complication here is that the set of DMUs do not explicitly form a homogeneous set of units. We develop a DEA-style model aimed at measuring efficiency in the presence of such nonhomogeneous two-group structures.

Keywords: DEA; multistage; supply chains; nonhomogeneous; context dependent.

1. Introduction

In the nearly 40 years since the development of the data envelopment analysis (DEA) methodology by Charnes et al. (1978), the original concepts have been applied to an enormous number of practical problem settings, and the model structure has been extended in many directions. Recent surveys include Cook et al. (2009), Paradi and Zhu (2013), and Liu et al. (2013). One important area of investigation in DEA over the past several years has been that relating to multistage or network structures (Färe and Grosskopf, 1996), and in particular two-stage DEA. This latter, and its extensions to multi-stage situations has been particularly influential in such settings as supply chain management. See, for example, Kao (2009), Tone and Tsutsui (2009), and Liang et al. (2006). A survey of network models is provided by Cook et al. (2010). A recent book by Cook and Zhu (2014) provides a thorough coverage of recent research on multistage situations.

A number of different methodologies have been proposed for evaluating the efficiency of a DMU when that DMU consists of multiple stages. One of the first approaches to two-stage processes was given by Seiford and Zhu (1993), and simply uses the standard DEA structure, applying the Constant Returns to Scale or CRS model (Charnes et al., 1978) to each of stage 1 and stage 2. In that context, the role of the intermediate variables is essentially ignored. One of the original models for two-stage processes that explicitly recognized the intermediate variables in their double role, was that put forward by Kao and Hwang (2008), and is a form of the cooperative or centralized approach. In their model, the overall efficiency is expressed as the product $e_o = e^1_o \cdot e^2_o$. It is important to point out that the Kao and Hwang (2008) approach is designed for those settings where a CRS methodology is appropriate. It is, as well, assumed in the Kao and Hwang model that the multistage setup is closed in the sense that the outputs from the first stage are pure intermediate variables, serving only as inputs to the next (second) stage. To
Efficiency Measurement of Multistage Processes

accommodate Variable Returns to Scale (VRS) settings [Banker et al., 1984], and those cases where some outputs at a given stage may leave the system, and not flow as inputs to the next stage. Chen et al. (2009) proposed an additive or arithmetic mean approach for combining the two-stages, as opposed to the geometric mean-type methodology suggested by Kao and Hwang (2008). Specifically, the proposal is to express the overall efficiency as a convex combination $e_o = w_1 e_{1o} + w_2 e_{2o}$ of the stage efficiency scores, where $w_1, w_2$ are chosen to reflect the relative importance of those scores. This is discussed below.

In the usual multistage problem settings to which DEA has been applied, it is generally assumed that all DMUs included in the comparison group possess the same numbers of inputs and outputs, and in particular, all DMUs have the same number of stages (each DMU consists of two-stages, three-stages, etc.) There are many examples in the real world, however, where this assumption does not hold. Consider the very common case in a supply chain setting where for some DMUs, a product sold by the associated retailer is shipped directly from the supplier to that retailer, while for other DMUs, the product flows from the supplier to a distribution center or warehouse, and from there it is sent to the retailer. So in this case some of the DMUs (supply chains) consist of a single-stage while others are of the two-stage type.

In Sec. 2, we investigate an efficiency measurement situation where the DMUs exhibit a mix of single and two-stage setups. The particular case examined involves a set of high technology firms that can be thought of as falling into two groups. Group 1 consists of firms where the outputs of interest are the annual revenue generated, and a soft factor representing innovation potential. Group 2 consists of firms that invest a portion of their annual revenues in R&D, the outcome of which is the creation of new products, patents and so on. Firms in group 1 can be viewed as being single-stage DMUs while those in group 2 are of the two-stage type. The modeling complication here is that the set of DMUs do not explicitly form a homogeneous set of units. Section 3 develops a DEA-style model aimed at measuring efficiency in the presence of such two-group structures. The model provides a measure of aggregate or overall efficiency for those DMUs of the two-stage type, while at the same time providing an efficiency score for each component or stage within a multistage setting. An important thing to recognize here is that the evaluation of any given stage of a DMU will be carried out by comparing that stage to similar stages in peer DMUs, even though the numbers of stages experienced by some peers differ from that of other peers. So in a supply chain setting, for example, this might mean that if some chains have two stages (e.g., supplier to central warehouse to retailer) while others have three stages (e.g., supplier to central warehouse to branch warehouse to retailer), all stages of the “supplier to central warehouse” type will be compared to one another even though some of them belong to two-stage DMUs and others to three-stage DMUs. Section 4 discusses the application of the proposed model to the
The indigenous innovation capability of high technology firms in China has become an important issue since the financial crisis of 2008. Accordingly, the conclusions of numerous research scholars are that science and technology inputs and outputs reflect the capability of such firms. This paper considers the problem of measuring the efficiencies of such decision-making units. The major issue is that while some of those firms reinvest a portion of earned revenues in research and development, with the promise of creating new products, awards and patents, other firms are purely interested in generating revenues. The members of the former group can be viewed as consisting of two-stages, while those of the latter group are single-stage processes. The issue, we address herein is how to evaluate the efficiencies of high technology companies as a group, when some behave as single-stage and some as two-stage processes.

One may, of course, postulate that a setting such as this can in fact be seen as constituting a three-stage setup. Specifically, one may take the position that payoffs in the form of awards and patents do not represent ends in themselves, but rather are a means to an end, with that end being further revenues. Furthermore, the generated follow up revenues can be viewed as a "feedback" input to stage 1.

In the problem setting herein, data on such “follow up payoffs” created by stage 2 outputs were not explicitly available. Hence, we examine only the two-stage setting.

Consider the following data on high tech firms where we have two groups of DMUs:

**Group 1** has $N_1$ DMUs where there is a single-stage with inputs

- $x_1$... Non-degree employees, unit: persons
- $x_2$... Undergraduate Degree employees, unit: persons
- $x_3$... Graduate Degree employees, unit: persons
- $x_4$... Registered capital: one hundred million

The outputs from this stage are

- $z^1$... Revenue generated: one hundred million
- $z^2$... Innovation ability: scores

We should point out here that in the case of DMUs in $N_1$, there is no portion of $z^1$ that is split off for R&D investment (i.e., there is no second stage). However, for DMUs in $N_2$, $z^1$ is split into two parts, $z^1_1$ and $z^1_2$, where $z^1_2$ is the revenue invested in R&D. Hence, DMUs in $N_1$ possess only 2 outputs $z^1_1, z^2$, while those in $N_2$ have 3 outputs $z^1_1, z^1_2, z^2$. For convenience, we define two revenue components $z^1_1$ and $z^1_2$ in both $N_1$ and $N_2$, noting that $z^1_2$ is zero in $N_1$. 

Conclusions and future research directions follow in Sec. 5.
Group 2 has $N_2$ DMUs, with each having 2 stages. See Fig. 1.

Stage 1, for DMUs in $N_2$, has the same input and output structure as DMUs in $N_1$. As indicated, on the output side, the revenue generated by any DMU in $N_2$ is split into 2 portions $z_1^1$ and $z_2^2$.

We note that $z_1^1$ leaves the DMU after stage 1 and only $z_2^2$ goes on as an input to stage 2.

Stage 2, for DMUs in $N_2$ has 2 inputs, namely $z_2^1$ and $z_2^2$, and outputs are $y_1$, ..., Patents: items $y_2$, ..., Awards: items $y_3$, ..., Inventions: items

3. Modeling Efficiency in a Mixed Multistage Situation

Some relevant literature

The idea of non-homogeneous DMUs in DEA is not new. Cook and Green (2004) presented a model which identifies for each DMU those outputs or groups of outputs that represent the core business of that DMU. As a result, outputs in some DMUs are transferred to other DMUs. Hence, one might view this as creating non-homogeneous DMUs from what are currently homogeneous units. In a more deliberate sense, Cook et al. (2013) look specifically at a set of DMUs where not all of the full set of outputs are produced at each DMU. Du et al. (2015) further extends this idea to examine DMUs with parallel subunits, whereby not all subunits have the same sets of inputs and outputs. Furthermore, there are intermediate factors connecting the subunits.

The above cited literature is primarily directed at single-stage problems where homogeneity is viewed from the output mix perspective as opposed to the number of stages exhibited by certain groups of DMUs versus other groups.

To derive an aggregate efficiency score and scores for each of the stages in two-stage DMUs, we proceed in two steps. Step 1 develops an approximate aggregate
efficiency score as well as approximate stage scores for each DMU. Step 2 seeks to
improve on those estimates.

**Step 1: Approximating efficiency scores in \( N_2 \)**

We first derive an aggregate efficiency score for each 2-stage DMU in \( N_2 \). An
important fact to recognize here, as pointed out above, is that we do not have a
closed serial system. In other words, part of the revenue \( z \) generated, namely \( z_1 \),
leaves the system, and therefore is not a pure intermediate variable to be passed
on as an input to stage 2. The R&D component \( z_1 \), is, however, an input to stage
2. This means that while one might normally apply the logic of Kao and Hwang
(2008), where one represents the aggregate efficiency as the product (geometric
mean) of the efficiency scores for the two-stages, that is \( e_{agg}^{o} = e_1^{o} \cdot e_2^{o} \), in the current
case with a non-closed system, such an approach leads to a highly nonlinear model.
Furthermore, in the case of the variable returns to scale (VRS) model of Banker
et al. (1984), the geometric mean methodology of Kao and Hwang (2008), which is
designed for the CRS model, is not immediately applicable.

An equally viable, but linear representation of aggregate efficiency in terms
of the stage level scores, can potentially be achieved by invoking the arithmetic-
mean methodology of Chen et al. (2009). The weights in that additive form are
endogenously chosen (within the model) as the proportions of total weighted inputs
allocated to the two-stages. However, rather than adopting the approach, which has been shown by Despotis et al. (2016) and by Ang and Chen (2016) to have weights biased toward the first stage, we choose instead to define the
aggregate efficiency as a weighted average (convex combination) of the efficiency
scores of the stages 1 and 2, namely,

\[
e_{agg}^{o} = w_1 e_1^{o} + w_2 e_2^{o}
\]  

with the weights being set exogenously.

While the approach of Chen et al. (2009) permits the resulting model to be
converted into a linear program, the above model does not have such a linear
representation. However, given the approaches of Liang et al. (2006), Lim and Zhu
(2013), and Ang and Chen (2016), one can solve (for an approximate solution to)
the resulting DEA model by treating it as a parametric linear program, provided
an appropriate set of bounds for that parameter can be derived. Specifically, near-
optimal solutions can be obtained via linear models when we search over a given
range for the parameter in question.

**Modeling Aggregate Efficiency:**

Given the above explanation, we begin by stating the stage 1 and stage 2 effi-
ciency representations under variable returns to scale (VRS) for DMUs in \( N_2 \),
specifically:
Efficiency Measurement of Multistage Processes

Stage 1 model:

\[
e_{1}^{o} = \max \frac{\eta_{1}^{1}z_{1o} + \eta_{2}^{1}z_{2o} + \eta^{2}z^{2} + \eta^{0}}{\sum_{i=1}^{4} \nu_{i}x_{io}}
\]

s.t. \[\eta_{1}^{1}z_{1j} + \eta_{2}^{1}z_{2j} + \eta^{2}z^{2} + \eta^{0} - \sum_{i=1}^{4} \nu_{i}x_{ij} \leq 0, \quad \forall j \in N_{1} \cup N_{2}\]

\[\eta_{1}^{1}, \eta_{2}^{1}, \eta^{2}, \nu_{i}, \nu_{r} \geq 0, \quad \forall i, r, \eta^{0} \text{ unrestricted in sign}\]

Comment: We point out that with the VRS technology, it is not necessary to include constraints such as \[\sum_{i=1}^{3} u_{r}y_{rj} + u^{0} - (\eta_{1}^{1}z_{2j} + \eta^{2}z^{2}) \leq 0, \forall j \in N_{2}\], connecting the inputs and outputs for stage 2. This is due to the presence of the unrestricted variable \(u^{0}\) that can take any value negative, positive or zero, hence these constraints are redundant. It is important, however, to point out that such may not be the case for the CRS technology, where this variable is dropped, and the constraints are thereby invoked. The reason for including these constraints in the CRS case is that the multipliers \(\eta_{1}^{1}, \eta^{2}\) appear on both the input and output sides (i.e., in both stages.) This same line of argument applies to stage 2.

Stage 2 model:

\[
e_{2}^{o} = \max \frac{\sum_{r=1}^{3} u_{r}y_{ro} + u^{0}}{\eta_{2}^{1}z_{2o} + \eta^{2}z^{2}}
\]

s.t. \[\sum_{r=1}^{3} u_{r}y_{rj} + u^{0} - (\eta_{2}^{1}z_{2j} + \eta^{2}z^{2}) \leq 0, \forall j \in N_{2}\]

\[\eta_{1}^{1}, \eta_{2}^{1}, \eta^{2}, \nu_{i}, \nu_{r} \geq 0, \quad \forall i, r, \nu^{0} \text{ unrestricted in sign}\]

From the previous discussion, the aggregate efficiency model for each DMU \(j_{o}\) in \(N_{2}\), as per (1), is then given by

\[
e_{agg}^{o} = \max \left[ w_{1} \frac{\sum_{d=1}^{2} \eta_{d}^{1}z_{do} + \eta^{2}z^{2} + \eta^{0}}{\sum_{i=1}^{4} \nu_{i}x_{io}} + w_{2} \frac{\sum_{r=1}^{3} u_{r}y_{ro} + u^{0}}{\eta_{2}^{1}z_{2o} + \eta^{2}z^{2}} \right]
\]

s.t. \[\sum_{r=1}^{3} u_{r}y_{rj} + u^{0} - (\eta_{2}^{1}z_{2j} + \eta^{2}z^{2}) \leq 0, \forall j \in N_{2}\]

\[\eta_{1}^{1}, \eta_{2}^{1}, \eta^{2}, \nu_{i}, \nu_{r} \geq 0, \quad \forall i, r, \nu^{0}, u^{0} \text{ unrestricted in sign}\]

The fractional programming problem (1) is highly nonlinear, and we propose a parametric programming model to derive an approximate solution. A number of approaches have been suggested in the literature for approximating the solution to problems of this form, including Liang et al. (2006), Lim and Zhu (2013), and Ang and Chen (2016). Along the lines of some of the previous approaches, we argue that if we knew the value of the efficiency score for one of the two-stages (say the
stage 1 score \( e^1 \), in the aggregate model (4), then the other stage score (for example stage 2) can be derived from a modified version of that model, namely

\[
e^o_{agg} = \max \left[ \frac{w_1 e^1 + w_2 \sum_{r=1}^{3} u_r y_{r_0} + u^0}{\eta_2^1 z_{2,0} + \eta_2^2 z_{2,0}^2} \right]
\]

s.t. \[
\sum_{r=1}^{3} u_r y_{r_j} + u^0 - (\eta_2^1 z_{2,0} + \eta_2^2 z_{2,0}^2) \leq 0 \quad \forall j \in N_2
\]

\[
\eta_1^1 z_{1,j} + \eta_2^1 z_{2,j} + \eta_2^2 z_{2,j}^2 + \eta^0 - \sum_{i=1}^{4} \nu_i x_{ij} \leq 0, \quad \forall j \in N_1 \cup N_2
\]

\[
\eta_1^1 z_{1,0} + \eta_2^1 z_{2,0} + \eta^2 z_{2,0}^2 + \eta^0 - e^1 \sum_{i=1}^{4} \nu_i x_{i0} = 0
\]

\[
\eta_2^1, \eta_2^2, \eta^0, \nu_i, \mu_r \geq 0, \quad \forall i, r, u^0, \eta^0, \text{unrestricted in sign}
\]

Note that the constraint \( \eta_1^1 z_{1,0} + \eta_2^1 z_{2,0} + \eta^2 z_{2,0}^2 + \eta^0 - e^1 \sum_{i=1}^{4} \nu_i x_{i0} = 0 \) in (5) requires that the stage 1 score be maintained in solving for the stage 2 score. Replacing the objective function in (5) by the reduced version

\[
e^o_2 = \max \frac{\sum_{r=1}^{3} u_r y_{r_0} + u^0}{\eta_2^1 z_{2,0} + \eta_2^2 z_{2,0}^2},
\]

that model can now be converted to a linear programming form utilizing the Charnes and Cooper (1962) transformation. Specifically, let \( t = \frac{1}{\eta_2^1 z_{2,0} + \eta_2^2 z_{2,0}^2} \) and define the variables \( \pi_1^1 = t \eta_1^1, \pi_2^1 = t \eta_2^1, \pi_2^2 = t \eta_2^2, \mu_r = t u_r, \nu^0 = t u^0, \omega_i = t \nu_i \).

Then (5) becomes the linear model (6)

\[
e^o_2 = \max \sum_{r=1}^{3} \mu_r y_{r_0} + \mu^0
\]

s.t. \[
\sum_{r=1}^{3} \mu_r y_{r_j} + \mu^0 - (\pi_2^1 z_{2,0} + \pi_2^2 z_{2,0}^2) \leq 0 \quad \forall j \in N_2
\]

\[
\pi_1^1 z_{1,j} + \pi_2^1 z_{2,j} + \pi_2^2 z_{2,j}^2 + \pi^0 - \sum_{i=1}^{4} \omega_i x_{ij} \leq 0, \quad \forall j \in N_1 \cup N_2
\]

\[
\pi_1^1 z_{1,0} + \pi_2^1 z_{2,0} + \pi_2^2 z_{2,0}^2 + \pi^0 - e^1 \sum_{i=1}^{4} \omega_i x_{i0} = 0
\]

\[
\pi_1^1, \pi_2^1, \pi_2^2, \omega_i, \mu_r \geq 0, \quad \forall i, r, \mu^0, \pi^0, \text{unrestricted in sign}
\]

\[
\pi_2^1 z_{2,0} + \pi_2^2 z_{2,0}^2 = 1
\]

Approximating the aggregate efficiency- a parametric approach:

To solve model (6), we treat, as a parameter, the efficiency score \( e^1 \) for stage 1 of the DMU “o” under investigation. (We could equally use the stage 2 score as a parameter.) The procedure we adopt involves establishing a range over which that
parameter can vary, and then solving (5) (or its linear version (6)) at a selected set of points in that range. First, we determine the maximum and minimum values of each of the stage 1 and stage 2 efficiency scores.

Model (2) yields, for any \( j = a \), the maximum value \( \bar{e}_1 \), of the stage 1 efficiency, and (3), the maximum value \( \bar{e}_2 \), of the stage 2 efficiency. To derive the minimum values for these two scores, we solve (7) for the minimum value for \( e^1 \), and (8) for the minimum of \( e^2 \), specifically

\[
\bar{e}_1 = \max \pi_1^1 z_{1o}^1 + \pi_2^1 z_{2o}^1 + \pi^2 z_o^2 + \pi^0 \\
\text{s.t.} \sum_{r=1}^{3} \mu_r y_{rj} + \mu^0 - (\pi_2^{1} z_{2j}^{1} + \pi^2 z_j^2) \leq 0 \quad \forall j \in N_2 \\
\pi_1^{1} z_{1j}^{1} + \pi_2^{1} z_{2j}^{1} + \pi^2 z_j^2 + \pi^0 - \sum_{i=1}^{4} \omega_i x_{ij} \leq 0, \quad \forall j \in N_1 \cup N_2 \\
\sum_{r=1}^{3} \mu_r y_{ro} + \mu^0 - \bar{e}_2 (\pi_2^{1} z_{2o}^{1} + \pi^2 z_o^2) = 0 \\
\sum_{i=1}^{4} \omega_i x_{io} = 1 \\
\pi_1^1, \pi_2^1, \pi^2, \omega_i, \mu_r \geq 0, \quad \forall i, r, \mu^0, \pi^0, \text{unrestricted in sign}
\]

\[
\bar{e}_2 = \max \sum_{r=1}^{3} \mu_r y_{ro} + \mu^0 \\
\text{s.t.} \sum_{r=1}^{3} \mu_r y_{rj} + \mu^0 - (\pi_2^{1} z_{2j}^{1} + \pi^2 z_j^2) \leq 0 \quad \forall j \in N_2 \\
\pi_1^{1} z_{1j}^{1} + \pi_2^{1} z_{2j}^{1} + \pi^2 z_j^2 + \pi^0 - \sum_{i=1}^{4} \omega_i x_{ij} \leq 0, \quad \forall j \in N_1 \cup N_2 \\
\pi_1^{1} z_{1o}^{1} + \pi_2^{1} z_{2o}^{1} + \pi^2 z_o^2 + \pi^0 - \bar{e}_1 \sum_{i=1}^{4} \omega_i x_{io} = 0 \\
\pi_2^{1} z_{2o}^{1} + \pi^2 z_o^2 = 1 \\
\pi_1^1, \pi_2^1, \pi^2, \omega_i, \mu_r \geq 0, \quad \forall i, r, \mu^0, \pi^0, \text{unrestricted in sign}
\]

We point out that unlike in the case of model (2), the comment concerning redundant constraints does not necessarily apply here. In other words, the constraints involving stage 2, namely, \( \sum_{r=1}^{3} \mu_r y_{rj} + \mu^0 - (\pi_2^{1} z_{2j}^{1} + \pi^2 z_j^2) \leq 0 \forall j \in N_2 \) are still retained in (7). We do this since the constraint that fixes the stage 2 score at its maximum, namely \( \sum_{r=1}^{3} \mu_r y_{ro} + \mu^0 - \bar{e}_2 (\pi_2^{1} z_{2o}^{1} + \pi^2 z_o^2) = 0 \) must be present, thereby constraining or having an effect on the “unrestricted” nature of the variable \( \mu^0 \). Similar comments apply to model (3).
It is noted that the minimum value of the stage 1 efficiency score is assumed when the stage 2 score is at its maximum, and vice versa, the stage 2 score assumes its minimum when the stage 1 score is at its maximum.

Having established the lower and upper limits \([\bar{e}_1, \hat{e}_1]\) for the optimal stage 1 efficiency score \(e_1\), we choose a small increment \(\delta\) (e.g., .01), select a set of values, say \(\bar{e}_1, \hat{e}_1 + \delta, \hat{e}_1 + 2\delta, \ldots, \hat{e}_1\), in that range, and then solve (6) at each of those points. At each value of the parameter, \(\bar{e}_1 + h\delta\), for each \(h = 1, 2, 3, \ldots\), compute \(e_{agg}^h = w_1 e_1^h + w_2 e_2^h\), where \(e_1^h = \bar{e}_1 + h\delta\) and \(e_2^h\) is the solution arising from (6). Let \(\hat{h}\) denote that value of \(h\) yielding the largest aggregate efficiency score \(e_{agg}^{\hat{h}} = w_1 e_1^\hat{h} + w_2 e_2^\hat{h}\).

Having obtained, by way of (6), an approximation \(e_{agg}^{\hat{h}} = w_1 e_1^\hat{h} + w_2 e_2^\hat{h}\), of the optimal aggregate efficiency of the DMU (in \(N_2\)) under analysis, and in the process, approximate optimal scores \(e_1^\hat{h}, e_2^\hat{h}\) for stages 1 and 2, respectively, we now seek to improve on those approximations.

**Step 2: Improving the efficiency scores:**

One approach to deriving possible improvements in the efficiency scores is to adopt a leader/follower methodology, and attempt to increase the stage 1 and stage 2 scores. Suppose, we choose stage 1 as the leader, and solve the following linear programming problem.

\[
e_1^* = \max \sum_{r=1}^{3} \mu_r y_{rj} + \mu^0 - (\pi_2 z_{2j} + \pi^2 z_{2j}^2) \quad \forall j \in N_2
\]

\[
\text{s.t.} \quad \sum_{r=1}^{3} \mu_r y_{rj} + \mu^0 - (\pi_2 z_{2j} + \pi^2 z_{2j}^2) \leq 0 \quad \forall j \in N_2
\]

\[
\pi_1 z_{1j} + \pi_2 z_{2j} + \pi^2 z_{2j}^2 + \pi^0 - \sum_{i=1}^{4} \omega_i x_{ij} \leq 0, \quad \forall j \in N_1 \cup N_2
\]

\[
\sum_{r=1}^{3} \mu_r y_{r0} + \mu^0 - e_2^{\hat{h}} (\pi_2 z_{20} + \pi^2 z_{20}^2) \geq 0
\]

\[
\pi_1 z_{10} + \pi_2 z_{20} + \pi^2 z_{20}^2 + \pi^0 - e_1^{\hat{h}} \sum_{i=1}^{4} \omega_i x_{io} \geq 0
\]

\[
\sum_{i=1}^{4} \omega_i x_{io} = 1
\]

\[
\pi_1, \pi_2, \pi^0, \omega_i, \mu_r, \geq 0, \quad \forall i, r, \mu^0, \pi^0 \text{ unrestricted in sign}
\]

The constraint \(\pi_1 z_{10} + \pi_2 z_{20} + \pi^2 z_{20}^2 + \pi^0 - e_1^{\hat{h}} \sum_{i=1}^{4} \omega_i x_{io} \geq 0\) guarantees that the optimal objective function value will be at least as large as \(e_1^{\hat{h}}\). The constraint \(\sum_{r=1}^{3} \mu_r y_{r0} + \mu^0 - e_2^{\hat{h}} (\pi_2 z_{20} + \pi^2 z_{20}^2) \geq 0\) warrants that the current best efficiency score for stage 2, namely \(e_2^{\hat{h}}\), will not be compromised in the optimization process.
Efficiency Measurement of Multistage Processes

Treating stage 2 as the follower, consider the following linear model:

\[ e_2^* = \max \sum_{r=1}^{3} \mu_r y_{r0} + \mu^0 \]

s.t. \( \sum_{r=1}^{3} \mu_r y_{rj} + \mu^0 - (\pi_1^1 z_{1j}^1 + \pi_2^2 z_{2j}^1) \leq 0 \quad \forall j \in N_2 \)

\[ \pi_1^1 z_{1j}^1 + \pi_2^1 z_{2j}^1 + \pi^2 z_{2j}^2 + \pi^0 - \sum_{i=1}^{4} \omega_i x_{ij} \leq 0, \quad \forall j \in N_1 \cup N_2 \]

(10)

\[ \pi_1^1 z_{1o}^1 + \pi_2^1 z_{2o}^1 + \pi^2 z_{2o}^2 + \pi^0 - e_2^h \sum_{i=1}^{4} \omega_i x_{io} \geq 0 \]

\[ \sum_{r=1}^{3} \mu_r y_{r0} + \mu^0 - e_2^h (\pi_1^1 z_{20}^1 + \pi^2 z_{20}^2) \geq 0 \]

\[ \pi_1^1 z_{20}^1 + \pi^2 z_{20}^2 = 1 \]

\[ \pi_1^1, \pi_2^1, \pi^2, \omega_i, \mu_r \geq 0, \quad \forall i, r, \mu^0, \pi^0, \text{ unrestricted in sign} \]

Along the same lines as above, the follower score \( e_2^* \) will be at or above the current best value \( e_2^h \).

**Deriving efficiency scores for DMUs in \( N_1 \)**

For the strictly stage 1 DMUs in \( N_1 \), we solve model (2) or its linear equivalent.

### 4. Application

Table 1 displays the input and output data for a sample of 40 high technology companies in China. As discussed above, the efficiency measurement problem in this setting is complicated by the fact that some of the companies (\( N_1 \)) behave as single-stage DMUs, while others (\( N_2 \)) exhibit a two-stage structure. The first 20 of the companies make up the set \( N_1 \), while the latter 20 companies comprise the set \( N_2 \).

Three of the principal inputs (to the first stage) are different levels of employees, namely non-degree, bachelors-degree and graduate-degree employees. The fourth input is registered capital held by the firm. Different employee levels can imply different core cultures for the firms, and different competitive capabilities. It would appear, as well, that firms with better educated workforces generally possess higher amounts of registered capital. It might be said, in addition, that firms with higher registered capital tend to exhibit greater social wealth that can help those firms advance in the marketplace, and can lead to stock market competitiveness. Another phenomenon appears to be that firms with higher educated work forces often have greater capability (and perhaps desire) to set aside a portion of revenues generated from stage 1 to invest in R&D, with the aim of earning other rewards in the form of patents, industry awards, and inventions.
Table 1. Data on high technology firms in China.

<table>
<thead>
<tr>
<th>DMU</th>
<th>x1</th>
<th>x2</th>
<th>x3</th>
<th>x4</th>
<th>z11</th>
<th>z21</th>
<th>y1</th>
<th>y2</th>
<th>y3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>0.050</td>
<td>0.220</td>
<td>0.000</td>
<td>53</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>57</td>
<td>3</td>
<td>1.000</td>
<td>19.356</td>
<td>0.000</td>
<td>48</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>20</td>
<td>2</td>
<td>0.100</td>
<td>1.284</td>
<td>0.000</td>
<td>55</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>32</td>
<td>12</td>
<td>0</td>
<td>0.100</td>
<td>0.535</td>
<td>0.000</td>
<td>60</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>292</td>
<td>142</td>
<td>11</td>
<td>6.500</td>
<td>75.629</td>
<td>0.000</td>
<td>48</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>0.084</td>
<td>0.618</td>
<td>0.000</td>
<td>47</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>799</td>
<td>20</td>
<td>0</td>
<td>10.100</td>
<td>93.039</td>
<td>0.000</td>
<td>51</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>0.200</td>
<td>0.589</td>
<td>0.000</td>
<td>32</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>36</td>
<td>262</td>
<td>8</td>
<td>7.400</td>
<td>22.385</td>
<td>0.000</td>
<td>49</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0.100</td>
<td>0.091</td>
<td>0.000</td>
<td>52</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>424</td>
<td>279</td>
<td>88</td>
<td>9.000</td>
<td>41.319</td>
<td>0.000</td>
<td>54</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>40</td>
<td>135</td>
<td>38</td>
<td>3.801</td>
<td>9.139</td>
<td>0.000</td>
<td>48</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>1191</td>
<td>312</td>
<td>47</td>
<td>17.890</td>
<td>339.743</td>
<td>0.000</td>
<td>37</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>219</td>
<td>10</td>
<td>0</td>
<td>3.200</td>
<td>2.900</td>
<td>0.000</td>
<td>49</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>110</td>
<td>10</td>
<td>0</td>
<td>1.166</td>
<td>157.173</td>
<td>0.000</td>
<td>55</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>280</td>
<td>186</td>
<td>9</td>
<td>1.000</td>
<td>85.115</td>
<td>0.000</td>
<td>37</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>3</td>
<td>12</td>
<td>2</td>
<td>0.100</td>
<td>1.804</td>
<td>0.000</td>
<td>55</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>32</td>
<td>188</td>
<td>50</td>
<td>2.175</td>
<td>28.512</td>
<td>0.000</td>
<td>53</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>1682</td>
<td>52</td>
<td>0</td>
<td>8.643</td>
<td>358.519</td>
<td>0.000</td>
<td>50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>97</td>
<td>145</td>
<td>19</td>
<td>3.100</td>
<td>22.771</td>
<td>0.000</td>
<td>50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>126</td>
<td>20</td>
<td>3</td>
<td>3.244</td>
<td>34.239</td>
<td>2.078</td>
<td>54</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>2.388</td>
<td>9.977</td>
<td>0.255</td>
<td>53</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>23</td>
<td>41</td>
<td>9</td>
<td>0</td>
<td>2.043</td>
<td>0.141</td>
<td>0.560</td>
<td>47</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>573</td>
<td>58</td>
<td>1</td>
<td>4.500</td>
<td>21.083</td>
<td>1.086</td>
<td>56</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>25</td>
<td>73</td>
<td>6</td>
<td>0</td>
<td>5.000</td>
<td>19.427</td>
<td>1.094</td>
<td>42</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>26</td>
<td>510</td>
<td>60</td>
<td>0</td>
<td>2.000</td>
<td>72.047</td>
<td>2.226</td>
<td>44</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>27</td>
<td>207</td>
<td>38</td>
<td>4</td>
<td>19.000</td>
<td>171.629</td>
<td>5.562</td>
<td>51</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>28</td>
<td>24</td>
<td>15</td>
<td>2</td>
<td>2.100</td>
<td>4.372</td>
<td>0.531</td>
<td>59</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>29</td>
<td>903</td>
<td>801</td>
<td>59</td>
<td>10.000</td>
<td>90.903</td>
<td>38.265</td>
<td>36</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>30</td>
<td>21</td>
<td>30</td>
<td>2</td>
<td>0.200</td>
<td>2.613</td>
<td>0.224</td>
<td>55</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>31</td>
<td>39</td>
<td>8</td>
<td>0</td>
<td>3.800</td>
<td>4.402</td>
<td>0.259</td>
<td>43</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>32</td>
<td>91</td>
<td>39</td>
<td>7</td>
<td>1.500</td>
<td>1.984</td>
<td>6.634</td>
<td>58</td>
<td>4</td>
<td>45</td>
</tr>
<tr>
<td>33</td>
<td>457</td>
<td>51</td>
<td>1</td>
<td>34.681</td>
<td>78.500</td>
<td>1.600</td>
<td>59</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>34</td>
<td>197</td>
<td>25</td>
<td>1</td>
<td>0.050</td>
<td>17.861</td>
<td>0.702</td>
<td>43</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>35</td>
<td>26</td>
<td>43</td>
<td>3</td>
<td>1.000</td>
<td>5.856</td>
<td>0.236</td>
<td>42</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>36</td>
<td>109</td>
<td>12</td>
<td>1</td>
<td>2.000</td>
<td>7.240</td>
<td>0.674</td>
<td>45</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>37</td>
<td>505</td>
<td>58</td>
<td>7</td>
<td>100.000</td>
<td>32.975</td>
<td>1.391</td>
<td>37</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>38</td>
<td>154</td>
<td>55</td>
<td>12</td>
<td>9.900</td>
<td>23.949</td>
<td>2.405</td>
<td>60</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>39</td>
<td>139</td>
<td>102</td>
<td>8</td>
<td>3.000</td>
<td>6.532</td>
<td>0.995</td>
<td>48</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>231</td>
<td>106</td>
<td>3</td>
<td>1.121</td>
<td>11.232</td>
<td>1.100</td>
<td>51</td>
<td>7</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2 displays efficiency results for single-stage and two-stage high technology firms. We point out that the efficiency scores for DMUs in $N_1$ are obtained by solving model (2), and both sets of DMUs make up the constraint set of that model. We assume that the revenue output is split into two parts, $z^1_1$ and $z^2_1$, although $z^3_1 = 0$ for those DMUs in $N_1$. We might, as an alternative, consider combing the two revenue variables into a single figure, although doing so in $N_2$ would seem to be problematic in that the variables play different roles. Specifically, one of those variables is passed on to stage 2, while the other revenue variable leaves the system. Hence, it would
Table 2. Efficiency scores.

<table>
<thead>
<tr>
<th>DMU</th>
<th>e1min</th>
<th>e1max</th>
<th>e2min</th>
<th>e2max</th>
<th>e1h</th>
<th>e2h</th>
<th>e1*</th>
<th>e2*</th>
<th>Aggregate</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>2</td>
<td>0.6035</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.6035</td>
<td>0.6035</td>
</tr>
<tr>
<td>3</td>
<td>0.8843</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.8843</td>
<td>0.8843</td>
</tr>
<tr>
<td>4</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>5</td>
<td>0.1844</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1844</td>
<td>0.1844</td>
</tr>
<tr>
<td>6</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>7</td>
<td>0.2821</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.2821</td>
<td>0.2821</td>
</tr>
<tr>
<td>8</td>
<td>0.8711</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.8711</td>
<td>0.8711</td>
</tr>
<tr>
<td>9</td>
<td>0.4770</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.4770</td>
<td>0.4770</td>
</tr>
<tr>
<td>10</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>11</td>
<td>0.0713</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0713</td>
<td>0.0713</td>
</tr>
<tr>
<td>12</td>
<td>0.2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.2013</td>
<td>0.2013</td>
</tr>
<tr>
<td>13</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>14</td>
<td>0.1063</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1063</td>
<td>0.1063</td>
</tr>
<tr>
<td>15</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>16</td>
<td>0.5888</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.5888</td>
<td>0.5888</td>
</tr>
<tr>
<td>17</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>18</td>
<td>0.6685</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.6685</td>
<td>0.6685</td>
</tr>
<tr>
<td>19</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>20</td>
<td>0.1798</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1798</td>
<td>0.1798</td>
</tr>
<tr>
<td>21</td>
<td>0.4475</td>
<td>0.7012</td>
<td>0.2725</td>
<td>0.6941</td>
<td>0.6612</td>
<td>0.6869</td>
<td>0.6612</td>
<td>0.6869</td>
<td>0.6740</td>
<td>0.6741</td>
</tr>
<tr>
<td>22</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>23</td>
<td>0.1688</td>
<td>1.0000</td>
<td>0.7781</td>
<td>0.8813</td>
<td>1.0000</td>
<td>0.7781</td>
<td>1.0000</td>
<td>0.7781</td>
<td>0.8891</td>
<td>0.8991</td>
</tr>
<tr>
<td>24</td>
<td>0.2777</td>
<td>0.5117</td>
<td>0.7270</td>
<td>0.7407</td>
<td>0.5117</td>
<td>0.7270</td>
<td>0.5117</td>
<td>0.7280</td>
<td>0.6193</td>
<td>0.6199</td>
</tr>
<tr>
<td>25</td>
<td>0.3455</td>
<td>1.0000</td>
<td>0.8602</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.8602</td>
<td>1.0000</td>
<td>0.8602</td>
<td>0.9301</td>
<td>0.9301</td>
</tr>
<tr>
<td>26</td>
<td>0.2804</td>
<td>1.0000</td>
<td>0.9221</td>
<td>0.9469</td>
<td>1.0000</td>
<td>0.9221</td>
<td>1.0000</td>
<td>0.9222</td>
<td>0.9610</td>
<td>0.9611</td>
</tr>
<tr>
<td>27</td>
<td>0.9449</td>
<td>1.0000</td>
<td>0.7668</td>
<td>0.7877</td>
<td>1.0000</td>
<td>0.7668</td>
<td>1.0000</td>
<td>0.7668</td>
<td>0.8834</td>
<td>0.8834</td>
</tr>
<tr>
<td>28</td>
<td>0.5859</td>
<td>1.0000</td>
<td>0.7899</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.7899</td>
<td>1.0000</td>
<td>0.7899</td>
<td>0.8924</td>
<td>0.8950</td>
</tr>
<tr>
<td>29</td>
<td>0.0713</td>
<td>1.0000</td>
<td>0.6396</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.6396</td>
<td>1.0000</td>
<td>0.6396</td>
<td>0.8198</td>
<td>0.8198</td>
</tr>
<tr>
<td>30</td>
<td>0.5995</td>
<td>0.6840</td>
<td>0.7880</td>
<td>1.0000</td>
<td>0.5995</td>
<td>1.0000</td>
<td>0.5995</td>
<td>1.0000</td>
<td>0.7997</td>
<td>0.7989</td>
</tr>
<tr>
<td>31</td>
<td>0.5019</td>
<td>0.5019</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.5019</td>
<td>1.0000</td>
<td>0.5019</td>
<td>1.0000</td>
<td>0.7509</td>
<td>0.7510</td>
</tr>
<tr>
<td>32</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>33</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.6526</td>
<td>0.6526</td>
<td>1.0000</td>
<td>0.6526</td>
<td>1.0000</td>
<td>0.6526</td>
<td>0.8263</td>
<td>0.8263</td>
</tr>
<tr>
<td>34</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.9345</td>
<td>0.9345</td>
<td>1.0000</td>
<td>0.9345</td>
<td>1.0000</td>
<td>0.9345</td>
<td>0.9672</td>
<td>0.9673</td>
</tr>
<tr>
<td>35</td>
<td>0.3660</td>
<td>0.3660</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.3660</td>
<td>1.0000</td>
<td>0.3660</td>
<td>1.0000</td>
<td>0.6830</td>
<td>0.6830</td>
</tr>
<tr>
<td>36</td>
<td>0.1217</td>
<td>0.4837</td>
<td>0.4852</td>
<td>0.9063</td>
<td>0.3747</td>
<td>0.7777</td>
<td>0.3747</td>
<td>0.7783</td>
<td>0.5757</td>
<td>0.5760</td>
</tr>
<tr>
<td>37</td>
<td>0.0564</td>
<td>0.1497</td>
<td>0.4425</td>
<td>1.0000</td>
<td>0.0564</td>
<td>1.0000</td>
<td>0.0564</td>
<td>1.0000</td>
<td>0.5282</td>
<td>0.5282</td>
</tr>
<tr>
<td>38</td>
<td>0.3283</td>
<td>1.0000</td>
<td>0.8166</td>
<td>0.9495</td>
<td>1.0000</td>
<td>0.8166</td>
<td>1.0000</td>
<td>0.8166</td>
<td>0.9083</td>
<td>0.9083</td>
</tr>
<tr>
<td>39</td>
<td>0.1462</td>
<td>0.1462</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.1462</td>
<td>1.0000</td>
<td>0.1462</td>
<td>1.0000</td>
<td>0.5731</td>
<td>0.5731</td>
</tr>
<tr>
<td>40</td>
<td>0.3481</td>
<td>0.3481</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.3481</td>
<td>1.0000</td>
<td>0.3481</td>
<td>1.0000</td>
<td>0.6741</td>
<td>0.6741</td>
</tr>
</tbody>
</table>

appear to be impractical to consider the two variables in combined form as outputs from stage 1, but then proceed to unbundle them to create two separate variables as inputs to stage 2. Column 3, labeled e1max, displays the efficiency scores for DMUs in \( N_1 \). These also appear in the last 2 columns of the table.

For DMUs in \( N_2 \), columns 2 and 3 display the minimum and maximum of stage one’s efficiency scores, respectively. Columns 4 and 5 present the minimum and maximum of the stage two scores, respectively. Column 6 and 7 present e1h and e2h, with the latter coming from model (6) to support the optimal overall solution.
displayed in Column 10. It is important to point out that in applying model (6), we used as a parameter, the stage 1 score, at a selected set of points in its range from minimum to maximum. We record the value of $h (\hat{h})$, for which the average of $e_1 h$ and $e_2 h$ (from (6)) is optimal. This average is recorded in Table 2 as column 10 (designated as the aggregate). We then apply models (9) and (10) to attempt to improve the stage 1 and 2 efficiency scores, and then compute their average as displayed in Column 11 and designated as the overall score.

It is observed that some small firms with low levels of capital generally exhibit better performance scores, such as is the case, for example, for DMUs 1, 4, 6, 10 and 19. In China, some of the small technology firms have very strong marketing competitiveness. Those firms are small in terms of their inputs, in regard to both employees and registered capital, hence explaining their high efficiency in an input-oriented sense. At the same time, DMUs 13, 17 and 19 have very good efficiency scores, and do not invest in R&D. In spite of having large numbers of employees with bachelors and graduate degrees in DMUs 5, 11, 14 and 20, those firms present very low scores. See, for example, DMUs 11 and 12 with efficiency scores of 0.0713 and 0.2013, respectively. This would seem to imply that a workforce with high academic qualifications is not necessarily a substitute for the management competitiveness, entrepreneurial skill and creativity that are necessary to succeed in generating new products, patents and awards. Of course, this may not be relevant for these two firms which happen to be in $N_1$ rather than in $N_2$. We can from the stage two note that DMUs 22, 25, 28, 29, 30, 31, 32, 35, 37, 39, 40 are efficient. Other DMUs have different levels of inefficiency at stage 2 in $N_2$. Only DMUs 22 and 32 are efficient in both stage one and stage two. In our analysis of this group of high-technology companies, one might argue that there is a need to include some more tangible way of capturing qualitative features such as entrepreneurship and creativity. While our variable, innovation ability, attempts to highlight features such as entrepreneurship, there is still a need to provide a more robust measure, without which the models may produce distorted measures of performance.

In addition to the above, another potential shortcoming of the data used herein is its failure to account for some form of segregation by industry type. For example, it might be anticipated that many of the highly capitalized firms are in some form of engineering and/or robotics design, as opposed to say the manufacture of high technology products with well-established specifications. The latter firms may simply not be in the business of design whatsoever, hence would likely fall into the $N_1$ category. Therefore, it may be an unfair comparison to apply performance measurement methodologies to a heterogeneous group of firms, in terms of “type” of industry.

5. Conclusions and Future Research Directions
The current paper focuses on an efficiency measurement problem setting wherein the DMUs exhibit a mix of single- and two-stages setups. We develop a DEA
methodology to handle the particular case involving two groups of high-technology firms, whose performance we wish to evaluate. Firms in group 1 \((N_1)\) exhibit single-stage behavior and define their business as aimed purely at revenue generation. Group two firms \((N_2)\) aim to earn revenues as well, but go further by investing a portion of those revenues in R&D, with a view toward generating patents, inventions and awards. This latter activity can be looked upon as the second stage in a two-stage process.

In this paper, we choose to define the aggregate efficiency of firms in \(N_2\) as the simple average of the efficiency scores of the stages 1 and 2. As discussed earlier, this approach is taken here, rather than choosing stage weights as in Cook et al. (2010). The latter approach was shown by Ang and Chen (2016), Despotis et al. (2016) and others, to have an inherent bias favoring stage 1. The computational disadvantage of the simple average approach is that the resulting model is nonlinear, with no direct linear transformation. This being the case, we solve for the aggregate efficiency by way of a parametric linear programming model. Specifically, approximate optimal solutions can be obtained for the nonlinear structure by way of a linear parametric model, when we search over a given bounded range for the stage one efficiency \(e_1\), where \(e_1 = e_1 \text{min} + kh\). Thus, the modeling complication here is not only that the DMUs do not constitute a homogeneous set, but as well one needs to derive DMU-specific ranges for the parameter \(e_1\).

As described above, the data set used herein could potentially be improved by delving more deeply into the industry types involved, as a way of exposing the rationale behind why certain firms choose to invest in R&D while other firms do not. Furthermore, studies of this nature could be improved if the data were enhanced such as to characterize inherent creativity of management.

An important extension on the methodology developed herein is the modeling of mixed stage supply chains. As discussed earlier, some supply chains for a product can have mixed numbers of stages depending, for example, on the logistics involved for particular retailers. It must be pointed out that the particular setup discussed herein happened to involve a case where it was the final (second) stage that was missing for certain DMUs. Supply chains, in general, can exhibit much more complex structures. Consider the simple example cited in the Introduction, where one retailer receives a product directly from a manufacturer, another receives the product by way of a manufacturer to a distribution center and then to the retailer, and still a third has the product shipped from a manufacturer to a central warehouse, then to a branch warehouse, and finally to the retailer. The mathematical modeling of efficiency in this setting is more complex than described herein. This is the subject of ongoing research.

Acknowledgments
Wade Cook acknowledges financial support from NSERC a8966.
References


Biography

Wade Cook is Emeritus Professor at the Schulich School of Business, York University, Toronto, Canada. He holds the title of University Professor. Professor Cook is an internationally recognized expert in Data Envelopment Analysis (DEA). He has published numerous books and more than 160 papers in a wide range of journals including Management Science, Operations Research, IIE Transactions, Naval Research Logistics, European Journal of Operational Research, Journal of Operational Research Society, Annals of Operations Research, OMEGA, and others.

Chuanyin Guo is a Ph.D. candidate in the School of Economics and Management, Beihang University. He publishes in journals such as European Journal of Operational Research and Computers & Industrial Engineering. His areas of specialty include data envelopment analysis and multi-criteria decision analysis.

Wanghong Li currently serves as a Post-Doctoral Fellow of Operations Management and Information Systems at the Schulich School of Business, York University, Toronto, Canada. She received her PhD in Management Science and Engineering from School of Management, University of Science and Technology of China. Her research interests draws on the resource based view and aims at creating decision support tools for identifying best benchmarks for organizational performance with Network Data Envelopment Analysis (DEA). Her works have been published in journals such as European Journal of Operational Research (EJOR), AOR and OMEGA.

Zhepeng Li is an Assistant Professor in the Area of Operations Management and Information Systems at the Schulich School of Business, York University, Toronto, Canada. He has received his Ph.D. in Information Systems from the University of Utah in 2013. His research focuses on business analytics, data mining, and machine learning. He has published in leading research outlets, such as Management Science and Information Systems Research.

Liang Liang is President of HFUT. His major research areas are data envelopment analysis, supply chain management and enterprise resource optimization. He has published over 90 papers in SCI and SSCI indexed international academic journals such as Operations Research and Journal of Operations Management.

Joe Zhu is Professor of Operations Analytics in the Foisie Business School at Worcester Polytechnic Institute, USA. He is the founding Dean of College of Auditing and Evaluation at Nanjing Audit University. He is an internationally recognized expert in methods of performance evaluation and benchmarking, and his research interests include operations and business analytics, productivity modeling, and performance evaluation and benchmarking. He has published more than 15 books related to data envelopment analysis, and 120 peer-reviewed articles in journals such as Management Science, Operations Research, Sloan Management.
W. D. Cook et al.

Review, IIE Transactions, Naval Research Logistics, European Journal of Operational Research, Journal of Operational Research Society, Annals of Operations Research, OMEGA, and others. He has more than 20,000 google scholar citations. He is an Area Editor of OMEGA, an Associate Editor of Journal of the Operational Research Society, and an Associate Editor of INFOR. He is also an Editorial Board member of European Journal of Operational Research and Computers & Operations Research. He is the Associate Series Editor for the Springer International Series in Operations Research and Management Science. He is a Japan Society for Promotion of Science (JSPS) fellow and a William Evans Visiting Fellow (University of Otago, New Zealand). His research has been supported by KPMG Foundation, National Institute of Health, and Department of Veterans Affairs. He is a Chang Jiang Scholar Chair Professor awarded by the Ministry of Education of China.