Production, Manufacturing and Logistics

Two-dimensional efficiency decomposition to measure the demand effect in productivity analysis

Chia-Yen Lee, Andrew L. Johnson

A R T I C L E   I N F O

Article history:
Received 24 July 2010
Accepted 4 August 2011
Available online 12 August 2011

Keywords:
Data envelopment analysis
Productivity and profitability change
Efficiency decomposition
Demand fluctuation
Airlines industry

A B S T R A C T

This paper proposes a two-dimensional efficiency decomposition (2DED) of profitability for a production system to account for the demand effect observed in productivity analysis. The first dimension identifies four components of efficiency: capacity design, demand generation, operations, and demand consumption, using Network Data Envelopment Analysis (Network DEA). The second dimension decomposes the efficiency measures and integrates them into a profitability efficiency framework. Thus, each component’s profitability change can be analyzed based on technical efficiency change, scale efficiency change and allocative efficiency change. An empirical study based on data from 2006 to 2008 for the US airline industry finds that the regress of productivity is mainly caused by a demand fluctuation in 2007–2008 rather than technical regression in production capabilities.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Analysts of production systems use a variety of techniques to assess performance and search for improvement alternatives. Singh et al. (2000) claim three main categories of performance measurement techniques: index measurement, linear programming, and econometric models. The first includes the concept of total factor productivity or financial ratio, while the latter two categories are based on production function. The production economics approach, Hackman (2008), can be used to estimate the frontier production function and characterize how efficient production processes use inputs to generate outputs. Consequently, given the same input resource, a system is termed inefficient when its outputs levels are lower than other potential production processes. However, the reduced actual output can be caused by inefficient demand, i.e. demand fluctuations can bias productivity measures and lead to a decrease in measured efficiency. Similarly, in panel data analysis, the Malmquist productivity index quantifies efficiency change and technology change over time. Technical regression is often attributed to production issues when in actuality it may result from lack of demand. This paper incorporates demand into the analysis and attributes some changes in production to demand.

The literature on the demand effect in productivity analysis divides into two streams. One stream uses parametric equilibrium models to measure total factor productivity (TFP) change. Nakiri and Schankerman (1981) discuss the reasons for productivity slowdown observed between 1965 and 1978. The authors propose a model for decomposing changes in TFP that identifies the contributions of factor-price effect, demand effect, R&D effect, and technical change. They conclude the productivity slowdown of American manufacturing was mainly due to the deceleration in demand growth. Appelbaum and Berechman (1991) provide a market equilibrium model that considers supply (cost), demand, and government regulatory conditions. The model builds an output demand function to represent the relationship between the supply-side provision of firms and the demand-side consumption of customers, calculates the cost growth rate, and decomposes it into changes in outputs scale, factor prices and technical efficiency. It also calculates the growth rate of cost efficiency to clarify the effects on demand and regulatory conditions. Good et al. (1999) further describe static and dynamic factor demand models to measure TFP growth and its decomposition.

The second literature stream models demand generation or consumption as a component of a production system. In other words, a firm uses its marketing or sales departments to change its demand level. Studying the performance evaluation of a transportation system, Fielding et al. (1985) distinguish between the production process and the consumption process, arguing that output consumption is substantially different from output production since transportation services cannot be stored. They propose three performance indicators for a transit system: cost efficiency, service effectiveness, and cost effectiveness. More specifically, they define service effectiveness as service consumption normalized by the service output. However, their study only considers a single factor...
productivity ratio which assumes that other resources are unlimited and other outputs are unrelated. Chen and McGinnis (2007) discuss the limitations of focusing on a single productivity indicator rather than attempting to model all important factors in a production system. Lan and Lin (2005) and Yu and Lin (2008) study similar transportation systems and use Data Envelopment Analysis (DEA) and Network DEA models to characterize a consumption process. Ertay and Ruan (2005) present a methodology with one efficiency measure to identify the most efficient number of operators and the efficient assignment of labor in a cellular manufacturing system. DEA is used to measure efficiency and a simulation model is used to model capacity design and demand generation. Ertay et al. (2006) present a DEA approach to evaluate a facility layout with both quantitative and qualitative metrics. They apply an analytic hierarchy process (AHP) to aggregate the qualitative data such as flexibility in volume and variety and quality, and quantitative criteria such as material handling cost, adjacency score, shape ratio, and material handling vehicle utilization. Although these and similar studies integrate demand factors and the variable levels of demand, they do not consider the network structure of production or a dynamic productivity analysis.

Noting that the demand and the production system characterized in service production systems typically differ from those in manufacturing due to the types of demand described below, this paper models the demand generation process (i.e. a marketing department) explicitly as a component of the production system. Fig. 1 describes a manufacturing production system in terms of a serial model where the input resource is transformed into actual output product (Lee and Johnson, 2011). The first component, capacity design process, identifies the maximal output level as the peak output or the historical best performance. The demand generation process attempts to generate sufficient demand to support maximal output without idle capacity. The operations process transforms raw material into final product. Finally, the demand consumption process measures realized demand – the amount of final product consumed by customers. Manufacturers tend to receive demand based on the contractual agreements made between a manufacturer and a customer with defined sales quantities and prices prior to production. Thus, the manufacturer commonly develops an internal demand-generating process due to long production lead times (unlike service production systems which tend to rely mainly on non-contract demand requested informally by customers after production). The result is an external demand-consumption process. Since services are typically non-storable commodities which must be immediately consumed by customers once transformed from inputs, i.e. demand consumption can be inefficient in that many service opportunities go unused.

In addition, changes in demand can also have an effect on the measurement of productivity or profitability changes over time as estimated through frontier shifts indicating either technical progress or regress. Nishimizu and Page (1982) decompose total factor productivity into technology progress and change in efficiency. Färe et al. (1992, 1994) develop the explicit measurement of productivity change based on the Malmquist productivity index (MPI) proposed by Cave et al. (1982), which uses Shephard’s input distance function to estimate inefficiency nonparametrically. The productivity change estimated via MPI can also be decomposed into two components: change in technology and change in efficiency. Färe et al. (1994) develop an additional component, change in scale. Alternatively, Ray and Mukherjee (1996) use the Fisher productivity index and propose a decomposition into efficiency change, technical change described by the cost function index, change in scale efficiency captured by the average cost index, change in allocative efficiency, and an adjustment index which captures the pure effect of a change in the attributes on the measured productivity index. However, their decomposition is restricted to the single-output technology and mixed-period measures, making interpretation difficult. Zofio and Prieto (2006) present a decomposition of the Fisher index into the MPI and an economic component consisting of allocative efficiency and a residual allocative term based on a generalized distance function which employs a relative weight of the input- and output-oriented projection paths to the frontier. Their decomposition also has some limitations, because the residual terms with mixed-period measures are difficult to interpret and weighting the projections is debatable. Recently, Kuosmanen and SIPILÄINEN (2009) propose an exact decomposition of the Fisher productivity index into five components: change in efficiency, technical change, change in scale efficiency, change in allocative efficiency, and price effect. Their decomposition reveals that the change in profitability efficiency is the product of only three components (change in efficiency, change in scale efficiency, and change in allocative efficiency) and is invariant to both technical change and price effect. Note that the price effect begins to integrate demand-side effects into the productivity analysis. This paper extends Kuosmanen and Sipiläinen work to make the effects of demand more explicit.

2. Literature review of productivity analysis in the airline industry

An airline’s production system is a hybrid of the manufacturing and service systems described above. Any individual airline’s production process is characterized by transforming capital, labor, energy, and materials into passenger and cargo services. The sources of uncertainty are capital utilization rates, changing technology, labor-intensive services, and demand diversity. Obviously, an airline
operates under enormous pressure to maintain the high service rates that give it a competitive edge. The existing academic literature discusses the productivity change in the global airline industry in light of price changes in crude oil and jet fuel, the introduction of e-commerce, rising interest rates, deregulation, etc.

Sickles et al. (1986) consider the passage of the US federal Air Deregulation Act of 1978 (ADA) in improving the ability of price adjustment and competition capability and identify the effect of a rapid increase in the price of jet fuel. The result of analyzing allocative inefficiency from 1970 to 1981 supports the common perception that deregulation reduces inefficiency and the total cost of distortions from cost-minimizing allocation. However, Sickles et al. attribute the largest benefits to administrative reforms in the early 1970s, including multiple route authorizations and show-cause proceedings to reduce cost and time in obtaining certificates, rather than ADA itself. Good et al. (1993a) investigate differences in productivity growth between European and US carriers during the period 1976–1986. Using a Cobb-Douglas stochastic frontier production model, potential efficiency gains of European liberalization are identified; however, while Great Britain favors liberalization, France and Italy oppose it since their airlines benefit from high levels of subsidies to cover operating losses. Ray and Mukherjee (1996) employ an efficiency decomposition of the Fisher productivity index in the US airline industry in 1983–1984 and quantify the productivity growth in each component. The comprehensive decomposition provides more detailed benchmarking information for productivity improvement.

Semenick Alam and Sickles (2000) use DEA and MPI to estimate the productivity growth of US airlines between 1970 and 1990 and employ second-stage regression with contextual variables to capture the efficiency difference caused by firm-specific characteristics. They use cointegration analysis to examine the existence of a stationary relationship between non-stationary variables over time and indicate that efficiency estimates of firms within the industry should be co-integrated since one firm’s efficiency-enhancing technology should be adopted by other firms, else all will be driven out of the industry. Semenick Alam and Sickles identify a narrowing of the differences in efficiency over time between the top performers and the other firms. Färe et al. (2007) employ MPI to estimate productivity growth after deregulation from 1979 to 1992 and show that service quality, such as direct routings and arriving on time indeed affects industry productivity. Nevertheless, slow productivity growth indicates a decline in the quality of service post-deregulation. The additional research regarding dynamic efficiency or deregulation issue in airline industry, see Good et al. (1993b, 1995), Sickles (1985), Sickles et al. (2002).

The method proposed in this paper provides an integrated decomposition of a production system and decomposes profitability change. The decomposition of a production system can characterize a typical manufacturing system where demand is realized and products are built-to-order, a typical service production system with spot demand, or a hybrid of the two. Some previous studies neglect demand fluctuations which can have a significant impact on productivity. To address this omission, we apply 2DED to an empirical study of the US airline industry from 2006 to 2008. We decompose the production system into capacity design, demand generation, operations, and demand consumption, while characterizing potential frontier shifts over time by decomposing profitability efficiency change into technical efficiency change, scale efficiency change, and allocative efficiency change.

This paper is organized as follows. Section 3 describes the modeling framework, illustrates the decomposition of the production system, and explicitly quantifies the role of demand in efficiency analysis. The 2DED model is presented for the purpose of productivity diagnosis and improvement. Section 4 describes a method to estimate production capacity via a sequential model, and then introduces a Network DEA model for efficiency decomposition of the production system. Section 5 focuses on profitability change and reviews both Shephard’s distance function and the Malmquist productivity index, while integrating demand into a decomposition of change in profitability efficiency. Section 6 discusses the results of the case study and Section 7 concludes.

3. Model description

3.1. Production system decomposition

The first component is the capacity design process which defines the physical capacity of the production system and represents a limitation on long-term system performance. Poor capacity design would include purchasing capital that is incompatible with existing or other purchased capital, selecting outdated technologies, etc. The inputs to this phase, Fixed input, are the resources used to generate the infrastructure of the production system and support the operations of the production process. Peak output is
the maximal output level the firm can achieve; it characterizes the production system’s physical capability. Section 4.1 explains how to estimate peak output.

The efficiency of the capacity design component is defined as the ratio of the fixed input resources used to the production capacity. A critical assumption at this stage is sufficient demand exists to use the firm’s current inputs completely. The design phase has a long-term impact on production performance.

3.1.2. Demand generation
The second component is the demand generation process in which the sales group attempts to generate enough demand to completely utilize the built-in production capacity. The output of this stage is *Expected demand*, which is the sum of contracted demand and expected spot demand. The firm might generate more actual output as a buffer to capitalize on potential spot demand. For the purposes of simplification the expected spot demand is characterized as a proportional expansion of the contract demand or an expected value calculated from a historical distribution. Section 6.1 explains expected demand and *Scheduled demand* as they apply to the US airline industry. Typically contract demand is tractable and fulfilled more easily than spot demand. It is based on an agreement between the firm and a customer and has a specific sales quantity and price associated with it. However, in some situations such as the airline industry, the number of passengers flown is highly stochastic. Passengers might change their flight routes or cancel the itineraries just before the flight takes off. This uncertainty leads to differences between the contracted demand and the realized demand. This issue will be discussed in Section 6.3 in terms of contextual variables.

The efficiency of the demand generation component is defined as the ratio of expected demand to peak output. Typical productivity analysis assumes all deviations from the efficient frontier are attributed to inefficiency in the production system. Under these standard assumptions, insufficient demand may bias productivity measures.

3.1.3. Operations
The third component is the operational process in which raw materials are transformed into final goods or services. Thus, *Actual output* is the number of final products generated from the production process. In the airline industry it is characterized by *Available output*, the number of passenger-miles and freight-ton-miles generated.

The efficiency of this component is defined as the ratio of actual output to a weighted aggregation of expected output and variable input. In general, observed output may be reduced by scheduling inefficiencies, machine breakdown, inconsistent operational performance, etc.

3.1.4. Demand consumption
The fourth component is the demand consumption process in which the sales group tries to sell any production beyond the contracted demand to maximize profit. *Realized demand* is the realized quantity of product or output customers actually consume at the market price after production. It is the sum of contract demand and realized spot demand. Our empirical study of the airline industry considers contract demand and spot demand as scheduled demand and non-scheduled demand, respectively.

The efficiency of this component is defined as the ratio of realized demand to actual output. This paper focuses on the scenario in which realized demand (contract demand plus spot demand) is less than actual output. If the realized demand exceeds actual output, some customer requests will be off-loaded to other providers, substituted with a similar but different product, filled from inventory, produced using overtime, renegotiated for delivery to a subset of customers, etc.

For the empirical study the linkage between the four components of efficiency and airlines service context is shown in Appendix C, where Table C1 indicates the subcomponent and its corresponding factors mapping to flight factors in application.

3.2. Two-dimensional efficiency decomposition (2DED)
As mentioned, our 2DED model is a tool for productivity diagnosis and improvement. The two-dimensions for decomposition are the network structure of the firm in each cross-section of time (described in detail in Section 4) and profit efficiency change between periods over time (described in detail in Section 5). In the empirical study we collect panel data with the necessary variables to analyze the four components of the hybrid production systems defined above (see Section 6.1 and Appendix C for an explicit definition).

The panel data will be analyzed as a series of cross-sectional analyses and then an index number approach will be used to investigate the change in profitability efficiency and its components. An index number is a metric to quantify productivity growth. If the index number is larger than 1, there is productivity growth, otherwise, productivity is constant or regresses; the details are described in Section 5.1. Decomposing an index number identifies the components of profitability change and can aid in identifying strategies a firm may use to improve.

Detail components of profitability allow us to scrutinize each airline firm’s technical innovation, scale of production, and resource allocation. Fig. 3 shows that we initially collect the cross-sectional data in period \( t_0 \) and add the new data collected in period \( t_1 \) to the data set. In this way we use Diewert’s sequential reference set method (Diewert, 1992) to estimate efficiency. Using Fisher’s index (Kuosmanen and Sipiläinen, 2009) allows us to estimate the profitability change between the two periods. In other words, for one dimension, the components of the production system are identified in a cross-section; in the second dimension, panel data provides a dynamic efficiency analysis of profitability change. Table 1 illustrates that efficiency change can be decomposed into 12 components to help managers further identify improvement strategies.

4. Efficiency decomposition of production process

4.1. Capacity estimation
To construct our Network DEA model, we first need to estimate the capacity level if no capacity data can be collected directly. For the purposes of this paper we use Johansen’s (1968) definition of physical capacity, which is the maximum amount that can be produced with existing plant and equipment (fixed inputs) given an unlimited availability of variable factors. Eilon and Soesan (1976) extend the concept from a single output to a multiple output case and propose a measure involving the radial expansion of the output vector given current technology and a fixed input vector. Based on Eilon and Soesan’s definition, Färe et al. (1989) employ a non-parametric approach to obtain the capacity measure with a cross-sectional dataset.

The capacity is not directly observable, thus we will estimate the peak observed output as a proxy for capacity. In order to estimate the peak observed output, we need to identify a reference set to which we compare each observation in each period of time. Diewert (1980, 1992) describes a sequential method which constructs the production reference set by adding new observations to augment each previous period’s reference set. The method assumes that a production process can be compared to any previ-
ously observed production process. Therefore, our empirical study uses Dievort’s sequential method to estimate the firm specific capacity via output-oriented variable returns to scale (VRS) data envelopment analysis (DEA) and the reference set constructed from all previous period’s observations of the firm’s production. $X_{it}^k$ is the $i$th fixed input resource of $k$th firm in $t$th period, $Y_{qkt}$ is the amount of actual output for the $q$th product of $k$th firm in $t$th period, and $\lambda_{kt}$ is the DEA envelopment multiplier variable of the $k$th firm in $t$th period. $\theta_{kt}$ is the efficiency estimate of firm $r$ in the current period $s$. For firm $r$, the linear programming formulation is:

$$\begin{align*}
\text{Max} & \quad \theta_{ts} \\
\text{s.t.} & \quad \sum_{k \in (0, \ldots, s)} \lambda_{kt} X_{it}^k \leq X_{it}^k, \quad \forall i, \\
& \quad \sum_{k \in (0, \ldots, s)} \lambda_{qt} Y_{qkt} \geq \theta_{r} Y_{qyt}, \quad \forall q, \\
& \quad \sum_{k \in (0, \ldots, s)} \lambda_{kt} = 1, \\
& \quad \lambda_{kt} \geq 0, \quad \forall k, \forall t.
\end{align*}$$

If the efficiency is equal to 1, the physical capacity is equal to the number of actual outputs in period $s$, otherwise, the physical capacity is equal to the actual output multiplied by the efficiency estimate $\theta_{ts}$ that is, $Y_{qy} = Y_{qyt} \times \theta_{ts}$. Then, the time shifts to the next period and the new observation is added into the reference set, and the process repeats.

### 4.2. Efficiency measurement by network DEA

We use Kao’s (2009) Network DEA model\(^2\) for efficiency decomposition because it accounts for the interrelationship of the components of the production system rather than estimating efficiencies independently. Kao’s model was developed under the constant return to scale (CRS) assumption. However, we relax this assumption independently. Kao’s model was developed under the constant returns of the production system rather than estimating efficiencies because it accounts for the interrelationship of the components.

The proposed VRS model estimates technical efficiency and provides scale efficiency estimation by means of Kao’s CRS model. The formulation (2) adds the variables $z_{qiy}$, $u_{qiy}$, $z_{qiy}$ and $u_{qiy}$, respectively. $z_{qiy}$, $u_{qiy}$, $z_{qiy}$ and $u_{qiy}$ are the intercept variables. We estimate the input-oriented efficiency $E_{i}^{u}$ of the production system of firm $r$ in period $s$ with a sequential reference set using the following formulation:

$$E_{i}^{u} = \text{Max} \sum_{q \in Q} u_{qiy} D_{qyt} - u_{qiy}$$

subject to:

$$\begin{align*}
\sum_{j \in J} \lambda_{jyt} Y_{qyt} & = \theta_{r} Y_{qyt}, \quad \forall q, \\
\sum_{i \in I} \lambda_{i} X_{it}^k & = 1, \quad \forall k, \forall t, \\
\lambda_{i} & \geq 0, \quad \forall k, \forall t.
\end{align*}$$

The proposed VRS model estimates technical efficiency and provides scale efficiency estimation by means of Kao’s CRS model. The formulation (2) adds the variables $z_{qiy}$, $u_{qiy}$, $z_{qiy}$ and $u_{qiy}$. These variables characterize the intercept and relax the condition that the production function must pass through the origin.

By solving this optimization model, the optimal multipliers, $\lambda_{i}$, $\lambda_{j}$, $z_{qiy}$, $u_{qiy}$, $z_{qiy}$ and $u_{qiy}$ are obtained and efficiency can be decomposed. Recall that estimating the efficiency of each component allows the firm to identify which component will give the largest system productivity gain if improved. Let $E_{i}^{u}$, $E_{i}^{u}$, $E_{i}^{u}$ and $E_{i}^{u}$ denote efficiency of production design, efficiency of demand generation, efficiency of operations and efficiency of demand consumption respectively:

$$\begin{align*}
E_{i}^{u} & = \left( \sum_{q \in Q} z_{qiy} Y_{qyt} - z_{0i} \right) \left( \sum_{j \in J} \lambda_{jyt} Y_{qyt} - z_{qiy} \right), \\
E_{i}^{u} & = \left( \sum_{q \in Q} u_{qiy} D_{qyt} - u_{qiy} \right) \left( \sum_{j \in J} \lambda_{jyt} Y_{qyt} - z_{qiy} \right), \\
E_{i}^{u} & = \left( \sum_{q \in Q} u_{qiy} D_{qyt} - u_{qiy} \right) \left( \sum_{j \in J} \lambda_{jyt} Y_{qyt} - z_{qiy} \right), \\
E_{i}^{u} & = \left( \sum_{q \in Q} u_{qiy} D_{qyt} - u_{qiy} \right) \left( \sum_{j \in J} \lambda_{jyt} Y_{qyt} - z_{qiy} \right).
\end{align*}$$

\(^2\) An important property of Kao’s Network DEA model is that the whole system is efficient only when all components are efficient in contrast to the traditional network DEA (Färe and Grosskopf, 2000).
5. Efficiency decomposition of profitability change

As mentioned, Kuosmanen and Sipiläinen (2009) propose to decompose the change in profitability efficiency into change in technical efficiency, change in scale efficiency, and change in allocative efficiency. More interestingly, change in profitability efficiency is invariant to technical change and change in price effect, i.e., competition and price fluctuation would not affect the change in profitability efficiency in their decomposition. Kuosmanen and Sipiläinen also assume that demand is beyond a firm's influence, however this assumption may not hold in many industries.

5.1. Efficiency decomposition of profitability efficiency change

Now, we describe how the network decomposition of the production process can provide additional information for a decomposition of profitability efficiency characterizing a broader set of production processes and including firms that influence their demand levels through sales, advertising, etc. Let $x' \in R^{n_i}$ denote an input factor of the production system in period $t$, and $y' \in R^{n_o}$ denote an output factor of the production system in period $t$. We estimate technology $T^t=\{ (x,y) \}$ can produce $y$ in period $t$ defining the production possibility set at period $t$ by a piece-wise linear convex function enveloping all observations. Model (2) has a dual formulation with dual multipliers, $\lambda_{kt}$, $\beta_{kt}$, $\gamma_{kt}$ and $\delta_{kt}$, and illustrates the feasible region of production possibility set $T^t$ and the multipliers associated with the four components of the production system decomposition:

$$
\overline{T}^t = \left\{ (x,y) : \sum_{k \in K} \sum_{t \in \{1,\ldots,s\}} \lambda_{kt} Y_{qkt}^{qkt} - \sum_{k \in K} \sum_{t \in \{1,\ldots,s\}} \beta_{kt} Y_{qkt} \leq 0, \ \forall q \right\}
$$

$$
\sum_{k \in K} \sum_{t \in \{1,\ldots,s\}} \lambda_{kt} X_{kt}^{kt} = \sum_{k \in K} \sum_{t \in \{1,\ldots,s\}} \beta_{kt} X_{kt} \leq 0, \ \forall t
$$

$$
\sum_{k \in K} \sum_{t \in \{1,\ldots,s\}} \gamma_{kt} D_{qkt}^{qkt} - \sum_{k \in K} \sum_{t \in \{1,\ldots,s\}} \delta_{kt} D_{qkt} \leq 0, \ \forall q
$$

$$
\sum_{k \in K} \sum_{t \in \{1,\ldots,s\}} \lambda_{kt} X_{kt}^{kt} \leq X_{kt} \leq \sum_{k \in K} \sum_{t \in \{1,\ldots,s\}} \gamma_{kt} X_{kt}^{kt}, \ \forall i
$$

Note that VRS is allowed through this characterization of the production possibility set. Defining $D^t(x,y)$ generated directly from the above model as the inverse of Shephard's input-oriented distance function allows us to measure the production efficiency of an observation at period $t$ relative to the production possibility set at period $t$. In other words, the input-oriented technical efficiency (ITE) is defined as $D^t_t(x,y) = \inf\{\theta | (x,y) \in \overline{T}^t\}$.

Similarly, the output-oriented technical efficiency (OTE) is defined as $D^t_t(x,y) = \inf\{\theta | (x,y) \in \overline{T}^t\}$.

Now, we obtain the cost function and revenue function as $C^t(w,y) = \min_x \{ w \cdot x | (x,y) \in \overline{T}^t \}$ and $R^t(p,y) = \max_{x \in \overline{T}^t} \{ p \cdot y | (x,y) \in \overline{T}^t \}$, given input price $w$ and output price $p$, respectively. Then, the profitability function $\rho^t(w,p) = \max_y \left\{ \frac{p^t}{p^t} | (x,y) \in \overline{T}^t \right\}$ presents the maximal return to Dollars achievable with the given input and output price. We define the profitability efficiency ($\rho_E$) as the ratio of the profitability of an observation and the maximum profitability given the specific input and output price $\rho_E^t(w',p',x',y') = \frac{\rho^t(w',p',x',y')}{\rho^t(w,p)}$.

While such envelopment models allow us to easily calculate the efficiency of the component, the use of dual multiplier models facilitates the analysis of the cost, revenue, and profitability functions. For example given a firm $r$ in period $s$, we can calculate the profitability function of production system by the following formulation, where $p^t_{qrs}, w^t_{qrs}$, and $w^t_{qrs}$ are the unit prices for the $q^t_{rs}$ product with realized demand $D^t_{qrs}$, the $r$th fixed input $X^t_{qrs}$, and the $jth$ variable input $X^t_{qrs}$, respectively. The formulation below is similar for the cost and revenue function except that we replace the objective function by minimizing $\sum w^t_{qrs} X^t_{qrs} + \sum w^t_{qrs} X^t_{qrs}$ and maximizing $\sum p^t_{qrs} D^t_{qrs}$, respectively:

$$
\text{Max} \quad \frac{\sum q^t_{qrs} D^t_{qrs}}{\sum w^t_{qrs} X^t_{qrs} + \sum w^t_{qrs} X^t_{qrs}}
$$

s.t. $\sum_{k \in K} \sum_{t \in \{1,\ldots,s\}} \lambda_{kt} Y_{qkt}^{qkt} - \sum_{k \in K} \sum_{t \in \{1,\ldots,s\}} \beta_{kt} Y_{qkt}^{qkt} \geq 0, \ \forall q$

$$
\sum_{k \in K} \sum_{t \in \{1,\ldots,s\}} \lambda_{kt} Y_{qkt}^{qkt} \leq \sum_{k \in K} \sum_{t \in \{1,\ldots,s\}} \sum_{k \in K} \sum_{t \in \{1,\ldots,s\}} \beta_{kt} Y_{qkt}^{qkt} \leq 0, \ \forall q$

$$
\sum_{k \in K} \sum_{t \in \{1,\ldots,s\}} \gamma_{kt} D_{qkt}^{qkt} - \sum_{k \in K} \sum_{t \in \{1,\ldots,s\}} \delta_{kt} D_{qkt}^{qkt} \leq 0, \ \forall q$

$$
\sum_{k \in K} \sum_{t \in \{1,\ldots,s\}} \lambda_{kt} X_{kt}^{kt} \leq X_{kt}^{kt} \leq \sum_{k \in K} \sum_{t \in \{1,\ldots,s\}} \gamma_{kt} X_{kt}^{kt}, \ \forall i$

Note that we can augment or replace the constraints in the dual model (8.1)–(8.12) with different equations and can adjust the objective function to estimate the profitability function of each component. Eqs. (8.13)–(8.18) for the profitability function of capacity design, (8.19)–(8.27) for the demand generation, (8.28)–(8.35) for operations, and (8.36)–(8.40) for the demand consumption appear in Appendix A.

As mentioned, Kuosmanen and Sipiläinen (2009) also propose an exact decomposition of the Fisher ideal TFP index. The Fisher ideal TFP is the product of the change in the components of tech-
nical efficiency ($\Delta TE$), technical change ($\Delta Tech$), change in scale efficiency ($\Delta SE$), change in allocative efficiency ($\Delta AE$), and change in price effect ($\Delta PE$). Interestingly, Kuosmanen and Sipiläinen show that the change in profitability efficiency ($\Delta PE$) is invariant to $\Delta Tech$ and $\Delta PE$, i.e. $\Delta PE$ has three parts: $\Delta TE$, $\Delta SE$, and $\Delta AE$. $\Delta PE$ already captures technical change and price change through the target point and the price change is characterized through the identification of the allocatively efficient benchmark. The formulation of change in profitability efficiency is shown in Appendix B.

5.2. Profitability efficiency and financial performance index

This section identifies a connection between profitability efficiency and financial performance indices in order to motivate the relevance of the less widely used metric profitability efficiency. Return on investment (ROI) is considered a crucial indicator of a firm's financial performance. The Dupont ROI formula (Brown, 1927) decomposes this index into two ratios. The first is the ratio of return on sales (ROS) which measures a firm's ability to generate profit related to its sales revenue. The second is the ratio of investment turnover which measures how effectively a firm can generate revenue using investments. The Dupont ROI formula is:

$$\text{ROI} = \frac{\text{profits}}{\text{investment}} = \frac{\text{revenue}}{\text{investment}} \times \frac{\text{revenue}}{\text{investment}} = \text{ROS} \times \text{Investment turnover}$$

(9)

The ROS component reveals the profitability ratio which measures the revenue to cost (Banker et al., 1993, 1996):

$$\text{ROS} = \frac{\text{profits}}{\text{revenue}} = \frac{\text{revenue} - \text{cost}}{\text{revenue}} = 1 - \frac{1}{\text{profitability}}$$

(10)

Under a fixed investment turnover rate the higher the profitability the higher the ROI. This illustrates a strong relationship between profitability and ROI.

There are three reasons for employing a profitability efficiency index to assess a firm's productivity performance. First, profitability is a more reasonable index to assess productivity than a profit index, because the profitability function is homogenous of degree zero in prices. Namely, while the price doubles, the profit doubles, but the profitability does not change. This unscaled nature of profitability is similar to productivity and represents the input-to-output performance. Second, profitability efficiency is a benchmarking technique that builds on the concept of the production possibility set and clearly identifies the frontier and facilitates for comparisons, in contrast to profitability or profit indices. Third, simple output-input ratios do not reflect all of the critical factors in performance evaluation (Chen and McGinnis, 2007), because partial productivity ratios relating a single output to a single input postulate that all other resources are always adequate and the production of any other outputs are irrelevant. Therefore, we select profitability efficiency and change in profitability efficiency as our indices.

6. Empirical study

Our empirical case study analyzes the US airline industry from 2006 to 2008 using a data set of 15 firms. The data was gathered from Air Carrier Financial Statistics and Air Carrier Traffic Statistics published by the Bureau of Transportation Statistics within the Research and Innovative Technology Administration (RITA, 2009). Each observation is one airline firm in a given year. The data definitions of input and output factors for the productivity analysis are described in Section 6.1. Section 6.2 gives a detailed analysis of each firm's production process employing a Network DEA model for process decomposition. Further profitability efficiency change is quantified for each component and the production system as a whole. Section 6.3 summarizes the efficiency differences between civil airlines and cargo airlines using a contextual variable approach.

6.1. Data description

We characterize the resources used in the production system as: aircraft fleet size as a fixed input, fuel and employees as variable inputs, and capacity, scheduled demand, and available output as intermediate factors with the two dimensions, passenger and freight; realized demand is the final output (see Appendix C for the raw data). We estimate capacity peak output by fixed input and scheduled demand data via the sequential method in Section 4.1. The following describes the resources.

6.1.1. Inputs

Airport fleet size (FS) is the average number of aircraft employed in a firm over a particular year. However, a firm may own different models of airplanes purchased in different years, giving rise to a vintage issue, Johansen (1968). To address the heterogeneity of capital issue, we transform the data based on number of seats per model type so that each fleet is measured in Boeing-737 equivalent units. In general, since a firm's fleet is the most significant component of capital and is difficult to change in the short-term, we model the capital as a fixed input. We obtain firm-specific prices by dividing the flight equipment capital reported in the firms' balance sheets by the average number of equivalent Boeing-737 aircraft.

Fuel (FU) is the number of gallons consumed annually, estimated by fuel expenses over the average jet fuel cost per gallon. Note that FU is a variable input because its usage can be controlled on a day-to-day basis.

Employee (EP) is defined as the number of employees during the year, which includes flight shipping staff, pilots, flight attendants, and managers but not ground shipping drivers. Average prices are calculated by salaries and benefits expenses over number of employees. EP is modeled as a variable input since firms can partially adjust this variable in the short-term.

6.1.2. Demand and output levels

Scheduled passenger demand (SPD) is the scheduled revenue passenger-miles for a particular year. We measure passenger service using revenue passenger-miles, the number of revenue-paying passengers aboard the airplane multiplied by the distance traveled measured in miles. The average price per passenger mile for SPD is calculated as the scheduled passenger revenue divided by passenger-miles.

Scheduled freight demand (SFD) is defined as the demand of scheduled revenue freight-ton-miles for a particular year. We measure freight service using revenue freight-ton-miles, the weight of freight and mail measured in tons multiplied by the distance flown measured in miles. The average price for SFD is calculated as the scheduled freight and mail revenue divided by ton-miles.

Available passenger output (APO) is the actual output of available seat-miles during the year. Available seat-miles is calculated as the number of seats including first class and economy on an airplane multiplied by the distance traveled measured in miles. The average price for APO is equivalent to the price used in scheduled passenger demand.

Available freight output (AFO) is the actual output of available freight-ton-miles during the year. Available freight-ton-miles is calculated as the number of available tons of freight and mail multiplied by the distance flown measured in miles. Note that it is calculated by subtracting revenue passenger-ton-miles from total...
available ton-miles. The average price for AFO is equivalent to the price employed in scheduled freight demand.

Realized passenger demand (RPD) is the realized demand of scheduled and nonscheduled revenue passenger-miles during the year. The realized demand is calculated as the sum of scheduled and nonscheduled revenue passenger-miles. The average price for RPD is calculated by total passenger revenue over scheduled and nonscheduled passenger-miles.

Realized freight demand (RFD) is the realized demand of scheduled and nonscheduled revenue freight-ton-miles during the year. The realized demand is calculated as the sum of scheduled and nonscheduled revenue freight-ton-miles. The average price for RFD is calculated by total freight and mail revenue over scheduled and nonscheduled ton-miles.

6.1.3. Capacity estimation

Peak passenger output (PPO) is the maximal output level of revenue passenger-miles during the year. We estimate it using the sequential frontier method described in Section 4.1 with aircraft fleet size as fixed input and available passengers as the output. The average price for PPO is equivalent to the price employed in scheduled passenger demand.

Peak freight output (FPO) is defined as the maximal output level of revenue freight-ton-miles during the year. We estimate it using the sequential frontier method described in Section 4.1 with aircraft fleet size as fixed input and available freight as the output. The average price for FPO is equivalent to the price employed in scheduled freight demand.

The flight data is ordered according to capacity design, demand generation, operations, and demand consumption. Table 2 shows the factor mapping table of the production process and the data set.

6.2. Productivity change analysis

Table 3 presents the results of our efficiency decomposition analysis based on network structure for a partial set of the firms in 2006 (the entire table appears in Table D1 of Appendix D). Note that the efficiency estimates and thus the decomposition are based on the production possibility set of all previous periods because Dievert’s sequential method is used.

Consider Alaska Airlines with an input-oriented technical efficiency (ITEff) of 0.89. Further investigation of the components of efficiency reveals that it is not an issue of poor capacity design or operational inefficiency, but rather that the system inefficiency is mainly caused by insufficient demand generation and consumption (both efficiencies are 0.94). We conclude that management should focus on raising demand rather than making operational changes, perhaps by asking sales and marketing to address the productivity concerns. In contrast, Continental Airlines’ system efficiency of 0.81 is largely due to poor capacity design and unfavorable operation process (both efficiencies are 0.90). We conclude that management should engage in capacity redesign and investigate operation behavior to improve overall productivity.

Tables 4 and 5 show how 2DED provides process and dynamic efficiency analysis. Recall that we separate the efficiency decomposition of profitability change into changes in technical efficiency, change in scale efficiency, and change in allocative efficiency and decompose them into our four components. Note that $\Delta E$, $\Delta SE$, and $\Delta AE$ are not mutually independent, but have different strategic interpretations. $\Delta E$ characterizes the firm’s change in efficiency and productivity, which is largely driven by process improvement. $\Delta SE$ measures a firm’s ability to adjust scale size in the long-term. $\Delta AE$ indicates a firm’s ability to allocate input and output resource to achieve maximal profitability with respect to a specific price.

Table 4 shows the weighted average profitability efficiency change for the production system and for each component. Within each component performance is decomposed into technical, scale, and allocative effect for the 15 airlines. The average is weighted by the dollar measure of peak output. Observe the overall progress in the average profitability change from 2006 to 2008, where the average profitability change of production system is 1.015. The capacity design component is 1.02, demand generation is 0.99, operational component is 1.00, and demand consumption is 0.99. Considering each component individually, the 2% improvement on average of capacity design efficiency over the time horizon indicates that the airlines have been proactive in improving their capacity installation. Note also that the profitability efficiency changes in demand generation and consumption components (around 0.99) indicate that some airlines are failing to generate sufficient demand and to stimulate product consumption. The operations component represents no significant average change in profitability. Further investigating the yearly effect, profitability regresses in 2007–2008 and nine firms experience profitability decline (67% regress in the design component, 89% in demand generation, 33% in operation, and 78% in demand consumption). These results indicate that most firms could improve productivity through stimulating demand and improved marketing. Table 4 also shows that the variation of capacity design is larger than the other three components in 2006–2008. We conclude that the design process is a significant component and will influence profitability.

Table 5 shows the detailed 2DED for a partial set of the airlines in 2006–2008; the full table which includes the average (geometric mean, GM) change of production system is summarized in Appendix E. The figure shown in Appendix F maps the average $\Delta E$ and $\Delta AE$ of the production system by each airline on a two-dimensional coordinate (figure F1). Thus, the four quadrants reveal the strategy of productivity improvement. Using SkyWest Airlines (point K) as an example, observe a high performance in profitability change $\Delta pE$ of 1.091, an above-average $\Delta E$ of 1.023 and an $\Delta AE$ of 1.071, and a relatively poor $\Delta AE$ of 0.996. Further drilling down into $\Delta AE$ via efficiency decomposition reveals an $\Delta AE$ value of capacity design of 1.00, demand generation of 0.98, operations of 1.00, and demand consumption of 0.98. Thus, SkyWest Airlines should strive to improve its resource allocation in demand generation and consumption process to catch up with its competitors.

Appendix F includes airlines that are largely cargo service carriers as indicated by triangle points N and H. Observation H is below the average of productivity growth. However, this result may seem counter-intuitive because it performs well in terms of the profitability efficiency levels shown in Table D1. Note that firms with high levels of efficiency initial tend to have small productivity increases.

---

**Table 2**

<table>
<thead>
<tr>
<th>Components</th>
<th>Factor (Ref. Fig. 2)</th>
<th>Flight factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacity design</td>
<td>Fixed input, Peak output</td>
<td>FS, PPO</td>
</tr>
<tr>
<td>Demand generation</td>
<td>Peak output, Expected demand</td>
<td>SPD, SFD</td>
</tr>
<tr>
<td>Operations</td>
<td>Variable input, Expected demand</td>
<td>FU, EP</td>
</tr>
<tr>
<td>Demand consumption</td>
<td>Actual output, Realized demand</td>
<td>APO, AFO, RPD, RFD</td>
</tr>
</tbody>
</table>

---

*The demand-related components have statistically significant differences from the design component by t-test with a value equal to 0.1.*
changes in the future. In general, this phenomenon is the result of the public good nature of technology that leads to spillover effects from leaders to followers as the laggards learn from the innovators and play catch-up (Semenick Alam and Sickles, 2000). Here, the position (2DED) model as a diagnostic tool for identifying the sources of production system and profitability efficiency change. Two main reasons support these results: the distinct nature of the passenger shipping network structure, and the more consistent demand of the cargo shipping structure.

Most passengers prefer direct flights and are generally unwilling to endure long travel times. In contrast packages may use a variety of routes to arrive at their final destination. Thus, fewer routing constraints and the possibility of consolidation at hub locations benefit cargo-shipping airlines. Often, passengers are flexible, choosing which airline to fly with, and even substituting driving or postponing travel by air. Thus, traveler's uncertainty can significantly reduce civil carriers' profitability. In contrast, the package delivery industry has fewer firms and substitutes for their services. Nevertheless, the slopes of the other three components show a less significant difference between civil and cargo services because both airline types rely on the performance of marketing forecasts, operations control, and sales effort rather than capacity design with respect to earning structure.

### 7. Conclusion

This paper has proposed a two-dimensional efficiency decomposition (2DED) model as a diagnostic tool for identifying the sources of production system and profitability efficiency change.

---

<table>
<thead>
<tr>
<th><strong>Table 3</strong></th>
<th>Technical, scale, allocative, and profitability efficiency decomposition.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Firm</strong></td>
<td><strong>Production system</strong></td>
</tr>
<tr>
<td></td>
<td>ITE</td>
</tr>
<tr>
<td>AirTran Airways</td>
<td>0.97</td>
</tr>
<tr>
<td>Alaska Airlines</td>
<td>0.89</td>
</tr>
<tr>
<td>American Airlines</td>
<td>1.00</td>
</tr>
<tr>
<td>American Eagle</td>
<td>0.92</td>
</tr>
<tr>
<td>Continental</td>
<td>0.81</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Table 4</strong></th>
<th>2DED of US airline industry.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overall components</strong></td>
<td><strong>2006–2007</strong></td>
</tr>
<tr>
<td></td>
<td>ΔρE</td>
</tr>
<tr>
<td>Production</td>
<td>1.015</td>
</tr>
<tr>
<td>Design</td>
<td>1.052</td>
</tr>
<tr>
<td>Generation</td>
<td>0.995</td>
</tr>
<tr>
<td>Operations</td>
<td>0.990</td>
</tr>
<tr>
<td>Consumption</td>
<td>0.996</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Table 5</strong></th>
<th>2DED of US airline firms.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Firm</strong></td>
<td><strong>Year</strong></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>AirTran Airways</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>07-&gt;08</td>
</tr>
<tr>
<td></td>
<td>GM</td>
</tr>
<tr>
<td>Alaska Airlines</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>07-&gt;08</td>
</tr>
<tr>
<td></td>
<td>GM</td>
</tr>
<tr>
<td>American Airlines</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>07-&gt;08</td>
</tr>
<tr>
<td></td>
<td>GM</td>
</tr>
</tbody>
</table>
Table 6

<table>
<thead>
<tr>
<th>Regression</th>
<th>Production</th>
<th>Design</th>
<th>Generation</th>
<th>Operations</th>
<th>Consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.59</td>
<td>0.68</td>
<td>0.85</td>
<td>0.81</td>
<td>0.85</td>
</tr>
<tr>
<td>Slope</td>
<td>0.21</td>
<td>0.31</td>
<td>0.10</td>
<td>0.02</td>
<td>0.10</td>
</tr>
</tbody>
</table>

A typical production system consists of four components: capacity design, demand generation, operations, and demand consumption. The efficiency efficiency change was decomposed into technical efficiency change, scale efficiency change, and allocative efficiency change. An empirical study of productivity change in the US airline industry from 2006 to 2008 illustrated and validated the proposed method.

We found that the regress of productivity was mainly caused by demand fluctuation in 2007–2008 rather than technical regression in production capabilities. Furthermore, our contextual variable analysis suggests that the profitability efficiency of the overall production system in cargo service was 21% more efficient than civil service and that the capacity design component significantly affected efficiency.

We believe that the proposed model can be generalized and applied to other production systems for which a network structure can be identified and decomposed. For example, a supply chain system is usually defined by its materials suppliers, manufacturers, distribution centers, and retailers, hence, Network DEA efficiencies could properly estimate these entities. We suggest that the use of such decomposition enhances the rapid identification of sources of inefficiency as well as providing support for managerial trouble-shooting.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/ejor.2011.08.004.

References


Cave, D.W., Christensen, L.R., Diewert, W.E., 1982. The economic theory of index numbers and the measurement of input, output, and productivity. Econometrica 50 (6), 1393–1414.


