Journal of Air Transport Management 59 (2017) 18–25

Contents lists available at ScienceDirect

Journal of Air Transport Management

journal homepage: www.elsevier.com/locate/jairtraman

Multi-period efficiency and productivity changes in US domestic airlines

Kanghwa Choi [1]

Hansung University, 116 Samseongyoro 16 gil, Seongbuk-gu, Seoul, 136-792, South Korea

A R T I C L E I N F O

Article history:
Received 30 July 2016
Received in revised form
12 November 2016
Accepted 13 November 2016

Keywords:
US airlines
Efficiency
Bootstrapping DEA
Double bootstrap regression
M&A

A B S T R A C T

This study tracked the static efficiency and dynamic productivity changes of 14 US airlines from 2006 to 2015. Moreover, we estimated the principal economic drivers of the environmental variables to increase the US domestic airlines' efficiency using the double bootstrap regression analysis. The major aspects of this study are as follows: First, network legacy carriers have the highest efficiency, whereas low-cost carriers are lowest. Nonetheless, network legacy carriers still have room to improve scale inefficiency. Second, the fluctuations in technical change, rather than in efficiency change, tended to have greater effect on the fluctuation of Malmquist productivity index for US domestic airlines. Third, M&A between US airlines have both positive and negative effects in terms of efficiency and economies of scale. Fourth, cost environmental factors have a negative effect on US airlines' efficiency, while revenue factor is a positive effect. The results of this study may help US airline industry practitioners to understand the US domestic airline environment from an operator's perspective.

1. Introduction

US airlines have experienced unprecedented turbulence over the past 15 years from the 9/11 terrorist attacks and subsequent drastic reduction in air travel volumes to the global financial crisis and skyrocketing oil prices in 2008–2009 (Belobaba et al., 2015; Jang et al., 2011). These sequences of major events have caused the efficiency and productivity of US airlines to fluctuate. This change in operational efficiency has induced mergers among US carriers in order to survive in the competitive airline industry and enhance competitiveness and efficiency (Barros et al., 2013; Lenartowicz et al., 2013; Merkert and Morrell, 2012). Indeed, over the past decade, several mergers among US airlines have occurred (e.g., *Delta–Northwest, United–Continental, and Southwest–AirTran*) to varying degrees of success.

A vast amount of previous studies employ data envelopment analysis (DEA) models to quantify the efficiency and productivity of US airlines (Assaf and Josiassen, 2012; Barros et al., 2013; Cheng, 2010; Duygun et al., 2016; Franke, 2004; Lee and Worthington, 2014; Li et al., 2015; Min and Joo, 2016). Furthermore, some of recent studies have suggested the successful implementation of mergers and acquisitions (M&A) based on annual static efficiency, while others have found dynamic productivity changes in the airline sector (Barbot et al., 2008; Barros and Couto, 2013; Belobaba et al., 2011; Pires and Fernandes, 2012).

The survival strategy of individual airlines is to respond actively to changes in the technology and market structure of the airline service industry. This study, therefore, suggests strategic operational plans to cope with the fluctuations in the internal and external environment and identify best-practice US airlines that others can emulate.

The objective of the study is threefold: First, this study investigates the efficiency and productivity of 14 US airlines from 2006 to 2015 and measures changes in the operational efficiency of each carrier in order to suggest tailored strategic initiatives. Second, this study analyzes the long-term effect of M&A between US airlines by incorporating bootstrapping efficiency scores and RTS (returns-to-scale) perspectives. Finally, we estimates the principal economic drivers of the environmental variables to increase the US domestic airlines' efficiency by double bootstrap regression analysis suggested by Simar and Wilson (2007). To reveal how external determinants impact on efficiency is essential for airline operation practitioners to identify performance improvement strategies.

This research offers quadruple main findings. First, the efficiency analysis by airline group shows that network legacy carriers (NLCs) have the highest efficiency followed by ultra low-cost

E-mail address: khchoi@hansung.ac.kr.
[1] Tel.: 82-2-760-8015; fax: 82-2-760-4442.

http://dx.doi.org/10.1016/j.jairtraman.2016.11.007
0969-6997/

carriers (ULCCs) and low-cost carriers (LCCs) under the variable returns-to-scale (VRS) assumption. Second, the comparison of the M&A performance of three merged airlines indicates that M&A have positive or negative effects on economies of scale and efficiency levels, which suggest that new service innovation is still required to enhance airline efficiency and achieve the optimum economies of scale. Third, the result of bootstrapped truncated regression suggest that environment factors have a positive or negative effect on US domestic airlines' efficiency. The cost such as fuel expense and number of full-time equivalent employee has a negative effect on efficiency, while operating revenue have a positive effect. Fourth, productivity change of US airlines mainly depends on a change of technological change (TC). Furthermore, ULCCs have the highest productivity growth, whereas LCCs have experienced a lowest efficiency change.

The remainder of this paper is organized as follows. Section 2 describes the methodologies used in our study. Section 3 defines the input/output variables necessary for DEA and explores the characteristics of the decision-making units (DMUs). Section 4 presents the empirical result, namely the analysis of annual efficiency and productivity change in US airlines as well as the M&A performance of airlines by using bootstrapping DEA And we investigate the main driver of environmental factor to increase the efficiency. Section 5 discusses and suggests managerial implications.

2. Methodology

In this study, we used output-oriented DEA to estimate and compare the contemporaneous efficiency score of US domestic airlines from 2006 to 2015 (Färe et al., 1994; Tulkens and Vanden Eeckaut, 1995). Moreover, this paper builds on a two-stage DEA to determine potential determinants of efficiency of US domestic airlines from 2006 to 2015. The first stage is concerned with bootstrapped DEA approaches to measure the efficiency of the US domestic airlines (Simar and Wilson, 2007). To measure the robustness of the data, Simar and Wilson (1998, 2000) introduced bootstrapping DEA as a tool to extract the sensitivity of DEA scores to the randomness attributed to the distribution of efficiency. Bootstrapping, a statistical method based on empirical data, employs the repeat sampling of correlation estimations in order to improve the estimates of confidence intervals and threshold accuracy (Staat, 2006). Therefore, we use an alternative bootstrapping method to improve the DEA efficiency estimates and thus evaluate the DMU, are described as follows:

- Step 1. Use DEA to calculate efficiency scores.
- Step 2. Draw with replacement from the empirical distribution of efficiency scores. Simar and Wilson (1998) suggest that smoothing the empirical distribution provides results that are more consistent.
- Step 3. Divide the original efficient input levels by the pseudo-efficiency scores drawn from the (smoothed) empirical distribution to obtain a bootstrap set of pseudo-inputs.
- Step 4. Apply DEA using the new set of pseudo-inputs and the same set of outputs and calculate the bootstrapped efficiency scores.
- Step 5. Repeat from steps 1–4 B times and use bootstrapped scores for statistical inference and hypothesis testing (B is a large number).

In the second stage of our analysis, we regress the bias-corrected efficiency scores $\widehat{\widehat{\theta}}_i$, derived from the bootstrap algorithm on a set of environmental factors using the following regression model (Barros and Peypoch, 2009; Hall, 1986; Lee and Worthington, 2014; Simar and Wilson, 2007):

$$\widehat{\widehat{\theta}}_i = \alpha + z_i\beta + \varepsilon_i \quad i = 1, \ldots, n \tag{1}$$

where $\varepsilon_i \sim N = (0, \sigma_\varepsilon^2)$ with left-truncation at $1 - z_i\beta$; α is a constant variable; z_i is a vector of environmental variables that is expected to affect bootstrapped efficiency score of US domestic airline i and β refers to a vector of parameters with some statistical noise ε_i. Simar and Wilson (2007) detail the bootstrap truncated regression algorithm, also described in a step-by-step approach in Lee and Worthington (2014) and Barros and Peypoch (2009).

While DEA measures annual efficiency by focusing on the optimal inputs and outputs, Malmquist index (MI) analysis concentrates on productivity change to investigate the input–output relationship during a specific period (Asmild and Tam, 2007). Thus, this study additionally adopts the output-oriented MI model suggested by Färe et al. (1994) to measure the change in total productivity. The reader is referred to Färe et al. (1994) and Lovell (1993) for standard conventions and details of DEA and MI.

3. Input and output data

To compare the static efficiency and dynamic productivity of the 14 US domestic airlines, financial and non-financial data were collected from the Bureau of Transportation Statistics (www.rita.dot.gov/bts) and Airline Data Project from MIT (www.web.mit.edu/airlinedata/) during 2006–2015. The air transportation industry is a large-scale service factory (Schmenner, 1986) and a service operation system generating maximum performance with limited resources for air transportation services. In airline analysis, five common industry metrics to measure the efficiency of an airline operation are the load factor, available seat miles (ASM), revenue passenger miles (RPM), cost per available seat mile (CASM), and yield per revenue passenger (Barbot et al., 2008; Barros and Peypoch, 2009; Lee and Worthington, 2014; Li et al., 2015; Mallikarjun, 2015). Based on the previous literature review and data availability, we obtain an input variable and three output variables. The CASM are significant input factor. In addition, revenue per ASM (RASM), passenger yield, and load factor (L/F) are useful indices for estimating the business competences of carriers (e.g., profitability and market share) as well as strategic importance of major service operations.

The definition of input/output variables is as follows (http://web.mit.edu/airlinedata):

- *CASM*: Measure of unit cost in the airline industry. CASM is calculated by dividing the operating expenses of an airline by ASM. In general, management uses CASM excluding fuel or transport-related expenses to better isolate and directly compare operating expenses.
- *RASM*: Also called “unit revenue,” it is obtained by dividing operating income by ASM.
- *Passenger Yield*: A measure of airline revenue derived by dividing passenger revenue by revenue passenger miles (RPMs). This measure is useful in assessing changes in fares over time.
- Load Factor (*L/F*): The percentage of available seats that are filled with revenue passengers. The load factor measures the capacity utilization of airline transport service.

Moreover, the 14 US domestic airlines can be classified into three group according to their business models, as follows:

- *NLCs* or *full service network carriers* (hub-and-spoke airlines) focus on providing a wide range of pre-flight and onboard services, including different service classes and connecting flights: American Airlines (AA), Alaska airlines (AS), Continental Air

Lines (CO), Delta Air Lines (DL), Northwest Airlines (NW), United Air Lines (UA), and US Airways (US)

- *LCCs* focus on cost reductions in order to implement a price leadership strategy in the markets they serve: JetBlue Airways (B6), AirTran Airways (FL), Virgin America (VX), and Southwest Airlines (WN)
- *ULCCs* generally has been used to differentiate some low-cost airlines whose model deviates further from that of a standard low-cost carrier, with ultra low-cost carriers having minimal inclusions in the fare and a greater number of add-on fees: Frontier Airlines (F9), Allegiant Air (G4), Spirit Air Lines (NK)

Table 1 presents the descriptive statistics related to the changes in input (CASM), output (RASM, yield, L/F) used in the first stage and environment factors (fuel expense, passenger revenue per employee, full-time equivalents, operating revenue) of US domestic airlines during 2006–2015 used in the second regression stage. Additionally, this study used the software package of MaxDEA 6 and STATA 14 to measure the DEA estimations and truncated bootstrapped second-stage regression suggested by Simar and Wilson (2007).

This table shows that the mean scores of all variables are stable until 2008 and then decrease after the global financial crisis in 2009, reaching a trough in 2009, which indicates that the aftermath of this crisis peaked in 2009. After 2010, the values of those variables steadily increase.

4. Empirical results

4.1. Efficiency of US airlines

This study measured the technical efficiency (TE), pure technical efficiency (PTE), and scale efficiency (SE) scores of 14 US airlines from 2006 to 2015 using the output-oriented DEA model, as seen in Table 2.

Northwest Airlines (NW) and Continental Air Lines (CO) were pure technical efficient before merger with Delta Air Lines (DL) and United Air Lines (UA). After merger with NW, DL has changed efficient DMU under the VRS assumption, showing that acquirer DL's PTE maintained 1. However, post-merger UA, has kept inefficient DMU. In addition, Alaska airlines (AS) was an efficient DMU with a TE score of 1 during 2012–2015.

Moreover, a close look at the main driver of inefficiency in NLCs reveals that the inefficiency is mostly attributed to scale inefficiency ($PTE > SE$) with decreasing returns-to-scale (DRS). A possible source of DRS in some periods is that larger carriers were unable to make use of their installed capacity in low demand. In this case, contraction in the size of airline service operations such as elimination of overcapacity and adjustment of overlapping route may increase their efficiency levels at the cost of a less than proportional reduction of achieved output levels.

Meanwhile, Southwest Airlines (WN) as LCCs maintained a pure technical efficient during overall analysis periods. On the contrary, JetBlue Airways (B6) and Virgin America (VX) had comparatively higher pure technical inefficiency with $PTE < SE$ during the analysis period. The extent of pure technical inefficiency in VX is the tune of 54.1%, whereas B6 is 40.0%. This result indicates that these two carriers failed to allocate service resources efficiently and had a poor input utilization (i.e. managerial inefficiency) (Kumar, 2011). Therefore, these two carriers need to strategic approach to improve managerial performance.

In general, ULCCs is much more efficient than NLCs and LCCs. Allegiant Air (G4), Frontier Airlines (F9), and Spirit Air Lines (NK)

Table 1
Input/output and environmental variables for US domestic airlines from 2006 to 2015.

Input/output variables			2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Input	*CASM (Cents per ASM)*	Max	12.1727	10.5960	9.2774	7.0078	7.2180	8.7282	8.6217	8.3429	9.5503	6.9098
		Min	4.7830	5.2096	6.2716	4.9873	4.5956	6.9166	7.1661	6.8554	6.8052	4.9831
		Ave.	6.6387	6.4632	7.7592	6.0072	6.3612	7.7463	7.8496	7.6207	7.6238	5.9513
		S.D.	1.7985	1.3375	0.8787	0.6324	0.7606	0.5443	0.5159	0.5259	0.7726	0.6480
Output	*RASM (Cents per ASM)*	Max	18.6971	16.2609	12.0320	10.2137	11.2845	12.2735	12.5047	13.2545	14.2855	14.2713
		Min	6.4607	6.9398	7.7326	7.0953	6.9150	7.8018	7.3968	7.6169	7.5010	5.9826
		Ave.	10.1731	10.0509	10.2107	9.0264	9.8630	10.9365	10.9171	11.3035	11.6798	10.8300
		S.D.	2.8448	2.2851	1.2470	1.0324	1.3417	1.3924	1.6519	1.7335	2.0219	2.6243
	Yield (Cents per RPM)	Max	23.6939	20.1122	14.1386	12.9117	14.2342	14.9938	15.0999	16.1452	16.5026	16.3869
		Min	7.9786	8.3124	8.6350	7.9089	8.3073	9.0643	8.5985	8.7279	8.6412	7.0686
		Ave.	12.8036	12.5150	12.5646	10.9970	11.8606	12.9735	12.8776	13.2907	13.6479	12.6912
		S.D.	3.6425	2.8197	1.6728	1.4467	1.7562	1.7897	2.0914	2.2564	2.5066	3.0730
	Load Factor	Max	0.8366	0.8389	0.8955	0.8971	0.8993	0.9094	0.8957	0.9149	0.8971	0.8709
		Min	0.7271	0.7258	0.7116	0.7596	0.7928	8077	0.7954	0.8009	0.8239	0.8234
		Ave.	0.7963	0.8040	0.8154	0.8232	0.8335	0.8449	0.8498	0.8534	0.8578	0.8536
		S.D.	0.0359	0.0353	0.0437	0.0328	0.0250	0.0279	0.0308	0.0333	0.0232	0.0155
Environment variables			**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**
Environment	*Fuel Expense (Million)*	Max	3700.99	3831.49	5036.20	3057.64	4185.29	4992.41	5798.40	5693.75	6416.95	3796.86
		Min	101.59	152.14	194.27	144.33	198.66	323.24	263.65	385.19	387.81	277.73
		Ave.	1526.13	1623.21	2099.96	1245.87	1557.98	2035.45	2325.13	2419.87	2455.69	1706.60
		S.D.	1118.11	1132.34	1572.08	1013.66	1343.35	1628.75	2104.86	1985.22	2079.56	1479.27
	Passenger_Revenue (Million)	Max	11,397.9	11,354.5	11,084.0	9560.84	11,671.8	12,775.9	15,309.7	16,522.3	17,443.2	19,185.7
		Min	178.35	12.23	330.97	374.70	454.67	532.12	631.12	708.08	800.96	823.79
		Ave.	4773.35	4598.02	4671.12	3898.30	4692.87	5204.07	6335.20	6691.53	7174.22	8088.97
		S.D.	3561.37	3654.49	3614.01	3024.65	3975.90	4281.43	5472.73	5772.23	6186.21	7510.39
	Full-time Employee Equivalents	Max	72,757	71,818	70,925	66,519	76,742	80,158	87,966	87,405	84,472	98,885
		Min	841	1133	980	1421	1585	1571	1799	1978	1938	2546
		Ave.	26,381.1	26,843.2	24,622.9	23,684.6	25,329.1	25,885.6	33,619.7	30,165.1	33,315.6	35,156.8
		S.D.	22,039.5	21,875.8	21,570.9	20,299.8	24,511.2	25,211.8	31,502.6	30,308.4	30,394.7	37,244.9
	Total Operating Revenue (Million)	Max	22,493.4	22,832.8	23,696.1	19,898.3	31,893.7	35,230.4	37,160.2	38,287.1	40,426.5	41,084.4
		Min	229.86	16.15	369.25	536.47	635.46	745.04	869.24	957.82	1099.67	1221.51
		Ave.	8808.8	8682.6	9263.9	7870.8	9640.1	10,753.0	13,144.8	13,679.9	14,372.6	15,809.1
		S.D.	7542.1	7869.7	8203.2	6781.1	9667.1	10,519.9	13,470.2	13,820.8	14,391.6	16,626.9

Table 2
Contemporaneous efficiency score of US domestic airlines during 2006–2015.

DMU		2006				2007				2008				2009				2010			
		TE	PTE	SE	RTS	TE	PTE	SE	RTS	TE	PTE	SE	RTS	TE	PTE	SE	RTS	TE	PTE	SE	RTS
NLCs	AS	0.870	0.963	0.904	DRS	0.925	0.978	0.945	DRS	0.910	1	0.910	DRS	0.991	1	0.991	DRS	0.970	1	0.970	DRS
	AA	0.895	0.978	0.915	DRS	0.903	0.988	0.914	DRS	0.879	0.977	0.899	DRS	0.874	0.998	0.875	DRS	0.891	0.999	0.892	DRS
	CO	0.947	1	0.947	DRS	0.944	1	0.944	DRS	0.933	0.988	0.944	DRS	0.934	1	0.934	DRS	1	1	1	CRS
	DL	0.896	0.967	0.927	DRS	1	1	1	CRS	1	1	1	CRS	0.859	0.994	0.864	DRS	0.869	0.999	0.870	DRS
	NW	0.930	1	0.930	DRS	0.983	1	0.983	DRS	0.851	1	0.851	DRS	0.825	1	0.825	DRS	N/A	N/A	N/A	N/A
	UA	0.936	0.989	0.946	DRS	0.919	0.992	0.926	DRS	0.842	0.972	0.866	DRS	0.943	0.998	0.945	DRS	0.876	1	0.876	DRS
	US	0.862	1	0.862	DRS	0.880	1	0.880	DRS	0.829	0.970	0.854	DRS	0.882	0.997	0.884	DRS	0.865	1	0.865	DRS
LCCs	FL	0.797	0.920	0.866	DRS	0.861	0.929	0.928	DRS	0.854	0.945	0.904	DRS	0.884	0.959	0.922	DRS	0.823	0.971	0.848	DRS
	B6	1	1	1	CRS	1	1	1	CRS	0.980	0.984	0.995	DRS	0.977	0.995	0.983	DRS	0.955	0.964	0.990	DRS
	WN	1	1	1	CRS	1	1	1	CRS	1	1	1	CRS	0.994	1	0.994	DRS	0.982	1	0.982	DRS
	VX	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.952	0.959	0.992	DRS	0.880	0.943	0.933	DRS
Ultra-LCCs	G4	0.854	0.979	0.873	DRS	0.927	1	0.927	DRS	0.908	1	0.908	DRS	1	1	1	CRS	1	1	1	CRS
	F9	0.840	0.954	0.881	DRS	0.897	0.957	0.938	DRS	0.890	0.963	0.924	DRS	1	1	1	CRS	0.990	0.994	0.996	IRS
	NK	0.752	0.992	0.758	DRS	0.950	0.994	0.955	DRS	1	1	1	CRS	0.999	1	0.999	IRS	0.759	0.932	0.814	DRS

DMU		**2011**				**2012**				**2013**				**2014**				**2015**			
		TE	**PTE**	**SE**	**RTS**	**TE**	**PTE**	**SE**	**RTS**	**TE**	**PTE**	**SE**	**RTS**	**TE**	**PTE**	**SE**	**RTS**	**TE**	**PTE**	**SE**	**RTS**
NLCs	AS	0.947	0.988	0.959	DRS	1	1	1	CRS	1	1	1	CRS	1	1	1	CRS	1	1	1	CRS
	AA	0.856	0.968	0.885	DRS	0.958	0.997	0.961	DRS	0.939	0.983	0.956	DRS	0.917	0.993	0.923	DRS	0.910	0.989	0.920	DRS
	CO	1	1	1	CRS	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	DL	0.889	0.979	0.908	DRS	0.931	1	0.931	DRS	0.982	1	0.982	DRS	0.836	1	0.836	DRS	0.960	1	0.960	DRS
	NW	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	UA	0.890	0.989	0.901	DRS	0.926	0.980	0.945	DRS	0.917	0.971	0.945	DRS	0.908	0.988	0.919	DRS	0.912	0.998	0.913	DRS
	US	0.831	0.986	0.843	DRS	0.937	0.999	0.938	DRS	0.981	0.998	0.983	DRS	0.953	1	0.953	DRS	N/A	N/A	N/A	N/A
LCCs	FL	0.768	0.940	0.817	DRS	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	B6	0.927	0.965	0.960	DRS	0.950	0.974	0.975	DRS	0.926	0.961	0.964	DRS	0.878	0.973	0.903	DRS	0.870	0.978	0.889	DRS
	WN	0.938	1	0.938	DRS	0.993	1	0.993	DRS	1	1	1	CRS	1	1	1	CRS	0.997	1	0.997	DRS
	VX	0.923	0.952	0.970	DRS	0.918	0.919	0.999	DRS	0.966	1	0.966	IRS	0.916	0.951	0.964	DRS	0.824	0.949	0.868	DRS
Ultra-LCCs	G4	0.914	1	0.914	DRS	1	1	1	CRS	1	1	1	CRS	0.990	1	0.990	DRS	0.992	1	0.929	DRS
	F9	0.941	1	0.941	DRS	0.967	1	0.967	DRS	1	1	1	CRS	0.972	1	0.972	DRS	0.900	1	0.900	DRS
	NK	1	1	1	CRS	0.970	1	0.970	IRS	0.969	0.976	0.992	DRS	1	1	1	CRS	1	1	1	CRS

have almost maintained a PTE score of 1 during 2011–2015, as seen in Table 2. Furthermore, the ration of constant returns-to-scale is 33.33%, compared to NLCs (7.2%) and LCCs (21.1%). Indeed, the overall technical inefficiency of ULCCs was caused by scale inefficiency (93.3%) compared with 6.7% attributed to pure technical inefficiency. The scale inefficiency of ULCCs comprised 56.67% for DRS and 10.0% for IRS, which implies indicating that ULCCs should achieve their optimal scale by downsizing in the same vein as NLCs. In addition, NK was in the increasing returns-to-scale regions in a specific period, indicating that NK needs to expand business size in order to increase TE.

4.2. Robustness test of US airlines

While standard DEA is relatively simple to estimate, it has long been criticized for being a non-statistical or deterministic technique given that it does not allow for random error in the efficiency estimation (Assaf and Josiassen, 2012; Lee and Worthington, 2014). To overcome these problems, we examined bootstrapping DEA score to verify the statistical significance of the efficiency scores by controlling for the bias in standard DEA and to suggest a statistical reliability range. We followed Simar and Wilson (1998) by using 2000 bootstrap replications to obtain the bootstrapping results with an adequate coverage of the confidence intervals, as seen in Table 3. According to Lee and Worthington (2014), this study mainly measures the bootstrapping VRS-DEA scores, because the assumption of VRS appears appropriate given that our study includes US domestic airlines of a range of sizes. Table 3 indicates the comparison on the PTE score changes in US domestic airlines by bootstrapping DEA.

NLCs had the highest efficiency score (0.9849) and LCC was the lowest (0.9612), as seen in Table 3. The bootstrapped VRS-DEA score of NW is on the top, while FL also had lowest score. Furthermore, we conducted ANOVA to compare the efficiency differences among US airlines based on the bootstrapped PTE values. According to Scheffe's multiple comparisons, the mean differences among airline groups are statistically significant (*Sig.* $=0.000 < 0.05$). NLCs have the highest efficiency with all positive numbers of mean difference (I-J) and LCCs the lowest efficiency with all negative numbers.

4.3. M&A effect based on the bootstrapping VRS-DEA

Table 5 compares the efficiency fluctuation of pre- and post-M&A between airlines, based on the bootstrapping VRS-DEA in Table 3. For instance, Northwest Airlines (NW) was a pure technical efficient DMU with DRS before merging with Delta Air Lines (DL) in 2010, while DL had scale inefficiency with DRS. However, after the merger with NW, DL changed to efficient DMU under VRS assumption, by eliminating overlapping flights, as seen in Table 2. In addition, the average bootstrapping PTE of post-merger DL (0.9872) was slightly higher than that of pre-merger (0.9799), as seen in Table 4. However, post-merger DL is still in the region of DRS, requiring downsizing their operations. Therefore, the merger between NW and DL was only a qualified success until now, indicating a room for efficiency improvement remains.

Meanwhile, the average bootstrapping PTE score of Continental Air Lines (CO) and United Air Lines (UA) was 0.9867 and 0.9841, respectively; however, the bootstrapping VRS-DEA score of acquirer UA slightly decreased to 0.9805 after the merger with CO. In general, the air routes of UA and CO are complementary with hubs in different US cities. UA and CO had fewer overlapping routes than case of DL and NW. Nonetheless, network synergies between UA and CO failed to lead to increased market share and efficiency

Table 3
Bootstrapped PTE scores of US domestic airlines during 2006–2015.

DMUs		2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	Mean	Group mean
NLCs	AS	0.9551	0.9723	0.9901	0.9911	0.9918	0.9794	0.9831	0.9878	0.9893	0.9895	0.9829	0.9849
	AA	0.9705	0.9843	0.9716	0.9948	0.9951	0.9617	0.9915	0.9789	0.9891	0.9858	0.9823	
	CO	0.9862	0.9934	0.9803	0.9946	0.9854	0.9802	N/A	N/A	N/A	N/A	0.9867	
	DL	0.9603	0.9860	0.9827	0.9907	0.9944	0.9710	0.9908	0.9887	0.9888	0.9894	0.9843	
	NW	0.9884	0.9894	0.9830	0.9954	N/A	N/A	N/A	N/A	N/A	N/A	0.9891	
	UA	0.9801	0.9880	0.9669	0.9950	0.9940	0.9808	0.9765	0.9664	0.9844	0.9947	0.9827	
	US	0.9766	0.9853	0.9647	0.9946	0.9961	0.9785	0.9936	0.9930	0.9946	N/A	0.9863	
LCCs	FL	0.9139	0.9241	0.9402	0.9560	0.9679	0.9360	N/A	N/A	N/A	N/A	0.9397	0.9612
	B6	0.9769	0.9852	0.9751	0.9913	0.9580	0.9589	0.9693	0.9571	0.9701	0.9757	0.9717	
	WN	0.9780	0.9840	0.9827	0.9953	0.9812	0.9792	0.9875	0.9884	0.9886	0.9890	0.9854	
	VX	N/A	N/A	N/A	0.9559	0.9391	0.9449	0.9128	0.9912	0.9465	0.9468	0.9482	
Ultra-LCCs	G4	0.9699	0.9917	0.9823	0.9906	0.9817	0.9795	0.9847	0.9880	0.9929	0.9921	0.9853	0.9809
	F9	0.9470	0.9526	0.9572	0.9911	0.9873	0.9824	0.9895	0.9874	0.9914	0.9958	0.9782	
	NK	0.9869	0.9881	0.9826	0.9915	0.9255	0.9818	0.9845	0.9706	0.9925	0.9883	0.9792	

Table 4
Scheffe's multiple comparisons of the post-hoc tests.

(I) DMU	(J) DMU	Mean difference (I-J)	Std. error	Sig.	95% Confidence interval	
					Lower bound	Upper bound
NLCs	LCCs	0.01928[a]	0.0035	0.000	0.01067	0.02789
	ULCCs	0.00342	0.0036	0.635	−0.00546	0.01230
LCCs	NLCs	−0.01928[a]	0.0035	0.000	−0.02789	−0.01067
	ULCCs	−0.01586[a]	0.0040	0.001	−0.02585	−0.00587
ULCCs	NLCs	−0.00342	0.0036	0.635	−0.01230	0.00546
	LCCs	0.01586[a]	0.0040	0.001	0.00587	0.02585

[a] The mean difference is significant at the 0.05 level.

Table 5
Comparison of the VRS efficiency scores in pre- and post-mergers between US airlines.

	Pre-M&A		Post-M&A	
	Code	*Mean of B_DEA*	*Code*	*Mean of B_DEA*
Between NLCs	DL	0.9799	DL	0.9872
	NW	0.9891		
	CO	0.9867	UA	0.9805
	UA	0.9841		
Between LCCs	FL	0.9397	WN	0.9884
	WN	0.9834		

for the acquirer. Moreover, UA still had scale inefficiency during entire analysis periods (see Table 2) from an RTS perspective. Consequently, this result demonstrates that post-merger UA needs to restructure each airline's original route and cost structure to maximize the network synergies and to achieve economies of scale.

In the case of mergers between LCCs, the average bootstrapping PTE score of AirTran Airways (FL) was 0.9397 before being merged with Southwest Airlines (WN) in 2012. WN also reported an average bootstrapping PTE of 0.9834 before the merger, as seen in Table 4. After the merger, WN experienced an increased efficiency score (0.9884) as well as scale efficiency in 2013–2014. Accordingly, the merger between FL and WN was a successful M&A strategy, reducing overall operating costs by the cost synergy and optimal economies of scale.

4.4. Truncated bootstrapped second-stage regression

To examine that environmental variables exert a significant impact on measured US domestic airline efficiency, we adopted the double bootstrap approach suggested by Simar and Wilson (2007). On the base of bootstrapped VRS-DEA score in the first stage, we calculated the regression coefficients through Simar and Wilson's bootstrap procedure in the second stage. Considering US domestic airlines' operating characteristics and data availability, four environmental variables were developed for the second-stage regression analysis, as seen in Table 1. An environmental data is obtained from the Form 10-K filed by each airline with the US Securities and Exchange Commission (SEC) and Airline Data Project from MIT. The estimated specification for the regression is:

$$\widehat{\widehat{\theta}}_i = \beta_0 + \beta_1 Fuel\_Exp_i + \beta_2 PR\_Emp_i + \beta_3 FTE_i + \beta_4 Op\_Rev_i + \varepsilon_i \tag{2}$$

where $\widehat{\widehat{\theta}}_i$ is the bootstrapped bias-corrected VRS-DEA score; $Fuel\_Exp_i$ is a natural logarithm of fuel expense; PR_Emp_i is a passenger revenue per employee; $FTEs_i$ is a full-time equivalent employee (FTE); and Op_Rev_i is an operating revenue. In this model, the all independent variable is in its log-transformed state to help fitting the variable into model, and the dependent variable is in its original metric. We apply a bootstrapped truncated regression with 2000 replications as proposed by Simar and Wilson (2007) to check for structural reasons for efficiency differences. The estimated coefficients and significance levels are shown in Table 6.

The signs of the coefficients show that passenger revenue per employee and FTE have a significant negative impact on efficiency, whereas operating revenue has a positive coefficient. In general, the airline business is labor intensive. Thus, salaries, wages and benefits for FTE were a largest expenses and represented approximately 25–31% of operating expenses (from 10-K of each airline). Additionally, pension plans and other postretirement benefit funding obligations for FTE might adversely affect liquidity and financial condition of US airlines. Consequently, the increasing employee led to rise the CASM, implying that the efficiency would be reduced.

Moreover, passenger revenue per employee affect has a negative relationship with efficiency, implying that higher labor productivity

Table 6
Truncated bootstrapped second-stage regression (*dependent variable: VRS score*).

Variable	Coefficient	Std. err.	z	p > |z|	95% Confidence interval	
					Lower bound	Upper bound
ln *Fuel_Exp*	−0.0186046	0.0156441	−1.19	0.234	−0.0492665	0.0120573
ln *PR_Emp*	−0.1024115*	0.0402665	−2.54	0.011	−0.1813323	−0.0234906
ln *FTEs*	−0.0992265*	0.0390563	−2.54	0.011	−0.1757754	−0.0226775
ln *Op_Rev*	0.1208237**	0.0368977	3.27	0.001	0.0485055	0.1931419
Constant	1.62336**	0.2678492	6.06	0.000	1.098386	2.148335
Sigma	0.024617	0.0036852	6.68	0.000	0.0173942	0.0318398

Number of observation = 120, Total number of bootstrap replication = 2,000, Wald χ^2 (5) = 14.727, Prob > χ^2 (5) = 0.005.
*$p < 0.01$, **$p < 0.05$.

could trigger lower efficiency of US airlines. This index is a ratio that is calculated as airline's passenger revenue divided by the current number of employees. Thus, we can estimate that this result was primarily due to offset in part by the number of FTE as one of primary components in CASM. This is a typical case that incremental CASM is beyond its passenger revenue per employee.

Meanwhile, airline fuel expense was not statistically significant. In general, airline business is dependent on the price and availability of aircraft fuel. High volatility in fuel costs and increased fuel prices could have a significant negative impact on airline's operating results and liquidity. However, the cost and availability of jet fuel is beyond airline's control, because it is subject to many economic and political factors. Although fuel expense was not statistically significant, it has a negative effect on efficiency, indicating that cost factors such as fuel expense and number of employee are the main driver of efficiency reduction.

4.5. MI change in US airlines

This study examined MI to investigate the multi-period productivity changes of 14 US airlines. It is important to evaluate changes in the total productivity of US airlines to understand whether the MI of individual airlines is improving or worsening during the analysis periods (Chen and Ali, 2004). Table 7 shows that the geometric mean of the technical change (TC) of US airlines for the 10-year period increased by 0.59%, and that of efficiency change (EC) increased by 0.33%. Consequently, MI increased by 0.92% on the strength of the uplift in technical change. Most US airlines except Continental Air Lines, AirTran Airways, and JetBlue Airways, maintained $TC > 1$, indicating technological advances throughout the analysis period. In particular, Virgin America had high productivity growth, showing 3.11% growth in MI with TC = 0.9988 and EC = 1.0322, whereas the MI of AirTran Airways dropped by 1.93% with TC = 0.9848 and EC = 0.9959.

The overall efficiency change of ULCCs had a value above 1, implying that these carriers are efficiently operated. On the contrary, most LCC had a low efficiency change score, indicating that these airlines must take action to increase internal operation efficiency, such as improving market competitiveness, cost structure, and capacity utilization.

Fig. 1 shows the yearly fluctuating patterns of technical change, efficiency change, and Malmquist index, indicating that each index repeatedly increased and decreased with a certain cycle. A fluctuation of efficiency change curve is stable, compared to technical change and Malmquist index, as seen in Fig. 1. Meanwhile, an unstable oscillation of Malmquist index has a similar pattern as technical change. This result indicates that Malmquist index for US domestic airlines was mainly due to a technical change.

Table 7
Changes in the TC, EC, and MI of US airlines from 2006 to 2015.

Group	DMU	EC	TC	MI
NLCs	AS	1.0156	1.0144	1.0302
	AA	1.0018	1.0174	1.0192
	CO	1.0061	0.9911	0.9972
	DL	1.0077	1.0158	1.0236
	NW	0.9867	0.9982	0.9850
	UA	0.9971	1.0173	1.0144
	US	1.0111	0.9955	1.0066
LCCs	FL	0.9959	0.9848	0.9807
	B6	0.9847	1.0047	0.9893
	WN	0.9996	1.0077	1.0073
	VX	0.9841	1.0166	1.0004
ULCCs	G4	1.0168	1.0032	1.0200
	F9	1.0077	1.0173	1.0251
	NK	1.0322	0.9988	1.0311
Geometric Mean		*1.0033*	*1.0059*	*1.0092*

5. Discussion and conclusion

In this study, we employed the DEA and MI model to measure the efficiency and productivity change of US airlines during 2006–2015 and seek sources of inefficiency within individual US airlines to provide insights for airline operators. In addition, we adopted bootstrapped DEA to estimate the M&A effects among US airlines. Finally, we use a double bootstrap regression analysis suggested by Simar and Wilson (2007) to reveal external determinants impact on efficiency. The results presented herein suggest the following three managerial implications.

First, most NLCs and ULCCs have relatively high bootstrapped VRS-DEA values compared with LCCs, as described in Fig. 2. In addition, most NLCs (except Northwest Airlines and Continental Air Lines) and ULCCs have the *Malmquist Index > 1* values, while LCCs (except Northwest Airlines and Virgin America) have a low Malmquist index values.

That is, most LCCs except Southwest Airlines have a relatively low efficiency and productivity. This result suggests that LCCs need

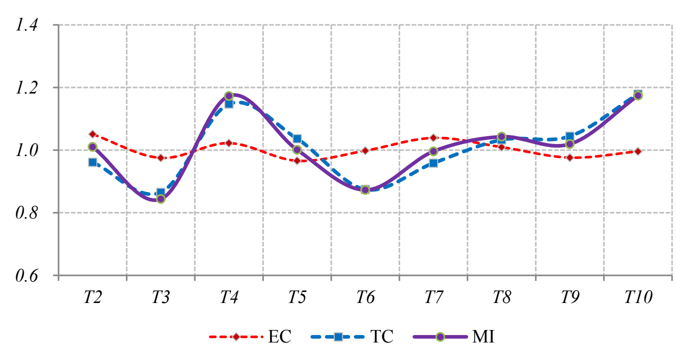

Fig. 1. Productivity changes for US domestic airlines during 2006–2015.

to renovate their service operation systems to make them more suitable for their business scale and achieve service innovations, which offers price advantages and customer convenience simultaneously.

For instance, JetBlue Airways as *mega-LCCs* has a low efficiency and productivity, while *ultra-LCCs* with lower base fares, such as Frontier Airlines, Allegiant Air, and Spirit Air Lines, have a relatively high efficiency and productivity, as shown in Fig. 2. Additionally, the efficiency change score of JetBlue airlines has dropped approximately 1.5% (Table 7), although JetBlue Airways has shown technical advances over the past 10 years. This result suggests that JetBlue Airways is required to make strategic approaches to enhance efficiency by diversifying into new markets and adjusting the cost structure.

Second, US airlines have sought new ways to survive in subsequent external influences such as the 9/11 terror attacks, global financial crisis, and skyrocketing oil prices over the past 15 years. To overcome these environmental risks, airlines tend to choose M&A to reduce unnecessary costs and overlapping routes, thereby raising competitiveness. That is, a motivation of mergers among airlines includes network synergies as well as cost synergies. Network synergies arise from expanding routes/destinations and efficient scheduling. In additional, mergers among airlines have been one alternative for reducing overall operating costs by maximizing the cost synergy and achieving economies of scale.

Despite its popularity and potential benefits, however, many M&A efforts have not achieved their desired results. For instance, Delta Air Lines and Northwest Airlines, which filed a Chapter 11 bankruptcy protection because of their accumulated deficit of $10 billion in 2005, have achieved a more increased efficiency than that before the merger, even though acquirer Delta Air Lines is still in decreasing returns-to-scale region. These two airlines had 12 overlapping nonstop routes, 597 overlapping connecting routes and 44 endpoint airports where both sides had more than 10% of the passengers. Therefore, Delta Air Lines needs to eliminate their overlapping network and to reduce their operating cost.

Since the 2008 financial crisis, additional M&A among US airlines with poor financial performance have continued. For example, following the merger between United Air Lines and Continental Air Lines, the efficiency of United Air Lines is lower than that beforehand. Furthermore, despite expecting to maximize network synergies by eliminating overlapping routes domestically and internationally (Lenartowicz et al., 2013), scale inefficiency remains, which indicates that United Air Lines needs to adjust its input resources to achieve optimal scale and implement multifaceted service innovation to improve the synergy effect. On the contrary, the merger between AirTran Airways and Southwest Airlines is an exemplary by maximizing a cost synergy effects, even though their airline service policies and business model differed (e.g., seat options, classes available).

In general, each airline merger created a different level of synergy, depending on each airline's original route and cost structure (Lenartowicz et al., 2013; Merkert and Morrell, 2012). Moreover, Mergers between airlines, however, often cause conflicts between organizational cultures and service operation systems in addition to employment issues. Therefore, strategic M&A initiatives among airlines with different business models should be developed to maximize the network and cost synergy effects of post-M&A, and eliminate waste in all aspects of operations (De Bondt and Thompson, 1992).

Third, a result of bootstrapped truncated regression suggest that cost factor such as fuel expense and number of full-time equivalent employee has led the US domestic airlines' efficiency to reduce, whereas operating revenue increase the airlines' efficiency in a

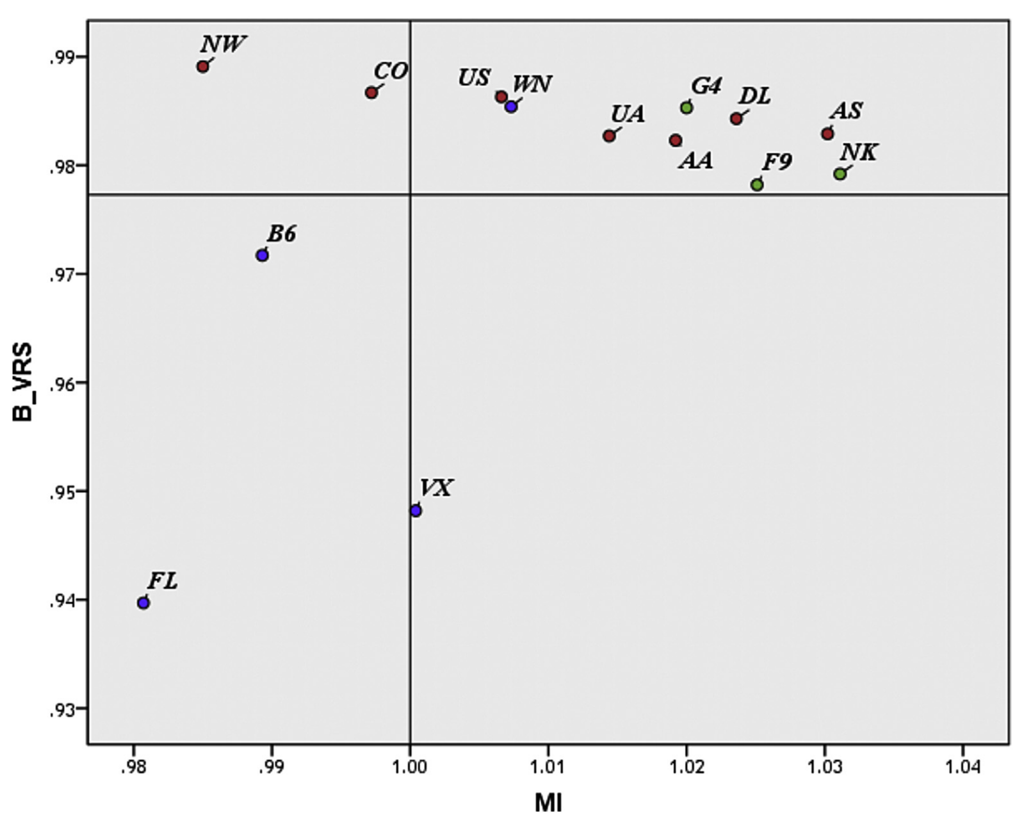

Fig. 2. Categorization for US domestic airlines based on MI and bootstrapped VRS-DEA.

revenue perspectives. Thus, these environmental variables have practical implications for US airline operations and are helpful in developing a sophisticated framework for executing the strategic operation management to maximize their efficiency.

For example, LCCs have reached their limits to growth even though they have enjoyed sustained growth over the past decade (De Wit and Zuidberg, 2012; Fu et al., 2015; Min and Joo, 2016). Specifically, *mega-LCCs* such as Southwest Airlines and JetBlue Airways are now in a vulnerable position in the face of the pressure of fixed costs, which might affect their existing point-to-point strategies and route density. Generally, a business model for LCCs and ULCCs is the cost reduction with low fares and fewer comforts compared to NLCs. In order to improve the efficiency, these airlines should focus their cost efforts on internal cost management initiatives for airline resource such as airplane and full-time equivalents, occupied major part of CASM. Moreover, JetBlue Airways has a relatively low revenue per available seat miles, and Virgin America has the lowest load factor among DMUs seen since 2012 (www.web.mit.edu/airlinedata/). Therefore, JetBlue Airways and Virgin America require a new revenue management to derive revenue from seat sales and to encourage improvement in load factors in the market.

This result addresses strategic operational plans tailored to individual US airlines to improve static efficiency and dynamic productivity in a rapidly changing environment. The results of this study may help US airline industry practitioners to understand the US domestic airline environment from an operator's perspective. Nonetheless, this study have a limitation that we failed to consider airline service quality in this study, despite its importance in the airline industry. Studies, therefore, have a limit on measuring efficiency and productivity without considering airline service quality. Further research must include airline-related other factors as well as airline service quality to confirm the findings of this study.

Funding

This research was financially supported by Hansung University.

Conflict of interest

The author has no conflict of interest to declare.

Acknowledgement

We thank the editor and two anonymous reviewers for their careful reading of our manuscript and their constructive comments and suggestions, which helped us to improve the manuscript. This research was financially supported by Hansung University.

References

Asmild, M., Tam, F., 2007. Estimating global frontier shifts and global Malmquist indices. J. Prod. Anal. 27, 137–148.
Assaf, A.G., Josiassen, A., 2012. European vs. U.S. airlines: performance comparison in a dynamic market. Tour. Manag. 33 (2), 317–326.
Barbot, C., Costa, Á., Sochirca, E., 2008. Airlines performance in the new market context: a comparative productivity and efficiency analysis. J. Air Transp. Manag. 14 (5), 270–274.
Barros, C.P., Liang, Q.B., Peypoch, N., 2013. The technical efficiency of US airlines. Transp. Res. Part A 50, 139–148.
Barros, C.P., Peypoch, N., 2009. An evaluation of European airlines' operational performance. Int. J. Prod. Econ. 122 (2), 525–533.
Barros, C.P., Couto, E., 2013. Productivity analysis of European airlines: 2000-2011. J. Air Transp. Manag. 31, 11–13.
Belobaba, P., Hernandez, K., Jenkins, J., Powell, R., Swelbar, W., 2011. Productivity trends in the US passenger airline industry 1978-2010. In: US Transportation Productivity Study. MIT.
Belobaba, P., Odoni, A., Barnhart, C., 2015. The Global Airline Industry. John Wiley and Sons, Hoboken, NJ.
Chen, Y., Ali, A.I., 2004. DEA Malmquist productivity measure: new insights with an application to computer industry. Eur. J. Oper. Res. 159 (1), 239–249.
Cheng, K., 2010. Evaluation of US legacy airline distribution strategies. J. Air Transp. Manag. 16 (6), 337–339.
De Bondt, W.F.M., Thompson, H.E., 1992. Is economic efficiency the driving force behind mergers? Manag. Decis. Econ. 13 (1), 31–44.
De Wit, G.J., Zuidberg, J., 2012. The growth limits of the low cost carrier model. J. Air Transp. Manag. 21, 17–23.
Duygun, M., Prior, D., Shaban, M., Tortosa-Ausina, E., 2016. Disentangling the European airlines efficiency puzzle: a network data envelopment analysis approach. Omega 60, 2–14.
Färe, R., Grosskopf, S., Lovell, C.A.K., 1994. Production Frontiers. Cambridge University Press, Cambridge.
Fu, X., Lei, Z., Wang, K., Yan, J., 2015. Low cost carrier competition and route entry in an emerging but regulated aviation market – the case of China. Transp. Res. Part A Policy Pract. 79, 3–16.
Franke, M., 2004. Competition between network carriers and low-cost carriers – retreat battle or breakthrough to a new level of efficiency? J. Air Transp. Manag. 10 (1), 15–21.
Hall, P., 1986. On the number of bootstrap simulations required to construct a confidence interval. Ann. Stat. 14 (4), 1453–1462.
Jang, S., Choi, K., Lee, K., 2011. External shocks and efficiency changes in the US airline industry. Serv. Ind. J. 31 (14), 2411–2435.
Kumar, S., 2011. State road transport undertakings in India: technical efficiency and its determinants. Benchmarking Int. J. 18 (5), 616–643.
Lee, B.L., Worthington, A.C., 2014. Technical efficiency of mainstream airlines and low-cost carriers: new evidence using bootstrap data envelopment analysis truncated regression. J. Air Transp. Manag. 38, 15–20.
Lenartowicz, M., Mason, K., Foster, A., 2013. Mergers and acquisitions in the EU low cost carrier market. A Product and Organisation Architecture (POA) approach to identify potential merger partners. J. Air Transp. Manag. 33, 3–11.
Li, Y., Wang, Y., Cui, Q., 2015. Evaluating airline efficiency: an application of virtual frontier network SBM. Transp. Res. Part E 81, 1–17.
Lovell, C.A.K., 1993. Production frontiers and productive efficiency. In: Fried, H.O., Lovell, C.A.K., Schmidt, S.S. (Eds.), The Measurement of Productive Efficiency. Oxford University Press, Oxford.
Mallikarjun, S., 2015. Efficiency of US airlines: a strategic operating model. J. Air Transp. Manag. 43, 46–56.
Merkert, R., Morrell, P.S., 2012. Mergers and acquisitions in aviation – management and economic perspectives on the size of airlines. Transp. Res. Part E 48, 853–862.
Min, H., Joo, S., 2016. A comparative performance analysis of airline strategic alliances using data envelopment analysis. J. Air Transp. Manag. 52, 99–110.
Pires, H.M., Fernandes, E., 2012. Malmquist financial efficiency analysis for airlines. Transp. Res. Part E 48, 1049–1055.
Staat, M., 2006. Efficiency of hospitals in Germany: a DEA-bootstrap approach. Appl. Econ. 38 (19), 2255–2263.
Schmenner, R.W., 1986. How can service businesses survive and prosper? Sloan Manag. Rev. 27 (3), 21–32.
Simar, L., Wilson, P.W., 2007. Estimation and inference in two-stage, semi-parametric models of production processes. J. Econ. 136, 31–64.
Simar, L., Wilson, P.W., 1998. Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models. Manag. Sci. 44, 49–61.
Simar, L., Wilson, P.W., 2000. A general methodology for bootstrapping in non-parametric frontier models. J. Appl. Stat. 27 (6), 779–802.
Tulkens, H., Vanden Eeckaut, P., 1995. Non-parametric efficiency, progress and regress measures for panel data: methodological aspects. Eur. J. Oper. Res. 80 (3), 474–499.