Journal of Air Transport Management 35 (2014) 87–91

Contents lists available at ScienceDirect

Journal of Air Transport Management

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

Validating delay constructs: An application of confirmatory factor analysis

Tony Diana*

Federal Aviation Administration, 800 Independence Avenue S.W., Washington, DC 20591, USA

Keywords:
Delay
Airport capacity
Arrival and departure demand
Confirmatory factor analysis

ABSTRACT

This paper proposes to use confirmatory factor analysis (CFA) to evaluate the relationship between six observed variables (arrival and departure counts, arrival and departure demand, taxi-out and airborne delays) and their underlying latent (unobserved) constructs (operations, demand, and delays) at six of the most delayed airports (EWR, JFK, LGA, MIA, ORD, and SFO) during the calendar years of 2006–2008. The CFA revealed a good fit between the six observed variables and the three factors that may explain on-time performance except in the case of JFK. The use of CFA can help analysts validate constructs when theory supports *a priori* predictions and relationships between observed and unobserved variables.

Published by Elsevier Ltd.

1. Introduction

In theory, delays represent the outcome of a trade-off between demand for arrivals and departures and available airport capacity. When an airport does not have enough capacity to satisfy demand, flights get delayed and sometimes cancelled. In this study, airport capacity is understood as the sum of an airport's arrival and departure rates. Delays usually increase to the point where congestion requires traffic management initiatives (TMI) such as ground delay programs or miles-in-trail in order to re-balance demand with capacity and to maintain a certain level of service (i.e. departures and arrivals within 15 min of airlines' published schedules[1]). In a context of constrained budgets and environmental scrutiny, airport operators do not have much time and resources to construct a new runway soon enough to improve on-time performance as demand grows. In fact, it takes more than a decade for an airport to open a new runway to traffic—often at a cost of over one billion dollars and lengthy litigations with surrounding communities opposed to airport expansion. Moreover, delays do not remain a local problem: They propagate throughout the National Airspace System (NAS) that constitutes a system of interdependent airports (Wu, 2010). The theoretical background that supports this article can be found in De Neufville and Odoni (2003), Janic (2009), Belobaba et al. (2009), and Horonjeff et al. (2010).

This paper proposes to use confirmatory factor analysis (CFA) to evaluate the relationship between six observed variables (i.e. arrivals and departures, arrival and departure demand, taxi-out and airborne delays) and their underlying latent (unobserved) constructs (i.e. operations, demand, and delays) at six of the most delayed airports during the 2006 to 2008 time period regardless of each airport's specificities (i.e., number of runways, percentage of operations in instrument meteorological conditions, among others). Table 1 of the appendix provides the details on on-time performance by sampled airport and calendar year. CFA allows analysts to determine whether the hypothesized structure provides a good fit to the data. Is there a relationship between the six manifest and the three latent factors which may explain poor on-time performance at the selected airports? The use of CFA can help aviation analysts validate constructs and eventually predict on-time performance.

Although CFA is popular in the fields of psychology and social work, it has not been widely utilized in research studies on aviation delays and airport capacity. Liedtka (2002) used CFA to investigate the information content of nonfinancial performance measures in the airline industry. Park (2007) evaluated the effects of airline service quality on airline image and passengers' future behavioral intentions. He examined passengers' perception of eleven factors that may influence their buying behavior and concluded that passenger perceptions are significantly different across airlines, seat classes, and usage frequencies. Babbar and Koufteros (2008) used a confirmatory approach to determine how much personal touch

This article does not reflect the official opinion of the Federal Aviation Administration.

* Tel.: +1 202 385 7311.
E-mail address: Tony.Diana@faa.gov.

[1] Under Title 14, Part 234 of the Code of Federal Regulations (14 CFR 234), a delay is defined as any gate departure or arrival time greater than 15 min in comparison with the airline's published schedule.

0969-6997/$ – see front matter Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.jairtraman.2013.11.014

Table 1
Variables and data sources.

Variable	Definition	Data sources
Total volume of arrivals and departures	The respective counts of individual incoming and outgoing flights	TFMS, ASPM
Arrival demand	The number of aircraft that have taken off but not yet landed at the destination airport	ASPM
Departure demand	The number of aircraft that have left the gate but not yet taken off	ASPM
Taxi-out delay	The difference between taxi-out time (from gate-out to wheels-off time) and an unimpeded taxi-out time.	ARINC's Out-On-Off-In data, ASPM, ASQP
Airborne delay	The difference between airborne time and the flight plan's estimated time enroute	ASPM

Note. TFMS, Traffic Flow Management System; ASPM, Aviation System Performance Metrics; ASQP, Airline Service Quality Performance.

Table 2
The variables at the six sampled airports (CY 2006–2008).

Variables	EWR			JFK			LGA			MIA			ORD			SFO		
	2006	2007	2008	2006	2007	2008	2006	2007	2008	2006	2007	2008	2006	2007	2008	2006	2007	2008
Average daily departures	609	594	594	524	610	605	551	538	521	516	516	493	1309	1269	1203	492	508	519
Average daily arrivals	572	575	571	502	598	591	551	547	528	443	443	433	1285	1275	1207	434	460	484
Average daily departure demand	984	952	970	894	1116	964	795	838	808	531	532	524	1557	1505	1407	512	546	553
Average daily arrival demand	1118	1160	1100	650	911	888	907	985	936	534	530	510	2055	1951	1861	629	665	749
Average minutes of taxi-out delay	15.4	15.2	15.6	17.1	18.9	14.3	13.0	15.0	14.7	3.4	3.5	3.4	7.8	7.6	7.0	4.4	5.3	4.5
Average minutes of airborne delay	6.2	5.5	4.8	6.0	7.2	6.3	6.0	6.9	7.3	2.2	1.9	2.9	3.6	3.8	3.9	3.1	3.4	3.6

displayed by contact employees is likely to have an impact on passenger satisfaction.

The six sampled airports include Newark Liberty (EWR), New York John F. Kennedy (JFK), New York LaGuardia (LGA), Chicago O'Hare (ORD), Miami International (MIA), and San Francisco International (SFO). Only LGA does not serve as a key international gateway among the sampled airports. The six observed variables were measured at a time when poor on-time performance led the FAA to intervene through demand management initiatives. In 2008, flights were capped at EWR and JFK during peak hours[2] (LGA already had limits on flights). From 2006 to 2008, the number of hourly arrivals at ORD could not exceed 88 from 07:00 to 19:59 local time. Moreover, the variables under consideration were observed prior to three events that have changed the U.S. aviation landscape: (1) the recession that hit the U.S. economy in 2008, (2) a decline in passenger demand that led to airlines' schedule reductions and (3) airline consolidations.

The next section defines the variables and their data sources. After briefly outlining the differences between exploratory and confirmatory factor analysis, the discussion will proceed with a description of the CFA model and then an explanation of its outputs.

2. The data

The CFA model includes the following variables aggregated at the monthly level for the calendar years 2006, 2007 and 2008. The data were collected from the Aviation System Performance Metrics (ASPM) data warehouse.[3] The variables, definitions and data sources are specified in Table 1.

Table 2 provides quantitative data about the six variables for the three years under consideration. The three New York airports and ORD show a strong imbalance between demand and recorded operations that is likely to result in poor on-time performance.

Although more variables such as taxi-in and block delays were initially introduced into the model, the number was reduced to the six identified ones in this paper on the basis of improvement in the Parsimonious Goodness-of-Fit Index (PGFI) and other goodness-of-fit statistics specified later in Table 3. Moreover, these six variables play a significant role in the theory of delay and delay modeling. Their interactions often explain airport delays and congestion. As demand increases, departing aircraft are likely to experience delays because air traffic control can only process so many planes per hour due to (1) the choice of runway configuration dictated by wind direction, ceiling, and visibility, (2) separation requirements that limit throughput (i.e. the volume of operations in periods of sustained demand), and (3) enroute congestion.

The next section focuses on the CFA methodology and its use to test whether there is a relationship between the observed variables and the latent factors as a predictor of poor on-time performance.

3. Methodology

3.1. Exploratory v. confirmatory factor analysis

Exploratory factor analysis (EFA) differs from confirmatory factor analysis (CFA) in that the former is mainly designed to explore the underlying factor structure of a set of observed variables without any preconceived model or structure of the outcome (see Child, 1990). EFA enables the determination of underlying constructs for a group of measured variables. EFA does not postulate any relationship pattern *a priori*, nor does it pertain to test hypotheses related to theoretical models.

On the other hand, CFA makes it possible to test whether a relationship between observed variables and underlying latent construct exists. CFA is grounded in theories and it requires the specification of an *a priori* model, the determination of a number of factors, as well as the identification of which variable loads on each

[2] At JFK, the total operations target was 80 operations per hour, except for the 15:00 through 19:59, when it was 81 total operations. The 30-min maximum was 44 operations and the 15-min maximum was 24 operations.

[3] The ASPM data warehouse is accessible at https://aspm.faa.gov.

Table 3
Goodness-of-fit statistics.

		EWR	JFK	LGA	MIA	ORD	SFO
Fit summary							
Modeling info	*N* observations	36	36	36	36	36	36
	N variables	6	6	6	6	6	6
	N moments	21	21	21	21	21	21
	N parameters	15	15	15	15	15	15
	N active constraints	0	0	0	0	0	0
	Baseline model function value	10.4916	10.8229	10.1948	11.4886	11.2428	11.2904
	Baseline model DF	15	15	15	15	15	15
Fit index	Fit function	0.1851	1.5356	0.0829	0	0.0057	0.017
	Model DF	6	6	6	6	6	6
	Root Mean Square Residual (RMSR)	230.92	133.091	115.3141	17.8914	2643.4899	691.2168
	Standardized RMSR (SRMSR)	0.0398	0.2083	0.026	0.0005	0.009	0.0157
	Goodness-of-fit index (GFI)	0.9904	0.9232	0.9956	1.0000	0.9997	0.9991
	Adjusted GFI (AGFI)	0.9665	0.7312	0.9847	1.0000	0.999	0.9969
	Parsimonious GFI	0.3962	0.3693	0.3982	0.4000	0.3999	0.3997
	Bentler–Bonett normed fit index	0.9824	0.8581	0.9919	1.0000	0.9995	0.9985

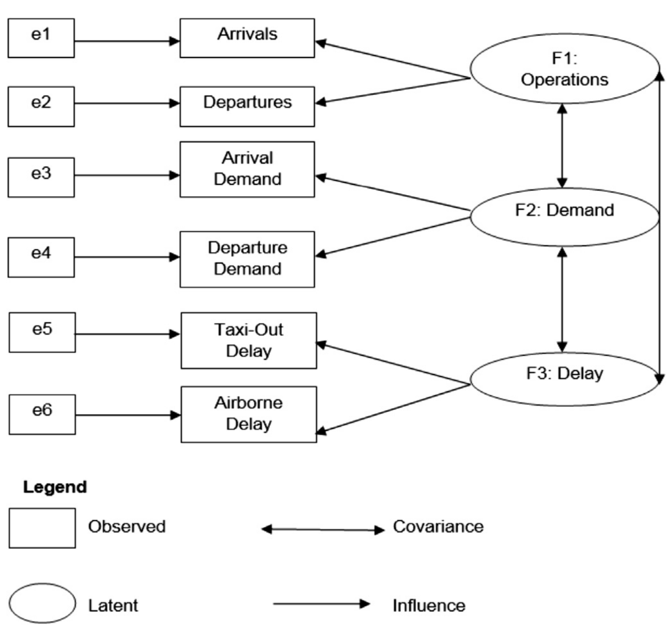

Fig. 1. The CFA model diagram.

factor. The relationship among the variables and latent construct is mapped in Fig. 1.

In this study, CFA was selected in order to achieve the following objectives:

- Determine whether there is a relationship between six key operational variables (arrival and departure counts, arrival and departure demand, taxi-out and airborne delays) and three latent variables (operations, demand and delays) based on previous research and theories mentioned in the introduction;
- Test the hypothesis of a relationship between the selected variables and the latent variables statistically;
- Establish whether the model fits in the case of six airports that differ in terms of traffic volume, location and runway configurations at three different time periods.

CFA is a special case of Structural Equation Model (SEM) that supports the investigation of causal relationship among latent variables. Readers interested in a clear exposition of CFA and SEM are referred to Jöreskog (1969, 1993), Brown (2006) and Harrington (2009).

3.2. *The confirmatory factor analysis model*

The sample data were processed with the SAS software. The CALIS procedure enabled the computation of CFA outputs. CALIS stands for 'Covariance Analysis of Linear Structural Equations'. Diagonally weighted least-squared (DWLS) estimation was used in this study because, as Mindrila (2010:60) put it, "DWLS provides more accurate parameter estimates, and a model fit that is more robust to variable type and non-normality" than Maximum Likelihood (ML) or Generalized Least Squares (GLS). All the six variables were positively skewed (the skewness coefficient was greater than zero). The variables also featured a peaked distribution (the kurtosis coefficient was close to 4.0). The exception was taxi-out delays at JFK that featured a distribution close to normal ($K = 0.08$). No data were trimmed from the airport samples so as to ensure generalization of the findings.

In the CFA model, the covariances among the latent factors were assumed to be correlated and free parameters. For identification purposes, the factor variances were fixed at 1.0 in order to make sure that the number of estimated parameters was less than or equal to the number of unique variances and covariances among the measured variables. Fig. 1 shows the three latent variables or factors labeled F1 to F3, the six endogenous variables under consideration (the variables that receive the arrows) as well as the exogenous error terms.

4. The outcomes of the confirmatory factor analysis

Table 3 provides the modeling information and the key statistics to assess whether the model fits in the case of six airports over the three calendar years. The standardized root mean square residuals (SRMSR) statistic represents the standardized average squared differences between the residuals of the sample covariances and the residuals of the estimated covariances. According to Hu and Bentler (1999), values as high as 0.08 for the SRMSR are deemed acceptable. The smaller the value of the SRMSR, the better the fit of the model. A value close to zero indicates a perfect fit. The exception is JFK (the index was 0.2083).

The goodness-of-fit indexes determine how closely the variances and covariances accounted for by the model fit the observed covariance matrix. Traditionally, a value higher than or equal to 0.90 is desirable. The statistics in Table 3 indicate acceptable goodness-of-fit for all the airport models except in the case of JFK. Whereas the GFI for JFK is above 0.90, both the AGFI and Normed Fit Index are lower than the desirable threshold. Several reasons may explain the lack of fit in the case of JFK.

The number of JetBlue and Delta's scheduled flights increased 24 percent[4] between 2006 and 2007. According to ASPM, while the volume of operations increased 15.8 percent from calendar year 2007 and 2006, the total airport capacity at JFK only increased 7.2 percent during the local hours of 07:00 to 21:59. Faced with an increase in the number of operations, JFK resorted to ground delays in order to reduce departure demand and airborne delays. Based on OPSNET data in Table 2 of the appendix, the total minutes of expected departure clearance time (EDCT) increased twofold between 2006 (288,780 min) and 2007 (873,833 min). It is also important to note that total number of traffic management initiatives to JFK defined by the FAA as "techniques used to manage demand with capacity in the NAS"[5] increased 176.9 percent between calendar year 2007 and 2006 and 36.9 percent between 2008 and 2007 (source: OPSNET). Traffic management initiatives refer to miles or minutes-in-trail, sequencing programs, enroute sequencing programs, among other initiatives aimed to pace or delay arrivals into an airport when demand exceeds available airport capacity.

5. Concluding remarks

Using the cases of six congested airports, the purpose of the CFA was to validate the theoretical assumption that there is a relationship between six observed variables (i.e. arrival and departure demand, arrival and departure counts, taxi-out and airborne delays) and their underlying latent constructs represented by three factors (i.e. operations, demand, and delay). Most of the models exhibited acceptable goodness-of-fit (regardless of airport specificities) except for JFK where an increase in operations in the face of airspace and airside constraints explained poor on-time performance in the summer of 2007. The outcomes of the CFA highlight the significance of delay as the trade-off between arrival and departure demand and airport capacity at the six most congested U.S. airports from 2006 to 2008. The balance between demand and capacity is precarious when airlines increase their operations while airport capacity cannot significantly change as the JFK example illustrated.

Airport operators and aviation regulators have limited options to maintain the balance between demand and capacity. A push by U.S. Department of Transportation in favor of slot auctions at congested airports faced opposition. At ORD, the FAA imposed a reduction in the number of scheduled arrivals and departures from 07:00 to 21:00 local time during weekdays and from 12:00 to 21:00 on Sundays. The FAA rules limiting flights were initiated in October 2006 and sunset in 2008 as the airport was ready to commission a new runway 9L | 27R on November 21, 2008. With the additional runway, arrival capacity increased from 96 to 112. According to the Chicago Airport System Commissioner Rosemarie Andolino, "the new runway reduced O'Hare's average delay to 16 minutes from 24, and when the entire modernization program is completed, it will take the delay factor down to six minutes."[6]

The Next Generation Air Transportation System (NextGen) aims to address delay and capacity issues in the NAS. NextGen programs are designed to transition the present air traffic control system (ATCS) utilizing ground-based radars to an air traffic management system (ATMS) relying on global positioning system (GPS) and satellite navigation, more data-centric communication systems between Air Traffic Management System (ATMS) and pilots, performance-based navigation (PBN), and Automatic Dependent Surveillance-Broadcast (ADS-B).

Since 2008, several initiatives have been introduced at JFK to mitigate delays: Departure queue management, taxiway improvements, and departure flow de-confliction. It may be of great interest to aviation practitioners to determine whether the construct and the outcomes described in this study are likely to change at the six airports as NextGen programs are implemented and the economy and passenger demand grow back up.

Appendix

Table 1
On-time performance.

	EWR			JFK			LGA			MIA			ORD			SFO		
	2006	2007	2008	2006	2007	2008	2006	2007	2008	2006	2007	2008	2006	2007	2008	2006	2007	2008
% On-time departures	74.03	70.51	71.34	76.52	71.31	77.79	77.65	75.52	78.83	79.16	74.24	72.66	71.34	69.59	71.58	77.17	77.1	76.77
% On-time arrivals	64.77	62.15	64.73	70.84	65.28	70.85	66.89	61.96	66.38	77.34	72.06	71.98	71.05	69.28	71.21	71.88	71.25	70.69

Table 2
Ground stops and expected delay clearance times.

	EWR			JFK			LGA			MIA			ORD			SFO		
	2006	2007	2008	2006	2007	2008	2006	2007	2008	2006	2007	2008	2006	2007	2008	2006	2007	2008
Ground stops (min)	195,750	181,238	58,615	48,171	73,691	33,842	88,792	82,908	50,000	8199	3887	2836	235,817	197,979	35,122	19,605	18,375	9309
EDCT (min)	1,991,946	2,589,522	3,148,127	288,780	873,833	1,312,537	1,300,524	1,967,965	2,408,079	129	2108	0	3,115,412	3,095,744	4,142,923	486,485	617,624	1,047,608

Sources: BTS for on-time performance, OPSNET for ground delay and EDCT minutes. The information is compiled in ASPM.

[4] Source: OAG data.

[5] Federal Aviation Administration, Order JO 7210.3X, Section6, Traffic Management Initiatives, February 9, 2012, accessed October 6, 2013, http://www.faa.gov/air_traffic/publications/atpubs/fac/1706.html.

[6] Scott McCartney, "How a New Runway At O'Hare Makes Travel Easier for All," The Wall Street Journal, July 22, 2009, accessed October 6, 2013, http://online.wsj.com/article/SB10001424052970203739404574290030754087254.html.

References

Babbar, S., Koufteros, X., 2008. The human element in airline service quality: contact personnel and the customer. Int. J. Operat. Product. Manage. 28 (9), 804–830.

Belobaba, P., Odoni, A., Barnhart, C. (Eds.), 2009. The Global Airline Industry. John Wiley and Sons, Chichester, U.K.

Brown, T.A., 2006. Confirmatory Factor Analysis for Applied Research. The Guilford Press, New York.

Child, D., 1990. The Essentials of Factor Analysis, Second Edition. Cassel Educational Limited, London.

De Neufville, R., Odoni, A., 2003. Airport Systems: Planning Design, and Management. McGraw Hill, New York.

Harrington, D., 2009. Confirmatory Factor Analysis. Oxford University Press, Oxford.

Horonjeff, R.M., McKelvey, F.X., Sproule, W.J., Young, S., 2010. Planning and Design of Airports, fifth ed. McGraw Hill, New York.

Hu, L.T., Bentler, P.M., 1999. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct. Eqn. Model.: Multidisciplin. J. 6 (1), 1–55.

Janic, M., 2009. Airport Analysis, Planning and Design: Demand, Capacity, and Congestion. Nova Publishers, Hauppauge, N.Y.

Jöreskog, K.G., 1969. A general approach to confirmatory maximum likelihood factor analysis. Psychometrika 34 (2), 183–202.

Jöreskog, K.G., 1993. Testing structural equation models. Sage Focus Ed. 154, 294.

Liedtka, S.L., 2002. The information content of nonfinancial performance measures in the airline industry. J. Business Finan. Account. 29 (7–8), 1105–1121.

Mindrila, D., 2010. Maximum likelihood (ML) and diagonally weighted least squares (DWLS) estimation procedures: a comparison of estimation bias with ordinal and multivariate non-normal data. Int. J. Digit. Soc. 1 (1), 93–102.

Park, J.W., 2007. Passenger perceptions of service quality: Korean and Australian case studies. J. Air Transport Manage. 13 (4), 238–242.

Wu, C.-L., 2010. Airline Operations and Delay Management: Insights from Airline Economics, Networks and Strategic Schedule Planning. Ashgate Publishing Company, Farnham, U.K.