# Assessing the effects of flight delays, distance, number of passengers and seasonality on revenue



Received 12 January 2018
Revised 2 November 2018
Accepted 10 December 2018

Murat Guven
*Department of Economics, Istanbul Technical University, Istanbul, Turkey*
Eyup Calik and Basak Cetinguc
*Yalova University, Yalova, Turkey*
Bulent Guloglu
*Department of Economics, Istanbul Technical University, Istanbul, Turkey, and*
Fethi Calisir
*Istanbul Technical University, Istanbul, Turkey*

## Abstract

**Purpose** – This study aims to investigate the effects of flight delays, distance, number of passengers and seasonality on revenue in the Turkish air transport industry.

**Design/methodology/approach** – The domestic return routes of a Turkish airline company were examined to address this issue. Among five cities and six airports, 14 major domestic return routes were selected. The augmented mean group (AMG) estimator and common correlated effects mean group (CCEMG) estimator were conducted with a two-way fixed effects (FE) robustness test in this study.

**Findings** – The results show that arrival flight delay and departure flight delay had negative effects on revenue, whereas the distance between airports, the number of air passengers and seasonality had positive effects on revenue.

**Research limitations/implications** – The data used in this study were retrieved from a Turkish airline company; for future research, other airline companies operating in Turkey may be included.

**Practical implications** – These findings could be evaluated by air transportation leaders to provide a guide to make strategic decisions to achieve greater performance in this competitive environment.

**Originality/value** – The originality of the paper comes from the facts that besides distance and number of passengers, the authors control for the seasonality when assessing the effects of flight delay on revenue; they use panel data techniques, which permit them to control for individual heterogeneity, and create more variability, more efficiency and less collinearity among the variables; they use two recent panel data techniques, CCEMG and AMG, allowing for cross-section dependence.

**Keywords** Distance, Seasonality, Revenue, Flight delays, Passenger

**Paper type** Research paper

## Introduction

The air transportation industry has a dynamic liquidity and is severely competitive. Revenue is one of the most important competition tools. The air transport industry made 5 per cent net profit on revenues of $705bn in 2016 [International Air Transport Association (IATA), 2017]. The annual revenue of this industry in Turkey was $23.4bn in 2015

Kybernetes
© Emerald Publishing Limited
0368-492X
DOI 10.1108/K-01-2018-0022

(SHGM, 2017). Furthermore, more than 3.8 billion passengers were flown safely in 2016, an increase of approximately 300 million compared with 2016 (IATA, 2017). A tremendous increase in both international and domestic passenger traffic in Turkey was observed. International passenger traffic increased to 70 million in 2016 from 25 million passengers in 2003, while domestic passenger traffic rose to 103 million in 2015 from 9 million passengers in 2003; moreover, there are six airline companies with 540 aircraft and a capacity of 100 thousand passengers (SHGM, 2017). Besides, there are several factors that affect the revenue of air transportation companies. In this study, the prominent factors are determined as number of passengers carried, flight delays, distance, and seasonality. The number of passengers carried represents commuters flown by the airline with the given route and time period. It is anticipated that a higher number of passengers may increase revenue. The number of passengers is also the first factor coming to mind while evaluating revenue in the airline industry. In the literature, available seat kilometers (ASK) and revenue passenger kilometers (RPK) are commonly used factors to measure the efficiency of not only an airline company but also airline industry. The relationship between ASK and RPK is handled by international airlines in given periods (Li *et al.*, 2015; Xu and Cui, 2017).

Furthermore, the flight delays is another factor effecting revenue. These delays occur for several reasons, including extreme weather conditions, airport operations, air traffic control, and security issues (Yablonsky *et al.*, 2014). Flight delays are economic, social, and environmental problems that cause inconvenience for both airline companies and passengers (Blackwood, 2012). Flight delays not only irritate air passengers and disrupt their schedules but also cause a decrease in efficiency, an increase in capital costs, reallocation of flight crews and aircraft, and additional crew expenses (Britto *et al.*, 2012; Yablonsky *et al.*, 2014). As flight delays require the consumption of extra labor, capital, and other inputs necessary in the process, higher operating costs are inevitable for airline companies (Zou and Hansen, 2014). Thus, the effect of flight delays on revenue is perceived as complicated due to the nature of high uncertainty. Moreover, flight delays cost air passengers and airline companies billions of dollars each year (Cook *et al.*, 2004). For example, Pegasus Airlines, an airline company in Turkey, has stated that flight delays cost them 12m Turkish Liras (TRY) every year; Atlas Global Airlines, another Turkish airline company, has also declared that costs increase by 20 per cent due to flight delays (Ozbek, 2015), indicating that these delays result in serious costs to airline companies, many millions of TRY each year. In air transportation systems, economic losses, safety issues, and flight delays are increased by heavy traffic congestion (Zhang *et al.*, 2016). Furthermore, airline companies deal with the consequences of flight delays by adjusting airfares and flight frequency. However, while making these adjustments, managers should seek equilibrium between recovering operating cost increases and maintaining passenger demand, because consumers may not be willing to pay higher airfares (Zou and Hansen, 2012). Moreover:

> [. . .] consumers may consider the potential for the delay before choosing to make a booking. As a result, on an aggregate basis, an airline's record of flight delays may have a negative impact on passenger demand (Britto *et al.*, 2012).

Thus, the decrease in passenger demand caused by flight delays directly affects revenue negatively.

At the same time, distance commonly affects revenue. Basically, distance is the kilometers between the departure and arrival airports. Generally, airfares may be expected to rise in direct proportion to distance (Hofer *et al.*, 2008) and may affect the revenue. Moreover, seasonality is another factor that affects revenue. Seasonal periods are considered as playing an important role in predicting airfares. Furthermore, the seasons are divided into two groups – peak and off-peak. The peak season includes the months from April to September, whereas the off-peak season is from November to March. Chiou and Liu (2016) show that air passengers, in hot seasons, prefer to buy their tickets earlier. Additionally, Hui and Dao-ming (2009) argue that off-season causes the single-leg seating problem in the revenue management of the airline industry. To observe peak and off-peak effect on revenue, in our study the data were collected for all monthly periods between the given years, which is different from previous studies for which the data were collected only quarterly (Britto *et al.*, 2012; Forbes, 2008; Hofer *et al.*, 2008; Scotti and Dresner, 2015; Zou and Hansen, 2014). At the same time, we used panel data which combines both cross-section and time series. Moreover, as argued by Baltagi (2013), panel data control for individual heterogeneity and create more variability, more efficiency, and less collinearity among the variables. Therefore, our study benefits from these advantages of using panel data. Furthermore, the domestic return routes of a Turkish airline company were examined to address this issue. Two recently developed panel data techniques that allow for slope heterogeneity and cross-section dependence, named the panel common correlated effects mean group (CCEMG) estimator and augmented mean group (AMG) estimator, are employed. As argued by O'Connell (1998) ignoring cross-sectional dependence can lead to over-rejection of the unit root hypothesis and creates a bias in the estimation of coefficients. Using CCEMG and AMG estimators and panel unit root tests allowing for cross sectional dependence as well. To check the robustness of the results we also employ two way fixed effects estimators. This is the novelty of the paper and we thus contribute to the empirical literature.

The following section discusses the conceptual background. The model and data are introduced in third section. The fourth section mentions descriptive statistics and correlations. Empirical findings and discussion are then presented in the fifth section. The conclusion and further research section is the last part of the study.

## Conceptual background

Several studies have explored the factors affecting revenue. Li *et al.* (2015) evaluated the efficiency of 22 international airlines from 2008 to 2012, and they found a strong positive relationship between RPK and ASK. Additionally, Xu and Cui (2017) analyzed 19 international airlines from 2008 to 2014 and found that there is a strong positive relationship between RPK and ASK.

Moreover, flight delays affect passenger demand and fares (Evans and Schäfer, 2013), which directly affect costs and revenue. Britto *et al.* (2012) examined the impact of flight delays on both passenger demand and airfares. Their data set covered quarters from 2006 to 2008 and included 57 short-haul routes in the US that were less than 500 miles and operated by 14 carriers (Full Service Network Carriers [FSNC]: 9, Low Cost Carriers [LCC]: 5). They also included control variables such as distance, competition type (FSNC or LCCs), market concentration (measured by the Herfindahl–Hirschman index [HHI]), route types, income, and population. They used the two-stage least square (2SLS) technique to estimate the model. They found that flight delays on a route reduced passenger demand and raised airfares. Moreover, in the passenger equation, distance, LCCs on the route, adjacent route LCCs, vacation route, population, and income were found to be significantly related to passenger demand, while in the fare equation, distance, competition type, LCCs on the route, and adjacent route LCCs were significantly related to fares. In the studies by Britto *et al.* (2012), Forbes (2008), Hofer *et al.* (2008), and Zou and Hansen (2014), the effects of the total number of passengers, distance, market concentration, and vacation route were the main variables examined, and significant effects on airfares were found. Furthermore, Zou and Hansen (2014) assessed the impact of flight delays on airfares and flight frequency in the US air transportation system. They focused on domestic flights by major LCCs such as American West Airlines, JetBlue Airways, Southwest Airlines, and Sun Country Airlines, and covered data from the first quarter of 2004 to the fourth quarter of 2008. Two different models were used: nonstop and one-stop route fare models. In their research models, they introduced some common variables such as flight delays at both origin and departure airports, the total number of passengers, market concentration, distance, vacation route, LCCs on the route, and adjacent route LCCs. The 2SLS method was used to estimate the models. They found that the variables mentioned above were significantly related to airfares in both models. Additionally, Forbes (2008) studied the impact of flight delays on airfares for 18 routes from LaGuardia Airport in New York. Forbes collected data quarterly between 1999 and 2000 and used ordinary least squares (OLS) for estimation. Forbes formed three panel groups for the study: direct passengers on LaGuardia routes, connecting passengers on LaGuardia routes, and direct passengers including routes to Reagan National Airport as the control group. According to the results, delays at LaGuardia led to a decrease in airline prices on routes to and from that airport and an increase in airfares on routes to and from other New York City airports. Forbes also found that connecting passengers were less sensitive to delays than direct passengers. Hofer *et al.* (2008) also investigated the factors affecting passenger demand and airfares. Data were collected from 1,000 US domestic-origin and destination-route markets for all quarters in 1992, 1997, and 2002. The model was tested using 2SLS. In the research model, variables such as market concentration, distance, route type, and LCCs on the route were introduced. At the same time, Scotti and Dresner (2015) examined the impact of baggage fees on passenger demand for US air routes. Their data set covered domestic airport-to-airport routes operated by Southwest Airlines between 2007 and 2010. They also included other variables such as the number of passengers, population, income, route type, distance, and market concentration. They conducted a three-stage least squares (3SLS) technique. Their analysis results showed that higher baggage fees caused a decrease in both the number of passengers and airfares. They also found that distance and market concentration were positively related to airline fares, and tourist routes, population, and income had a positive impact on passenger demand.

As mentioned above, the effects of different factors on revenue were discussed in the literature. Based on the aggregate discussions, the adopted model will be introduced in the next section.

## Model and data

As discussed in the previous section, different models and estimation techniques have been used to examine the impact of revenue on air transportation. In this study, we established a regression model given in equation (1). The variables retained in equation (1) are selected following the recent literature on revenue. We also included a seasonality variable to control for the effect of the peak season in addition to the variables in equation (1). Seasonality can be deterministic or random. Actually, time series models such as SARMA (Seasonally Autoregressive Moving Average) are more apt to take into account random seasonality. However, in the airline context seasonality is more deterministic than random as the peak (low) seasons can occur in the same period each year. Both in a panel and a time series model deterministic seasonality is usually modeled using the dummy variables. Thus, there is no difference between a panel and a time series in modeling seasonality. As explained in the above, panel data techniques have some advantage over time series and cross section models. As panel models take into account both time and cross section dimension of the data and they provide more information, more efficiency, more degrees of freedom, less multicollinearity, and they allow for individual heteroscedasticity. Besides panel data, models identify and measure effects that are simply not detectable in pure cross-section or pure time-series data (Baltagi, 2013).

Regression equation and its terms are given as follows:

$$\log(revenue)_{it} = \beta_{0i} + \beta_{1i}\log(arrdelay_{it}) + \beta_{2i}\log(depdelay_{it}) + \beta_{3i}\log(dis_i) + \beta_{4i}\log(routepax_{it}) + \beta_{5i}season_{it} + u_{it} \tag{1}$$

with:

$$u_{it} = \mu_i + \varphi_i v_t + \varepsilon_{it} \tag{2}$$

where t = 2008:01. . .2013:12 and i = 1, 2. . .14

In the above panel data regression, $t$ represents the time period and $i$ stands for individuals.

In this equation, the terms $\mu_i$, $\varphi_i$, $\nu_t$ and $\varepsilon_{it}$ represent unobserved individuals effects, company-specific factor, common effects and individual specific errors, respectively.

Besides, the description of the variables is as follows:

- *Log (revenue$_{it}$)*: logarithm of unit revenue passenger kilometer (RPK) for the given route and time period;
- *Log (arrdelay$_{it}$)*: logarithm of the average arrival flight delay at the origin airport for the given route and time period (in minutes);
- *Log (depdelay$_{it}$)*: logarithm of the average departure flight delay at the origin airport for the given route and time period (in minutes);
- *Log (routepax$_{it}$)*: logarithm of the number of air passengers carried for the given route and time period;
- *Log (distance$_i$)*: logarithm of the distance measured between airports (in kilometers); and
- *season*: dummy variable taking 1 if the season is peak and 0 otherwise.

Furthermore, the null and alternative hypotheses on model parameters can be formed as follows:

$$\begin{array}{lllll} H_o : \beta_1 = 0 & H_o : \beta_2 = 0 & H_o : \beta_3 = 0 & H_o : \beta_4 = 0 & H_o : \beta_5 = 0 \\ H_a : \beta_1 \prec 0 & H_a : \beta_2 \prec 0 & H_a : \beta_3 \succ 0 & H_a : \beta_4 \succ 0 & H_a : \beta_5 \succ 0 \end{array}$$

The first two hypotheses above postulate that average arrival flight delay and average departure flight delay at the origin airport affect negatively the revenue. The remaining ones suggest that distance, number of air passengers, number of air passengers and peak season would have a positive impact on revenue respectively.

In our study, data were collected from a Turkish airline company. From five cities and six airports, we selected 14 major domestic return routes. These routes were selected based on market structure and passenger demands. The data set included 1,008 observations from 14 routes for monthly periods between 2008 and 2013. Thus, we a have a panel dataset of 14 cross-sections ($N$ = 14) and 72 (T = 72) time points.

It is worth noting that in the study by Britto *et al.* (2012), 57 routes in the USA operated by 14 carriers were used. In the study of Zou and Hansen (2014), domestic flights in the US which were operated by Low Cost Carriers (LCCs) were taken into account. In the study by Forbes (2008), 18 routes from LaGuardia Airport in New York were included. Scotti and Dresner (2015) used domestic routes which were operated by Southwest Airlines. However, in our study, we chose major domestic flights in Turkey operated by a Turkish airline company.

## Findings

### *Descriptive statistics and pairwise correlations*

Before proceeding to the estimation of the model, descriptive statistics for the variables were obtained and are given in Table I.

Pairwise correlations between variables are also illustrated in Table II. This table indicates that the highest correlation is −0.34. Moreover, VIF values are between 1.04 and 1.30, of which tolerance varies from 0.77 to 0.96. The values of VIF are very low and tolerance values are close to unity. These values apparently show that there is no serious multicollinearity problem among the explanatory variables[1].

### *Unit root tests*

Since the seminal paper by Granger and Newbold (1974), it is very common in an empirical study to test the stationarity of the variables. They mentioned that if all variables or their linear combination are not stationary the spurious regression problem occurs. The problem is also mentioned by Baltagi (2013) for panel data. In this study, we employ the CIPS panel unit root test (Pesaran, 2007) to find out if the variables are stationary or have a unit root. Nonetheless, before starting the panel unit root tests, we also checked for the absence of cross-sectional correlations by employing the bias-adjusted LM test (details are given in the Appendix) introduced by Pesaran *et al.* (2008). The results of the tests are presented in Table III. Accordingly, we can reject strongly the absence of cross-section dependence. This justifies the use of the CIPS panel unit root test which allows for cross-sectional dependence. As for the panel unit root tests, as seen in Table III, the CIPS statistics reject the presence of

| Variables | Mean | SD | Maximum | Minimum |
|---|---|---|---|---|
| Revenue (TRY) | 104.36 | 22.43 | 184.69 | 60.62 |
| Arrdelay (minutes) | 7.37 | 7.06 | 39.52 | 1.45 |
| Depdelay (minutes) | 7.56 | 5.51 | 39.76 | 0.00 |
| Routepax (passengers) | 3058.79 | 1285.33 | 4469 | 86 |
| Distance (kilometers) | 431.94 | 129.76 | 716.52 | 325.00 |
| Season | 0.58 | 0.49 | 1 | 0 |

**Table I.**
Descriptive statistics ($N = 1008$)

| Variables | arrdelay | depdelay | distance | routepax | season |
|---|---|---|---|---|---|
| arrdelay | 1.0000 | | | | |
| depdelay | −0.2581 | 1.0000 | | | |
| distance | −0.0983 | 0.0750 | 1.0000 | | |
| routepax | −0.3444 | −0.0322 | 0.1793 | 1.0000 | |
| season | −0.0461 | −0.1297 | −0.0002 | 0.3022 | 1.0000 |

**Table II.**
Pairwise correlations between variables

a unit root and lend support to the hypothesis of stationarity in the variables. This allows us to use the level rather than the differences between the variables in the estimations.

The results of the bias-adjusted LM (LM_AD) test illustrate that there are significant cross-correlations among routes; hence, we are able to use the panel unit root test. The CIPS test results show that there is no unit root in the variables in this model.

*Slope homogeneity tests*

In panel data models, one of the important assumptions concerns the homogeneity of slope coefficients across cross-section units. Most studies assume homogeneity of the slope, yet the effects of an explanatory variable could substantially vary from one cross-section to another. Thus, estimations based on the homogeneity of coefficients could lead to misleading results. In this study, we use the slope homogeneity tests advanced by Pesaran *et al.* (2008). Table IV presents the slope homogeneity test results. All test results apart from $\hat{\Delta}_{adj}$ rejecting the null hypothesis of the slope homogeneity test for this model. As a result, we estimate our model allowing for slope heterogeneity. More precisely, we permit all coefficients to vary across routes. Then we obtain the mean group estimator for each of the coefficients as explained in the next section.

*Estimation method*

In this study, we employ two recently developed panel data econometric techniques. The first is the CCEMG developed by Pesaran (2006) and Kapetanios *et al.* (2011). This technique allows for slope heterogeneity and cross-section dependence, as in Equations 1 and 2. The cross-section dependence is taken into account by adding the cross-section averages of all

| Variables | CIPS intercept | CIPS intercept + trend | LM_AD intercept | LM_AD intercept + trend |
|---|---|---|---|---|
| Log (arrdelay) | −2.3956** | −2.823** | 436.502*** | 423.508*** |
| Log (depdelay) | −4.4399*** | −4.8073*** | 455.200*** | 448.453*** |
| Log (revenue) | −3.3576*** | −4.0514*** | 402.814*** | 394.250*** |
| Log (routepax) | −4.219*** | −4.5603*** | 407.814*** | 397.330*** |

**Notes:** Null hypothesis for CIPS is nonstationarity. The CIPS test Critical Values are taken from Pesaran (2007); *shows that statistics are significant at the 1% level of significance; **shows that statistics are significant at the 5% level of significance; ***shows that statistics are significant at the 10% level of significance

**Table III.**
Panel unit root test results

| Test | *F* |
|---|---|
| $\hat{S}$ | 278.38*** |
| $\tilde{\Delta}$ | 87427393.30*** |
| $\tilde{\Delta}_{adj}$ | 826099.30*** |
| $\hat{\Delta}$ | 92651463.07*** |
| $\hat{\Delta}_{adj}$ | 0.258 |

**Notes:** ***Shows that statistics are significant at the 1% level of significance; **shows that statistics are significant at the 5% level of significance; *shows that statistics are significant at the 10% level of significance

**Table IV.**
Slope homogeneity test results

dependent and explanatory variables into model 1. Then Model 1, augmented with the cross-section averages, is estimated for each of the cross-section units by OLS, and the CCEMG estimate of each coefficient is obtained by means of their arithmetic mean over all cross-sections.

The second technique is the AMG developed by Eberhardt and Teal (2010). Similar to CCEMG, the AMG technique allows for cross-section dependence and slope heterogeneity. However, it differs from the CCEMG in the way that cross-section dependence is taken into account. In the first step, time dummies are added to the original model and the first differences of the variables are taken to eliminate individual effects. In the second step, a new variable is constructed from the coefficients of time dummies obtained from the first step and included in the original model. In the second step, model 1 (including this new variable) is estimated for each of the cross-section units. As in CCEMG, the AMG estimator is obtained by the arithmetic mean of the coefficients across cross-sections.

Finally, we compare the results obtained by panel CCEMG and AMG with the standard two-way fixed effects (FE) regression. It is worth noting that in two-way FE a "weak form" of cross-section dependence is taken into account using time dummies. That is, in the two-way FE method it is assumed that the shocks affect all cross-sections in the same way over time, which contrasts with the assumption that the effects of a shock differ across cross-sections and over time.

*Estimation results*. The estimation results obtained from the panel AMG, CCEMG and two-way FE estimators are illustrated in Table V. As shown, the coefficient of arrival delay is negative and significant in all regressions. In the AMG and CCEMG regressions, the coefficients of arrival delay are very close to each other. The coefficient of departure delay is also negative and significant, and its magnitude does not vary much across regressions. The coefficient associated with route per passenger is positive and highly significant in all regressions. However, its size is very different in the FE regression, though it is close to 0.08 in AMG and CCEMG. The coefficient of seasonality is almost the same in all regressions and is positive. The coefficient of distance is negative in the three regressions, though it is not insignificant in AMG and CCEMG.

The estimation results show, first, that Log (arrdelay), Log (depdelay) and Log (distance) have significant negative effects on revenue as expected, whereas Log (routepax) and season have significant positive effects on it as expected.

| Variables | AMG | CCEMG | FE |
|---|---|---|---|
| Logarrdelay | −0.200*** [0.074] | −0.122*** [0.038 ] | −0.059*** [0.012 ] |
| Logdepdelay | −0.017*** [0.006] | −0.022*** [0.006] | −0.022*** [0.004] |
| Logroutepax | 0.083*** [0.019] | 0.079*** [0.023] | 0.012** [0.005 ] |
| season | 0.128*** [0.016] | 0.130*** [0.018 ] | 0.138*** [0.007 ] |
| Logdistance | −8.368* [5.37] | −6.732 [5.40] | −1.788 [1.77 ] |
| Constant | 57.169* [32.613] | 68.71 [70.493 ] | 15.413*** [10.685 ] |
| $R^2$ | 0.493 | 0.52 | 0.33 |
| $F$ | $F(10,998) = 196.94$*** | F(5,1003) = 193.97*** | $F(76,918) = 59.06$*** |

**Notes:** The figures in brackets are standard errors; ***; **; *show that statistics are significant at the 1, 5 and 10% level of significance respectively. AMG is the AMG estimator, CCEMG is the Common Correlated Effects, Mean Group Estimator and FE represents Fixed effects estimator. The coefficients of time dummy variables for fixed effects estimators are not reported to save space

**Table V.**
Estimation results

## Discussion

The results indicate that the revenue of a flight decreases due to arrival flight delay, departure flight delay at the origin airport and the distance between two different airports for any route and time period, whereas high numbers of passengers for any route and period and peak season in a year both increase it. These results are in line with previous literature which found that flight delays negatively impact ticket prices in the airline industry. Consistent with our findings, Britto *et al.* (2012), Forbes (2008), Hofer *et al.* (2008), and Zou and Hansen (2012) have found that delays are significantly negative estimators of airfares. The longer flight delays cause a decrease in airline prices. This finding implies that passengers are sensitive to flight delays. Therefore, managers should investigate the root causes of flight delays and determine all controllable risk factors. Then, they should make a risk assessment for each factor by taking probability and severity into consideration. Evaluating risk assessment results, managers should strategically plan flight routes in case of flight delays. Second, log (distance) has a significant negative effect on revenue. There may be several reasons for this, such as demand and type of flight (domestic or international). All of the flights included in the study are domestic flights. Although some distances between two destinations are closer than others, revenue was affected negatively due to the high demand for the closer destination. Moreover, lroutepax has a positive impact on revenue, showing that if an airline company carries high numbers of passengers for any route and time, it gains more revenue for that flight. Thus, managers should develop their service quality and airfare policy to appeal to customers. In addition, to increase revenue, low demand routes should be optimized by resource planning. Finally, as expected, seasonality was found to be a significant determinant in predicting revenue, suggesting that tickets are priced higher in the peak season. To increase the number of passengers in peak-off season, managers should attract potential customers by utilizing promotion policy.

## Conclusion and further research

This study examines the influence of several factors on revenue in the airline industry. The data set included 1,008 observations from 14 routes for monthly periods between 2008 and 2013. The AMG, CCEMG and two-way FE estimators have been used to analyze the relationships defined in the proposed research model. Based on the results, it was found that flight delays and distance affect revenue negatively, while the number of passengers and seasonality have a positive influence on revenue. The estimators used in this paper allow us to address the impact of each variable on revenue for the cross section. Accordingly, the impact of most of the variables on revenue does not vary across routes.

The main contribution of this study is to employ the newly introduced CCEMG and AMG panel data techniques. The other contribution is to collect data for all the monthly periods between the given years as distinct from previous studies that used only quarterly data. The reason for this choice was to categorize time periods on the basis of the selected airline's strategy, in terms of which seasons are divided into two groups – peak and off-peak. Although there are contributions here that provide a better understanding of the variables affecting revenue, this study has several limitations. However, further research ideas can be suggested to eliminate the barriers of each limitation. First, while the different estimators explained a significant portion of the variance of the dependent variable, a large percentage of the dependent variable remains unexplained. Thus, some additional factors such as market concentration, income, population, and vacation route introduced in previous sections may be included in a model for further studies. Second, this study used data retrieved from a Turkish airline company. The data from other airline companies operating in Turkey may be included for future research. Third, it is also important for group differences between FSNCs and LCCs to be analyzed in future studies. Finally, in this study, major domestic routes were taken into account, and as a result, each route was also operated by another airline company in Turkey. Therefore, routes that are operated by only one company might be chosen for a future study to examine the effects of delays on airfares in a monopolistic market.

## Note

1. Variance inflation factors (VIF) table is available from authors upon request.

## References


Baltagi, B.H. (2013), *Econometric Analysis of Panel Data*, edited by 5th Ed., John Wiley and Sons, Hoboken, NJ.

Blackwood, P. (2012), "Understanding flight delays at US Airports in 2010, using Chicago O'Hare international airport as a case study."

Britto, R., Dresner, M. and Voltes, A. (2012), "The impact of flight delays on passenger demand and societal welfare", *Transportation Research Part E: Logistics and Transportation Review*, Elsevier Ltd, Vol. 48 No. 2, pp. 460-469.

Chiou, Y.C. and Liu, C.H. (2016), "Advance purchase behaviors of air tickets", *Journal of Air Transport Management*, Elsevier Ltd, Vol. 57, pp. 62-69.

Cook, A., Tanner, G. and Anderson, S. (2004), *Evaluating the True Cost to Airlines of One Minute of Airborne or Ground Delay; Final Report*, Review Literature And Arts of The Americas.

Eberhardt, M. and Teal, F. (2010), "Productivity analsyis in global manufacturing production", No. 515.

Evans, A. and Schäfer, A. (2013), "The rebound effect in the aviation sector", *Energy Economics*, Vol. 36, pp. 158-165.

Forbes, S.J. (2008), "The effect of air traffic delays on airline prices", *International Journal of Industrial Organization*, Vol. 26 No. 5, pp. 1218-1232.

Granger, C.W.J. and Newbold, P. (1974), "Spurious regression in econometrics", *Journal of Econometrics*, Vol. 2 No. 2, pp. 111-120.

Hofer, C., Windle, R.J. and Dresner, M.E. (2008), "Price premiums and low cost carrier competition", *Transportation Research Part E: Logistics and Transportation Review*, Vol. 44 No. 5, pp. 864-882.

Hui, Y. and Dao-ming, Z. (2009), "Airline's optimal seating control strategy under the callable mechanism", *Systems Engineering - Theory and Practice, Systems Engineering Society of China*, Vol. 29 No. 6, pp. 32-38.

International Air Transport Association (IATA) (2017), "IATA annual review 2017", *Annual Report*.

Kapetanios, G., Pesaran, M.H. and Yamagata, T. (2011), "Panels with non-stationary multifactor error structures", *Journal of Econometrics*, Elsevier B.V., Vol. 160 No. 2, pp. 326-348.

Li, Y., Wang, Y. and zhang Cui, Q. (2015), "Evaluating airline efficiency: an application of virtual frontier network SBM", *Transportation Research Part E: Logistics and Transportation Review*, Vol. 81, pp. 1-17.

O'Connell, P.G.J. (1998), "The overvaluation of purchasing power parity", *Journal of International Economics*, Vol. 44 No. 1, pp. 1-19.

Ozbek, T. (2015), *THY Gecikmelerinin Perde Arkasında DHMİ Var*, Dogan News Agency, Istanbul.

Pesaran, M.H. (2006), "Estimation and inference in large heterogeneous panels with cross section dependence", *Econometrica*, Vol. 74 No. 4, pp. 967-1012.

Pesaran, M.H. (2007), "A simple panel unit root test in the presence of cross-section dependence", *Journal of Applied Econometrics*, Vol. 22 No. 2, pp. 265-312.

Pesaran, M.H., Ullah, A. and Yamagata, T. (2008), "A bias-adjusted LM test of error cross-section independence", *The Econometrics Journal*, Vol. 11 No. 1, pp. 105-127.

Scotti, D. and Dresner, M. (2015), "The impact of baggage fees on passenger demand on US air routes", *Transport Policy, Elsevier*, Vol. 43, pp. 4-10.

SHGM (2017), "Sivil havacilik faaliyet raporu, 2016".

Xu, X. and Cui, Q. (2017), "Evaluating airline energy efficiency: an integrated approach with network epsilon-based measure and network slacks-based measure", *Energy*, Elsevier Ltd, Vol. 122, pp. 274-286.

Yablonsky, G., Steckel, R., Constales, D., Farnan, J., Lercel, D. and Patankar, M. (2014), "Flight delay performance at Hartsfield-Jackson Atlanta international airport", *Journal of Airline and Airport Management*, Vol. 4 No. 1, pp. 78-95.

Zhang, M., Shan, L., Liu, K., Yu, H. and Yu, J. (2016), "Kybernetes terminal airspace sector capacity estimation method based on the ATC dynamical model for authors terminal airspace sector capacity estimation method based on the ATC dynamical model", *Kybernetes*, Vol. 45 No. 6, pp. 884-899.

Zou, B. and Hansen, M. (2012), "Flight delays, capacity investment and social welfare under air transport supply-demand equilibrium", *Transportation Research Part A: Policy and Practice*, Elsevier Ltd, Vol. 46 No. 6, pp. 965-980.

Zou, B. and Hansen, M. (2014), "Flight delay impact on airfare and flight frequency: a comprehensive assessment", *Transportation Research Part E: Logistics and Transportation Review*, Elsevier Ltd, Vol. 69, pp. 54-74.

## Appendix

The bias-adjusted LM (Pesaran *et al.*, 2008) test statistic is:

$$LM_{adj} = \sqrt{\frac{2}{N(N-1)}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{(T-k)\hat{\rho}_{ij}^{2}-\mu_{Tij}}{\upsilon_{Tij}} \rightarrow_{d} N(0,1)$$
$$i,j = 1\ldots..N$$

where $\hat{\rho}_{ij}$ shows the sample estimate of the pair-wise correlation for the residuals, k is the number of parameters of the model under consideration, T is the number of periods and N is the number of cross sections in panel data. The terms $\mu_{Tij}$ and $\upsilon_{Tij}$ represent the mean and the variance of the series respectively.

The $LM_{adj}$ test statistic in the above equation is used to test for cross sectional dependence in panel data models. The null hypothesis is the absence of cross section dependence. When the test statistic is greater than the critical value the null is rejected and then we can conclude that cross-section dependence is present and should be taken into account.

The CIPS statistic (Pesaran, 2007) is defined as:

$$CIPS = \frac{\sum_{I=1}^{N} CADF_i}{N}$$

where CADFi presents the simple averages of the individual cross-sectionally augmented ADF test statistics.

CIPS statistic is used to test for unit root in panel data. The critical value of the CIPS statistic is tabulated in Pesaran (2007). The null hypothesis is unit root and the alternative is stationarity.

**Corresponding author**
Fethi Calisir can be contacted at: calisirfet@itu.edu.tr