Modeling US Air Passenger Traffic Demand: Dynamic Data

David H. Hopfe
Embry-Riddle Aeronautical University, hopfed@erau.edu

Kiljae Lee
Embry-Riddle Aeronautical University, leek27@erau.edu

Chunyan Yu
Embry-Riddle Aeronautical University, yuc@erau.edu

Modeling US Air Passenger Traffic Demand: Dynamic Data

Completed Research

David H. Hopfe
Embry-Riddle Aeronautical University
hopfed@erau.edu

Kiljae Lee
Embry-Riddle Aeronautical University
Kiljae.lee@erau.edu

Chunyan Yu
Embry-Riddle Aeronautical University
yuc@erau.edu

Abstract

Conventional demand models (e.g., gravity model) in air transport literature tend to rely heavily on the mainstream econometric variables (e.g., distance, population, and GDP), which cannot be dynamically measured or used for short-term predictions. This study seeks to complement the short-term predictability of such conventional approaches by introducing dynamic predictors while alleviating the endogeneity by implementing panel data modeling analysis. Utilizing 40,072 air passenger data stacked in 3,344 city pairs over twelve months in 2020, we demonstrate that a large variability in demand can be explained by a handful of non-conventional variables such as internet search volume and geometric mobility indicators. The performance of our fixed effect model was dramatically improved by adding the regional intensity of google search for "airport" and "flight" and by adding the measure of people's time spent at residential areas in the origin and destination state (Adj. R^2 to .74).

Keywords

Google Search, Dynamic Data, Geometric Mobility, Demand Modeling.

Introduction

There is much uncertainty among the experts in the aviation industry over how long it will be before the air transportation sector recovers and passenger airlines are once again profitable. Suau-Sanchez et al. (2020) conducted a series of in-depth interviews with 16 senior aviation industry executives to gain an industry perspective on the impact of COVID-19 on commercial aviation. These interviews revealed perceived serious long-term consequences including airline industry consolidation, adverse effects from government aid, and the slow recovery of business travel due to teleworking and the digital transformation of the business world (Suau-Sanchez et al. 2020). Aviation experts believe the recovery of the U.S. passenger air transportation industry will take between three and six years (Hotle & Mumbower 2021; Sobieralski 2020). In the meantime, aviation industry stakeholders need to reconsider existing methods of air travel demand forecasting, as these models are slow to respond, and seek a modeling approach that includes dynamic data.

The process of air travel demand estimation involves model identification, parameter estimation and prediction using a specified model. Accurate air travel demand estimation is important for airlines and airports to predict future levels of passengers traveling via air transportation. A problem with traditional air travel demand forecasting methods is that researchers generally utilize solely historical data to predict future air travel demand. Current models for air travel demand do not incorporate data that is readily available in real-time thus making it difficult to make estimations in the presence of industry wide "shocks." Using data from daily updated repositories, mobile device location reporting and search engine interest results, this paper seeks to fill the gap in the literature that exists and bridge the gap between big data and traditional air travel demand estimation.

Theoretical Foundations

Air travel demand forecasting models have been considered extensively in the literature over the years. Early studies have implemented the gravity model to investigate different factors that have an effect on the demand for air transportation (Doganis 1966; Brown & Watkins 1968; Verleger 1972; Moore and Soliman 1981; Fotheringham 1983b; Rengaraju & Arasan 1992; Russon and Riley 1993; O'Kelly et al. 1995; Jorge-Calderón 1997; Shen 2004; Grosche et al. 2007). Since the work of Fotheringham (1981; 1983a) on spatial interaction models, spatial structures and distance-decay parameters, numerous empirical studies set out to model that of air travel demand (Abrahams 1983; Suryani et al. 2010; Carson et al. 2011; Li and Wan 2019; Suh & Ryerson 2019; and Birolini et al. 2020) employing a wide variety of methodologies to develop different forecasting models. Air travel demand modeling methodologies have included multiple regression (Ba-Fail et al 2000; Valdes 2015; and Wilken et al. 2016), stepwise regression (Abed et al. 2001), semi-logarithmic regression (Bhadra 2003; Sivrikaya & Tunç 2013), and multivariate neural forecasting (Alekseev & Seixas 2009; Blinova 2007). Further related artificial neural network (ANN) studies improved forecasting accuracy (Chen et al. 2012; Srisaeng et al., 2015c). Since, empirically tested genetic algorithms for predicting aviation demand have been successfully implemented showing better results than ANNs (Sineglazov et al. 2013; Srisaeng et al. 2015a; Srisaeng et al. 2016). AI methods have been employed for predicting air travel demand by using adaptive neuro-fuzzy inference systems (ANFIS) for even higher predictive power (Srisaeng et al. 2015b).

The vast majority of these empirical models recognize the aforementioned principles of push and pull factors involving repulsion and attraction. The demand between the two destinations is directly proportional to their size (population) and economic benefits (GDP) while inversely proportional to the cost of travel and the geographic distance between them. However, a review of the air travel demand forecasting literature uncovers very few empirical studies that take into consideration such dramatic and significant impacts as a pandemic on the air transportation industry. At the time of this writing, there is currently no body of knowledge of air transportation demand during a pandemic. Studies have yet to include influential factors that are specific to the estimation of air passenger volume under precipitous economic disruptions.

Empirical Strategy

First, this study seeks to provide detailed instruction in the methods of dynamic data collection of new explanatory factors for the inclusion in an analytical model to help predict air transportation volumes in times of unanticipated economic disruption. Second, the study hopes to develop a better method of air transportation volume estimation during times of unprecedented economic disruption. This is done through first starting with a traditional gravity model and then using a stepwise regression approach to show how dynamic data strengthens the predictive power of the air passenger traffic demand model.

Variables and Data

The dependent variable used in this study is the natural logarithm of number of passengers traveling from $city_i$ to $city_j$. Utilizing a data sample of passenger volumes from 3,344 city pairs over a twelve-month period in 2020, we examine the feasibility of parameterizing the influence of haphazard events such as government-imposed lockdowns, supply chain disruptions, and unpredictable social turmoil. Although multiple independent variables such as COVID cases, human mobility device location tracking, and volume of internet searches for "airline" were tested, only eleven important determinants for air travel demand were identified. The following sections define the variables used in the empirical analysis as well as the descriptive statistics of the factors (see Table 1).

Variable	N	Mean	SD	SE	Min	Median	Max
lnpax	40072	7.9606	1.6544	0.0083	0	8.1681	11.861
lnpipj	40072	27.152	2.1315	0.0106	13.512	27.458	31.592
lngigj	40072	35.523	2.2687	0.0113	22.796	35.858	40.154
cov_casei	40072	20278	42065	209.99	0	3853	268758
cov_casej	40072	20318	42078	210.05	0	3874	268758
cov_dthi	40072	132.62	311.8	1.56	0	35	3906
cov_dthj	40072	132.87	311.82	1.56	0	35	3906
gcovidi	40072	2908.8	1595.6	7.9652	0	3215.4	7460
gcovidj	40072	2908	1595.5	7.9646	0	3176.3	7460
gairporti	40072	2743.9	1047.8	5.2306	960.5	2392	9200
gairportj	40072	2747.6	1049.9	5.241	960.5	2392	9200
gflighti	40072	2989.9	1168.3	5.8322	1168.5	2692.8	7925
gflightj	40072	2992.3	1169.4	5.8374	1168.5	2692.8	7925
resij	40072	91.946	89.917	0.4489	-1.04	69.751	568.03

Table 1

Passengers

In this study, the variable representing passengers (pax) is defined as the total volume of people traveling from the origin to the destination or the city-pair in a given month. Data were collected from the Bureau of Transportation Statistics' TranStats pages under passengers form All Carriers and All Airports. This is represented by PAXij and indicates the quantity of one-way trips from the origin airport's city i and concluding in city j, whether a layover was made or not. In the econometric model, total demand between airports is represented by the natural logarithm of the passenger data, it is the dependent variable and it is provided for all city-pairs for all twelve months (see Figure 1, Figure 2).

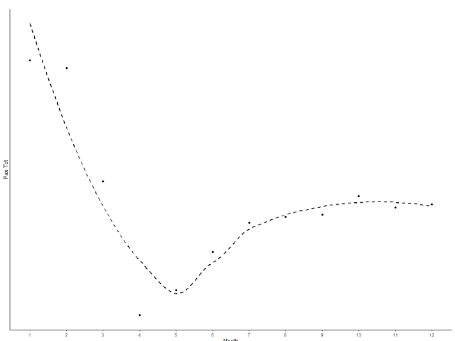

Figure 1

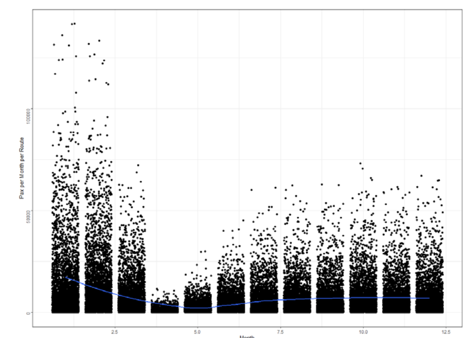

Figure 2

Population

US population data by county were obtained from the US Census Bureau. The 2010-2019 population estimates by county dataset was the most recent dataset at the time of data collection. The data were downloaded and used to estimate the monthly population numbers for 2020. By calculating the exponential growth rate for the previous year monthly populations for the 2020 year were estimated.

Distance

Airport distances were calculated using the haversine formula. Given the latitude and longitude the great circle distance between two points on a sphere can be determined using the law of haversines. Using the pracma R library this calculation can be completed in R with the haversine package [haversine()]. All coordinates for airports were pulled from the FAA's Airport Data and Information Portal (ADIP). To access, go to the adip.faa.gov website, click on advanced facility search and then click download all data. On the all-airport-data.xlsx spreadsheet, each airport's latitude and longitude is provided. Alternatively, one could navigate to the FAA Aviation Data Portal's National Flight Data Center (NFDC) on GitHub and follow the instructions to utilize the Application Programming Interface (API) call [airportData.facilities(options)].

Gross Domestic Product

The US Department of Commerce and the Bureau of Economic Analysis (BEA) in collaboration with the European Commissions' DG CONNECT and Eurostat provide a Transatlantic Open Data Partnership that focuses on economic data and open access to that data through an API. This partnership created the eu.us.opendata R library to enable open access to datasets from both the Eurostat API and BEA API. Using a Linked Open Data design, this R library makes it simple to obtain economic data from the BEA API. While US and European economic data can be assessed through the eu.us.opdendata API, we only need US data and thus chose to use the BEA API and bea.R library instead.

In order to access the BEA API through the bea.R library, we first registered for an API key through the apps.bea.gov website. Once registered we were able to install bea.R packages, open the bea.R library, and load the API key from the email we received. Real GDP in thousands of chained 2012 dollars data were obtained from the BEA and for brevity will be referred to as simply GDP.

The data files that were downloaded contained state and county level real GDP numbers and percent changes. Since monthly data is not available, it was created using both annual numbers and quarterly percent changes. Taking the 2019 ending GDP and calculating the Q1 percent change for 2020, the difference was divided evenly for each month of Q1. These same steps were used to determine Q2, Q3 and Q4.

COVID-19 Related Deaths

Data for COVID-19 related deaths were obtained from the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. This time series data provided the number of deaths per day for each county during 2020. The daily death toll was converted to a monthly report simply by subtracting the number of the deaths on the first day of the month from the last day of the month for each month and for each county. The data was collected from the CSSEGISandData repository on GitHub.

COVID-19 Related Cases

Data for COVID-19 confirmed cases were obtained from the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. This time series data provided the number of cases per day for each county during 2020. The confirmed COVID-19 cases were converted to a monthly report simply by subtracting the number of the cases on the first day of the month from the last day of the month for each month and for each county. The data was collected from the CSSEGISandData repository on GitHub.

Community Mobility Reports

These time series datasets reveal how outings to locations, such as grocery stores and parks are changing in each region. The datasets indicate how visits and length of stay at different locations change when compared to the baseline. Baseline days represent a normal value for that day of the week and is the median value from January 3, 2020 to February 6, 2020. Anonymized data comes from users who have opted-in to location history on their mobile device. Although the datasets contain county information, data has been omitted when there are too few datapoints to ensure anonymity, thus much of the county data is missing. Therefore, state mobility data were converted from daily time series to monthly time series by averaging all of the days of the month together for each state. The mobility reports are described in more detail and include the following place categories: Grocery & pharmacy, Parks, Transit stations, Retail & recreation, Residential, Workplaces.

Retail and recreation data shows how the number of visitors to places of retail and recreation has changed relative to the period before the pandemic. This includes places like restaurants, cafés, shopping centers, theme parks, museums, libraries, movie theaters. Grocery and pharmacy stores data shows how the number of visitors to grocery and pharmacy stores has changed relative to the period before the pandemic. This includes places like grocery markets, farmers markets, specialty food shops, drug stores, and pharmacies. Public transport stations data shows how the number of visitors to transit stations has changed relative to the period before the pandemic. This includes public transport hubs such as subway, bus, and train stations.

Parks and outdoor spaces data shows how the number of visitors to parks and outdoor spaces has changed relative to the period before the pandemic. This includes places like local parks, national parks, public beaches, marinas, dog parks, plazas, public gardens. Workplace visitors' data shows how the number of visitors to workplaces has changed relative to the period before the pandemic. Residential mobility trends provide information on time spent at home and the data shows how the number of visitors to residential areas has changed relative to the period before the pandemic.

Internet Search Engine Query Volume

There are several internet search engines on the web, but none have reached the popularity of Google. Top search queries can be analyzed for popularity across various regions and can be compared over time. Google Trends data provides access to actual search requests made to Google. This type of analysis provides anonymized, aggregated, and normalized data that can be either real-time or historic. Google's normalized data is shown on a scale of 0 to 100 based on the search volume.

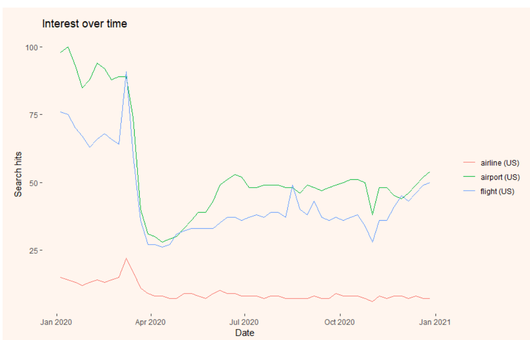

Figure 3

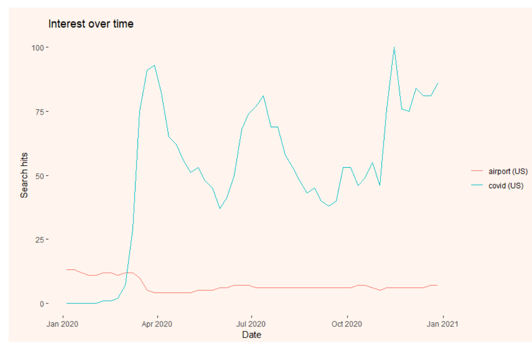

Figure 4

Normalized search engine trend query data has been successfully utilized by researchers to explore both travel demand and financial markets (Li Long, Guleria & Alam, 2021; Baig et al., 2013). Google trends provides an API that can be accessed through the gtrendsR package for R. Through this package, data on frequency of search terms in a time series format can be downloaded. This data can potentially have predictive power as the data represent the search interest for the general population.

In an effort to capture the interest intensity in air travel during 2020, three similar search queries were compared. "Airline" appeared to be the least intense search query of the three. Therefore, "airport" and "flight" were selected for the study (see Figure 3). Next, we looked to see how "airport" compared to the intensity of "covid." The data shows the intensity for the keyword "covid" was much greater (see Figure 4). Data on Google searches for "covid", "airport", and "flight" were obtained using the gtrendsR package. The normalized keyword search intensity data for each state were multiplied by the normalized keyword search intensity data for the overall nation.

Google searches for "covid" and "airport" were obtained using the gtrendsR package. The quantity of searches for "covid" is expected to have an inverse relationship with air transportation demand while searches for "airport" are expected to be an attraction factor.

We could see that google search on certain keyword is highly correlated with the passenger volume of each city-pair (see Figure 5).

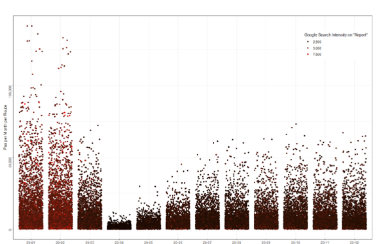

Figure 5

Correlation Matrix

The correlation matrix (see Figure 6) indicates that the conventional explanatory variables (Distance:lndist, Population:lnpipj, GDP:lngigj) are indeed strongly correlated with passenger volume (lnpax). However, it is notable that Google search (gcovidi, gcovidj, gairprti, gairprtj, gflighti, gflightj) and mobility (resi, resj) shows significant correlations with passenger volume.

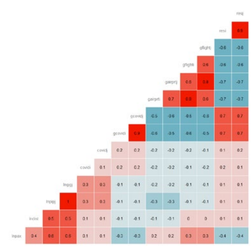

Figure 6

Sources of the Data

US population data by county were obtained from the US Census Bureau. Gross Domestic Product (GDP) data were obtained from the U.S. Bureau of Economic Analysis (BEA). Airport coordinates and information were collected from the FAA. Airport distances were calculated using the Haversine Formula. Data for COVID-19 related cases and deaths were obtained from the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. Community Mobility Reports were obtained from Google.

Treatment of Data

We downloaded all flights from the US DOT T-100 report for each month in 2020. From this list, we compiled all 9,401 of the individual origin destination airport pairs. We removed all service to international locations, leaving only domestic US flights. Next, we removed all locations without regularly scheduled service. We then removed all seaplane bases. We also removed all locations that did not offer scheduled service for all 12 months. The balanced panel data includes 3,344 individual city-pairs observed at monthly time periods for all months of 2020.

Research Approach

Given our theoretical interest in both the route-specific effects and the changes over time, we opted for a panel data analysis study. Analysis will be conducted using panel data regression analysis in the period of 2020. An appropriate panel regression model will be chosen with the assistance of the Hausman test and the redundant fixed-effect or log-likelihood test. In an effort to determine which factors of determining air travel demand is suitable to data, the regression coefficients will be calculated and inspected.

To interpret our results as elasticities, all the numerical variables, except the dummies and normalized data, will be transformed into natural logarithms. For the static model, the equation is specified as follows:

$$Pax_{k,t} = \beta' x_{k,t} + \alpha_k + \varepsilon_{k,t} \quad (1)$$

where t = {January 2020...December 2020} and k = {all origin destination city-pairs}.

$Pax_{k,t}$ is the log of the k-th city-pair's total passengers carried at time t; α_k is the city-pair effect invariant to time; and $\varepsilon_{k,t}$ is the error term. $x_{k,t}$ is the vector of explanatory variables.

$$x_{k,t} = \{\text{lnpipj, lngigj, covidi, covidj, groci, grocj, retaili, retailj, parki, parkj, transi, transj, worki, workj, resi, resj, gcovidi, gcovidj, gairprti, gairprtj, gflighti, gflightj}\} \quad (2)$$

where lnpipj is the log of the product of *k*-th city-pair's population at time *t*, lngigj is the log of the product of *k*-th city-pair's GDP at time *t*, covidi is the number of COVID-19 deaths in the origin city at time *t*, covidj is the number of COVID-19 deaths in the destination city at time *t*. Whereas groci, retaili, parki, transi, worki, resi, are the normalized mobility data for number of people visiting grocery and pharmacy stores, places of retail and recreation, parks and outdoor spaces, transit stations, workplaces and number of people staying home, respectively, in the origin city at time *t*, while grocj, retailj, parkj, transj, workj, resj, are the normalized mobility data for number of people visiting grocery and pharmacy stores, places of retail and recreation, parks and outdoor spaces, transit stations, workplaces and number of people staying home, respectively, in the destination city at time *t*. gcovidi is the normalized internet search data for the number of people searching for the word "covid" in the origin city at time *t*, gcovidj is the normalized internet search data for the number of people searching for the word "covid" in the destination city at time *t*, gairprti is the normalized internet search data for the number of people searching for the word "airport" in the origin city at time *t*, gairprtj is the normalized internet search data for the number of people searching for the word "airport" in the destination city at time *t*, gflighti is the normalized internet search data for the number of people searching for the word "flight" in the origin city at time *t*, gflightj is the normalized internet search data for the number of people searching for the word "flight" in the destination city at time *t*, lndist was omitted because the distance between the cities is a time invariant variable. Although each of these mobility reports were tested, only the residential mobility trends proved to be statistically significant.

The next step was to remove the variables that were not significant. Then we run a pooled OLS ignoring the time and panel information. The results of the pooled model (model 1) suggest that while the three conventional key variables for the gravity model (lndist, lnpipj, and lngigj) are still highly significant, during the inception of the pandemic, they can only explain 36 percent of variances in passenger (lnpax) (see Table 2). This is a significant drawback of the model considering that in normal times they usually explain more than 70% of the variance in passenger traffic (Grosche et al., 2007).

Introducing our dynamic variables in the pooled model shows a meaningful increase in performance (adj R^2 =.60). However, the pooled model could not take advantage of panel data, and it suffers from endogeneity. Further, Breusch–Pagan test results were significant (BP = 2762.8, df = 21, $p < 0.01$), suggesting that the homoscedasticity assumption does not hold, and thus, the pooled OLS cannot be used for modeling the data. Hausman test results (χ^2=1107.8, df=18, p<0.01) rejected the null hypothesis (i.e., random effect model is consistent) suggesting that fixed effect model (vs. random effect model) is the best choice for our data.

	Dependent variable:	
	Log(Passengers)	
	(1)	(2)
lndist	0.200*** (0.009)	0.130*** (0.007)
lnpipj	-0.200*** (0.013)	-0.074*** (0.011)
lngigj	0.570*** (0.012)	0.560*** (0.010)
cov_casei		0.00000*** (0.00000)
cov_casej		0.00000*** (0.00000)
cov_dthi		-0.0002*** (0.00002)
cov_dthj		-0.0002*** (0.00002)
gcovidi		0.0001*** (0.00001)
gcovidj		0.0001*** (0.00001)
gairprti		0.0002*** (0.00001)
gairprtj		0.0002*** (0.00001)
gflighti		0.0001*** (0.00001)
gflightj		0.0001*** (0.00001)
resij		-0.006*** (0.0001)
Constant	-8.100*** (0.120)	-12.000*** (0.120)
Observations	40,072	40,072
R^2	0.360	0.600
Adjusted R^2	0.360	0.600
F Statistic	7,401.000*** (df = 3; 40068)	4,322.000*** (df = 14; 40057)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 2

Estimation Method and Results

Model Description and Variables

As the fixed effect (FE) estimation eliminates all the time-invariant effects through de-meaning, it automatically drops constant variables (e.g., distance) from the model and highlights the effect of within route variations. We ran the analysis first with the mainstream key variables from the gravity model, distance, Population, and GDP (model 1). Next, we brought in the data from the COVID cases and COVID deaths. While significant, COVID cases and death did not substantially improve the model (model 2). It appears that the effect of these two variables may exhibit a "delayed" effect on the number of passengers. This indicates a need for a future study that explicitly factor in the time-differencing technique. Next, we include the internet keyword search intensity for "covid", "airport", and "flight." This results in a substantial improvement in model performance(model3). Finally, effects of the geographic mobility reporting from electronic devices have been included in the model (model 4); Adding the intensity of people's time spent at residential areas in the origin and destination state (i.e., resi * resj) results in a substantial improvement in model performance.

Regression Results

Table 3 presents the results from the fixed-effect model.

	Dependent variable:				
	Log(Passengers)				
	(1)	(2)	(3)	(4)	(5)
lnpipj	6.600*** (0.630)	5.800*** (0.640)	36.000*** (0.520)	9.900*** (0.450)	9.900*** (0.440)
lngigj	6.400*** (0.070)	5.000*** (0.068)	-0.580*** (0.067)	2.400*** (0.056)	2.400*** (0.056)
cov_casei		0.00001*** (0.00000)	0.00000*** (0.00000)	-0.00000 (0.00000)	
cov_casej		0.00001*** (0.00000)	0.00000*** (0.00000)	0.00000 (0.00000)	
cov_dthi		-0.001*** (0.00002)	-0.001*** (0.00001)	0.0001*** (0.00001)	0.0001*** (0.00001)
cov_dthj		-0.001*** (0.00002)	-0.0005*** (0.00001)	0.0001*** (0.00001)	0.0001*** (0.00001)
gcovidi			0.0001*** (0.00001)	0.0001*** (0.00000)	0.0001*** (0.00000)
gcovidj			0.00002*** (0.00001)	0.0001*** (0.00000)	0.0001*** (0.00000)
gairprti			0.001*** (0.00001)	0.0001*** (0.00001)	0.0001*** (0.00001)
gairprtj				0.0001*** (0.00001)	0.0001*** (0.00001)
gflighti			-0.00000 (0.00001)	0.0001*** (0.00001)	0.0001*** (0.00001)
gflightj			0.0003*** (0.00001)	0.0001*** (0.00001)	0.0001*** (0.00001)
resij				-0.007*** (0.0001)	-0.007*** (0.00005)
Observations	40,072	40,072	40,072	40,072	40,072
R^2	0.180	0.300	0.610	0.760	0.760
Adjusted R^2	0.110	0.230	0.580	0.740	0.740
F Statistic	4,118.000*** (df = 2; 36726)	2,577.000*** (df = 6; 36722)	5,309.000*** (df = 11; 36717)	8,938.000*** (df = 13; 36715)	10,564.000*** (df = 11; 36717)
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 3

Conclusions

As seen in the results the traditional gravity model variables perform very poorly in response to sudden catastrophic changes in socioeconomic fabric of our society. In this model the variables population and GDP are statistically significant at the 0.01 significance level. However, the coefficient of determination indicates that only 18% of the variation in the passenger numbers can be explained by the model.

Daily supply chain disruptions, weekly government-imposed restrictions, and an ongoing fear of leaving one's home are unaccounted for in these traditional models. It was this realization that led for the need of discovery into the underlying factors that contribute to the demand for air travel in the United States, more specifically, air transportation during times of major socioeconomic disruption.

In an effort to explain these effects several determinants were selected for investigation. The intensity of keyword searches for "flight" and "airport" as expected were positively correlated with the quantity of air passengers. Nevertheless, the searches for "covid" also had a positive effect on the dependent variable. Initially, it was hypothesized that as searches for "covid" increased it would represent the perception of fear and would be negatively corelated with the number of air passengers. The statistically significant results for the internet keyword search results indicate otherwise. This could potentially be explained by travelers searching for covid restrictions prior to traveling. Travelers need to know what to expect when walking through the airport, boarding the plane, and reaching their destination. Therefore, these results make intuitive sense and show that the model can be strengthened using dynamic data. These three keywords helped to increase the explanatory power of the model (adj. $R^2 = 0.61$).

Mobility reporting from personal electronic devices is another determinant that was investigated due to the almost real time results and dynamic data capabilities. Unfortunately, of the eight categories tested for this model, only the residence mobility variable was statistically significant ($p<0.01$). Time spent at home data shows how the number of visitors to residential areas has changed relative to the period before the pandemic. This explains the inverse relationship with the number of passengers using air transportation as indicated by the -0.007 coefficient for resij that is statistically significant at the 0.01 significance level. The addition of the mobility reporting increased the explanatory power of the model (adj.R^2 = 0.74).

During the ongoing pandemic of 2020, data of confirmed COVID-19 cases and deaths were being reported on a daily basis. The availability of this dynamic data and the socioeconomic implications of the information it contains made it convenient for further exploration. Surprisingly, the number of confirmed cases of COVID-19 were not statistically significant and were dropped from the model. The number of confirmed deaths as a result of COVID-19 were significant at the point of origin but not at the destination location. Thus, the variables were dropped from the model with no loss of explanatory power of the model.

REFERENCES

Abed, S. Y., Ba-Fail, A. O. and Jasimuddin, S. M. 2001. "An Econometric Analysis of International Air Travel Demand in Saudi Arabia," *Journal of Air Transport Management* (7:3), pp. 143-148 (doi: 10.1016/s0969-6997(00)00043-0).

Abrahams, M. 1983. "A Service Quality Model of Air Travel Demand: An Empirical Study," *Transportation Research Part A: General* (17:5), pp. 385-393. (doi: 10.1016/0191-2607(83)90007-9).

Alekseev, K. and Seixas, J. 2009. "A Multivariate Neural Forecasting Modeling for Air Transport – Preprocessed by Decomposition: A Brazilian Application," *Journal of Air Transport Management* (15:5), pp. 212-216. (doi: 10.1016/j.jairtraman.2008.08.008).

Ba-Fail, A. O., Abed, S. Y. and Jasimuddin, S. M. 2000. "The Determinants of Domestic Air Travel Demand in The Kingdom of Saudi Arabia," *Journal of Air Transportation World Wide* (5:2), pp. 72-86.

Baig, A., Butt, H. A., Haroon, O., & Rizvi, S. A. 2020. "Deaths, panic, lockdowns and US equity markets: The case of COVID-19 pandemic," *SSRN Electronic Journal* (doi:10.2139/ssrn.3584947).

Bhadra, D. 2003. "Demand for Air Travel in The United States: Bottom-Up Econometric Estimation and Implications for Forecasts by Origin-Destination Pairs," *Journal of Air Transportation* (8:2), pp. 19-56.

Birolini, S., Cattaneo, M., Malighetti, P. and Morlotti, C. 2020. "Integrated Origin-Based Demand Modeling for Air Transportation," *Transportation Research Part E: Logistics and Transportation Review* (142:102050). (doi: 10.1016/j.tre.2020.102050).

Blinova, T. O. 2007. "Analysis Of Possibility of Using Neural Network to Forecast Passenger Traffic Flows in Russia," *Aviation* (11:1), pp. 28-34. (doi: 10.3846/16487788.2007.9635952).

Brown, S. L. and Watkins, W. S. 1968. "The Demand for Air Travel: A Regression Study of Time-Series and Cross-Sectional Data in The U.S. Domestic market," *Highway Research Board* (213:1), pp. 21-34.

Carson, R. T., Cenesizoglu, T. and Parker, R. 2011. "Forecasting (Aggregate) Demand for Us Commercial Air Travel," *International Journal of Forecasting* (27:3), pp. 923-941. (doi: 10.1016/j.ijforecast.2010.02.010).

Chen, S., Kuo, S., Chang, K. and Wang, Y. 2012. "Improving The Forecasting Accuracy of Air Passenger and Air Cargo Demand: The Application of Back-Propagation Neural Networks," *Transportation Planning and Technology* (35:3), pp. 373-392. (doi: 10.1080/03081060.2012.673272).

Doganis, R. 1966. "Traffic Forecasting and The Gravity Model," *Flight International,*
(29:1), pp. 547-549.

Fotheringham, A. S. 1983b. "Some Theoretical Aspects of Destination Choice and Their Relevance to Production-Constrained Gravity Models," *Environment and Planning A: Economy and Space* (15:8), pp. 1121-1132. (doi: 10.1068/a151121).

Fotheringham, A. S. 1981. "Spatial Structure and Distance-Decay Parameters," *Annals of the Association of American Geographers (*71:3), pp. 425-436.

Fotheringham, A. S. 1983a. "A New Set of Spatial-Interaction Models: The Theory of Competing Destinations," *Environment and Planning A* (15:1), pp. 15-36. (doi: 10.1068/a150015).

Grosche, T., Rothlauf, F. and Heinzl, A. 2007. "Gravity Models for Airline Passenger Volume Estimation," *Journal of Air Transport Management* (13:4), pp. 175-183. (doi: 10.1016/j.jairtraman.2007.02.001).

Hotle, S. and Mumbower, S. 2021. "The Impact of Covid-19 On Domestic U.S. Air travel
Operations and Commercial Airport Service," *Transportation Research Interdisciplinary Perspectives* (9:100277). (doi: 10.1016/j.trip.2020.100277).

Jorge-Calderón, J. 1997. "A Demand Model for Scheduled Airline Services on International European Routes," *Journal of Air Transport Management* (3:1), pp. 23-35. (doi: 10.1016/s0969-6997(97)82789-5).

Li, T. and Wan, Y. 2019. "Estimating The Geographic Distribution of Originating Air Travel Demand Using a Bi-Level Optimization Model," *Transportation Research Part E: Logistics and Transportation Review* (131:1), pp. 267-291. (doi: 10.1016/j.tre.2019.09.018).

Li Long, C., Guleria, Y., & Alam, S. 2021. "Air passenger forecasting using neural Granger causal Google trend queries," *Journal of Air Transport Management* (95:102083). (doi: 10.1016/j.jairtraman.2021.102083).

Moore, O. E. and Soliman, A. H. 1981. "Airport Catchment Areas and Air Passenger Demand," *Transportation Engineering Journal of ASCE* (107:5), pp. 569-579. (doi: 10.1061/TPEJAN.0000949).

O'Kelly, M. E., Song, W. and Shen, G. 1995. "New Estimates of Gravitational Attraction by Linear Programming," *Geographical Analysis* (27:4), pp. 271-285. (doi: 10.1111/j.1538-4632.1995.tb00911.x).

Rengaraju, V. R. and Arasan, V. T. 1992. "Modeling For Air Travel Demand," *Journal of Transportation Engineering* (118:3), pp. 371-380. (doi: 10.1061/(ASCE)0733-947X(1992)118:3(371)).

Russon, M. G. and Riley, N. F. 1993. "Airport Substitution in A Short Haul Model of Air Transportation," *International Journal of Transport Economics* (20:2), pp. 157-174.

Shen, G. 2004. "Reverse-Fitting the Gravity Model to Inter-City Airline Passenger Flows by An Algebraic Simplification," *Journal of Transport Geography* (12:3), pp. 219-234. (doi: 10.1016/j.jtrangeo.2003.12.006).

Sineglazov, V., Chumachenko, E. and Gorbatyuk, V. 2013. "An Algorithm for Solving the Problem of Forecasting," *Aviation* (17:1), pp. 9-13. (doi: 10.3846/16487788.2013.777219).

Sivrikaya, O. and Tunç, E. 2013. "Demand Forecasting for Domestic Air Transportation in Turkey," *The Open Transportation Journal* (7:1). (doi: 10.2174/ 1874447820130508001).

Sobieralski, J. B. 2020. "COVID-19 And Airline Employment: Insights from Historical Uncertainty Shocks to The Industry," *Transportation Research Interdisciplinary Perspectives* (5:100123). (doi: 10.1016/j.trip.2020.100123).

Srisaeng, P., Baxter, G., Richardson, S. and Wild, G. 2015a. "A Forecasting Tool for Predicting Australia's Domestic Airline Passenger Demand Using a Genetic Algorithm," *Journal of Aerospace Technology and Management* (7:4), pp. 476-489.

Srisaeng, P., Baxter, G. S. and Wild, G. 2015b. "An Adaptive Neuro-Fuzzy Inference System for Forecasting Australia's Domestic Low Cost Carrier Passenger Demand," *Aviation* (19:3), pp. 150-163. (doi: 10.3846/16487788.2015.1104806).

Srisaeng, P., Baxter, G. S. and Wild, G. 2015c. "Forecasting Demand for Low Cost Carriers in Australia Using an Artificial Neural Network Approach," *Aviation* (19:2), pp. 90-103. (doi: 10.3846/16487788.2015.1054157).

Srisaeng, P., Richardson, S., Baxter, G. S. and Wild, G. 2016. "Forecasting Australia's Domestic Low Cost Carrier Passenger Demand Using a Genetic Algorithm Approach," *Aviation* (20:2), pp. 39-47. (doi: 10.3846/16487788.2016.1171798).

Suau-Sanchez, P., Voltes-Dorta, A. and Cugueró-Escofet, N. 2020. "An Early Assessment of The Impact of Covid-19 On Air Transport: Just Another Crisis or The End of Aviation as We Know It?" *Journal of Transport Geography* (86:102749). (doi: 10.1016/j.jtrangeo.2020.102749).

Suh, D. Y. and Ryerson, M. S. 2019. "Forecast To Grow: Aviation Demand Forecasting in An Era of Demand Uncertainty and Optimism Bias," *Transportation Research Part E: Logistics and Transportation Review* (128:1), pp. 400-416. (doi: 10.1016/j.tre.2019.06.016).

Suryani, E., Chou, S. and Chen, C. 2010. "Air Passenger Demand Forecasting and Passenger Terminal Capacity Expansion: A System Dynamics Framework," *Expert Systems with Applications* (37:3), pp. 2324-2339. (doi: 10.1016/j.eswa.2009.07.041).

Valdes, V. 2015. "Determinants Of Air Travel Demand in Middle Income Countries," *Journal of Air Transport Management* (42:1), pp. 75-84. (doi: 10.1016/j.jairtraman.2014.09.002j.jairtraman.2014.09.002).

Verleger, P. K. 1972. "Models Of the Demand for Air Transportation," *The Bell Journal of Economics and Management Science* (3:2), pp. 437-457. (doi: 10.2307/3003032).

Wilken, D., Berster, P. and Gelhausen, M. C. 2016. "Analysis Of Demand Structures on Intercontinental Routes to And from Europe with A View to Identifying Potential for New Low-Cost Services," *Journal of Air Transport Management* (56:1), pp. 79-90. (doi: 10.1016/j.jairtraman.2016.04.018).