INTERNATIONAL TRAVEL TO AND FROM THE UNITED STATES: DESTINATION CHOICES AND AIRLINE FARES

Priyanka Paithankar
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
priyanka.paithankar@utexas.edu

Fatemeh Fakhrmoosavi, Ph.D.
Postdoctoral Fellow
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
moosavi@austin.utexas.edu

Kara M. Kockelman, Ph.D., P.E.
(Corresponding Author)
Dewitt Greer Professor in Engineering
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
301 E. Dean Keeton St, Stop C1761, Austin, TX, 78712
kkockelm@mail.utexas.edu

Kenneth A. Perrine
Research Associate
Center for Transportation Research
The University of Texas at Austin
kperrine@utexas.edu

Word Count: ~ 4813 words + 8 tables (each 230 words) + 2 Figures = 6652 words

Submitted for presentation in the 102nd Annual Meeting of the TRB and for publication in TRR

ABSTRACT

Approximately one quarter of all U.S. air-passenger trips (involving US airlines only) are to and from foreign destinations, accounting for around 4.5 percent of total US person miles in 2019 (1). Few studies have investigated details of Americans' international travel, so this paper assesses travel demand, patterns, and costs (in time and money) between major US airports and foreign airports worldwide, as well as ground trips to Mexico and Canada using 2019 DB1B flight ticket data, the 2016/17 National Household Travel Survey (NHTS), and border crossing data. A model of trip distribution, from 334 US airports to 1028 foreign airports shows how trip flows fall about 41% with every 7-hour increase in flight start to end time . Destinations hosting tourist attractions (like London, Barcelona, Milan, Paris, and Dubai etc.) are also practically significant, increasing flows by 48%. Feasible generalized least-squares models quantify how flight fares (for one-way itineraries) rises by \$0.078 per mile for coach class and \$0.163 per mile for business class or higher. These fares are higher for English-speaking destinations as compared to other destinations (not English speaking), as well as for trips during April to June as compared to January to March with the similar distance, flight class, etc. Understanding international travel is important for local and global economics, the evolution of transportation technology and social networks, and the future of global climate and air quality. **Keywords:** International Travel, Demand Modeling, Air Fare, Destination Choice

MOTIVATION

According to the US Bureau of Transportation Statistics (BTS), person-miles travelled (PMT) in 2019 were 7.7 trillion, of which 4.5% were international person-miles involving US air carriers (2). International travel handled by US air carriers constitutes 15.5% of Americans' long-distance person-miles travelled (PMT) (where trips are classified as "long distance" here if they exceed 75-miles one-way) (1,3). BTS air passenger data suggest that the "average" American makes 0.37 international trips (inbound and outbound) per year by air, boat, road, and train. This implies an international departure by air every 5.43 years (4,5). The major purposes of these trips (84% of the total) were leisure or visiting friends and relatives (4). Air travel accounts for 60% of the international travel from US, land travel accounts for more than 39% (to Mexico and Canada) and travel through water (to Canada) is less than 1% (5,6). In addition to being a key source of passenger-miles traveled, air travel is a significant source of travel-based emissions and household expenditures. According to the US Travel Association, US international travel spending in 2019 was approximately $181 billion, which was reduced by 71.2% in 2021 owing to the COVID-19 pandemic (7). Although COVID-19 mutations may continue to impact international travel, increased vaccination rates, controlled infection cases, and loosened travel restrictions are projected to rebound the overseas journeys. The travel expenditure prediction indicates that spending will exceed 2019 levels in 2025 and raise by 9.4% in 2026 (7). Thus, to better prepare for future transportation requirements, it is necessary to estimate Americans' foreign travel demand and destination preferences.

Travel demand modeling studies and US travel surveys regularly miss international travel. Most studies focus on domestic trips, and very few include questions on long-distance trips (from the past month or year, rather than simply catching the few that happen on the survey day). However, international trips are a notable source of travel cost and emissions, with 8% of global greenhouse gas (GHG) emissions produced via tourism, and 40% of those emission coming via aviation (8). American tourists spent $113B on international travels in 2015 (9), which averages to several hundred dollars per year per capita. In an unusual but relatively recent example, Llorca et al. (8) developed a model for generation, distribution, and mode choice in person-trips over 40 km (25 miles, one-way) – but only from Ontario province, in Canada. They observed that land use attributes and trip purpose (or destination activities) are important for destination choice probabilities. For international trip generation, they estimated only the total number of trips (not destination or mode or timing), due to the lack of data. As noted earlier, international travel is a relatively rare event for most people.

Transport planning heavily relies on forecasts of travelers' trip decisions, including international travel. Tourism flows and international trade volumes do show up in the literature, especially for specific market pairs. For example, Qu and Lam (10) used ordinary least squares models to estimate travel demand for mainland Chinese tourists to Hong Kong. They identified income and visa requirements as key predictors. Keum (11) used a gravity model and Linder (economic) hypothesis to predict trade patterns and tourism flows across Korea. The study confirmed the robustness of the gravity model in estimating international flows. Wu et al. (12) explored tourism flows between Chinese regions and offered suggestions for tourism improvements. Most studies focus only on international flows between specific places (e.g., like US to Canada (13)) or all in-coming flows (14). International business trips are regularly overlooked.

Air is a major mode for trips over 500 miles, and international travel. Americans made 100 million international trips to other nations in 2019 (including one-way and round trips from US by all modes). BTS (2) reported that US airlines handled 115 million air passengers in the same year - including both Americans and non-American passengers. Airfare and duration are expected to be important indicators of international travel mode and destination choices. Flight price fluctuates depending on purchase time, number of stops, flight date, and seat class. Recent studies have used machine learning algorithms to predict flight fare using different datasets (e.g., 15,16). Ratnakanth (17) analyzed different methods presented in the literature for flight price prediction and indicated that random forest and gradient boosting techniques outperform other machine learning approaches for flight fare prediction. The study stated airline company, travel time, number of stops, and destination as effective factors in flight price. Flight fare prediction studies are mostly used for defining prices in the future for various airlines. In this study, flight fare and duration are inputs of the trip distribution model and its application.

All above-mentioned studies either focused on tourism attractions or price predictions for airlines. Furuichi and Koppelman (18) used a nested logit model to predict the departure and destination airports of air travelers. They used a survey of international air travelers from Japan and indicated that a joint departure airport and destination choice better predicts leisure and business international travels than multinomial logit models. A major concern in modeling international travel demand is data availability. Tourism studies mostly focus on aggregated data (e.g., number of tourists departing from/arriving in a country). Data on disaggregated international trips made by individuals were not available. Thus, this study uses the airline and border crossing data to distribute trips from US airports. This study aims to increase the understanding of international travel behavior, especially air travel, from the US. For this purpose, this study uses travel demand models to predict airfare, travel time, and trip distribution among major US and international airports across the world. This paper uses a 10% sample of the 2019 DB1B data (specifically targeting a time that predates the COVID pandemic), containing 2.6 million itineraries for 3.9 million passengers. It is comprised of passenger airline ticket sale data collected by BTS. Feasible generalized least square (FGLS) models are used to estimate flight fares per paid itinerary per passenger and their variations for international round trips from US and one-way trips with US origins. A binomial logit model is also employed to find Americans' choice of having an international trip relative to a domestic trip. A gravity model is estimated to distribute trips among different origins in the US to destinations across the world.

The remainder of this paper is organized as follows. The next section summarizes background studies on international travel demand modeling and flight price and duration estimation. Then, the datasets used in this study to estimate international trip distribution, flight fare, and flight duration models are explained. The fourth section elaborates on the estimated models and is followed by the main observations from those models. The last section summarizes the conclusions, limitations, and future applications of this study.

BACKGROUND

Tourism studies constitute the majority of existing literature on travel demand modeling of international trips. Seddighi and Theocharous (19) investigated the impacts of different factors, such as quality of service and political stability of a destination, on tourism attractiveness using the data of tourists visiting Cyprus. Divisekera (20) estimated travel demand models for international tourism in Australia from the US, the UK, Japan, and New Zealand. The study found a correlation between economic factors and international tourism. To find the tourists' utility function, the study used a consumer preferences model known as the Price Independent Generalized Log-Linear model, which represents market demands as the outcome of rational consumers' decisions. Keum (11) used a gravity model, and an economic hypothesis called the Linder hypothesis to analyze trade and tourism flows in Korea. The study confirmed the robustness of the gravity model in estimating the international flows. Wu et al. (21) presented a multi-level destination choice model for tourists in Japan. They indicated that travel time, diversity of destination, and variety-seeking affect destination choice behavior. Wu et al. (12) explored the spatial distribution of tourism flow in China and provided suggestions for tourism improvement. Their study mentioned that tourism flow is deeply influenced by transportation mode developments, regional economies, and quality of service at the destination. Most of the above-mentioned studies focus on international leisure trips (not all trip purposes) or attracting more tourists to specific destinations (aggregate trips to a city) and ignore the importance of international trips by Americans in planning purposes.

Due to the fluctuation in flight fares, it is important to estimate fares based on the available information. Wang et al. (22) used the Airline Origin and Destination Survey (DB1B), the Air Carrier Statistics database (T-100), and machine learning algorithms to predict flight fares. They indicated distance between origin and destination, seat class, passenger volume, quarter of the trip, and crude oil price as important factors in flight fare. Boruah et al. (23) used the Kalman filter technique to predict flight fare based on previous observations. They indicated day of week as the most significant factor in flight fare fluctuations. Zhang et al. (24) presented a neural network model for flight fare prediction as a function of flight start to end time, airline, and service. They applied their model to three months of flight fare data. The above-mentioned studies focus either on tourism or predict flight fares for airlines. Thus, there is a need to model international travel demands to better plan for the future of transportation, as new technologies emerge and become widely available.

DATA SETS USED

Using international travel datasets, this research examines the overseas destination preferences of Americans and models the international travel demand to better prepare for future transportation. This study uses 2019 DB1B flight ticket data, the 2016/17 NHTS as well as publicly available international travel data collected by the National Travel and Tourist Agency (NTTO), Survey of International Air Travelers (SIAT), and Travel and Tourism Satellite Account (TTSA). According to past annual passenger miles recorded by NTTO, international travel accounted for 40% of all revenue passenger miles travelled by US airlines in 2019 wherein US flagged carriers handled 47% of total international air passenger to and from the United States (23, 11). The SIAT survey on US residents visiting overseas countries revealed that European (19.1%) and Caribbean countries (9.4%) accounted for a large proportion of overseas destinations from US, after Canada and Mexico (54.9%) (12). Figure 1 show Americans' rate of travel to different overseas regions in 2019 by air.

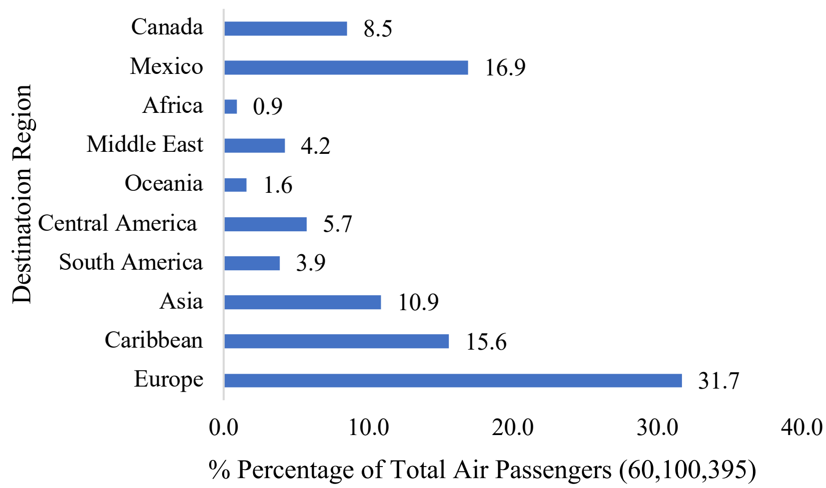

Figure 1. Americans' outbound travel by air in 2019 (SIAT, 2019)

The main data source in this study is the DB1B ticket data collected by the BTS Office of Airline Information. This data is a 10% random sample of US airline passenger ticket itineraries reported by US flag carriers only. It includes trip origin and destination data, yearly and quarterly indicators, number of passengers, number of legs, and distance and fare information for each itinerary. The dataset producers began publishing records in 1993, providing 28 years of available data. This study uses a 10% sample of the 2019 data (before the COVID-19 pandemic), which contains 2.6 million itineraries for 3.9 million passengers. Table 1 and Table 2 summarizes one-way itineraries to and from the US in the 2019 DB1B data.

Table 1. Summary statistics for the DB1B round-trip air ticket data - 2019

	Mean	Median	Std dev	Max	Min
Quarter 1, $N = 246{,}168$					
Flight Fare per Itinerary ($)	953	635	1175	16427	0
Distance Flown (miles)	6669	5232	4313	26051	196
Fare per mile ($)	0.171	0.127	0.16	2.918	0
Passengers	1.446	1	2.58	311	1
Segments	3.058	3	0.96	4	2
Quarter 2, $N = 318{,}033$					
Flight Fare per Itinerary ($)	1022	702	1151	17177	0
Distance Flown (miles)	7150	7298	4244	25870	196
Fare per mile ($)	0.173	0.128	0.16	3.209	0
Passengers	1.414	1	2.57	427	1
Segments	3.041	3	0.96	4	2
Quarter 3, $N = 309{,}842$					
Flight Fare per Itinerary ($)	1033	733	1100	18491	0
Distance Flown (miles)	7318	7662	4167	26950	196
Fare per mile ($)	0.171	0.128	0.15	2.883	0
Passengers	1.374	1	2.15	229	1
Segments	3.010	3	0.96	4	2
Quarter 4, $N = 174{,}532$					
Flight Fare per Itinerary ($)	1055	724	1226	17272	0
Distance Flown (miles)	6921	5331	4504	27338	196

Fare per mile ($)	0.186	0.144	0.16	2.617	0
Passengers	1.307	1	2.37	322	1
Segments	3.327	4	0.90	4	2

Table 2. Summary statistics for the DB1B one-way trip air ticket data - 2019

	Mean	Median	Min	Max	Std Dev
Quarter 1, *N* = 371,334					
Flight Fare per Itinerary ($)	494	304	0	11703	658.5
Distance Flown (miles)	3260	2129	98	21943	2621.6
Fare per mile ($)	0.191	0.138	0	3.795	0.196
Passengers	1.539	2	1	368	3.571
Segments	1.899	2	1	4	0.648
Quarter 2, *N* = 287,751					
Flight Fare per Itinerary ($)	495	316	0	12743	637.7
Distance Flown (miles)	3329	2228	98	22833	2586.1
Fare per mile ($)	0.194	0.143	0	3.867	0.196
Passengers	1.595	2	1	335	4.179
Segments	1.900	2	1	4	0.644
Quarter 3, *N* = 221,507					
Flight Fare per Itinerary ($)	534	342	0	11692	622.9
Distance Flown (miles)	3450	2306	98	20248	2656.4
Fare per mile ($)	0.201	0.153	0	3.5	0.193
Passengers	1.524	2	1	440	3.714
Segments	1.913	2	1	4	0.650
Quarter 4, *N* = 167,983					
Flight Fare per Itinerary ($)	500	318	0	11477	642.4
Distance Flown (miles)	3306	2165	98	20754	2639.1
Fare per mile ($)	0.197	0.144	0	3.469	0.191
Passengers	1.570	2	1	483	4.383
Segments	1.887	2	1	4	0.649

This study also uses the 2016/17 NHTS dataset to model Americans' international trip-making choice versus a domestic long-distance trip. Trip frequency model for long-distance trips (over 75-miles one-way) is estimated using this NHTS dataset leveraged in the study done by Fakhrmoosavi (3). Having this model, travelers' decision to make a long-distance international trip will be modeled using the 2016/17 NHTS dataset. The 2016/17 NHTS data includes 923,572 trips records, which sum to 371 billion trips using NHTS expansion factors. In this dataset, 134.46 million expanded trips are reported as international trips, which account for only 1 percent of the total long-distance trips (~7 billion weighted). The population of 2019 destination nations, as well as information about the languages spoken in the destination countries, were collected from the United Nations website (25). If English is one among major languages spoken, this study assumes the nation is significantly English-speaking. Additionally, the major tourist attractions in 2019 were obtained from the 2019 edition of Euromonitor International's city tourist arrivals (26) (Tourist is defined as an international tourist who visits another country for at least 24 hours and resides in paid or unpaid, group or private lodging for a period not exceeding 12 months) research report that covers over 400 cities worldwide. Mexico and Canada accounted for 40% (39.9 million) and 15% (14.9 million) of total outbound travel in the United States (99.7 million) (5). The STATS Canada and Banco de Mexico websites were used to obtain data on Americans' international visits to Canada and Mexico by land (staying 1 or more nights). Both records show that land travels to Mexico and Canada (39.6 million) account for a significant portion of overall outbound travel to these countries (54.9 million), representing 74.5 percent and 65.7 percent of total trips to respective countries. Figure 2 show Americans' rate of international travel in 2019 by land.

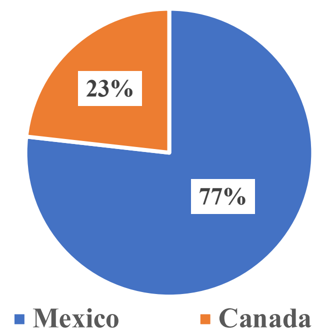

Figure 2. Americans' long-distance outbound trip share by land in 2019 (NTTO, 2019)

MODELS

Figure 3 illustrates the modeling framework for international trip distribution by Americans. The 2016/17 NHTS data is used to estimate Americans' international trip-making. The trip frequency model for long-distance trips (over 75-miles one-way) is estimated using this NHTS dataset in Fakhrmoosavi et al. (3). A travelers' decision to have a long-distance international trip is estimated using a binomial logit model. Trips are then distributed between each US origin airport and other countries' international airports using an origin-constrained gravity model and DB1B data. Flight duration and fare, English language country indicator, tourism attraction country indicator, and population of the country are used as the inputs of this model. Flight duration is not provided in the DB1B data. Thus, it is estimated here based on an average speed and average delay for each stop. In addition, flight fares and their variation are estimated using FGLS models for US outbound and round trips to be considered for model applications.

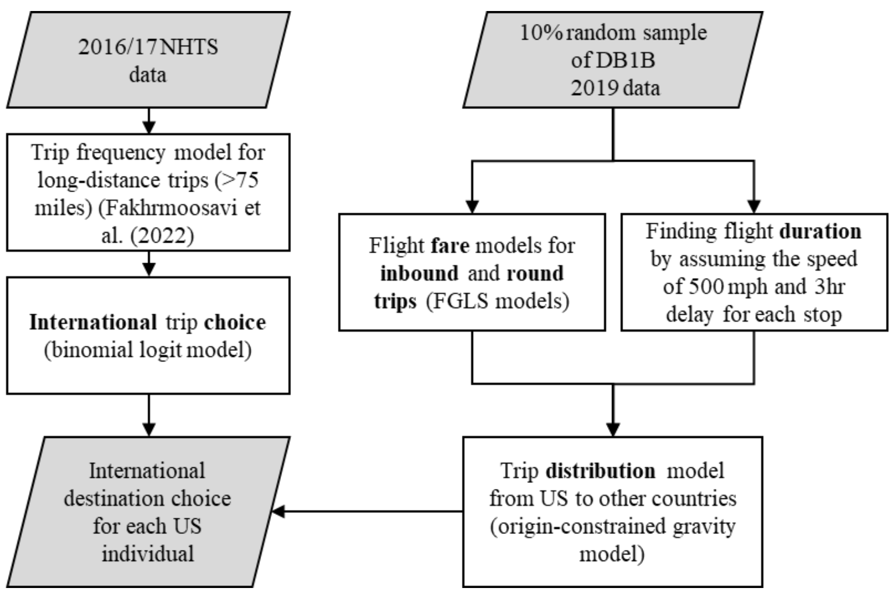

Figure 3. Modeling framework to predict destinations for international trips from US

Flight Fare and Duration Models

International Round Trips - US Origin

Due to the large sample size and unknown nature of heteroscedasticity, we employed feasible generalized least square models to predict the flight fare for international trips. FGLS models for 2019 round-trip itineraries (Table 3) indicate that the flight fare decreases whenever 1) an intermediate stop is included in an otherwise uninterrupted trip, and 2) the number of passengers on the itinerary increases. Trips taken from October to December are more expensive than those taken during other months of the year. Traveling to an English-speaking nation is less expensive than traveling to or from a non-English-speaking country if other variables are kept constant. Shifting all samples towards business or higher class and towards United Airlines increases the flight fare by 150% and 7.5%, respectively. Table 4 presents the model estimates when the log of linear model residuals is regressed on all dependent variables. The results show that the flight prices of the itineraries with more than 1 stop vary significantly in price compared to those without stops.

Table 3. FGLS model Estimates for international round trips to and from US (DB1B, 2019)

Y: Fare (\$) per paid Itinerary per passenger, N = 1,048,268, Adj. R^2: 0.3026			
Variable Name	**Estimate**	**t-stat**	**P-value**
(Intercept)	337.60	105.136	0.000
Distance Flown (miles)	0.058	208.72	0.000
Distance Flown (miles)*Business class or higher	0.281	108.32	0.000
Trip made during April to June	20.41	12.529	0.000
Trip made during July to September	18.54	11.364	0.000
Trip made during October to December	69.04	54.304	0.000
Restricted Coach Class	56.69	35.829	0.000
Business class or higher	-118.2	-5.649	0.000
#Passengers on the Itinerary	-8.388	-63.497	0.000
ln(Population of Destination Country)	-8.204	-36.758	0.000
Itinerary with 1 stop	-99.16	-59.481	0.000
Itinerary with 2 stops	-69.97	-67.127	0.000
Outbound Trip	125.0	124.255	0.000
Destination is English Speaking	-8.798	-5.588	0.000
Alaska Airlines	-53.25	-9.542	0.000
JetBlue Airlines	-15.52	-5.375	0.000
Delta Airlines	51.71	23.759	0.000
United Airlines	88.66	63.178	0.000
SkyWest Airlines	76.96	17.412	0.000
Endeavor Air	37.65	8.713	0.000
Canadian Pacific Airlines	14.90	2.92	0.004
Hawaiian Airlines	775.3	26.819	0.000
GoJet Airlines	-41.06	-11.981	0.000
Southwest Airlines	-159.9	-43.468	0.000

Spirit Airlines	143.6	47.993	0.000
Mesa Airlines	-17.45	-4.173	0.000
Republic Airlines	110.3	17.973	0.000
Eva Airlines	39.46	6.438	0.000
PSA (American Eagle) Airlines	148.3	13.788	0.000
Frontier Airlines	-215.3	-10.231	0.000
Sun Country Airlines	-227.8	-3.188	0.001
Horizon Air	-5.249	-0.67	0.503
Distance Flown (miles)*Destination is English Speaking	-0.002	-6.894	0.000
Distance Flown (miles)*Trip made during April to June	0.004	12.322	0.000
Distance Flown (miles)*Trip made during July to September	0.006	18.248	0.000
Business class or higher*Destination is English Speaking	131.2	6.995	0.000
Trip made during April to June*Business class or higher	-83.24	-4.228	0.000
Trip made during July to September*Business class or higher	-157.5	-7.571	0.000
Business class or higher*Alaska Airlines	-209.9	-3.888	0.000
Business class or higher*JetBlue Airlines	-524.0	-3.506	0.000
Business class or higher*Delta Airlines	-858.6	-38.11	0.000
Business class or higher*United Airlines	225.0	9.995	0.000
Business class or higher*SkyWest Airlines	-276.1	-4.241	0.000
Business class or higher*Endeavor Air	-578.5	-6.903	0.000
Business class or higher*Canadian Pacific Airlines	-393.6	-3.367	0.001
Business class or higher*Hawaiian Airlines	-588.2	-7.498	0.000
Business class or higher*Itinerary with 1 stop	-143.9	-4.367	0.000
Business class or higher*Itinerary with 2 stops	-318.5	-17.82	0.000
Distance Flown (miles)*Alaska Airlines	-0.008	-5.609	0.000
Distance Flown (miles)*JetBlue Airlines	0.009	13.158	0.000
Distance Flown (miles)*Delta Airlines	0.014	39.903	0.000
Distance Flown (miles)*Southwest Airlines	-0.011	-10.395	0.000
Distance Flown (miles)*Spirit Airlines	-0.031	-28.813	0.000
Distance Flown (miles)*SkyWest Airlines	0.017	15.968	0.000
Distance Flown (miles)*Republic Airlines	0.027	26.092	0.000
Distance Flown (miles)*Endeavor Air	0.029	25.92	0.000
Distance Flown (miles)*Eva Airlines	0.027	15.114	0.000
Distance Flown (miles)*PSA Airlines	0.013	7.713	0.000
Distance Flown (miles)*Horizon Air	-0.010	-3.433	0.001
Distance Flown (miles)*Hawaiian Airlines	-0.066	-18.948	0.000
Distance Flown (miles)*GoJet Airlines	0.007	2.596	0.009
Distance Flown (miles)*Frontier Airlines	-0.015	-2.116	0.034

Distance Flown (miles)*Sun Country Airlines	0.047	2.255	0.024

Table 4. Variance model estimates for international round trips to and from US (DB1B,2019)

Y = log(Residuals^2) , N =1,048,268, Adj. R^2: 0.2947			
Variable Name	**Estimate**	**t-stat**	**P-value**
(Intercept)	9.371	513.9	0.000
Distance Flown (miles)	0.000	325.3	0.000
Trip made during April to June	-0.018	-2.97	0.003
Trip made during July to September	-0.023	-3.76	0.000
Trip made during October to November	0.253	34.45	0.000
Alaska Airlines	0.073	3.90	0.000
JetBlue Airlines	-1.017	-76.26	0.000
Delta Airlines	0.211	34.25	0.000
Southwest Airlines	-0.609	-48.24	0.000
United Airlines	0.116	17.41	0.000
Spirit Airlines	-0.970	-61.27	0.000
Mesa Airlines	0.110	5.89	0.000
SkyWest Airlines	0.156	9.90	0.000
Republic Airways	-0.081	-4.58	0.000
Endeavor Air	0.187	10.30	0.000
Canadian Pacific Air Lines	0.136	4.46	0.000
Eva Air	0.166	6.50	0.000
PSA (American Eagle) Airlines	-0.056	-2.31	0.021
Horizon Air	0.428	11.37	0.000
Hawaiian Airlines	-0.129	-6.64	0.000
GoJet Airlines	0.262	6.05	0.000
Frontier Airlines	-0.993	-18.90	0.000
Sun Country Airline	1.336	21.93	0.000
Itinerary with 2 stops	-0.207	-23.03	0.000
Itinerary with 3 stops	-0.307	-53.60	0.000
Restricted Coach Class	-0.503	-63.38	0.000
Business class or higher	2.840	270.69	0.000
Outbound Trip	0.351	62.39	0.000
Destination is English Speaking	0.217	41.89	0.000
ln (Population of Destination Country)	0.010	7.06	0.000
#Passengers on the Itinerary	0.006	6.11	0.000

International One-way Trips - to and from US

The FGLS model results for air fares of international one-way trips to and from US is shown in Table 5 and variance model estimates are shown in Table 6. The estimated model coefficients reveal that a flight price cost $0.078 per mile flown for coach class and $0.163 per mile flown for business class or higher. The flight fare decreases as number of passengers, numbers of stops on the itinerary increases. Trip made during April to June shows high variation as compared to other days of the year. Shifting the sample towards business or higher class increases the flight fare by 125% while same shift towards Southwest Airlines decreases the cost by 58.5%.

Table 5. FGLS model estimates for international one-way trips - to and from US (DB1B,2019)

Y: Fare ($) per paid Itinerary per passenger, N = 1,048,575, Adj. R^2: 0.2446			
Variable Name	**Estimate**	**t-stat**	**P-value**
(Intercept)	320.0	154.7	0.000
Distance Flown (miles)	0.078	179.0	0.000
Distance Flown (miles)*Business class or higher	0.163	61.53	0.000
#Passengers on the Itinerary	-3.602	-70.87	0.000
Outbound Trip?	-34.20	-58.49	0.000
Restricted Coach Class	-7.743	-7.78	0.000
Business class or higher	-66.53	-2.40	0.016
Trip made during April to June	8.645	8.72	0.000
Trip made during July to September	1.907	1.76	0.079
Trip made during October to December	4.992	4.24	0.000
Itinerary with 1 stop	-40.23	-76.41	0.000
Itinerary with 2 stops	-23.69	-21.29	0.000
Itinerary with 3 stops	117.5	24.43	0.000
Destination is English Speaking	-24.75	-27.51	0.000
Ln (Population of Destination Country)	-7.668	-53.18	0.000
Alaska Airlines	-29.55	-13.63	0.000
JetBlue Airlines	-42.76	-24.86	0.000
Delta Airlines	-11.08	-7.54	0.000
United Airlines	-16.53	-11.23	0.000
SkyWest Airlines	24.71	11.15	0.000
Canadian Pacific Airlines	-18.48	-13.53	0.000
Horizon Air	-4.890	-1.98	0.048
Hawaiian Air	253.2	16.08	0.000
SunCountry Airline	-72.60	-12.45	0.000
Southwest Airlines	-7.897	-3.21	0.001
Spirit Airlines	-92.82	-53.99	0.000
Mesa Airlines	56.18	36.24	0.000
Republic Airline	1.783	0.75	0.453
Endeavor Airline	15.17	6.92	0.000
Eva Airline	21.46	6.29	0.000
PSA Airline	23.54	6.84	0.000

GoJet Airline	60.15	12.90	0.000
Frontier Airline	-118.5	-18.18	0.000
Distance Flown (miles)*Trip made during April to June	-0.004	-7.81	0.000
Distance Flown (miles)*Trip made during July to September	0.008	16.50	0.000
Distance Flown (miles)*Trip made during October to December	-0.004	-6.92	0.000
Distance Flown (miles)*Alaska Airlines	-0.008	-8.38	0.000
Distance Flown (miles)*JetBlue	0.014	21.68	0.000
Distance Flown (miles)*Delta Airlines	0.015	24.10	0.000
Distance Flown (miles)*Southwest Airlines	-0.031	-21.67	0.000
Distance Flown (miles)*United Airlines	0.018	35.57	0.000
Distance Flown (miles)*Spirit Airlines	-0.032	-35.40	0.000
Distance Flown (miles)*SkyWest Airlines	0.008	7.46	0.000
Distance Flown (miles)*Republic Airline	0.017	12.07	0.000
Distance Flown (miles)*Endeavor Airline	0.007	5.47	0.000
Distance Flown (miles)*Eva Airline	0.032	15.65	0.000
Distance Flown (miles)*PSA Airline	-0.009	-3.74	0.000
Distance Flown (miles)*Horizon Air	-0.004	-2.54	0.011
Distance Flown (miles)*Hawaiian Air	-0.039	-12.92	0.000
Distance Flown (miles)*GoJet Airline	-0.005	-1.72	0.085
Distance Flown (miles)*Frontier Airline	-0.012	-3.09	0.002
Business class or higher*Alaska Airlines	-123.4	-7.23	0.000
Business class or higher*JetBlue Airlines	505.0	12.18	0.000
Business class or higher*Delta Airlines	53.82	3.13	0.002
Business class or higher*United Airlines	-45.94	-4.17	0.000
Business class or higher*SkyWest Airlines	-43.62	-1.80	0.071
Business class or higher*Canadian Pacific Airlines	-95.83	-3.32	0.001
Business class or higher*Horizon Air	-68.84	-2.46	0.014
Business class or higher*Hawaiian Air	262.1	3.22	0.001
Business class or higher*SunCountry Airline	-360.0	-6.10	0.000
Business class or higher*Itinerary with 1 stop	-105.5	-11.57	0.000
Business class or higher*Itinerary with 2 stops	-353.4	-18.70	0.000
Business class or higher*Itinerary with 3 stops	-520.1	-7.41	0.000
Business class or higher* Destination is English Speaking	52.91	5.56	0.000
Distance Flown (miles)* Destination is English Speaking	-0.010	-25.76	0.000
Business class or higher* Ln (Population of Destination Country)	11.86	4.73	0.000
Trip made during April to June*Business class or higher	-50.15	-4.93	0.000
Trip made during July to September*Business class or higher	-126.1	-9.99	0.000
Trip made during October to December*Business class or higher	-35.60	-2.69	0.007

Table 6. Variance model estimates for international one-way trips - to and from US (DB1B, 2019)

Y = log(Residuals^2) , N = 1,048,575, Adj. R^2*: 0.2896*

Variable Name	Estimate	t-stat	P-value
(Intercept)	9.880	628.2	0.000
Distance Flown (miles)	0.000	364.8	0.000
#Passengers on the Itinerary	0.006	9.6	0.000
Itinerary with 2 stops	-0.203	-26.8	0.000
Itinerary with 3 stops	0.088	3.80	0.000
Restricted Coach Class	-1.037	-153.0	0.000
Business class or higher	1.959	192.8	0.000
Destination is English Speaking?	-0.214	-41.1	0.000
Ln (Population of Destination Country)	-0.046	-34.7	0.000
Trip made during April to June	-0.066	-13.0	0.000
Alaska Airlines	-0.174	-12.9	0.000
JetBlue Airlines	-1.324	-124.6	0.000
Delta Airlines	0.113	16.1	0.000
Southwest Airlines	-0.642	-48.4	0.000
United Airlines	-0.126	-18.7	0.000
Spirit Airlines	-1.351	-112.3	0.000
Mesa Airlines	-0.219	-13.4	0.000
SkyWest Airlines	-0.094	-6.9	0.000
Republic Airways	-0.119	-6.8	0.000
Endeavor Air	0.040	2.5	0.014
Canadian Pacific Air Lines	-0.446	-24.7	0.000
PSA Airlines	-0.162	-5.9	0.000
Horizon Air	-0.158	-7.5	0.000
Hawaiian Airlines	-0.228	-8.3	0.000
GoJet Airlines	0.164	5.1	0.000
Frontier Airlines	-1.145	-32.3	0.000

International Trip Choice

The specifications of the logistic regression model to estimate international trip choice for Americans are shown in Table 7. The model indicates that international trip frequency (per person) rises by about 16% with a 1 standard deviation increase in the respondent's household income (i.e., $62,000). Increasing the summer trip and spring trip indicators by 1 standard deviation also increases the frequency of international trips by 19% and 14%, respectively. International trips fall 23% when the female indicator increases by 1 standard deviation and 31% when the full-time employed indicator increases by 1 standard deviation. Religious and personal business trips are also less likely to be international.

Table 7. Specifications of the logistic regression model for international versus domestic trips using the 2016/17 NHTS data

	Coefficient Estimates	t-Stat	P-Value	Practical Significance
(Intercept)	-5.594	-7.14	0.000	-
Household income (1000$)	0.006	1.63	0.103	0.161
Female	-1.067	-2.42	0.016	-0.228
Hispanic	1.424	2.67	0.008	0.148
White	1.114	2.27	0.023	0.159
Full-time employed	-1.501	-3.65	0.000	-0.315
Summer trip	0.988	1.78	0.075	0.193
Spring trip	0.907	1.68	0.094	0.140
Personal business trip	-1.066	-1.44	0.150	-0.104
Religious community trip	-14.232	-47.88	0.000	-0.869

R-squared: 0.1344, *n*= 13,966

Trip Distribution Model

An origin-constrained gravity model was used to distribute trips among different origins and destinations. Gravity models in their traditional form consist of the production, attraction (e.g., tourism attractions, population, and language of the destination), friction (i.e., travel time and/or fare), and a gravity constant term. A logarithmic operator was applied to form a log-linear gravity model and an ordinary-least-squares (OLS) model was estimated to find the number of trips distributed between each origin and destination pair. Friction factor here is a function of impedance incorporating auto and air travel times and cost (i.e., flight fare, highway toll) normalized by value of time. Value of travel time for air travelers is assumed to be $30 per hour and $20 per hour for auto users. Table 8 shows the specifications of this log-linear model as well as the practical significance of different statistically significant variables. This model was estimated using data from multiple sources indicating trip production for 334 major US airports and attractions of country locations for 1028 international airports in countries other than the US. Due to the lack of data for origins and destinations of land travelers to Canada and Mexico, major airports in most touristic cities in Canadian provinces that are accessible with from US (e.g., Ontario, Quebec, British Columbia, Alberta, Nova Scotia) are considered as the destination locations. Origins are also assumed to be the major airport of the closest state in the US. For Mexico, all trips are aggregated in one origin and destination pair from Texas to the Sinaloa state in Mexico. The trip distribution model indicated that trips headed to a foreign destination from an American origin fall 41% when the travel start to end time raises by 7 hours or air ticket increases by $210. Destinations hosting tourist attractions increase origin-destination flow by 48%. when this indicator variable goes from 0 to 1. The population and English-speaking indicator at the destination country are neither practically nor statistically significant.

Table 8. Specifications of the log-linear gravity model to estimate the number of trips between US major airports and other countries' airports

	Estimate	t-stat	P-Value	Practical Significance
(Intercept)	9.796	104.65	0.000	
Trip Production in Origin Airport	0.238	81.62	0.000	0.969
Travel Time & Cost	-1.578	124.11	0.000	0.409
Population of Destination Country	0.0013	0.50	0.616	0.0012
Tourism Indicator in Destination Country	0.907	51.60	0.000	0.136
English Speaking Country (Destination)	0.0024	0.17	0.864	0.0004

CONCLUSIONS

This research contributes to a better knowledge of Americans' overseas travel by estimating travel demand and expenses (in time and money) between major US airports and international airports globally, as well as land trips to Mexico and Canada. The study uses 2019 DB1B aircraft ticket data, the 2016/17 NHTS, US outbound passenger travel aggregate estimates of 2019 NTTO, destination country characteristics from UN world information and major attraction cities data for tourists in 2019 from Euromonitor international report. The main data source of this study, 2019 DB1B provided by BTS, revealed that the flight fare for international travel falls as the number of passengers on the itinerary rises. Round trips made in October to December are more expensive than those taken during the other months of the year while one-way trips made during April to June show high variation as compared to other times of the year. A round trip to an English-speaking nation is less expensive than traveling to or from a non-English-speaking country if other variables are kept constant. The international round-trip air fares cost \$0.058 per mile flown for coach class and \$0.281 per mile flown for business class or higher. Shifting the sample towards business or higher class increases the one-way flight fare by 125% and the round-trip fare by 151%.

The international trip choice model reveals that the probability of taking international trips rises 16% when household income is increased by 1 standard deviation (i.e., \$62,000). Employment status, race, female indicator, trip season, and trip purpose are other significant variables affecting international trip choice by Americans. A log-linear model was used to distribute international trips among various major airports in the US and other countries. The trip distribution model indicated that travel time and cost, and tourism attraction in the destination are the statistically significant variables affecting the number of trips going to an international location. This model also suggested that trips headed to a foreign destination from an American airport fall 41% when the friction factor (i.e., travel time and normalized cost by value of time for different modes) raises by 7 hours and increases 48% when all destination shift to a tourist attraction from not being an attraction.

This study has some limitations that should be considered, prompting areas for future research. To the knowledge of the authors, there is no public data that thoroughly reports upon international ground trip counts from US cities to cities in Canada or Mexico. Data used in this study were comprised of aggregated border crossing travel counts; these were then used for distributing trips among different destinations based on their tourism attraction. Further work is required to more precisely account for ground trips for use in international travel demand models.

ACKNOWLEDGMENTS

The authors thank the Texas Department of Transportation (TxDOT) for financially supporting this research, under research project 0-7081, "Understanding the Impact of Autonomous Vehicles on Long Distance Travel Mode and Destination Choice in Texas". The authors thank Jade (Maizy) Jeong for editorial and submission support.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: Fakhrmoosavi, F., Kockelman, Paithankar, P., Perrine, K.; data collection: Perrine, K., Kockelman; analysis and interpretation of results: Paithankar, P., Fakhrmoosavi, F., and Kockelman; draft manuscript preparation: Paithankar, P., Fakhrmoosavi, F., Kockelman, Perrine, K; All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. BTS (2022). *U.S. Passenger-Miles* | Bureau of Transportation Statistics. [online] Available at: <https://www.bts.gov/content/us-passenger-miles> [Accessed 11 July 2022].
2. BTS (2022). 2019 *Traffic Data for U.S. Airlines and Foreign Airlines U.S. Flights - Final, Full-Year* | Bureau of Transportation Statistics. [online] Available at: <https://www.bts.gov/newsroom/final-full-year-2019-traffic-data-us-airlines-and-foreign-airlines> [Accessed 11 July 2022].
3. Fakhrmoosavi, F., Paithankar, P., Kockelman, K., Huang, Y., Hawkins, J., (2022). Self-Driving Vehicles' Impacts on Americans' Long-Distance Domestic Travel Choices. Submitted to *Transportation Research Record.* https://www.caee.utexas.edu/prof/kockelman/public_html/TRB23LDAVDomesticTrips.pdf
4. International Trade Administration | Trade.gov. (2022). APIS/I-92 Monitor. [online] Available at: <https://www.trade.gov/data-visualization/apisi-92-monitor> [Accessed 11 July 2022].
5. International Trade Administration | Trade.gov. (2022) *Travel and Tourism Research*. [online] Available at: < https://www.trade.gov/travel-and-tourism-research > [Accessed 23 July 2022].
6. www150.statcan.gc.ca. (2022). *International travellers entering or returning to Canada, by type of transportation and traveller type*. [online] Available at: <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=2410005301> [Accessed 26 July 2022].
7. U.S. Travel Association. (2022). *Travel Forecast*. [online] Available at: <https://www.ustravel.org/research/travel-forecasts> [Accessed 23 July 2022].
8. Llorca, C., Molloy, J., Ji, J., & Moeckel, R. (2018). Estimation of a long-distance travel demand model using trip surveys, location-based big data, and trip planning services. *Transportation Research Record*, *2672*(47), 103-113. https://doi.org/10.1177/0361198118777064
9. Masiero, L., & Qiu, R. T. (2018). Modeling reference experience in destination choice. *Annals of Tourism Research*, *72*, 58-74. https://doi.org/10.1016/j.annals.2018.06.004
10. Qu, H., & Lam, S. (1997). A travel demand model for Mainland Chinese tourists to Hong Kong. *Tourism Management*, *18*(8), 593-597. https://doi.org/10.1016/S0261-5177(97)00084-8

11. Keum, K. (2010). Tourism flows and trade theory: a panel data analysis with the gravity model. *The Annals of Regional Science*, *44*(3), 541-557. https://doi.org/10.1007/s00168-008-0275-2
12. Wu, S., Wang, L., & Liu, H. (2021). Study on tourism flow network patterns on May Day holiday. *Sustainability*, *13*(2), 947. https://doi.org/10.3390/su13020947
13. Qu, H., & Or, Y. S. (2006). Determinants of the travel demand model for Canadian tourists to the US. *International journal of hospitality & tourism administration*, *7*(4), 1-19. https://doi.org/10.1300/J149v07n04_01
14. Zhang, Y., Qu, H., & Tavitiyaman, P. (2009). The determinants of the travel demand on international tourist arrivals to Thailand. *Asia Pacific Journal of Tourism Research*, *14*(1), 77-92. https://doi.org/10.1080/10941660902728080
15. Lenzen, M., Sun, Y. Y., Faturay, F., Ting, Y. P., Geschke, A., & Malik, A. (2018). The carbon footprint of global tourism. *Nature climate change*, *8*(6), 522-528. https://www.nature.com/articles/s41558-018-0141-x
16. Tziridis, K., Kalampokas, T., Papakostas, G. A., & Diamantaras, K. I. (2017). Airfare prices prediction using machine learning techniques. In *2017 25th European Signal Processing Conference (EUSIPCO)* (pp. 1036-1039). IEEE. https://ieeexplore.ieee.org/abstract/document/8081365
17. Ratnakanth, G. (2022). Prediction of Flight Fare using Deep Learning Techniques. In *2022 International Conference on Computing, Communication and Power Technology (IC3P)* (pp. 308-313). IEEE. https://ieeexplore.ieee.org/abstract/document/9793411/metrics
18. Furuichi, M., & Koppelman, F. S. (1994). An analysis of air travelers' departure airport and destination choice behavior. *Transportation Research Part A: Policy and Practice*, *28*(3), 187-195. https://doi.org/10.1016/0965-8564(94)90016-7
19. Seddighi, H. R., & Theocharous, A. L. (2002). A model of tourism destination choice: a theoretical and empirical analysis. *Tourism management*, *23*(5), 475-487. https://doi.org/10.1016/S0261-5177(02)00012-2
20. Divisekera, S. (2003). A model of demand for international tourism. *Annals of tourism research*, *30*(1), 31-49. https://doi.org/10.1016/S0160-7383(02)00029-4
21. Wu, L., Zhang, J., & Fujiwara, A. (2012). A tourist's multi-destination choice model with future dependency. *Asia Pacific Journal of Tourism Research*, *17*(2), 121-132. https://doi.org/10.1080/10941665.2011.616902
22. Wang, T., Pouyanfar, S., Tian, H., Tao, Y., Alonso, M., Luis, S., & Chen, S. C. (2019, July). A framework for airfare price prediction: a machine learning approach. In *2019 IEEE 20th international conference on information reuse and integration for data science (IRI)* (pp. 200-207). IEEE. 10.1109/IRI.2019.00041
23. Boruah, A., Baruah, K., Das, B., Das, M. J., & Gohain, N. B. (2019). A Bayesian approach for flight fare prediction based on Kalman filter. In *Progress in Advanced Computing and Intelligent Engineering* (pp. 191-203). Springer, Singapore. https://link.springer.com/chapter/10.1007/978-981-13-0224-4_18
24. Zhang, Q., He, Y., & Jing, X. (2020). Mask Neural Network for Predicting Flight Ticket Price. In *Signal and Information Processing, Networking and Computers* (pp. 114-120). Springer, Singapore. https://link.springer.com/content/pdf/10.1007/978-981-15-4163-6_14.pdf

25. Data.un.org. (2022). UN data | record view | Total population, both sexes combined (thousands). [online] Available at: <https://data.un.org/Data.aspx?d=PopDiv&f=variableID%3A12> [Accessed 26 July 2022].
26. Yasmeen, R., (2022). Top 100 City Destinations 2019 Edition. [online] Go.euromonitor.com. Available at: < https://go.euromonitor.com/rs/805-KOK-719/images/wpTop100Cities19.pdf?mkt_tok=ODA1LUtPSy03MTkAAAGFNWvcuiZu38R-mDzlIa4FCCOhge7y7NFkVCbegXxdJiGDohBB9FKnR7JRzBIYmqmm8g7Zjjv9OX83jQQrEdpL1FYbBAepy5WBlmvz_87b2aUzGbY> [Accessed 26 July 2022].