Airline itinerary case (Boeing)

Michel Bierlaire

March 1, 2019

Context

This data comes from an Internet choice survey conducted by Boeing Commercial Airplanes in 2004 and 2005. The survey, which was designed with the assistance of Jordan Louviere of the University of Technology, Sydney, is described in Garrow et al. (2007). Boeing was interested in understanding the sensitivity that air passengers have toward the attributes of an airline itinerary, such as fare, travel time, transfers, legroom, and aircraft.

The survey was conducted by intercepting customers of an Internet airline booking service that searches for low-cost travel deals. While waiting for the search engine to return the real itineraries associated with the specific user request for a travel in a city pair, randomly selected customers were asked to complete a survey tailored to their origin and destination.

Data collection

A typical page of the survey instrument is shown in Figure 1. The respondent was offered three choices based on the origin-destination market request that was entered into the itinerary search engine. The first alternative is always a non-stop flight, the second always a flight with 1 stop on the same airline, and the third is always a flight with 1 stop and a change of airline. The respondent was asked to rank the available choices and was also given the option to decline all of the presented options. There are 3609 respondents, each providing 1 SP response.

Demographic data collected included age, gender, income, occupation, and education. Situational variables that were identified included the desired departure time, trip purpose, who is paying for the trip, and the number in the travel party. All trips were for origin-destination city pairs in the United States.

Variables and descriptive statistics

Descriptions of the available variables are reported in Tables 1 and 2 and some descriptive statistics are provided in Table 3.

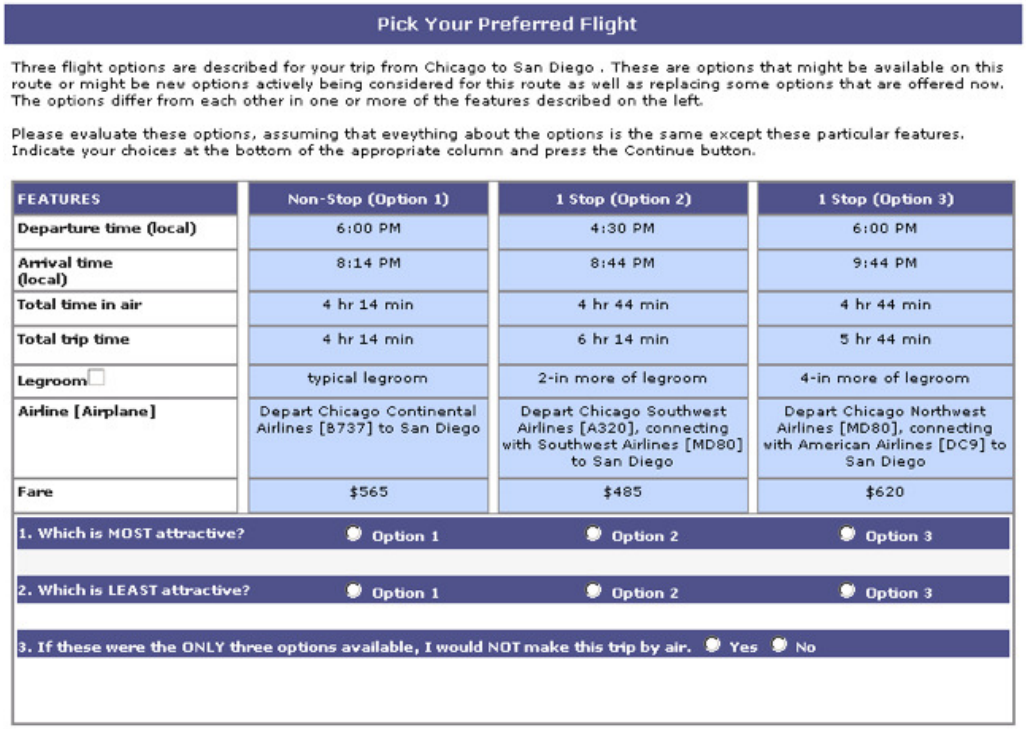

Figure 1: Example of survey instrument

Variable	**Description**
SubjectId	Unique identifier for each respondent.
q17_Gender	1 if male, 2 if female, 99 or -1 if missing.
q15_Age	Age (1 = Less than 18 years, 2 = 18-24 years, 3= 25-34 years, 3.5 = 25-44 years, 4 = 35-44 years, 5 = 45-54 years, 6 = 55-64 years, 7 = 65-74 years, 8 = 75 years or older, 99 or -1 if missing)
q19_Occupation	Occupation (01 = Executive and Managerial, 02 = Professional, 03 = Technicians and related support, 04 = Sales, 05 = Administrative support, 06 = Services, 07 = Precision production, craft, repair, 08 = Machine operators, assemblers, inspectors, 09 = Transportation and material moving, 10 = Handlers, cleaners, helpers, 11 = Farming, forestry, and fishing, 12 = Armed forces, 99 or -1 if missing)
q16_Income	Discrete income (14 different levels); -1 or 99 if missing information
Cont_Income	Annual income in $1000; -1 if missing information
q20_Education	Education (01 = Less than High School Diploma, 02 = High School Graduate, 03 = Some college, No Degree, 04 = Associate Degree - Occupational, 05 = Associate Degree - Academic, 06 = Bachelors Degree, 07 = Masters Degree, 08 = Professional Degree, 09 = Doctorate Degree, 99 or -1 if missing)

Table 1: Description of respondent specific variables

Variable	Description
q11_DepartureOrArrivalIsImportant	Importance of punctuality of departure or arrival (1 = departure is important; 2= arrival is important; otherwise, not important)
q02_TripPurpose	Trip purpose (1=business, 2=leisure, 3=attending conference/seminar/training, 4=both business and leisure, 0=trip purpose missing)
q03_WhoPays	1 if the traveler is paying for the trip, 2 if it is his employer, 3 if it is a third party, 0 if missing
q12_IdealDepTime	Respondents ideal departure time (minutes after midnight), -1 indicates a missing value
q13_IdealArrTime	Respondents ideal arrival time (minutes after midnight), -1 indicates a missing value
q14_PartySize	Number of persons traveling, -1 and 99 indicate missing values
OriginGMT	Origin city time zone (minutes from GMT (Greenwich Mean Time))
DestinationGMT	Destination city time zone (minutes from GMT)
Direction	Direction of itinerary (1=East to West, 2=West to East, 3=North-South, 0=missing)
DepartureTimeHours_X*	Option X: Departure time, local (hours after midnight)
ArrivalTimeHours_X*	Option X: Arrival time, local (hours after midnight)
FlyingTimeHours_X*	Option X: Total time in air (hours)
TripTimeHours_X*	Option X: Total trip time (hours)
Legroom_X	Option X: Legroom , 1 = 2 inches less than typical, 2 = typical, 3 = 2 inches more than typical, 4 = 4 inches more than typical
AirlineFirstFlight_X	Option X: Airline for first leg (only known to arbitrary airline number for proprietary reasons)
AirlineSecondFlight_X	Option X: Airline for second leg (if there exists a second leg) (only known to arbitrary airline number for proprietary reasons)
AirplaneFirstFlight_X	Option X: Airplane for first leg (only known to arbitrary airplane number for proprietary reasons)
AirplaneSecondFlight_X	Option X: Airplane for second leg (if there exists a second leg) (only known to arbitrary airplane number for proprietary reasons)
Fare_X	Option X: Fare ($)
BestAlternative_X	The chosen alternative is X

Table 2: Description of alternative specific attributes where X corresponds to choice option (1),(2) and (3)

* For the variables DepartureTimeHours_X, ArrivalTimeHours_X, FlyingTimeHours_X and TripTimeHours_X there is an additional variable measuring the time in minutes where the variable name in each case substitutes '...Hours_X' for '...Mins_X'.

Variable	Average	St. Dev.	Min	Max
SubjectId	1807.50	1043.41	1.00	3613.00
q17_Gender	1.46	0.50	1.00	2.00
q15_Age	3.95	1.15	1.00	8.00
q19_Occupation	2.54	1.90	1.00	12.00
q16_Income	8.09	3.53	1.00	14.00
Cont_Income	108.20	87.63	10.00	350.00
q20_Education	5.88	1.71	1.00	9.00
q02_TripPurpose	2.04	0.76	1.00	4.00
q03_WhoPays	1.20	0.46	1.00	3.00
q14_PartySize	1.70	0.99	1.00	5.00
OriginGMT	382.18	82.08	300.00	480.00
DestinationGMT	397.34	82.87	300.00	480.00
Direction	1.59	0.49	1.00	2.00
BestAlternative_1	0.69	0.46	0.00	1.00
BestAlternative_2	0.16	0.37	0.00	1.00
DepartureTimeHours_1	11.72	3.34	6.00	18.00
ArrivalTimeHours_1	15.21	3.35	7.67	21.63
FlyingTimeHours_1	3.74	1.59	0.67	6.35
TripTimeHours_1	3.74	1.59	0.67	6.35
Legroom_1	2.46	1.12	1.00	4.00
AirlineFirstFlight_1	4.61	2.56	1.00	11.00
AirlineSecondFlight_1	0.00	0.00	0.00	0.00
AirplaneFirstFlight_1	4.52	2.30	1.00	8.00
AirplaneSecondFlight_1	0.00	0.00	0.00	0.00
Fare_1	405.66	199.87	80.00	1330.00
DepartureTimeHours_2	11.67	3.35	6.00	18.00
ArrivalTimeHours_2	16.92	3.36	9.17	24.10
FlyingTimeHours_2	4.24	1.59	1.17	6.85
TripTimeHours_2	5.50	1.68	1.83	8.85
Legroom_2	2.48	1.13	1.00	4.00
AirlineFirstFlight_2	4.68	2.65	1.00	11.00
AirlineSecondFlight_2	0.00	0.00	0.00	0.00
AirplaneFirstFlight_2	4.51	2.29	1.00	8.00
AirplaneSecondFlight_2	0.00	0.00	0.00	0.00
Fare_2	407.07	200.96	80.00	1390.00
DepartureTimeHours_3	11.66	3.34	6.00	18.00
ArrivalTimeHours_3	16.89	3.41	9.25	24.03
FlyingTimeHours_3	4.24	1.59	1.17	6.85
TripTimeHours_3	5.48	1.67	1.92	8.85
Legroom_3	2.53	1.13	1.00	4.00
AirlineFirstFlight_3	4.65	2.59	1.00	11.00
AirlineSecondFlight_3	4.65	2.65	1.00	11.00
AirplaneFirstFlight_3	4.50	2.31	1.00	8.00
AirplaneSecondFlight_3	4.50	2.28	1.00	8.00
Fare_3	405.20	197.68	80.00	1275.00

Table 3: Descriptive statistics of variables

References

Garrow, L. A., Jones, S. P. and Parker, R. A. (2007), 'How much airline customers are willing to pay: An analysis of price sensitivity in online distribution channels', *Journal of Revenue and pricing Management* **5**(4), 271–290.