Modeling Competition among Airline Itineraries

Virginie Lurkin

Laurie A. Garrow

Transport and Mobility Laboratory (TRANSP-OR),
École Polytechnique Fédérale de Lausanne (EPFL)

hEART 2017: 6th Symposium of the European Association for Research in Transportation
September 12-14, Haifa (Israel)

Overview

1. **Motivation and research objective**
2. Data and methodology
3. Results
4. Conclusions and limitation
5. New research objective
6. Preliminary results
7. Next steps

Itinerary choice model

Itineraries from ZRH-TLV

LX 252 LX 256 LY 347 AF 1415/AF 1620

Itinerary choice model

Itineraries from ZRH-TLV

LX 252 LX 256 LY 347 AF 1415/AF 1620

$$U_i = V_i + \varepsilon_i$$

Itinerary choice model

Itineraries from ZRH-TLV

LX 252 | LX 256 | LY 347 | AF 1415/AF 1620

$$U_i = V_i + \varepsilon_i$$

$$V_i = \alpha_i + \beta_1 Cost_i + \beta_2 Time_i + ...$$

Itinerary choice model

Itineraries from ZRH-TLV

LX 252 LX 256 LY 347 AF 1415/AF 1620

$$U_i = V_i + \varepsilon_i$$

$$V_i = \alpha_i + \beta_1 Cost_i + \beta_2 Time_i + ...$$

$$P_i = \frac{e^{V_i}}{\sum_j e^{V_j}}$$

Factors influencing itinerary choice

The fundamental problem

$$demand = \beta \times price + ... + \varepsilon$$

$$\text{Demand} \rightleftarrows \text{Supply}$$

Price endogeneity

Research objective

Develop state-of-the-art **itinerary choice models** for **industry applications** that **correct for price endogeneity** using a database of **more than 10 million tickets** from the Airlines Reporting Corporation.

Overview

1. Motivation and research objective
2. **Data and methodology**
3. Results
4. Conclusions and limitation
5. New research objective
6. Preliminary results
7. Next steps

Data

- ARC is a ticketing clearinghouse that maintains financial transactions for all tickets purchased through travel agencies worldwide

Data

- ARC is a ticketing clearinghouse that maintains financial transactions for all tickets purchased through travel agencies worldwide
  - May 2013 departures
  - Continental U.S. markets
  - Simple one-way and round-trip tickets with at most 2 connections
  - Eliminate tickets with fares $<$ \$50 or in top 0.1%

Data

- ARC is a ticketing clearinghouse that maintains financial transactions for all tickets purchased through travel agencies worldwide
  - May 2013 departures
  - Continental U.S. markets
  - Simple one-way and round-trip tickets with at most 2 connections
  - Eliminate tickets with fares $<$ \$50 or in top 0.1%
- 3,265,545 directional itineraries, representing 10,034,935 passenger trips

Explanatory variables

Carrier characteristics

- Carrier preferences
- Marketing relationships: online, code-share, interline

Explanatory variables

Carrier characteristics

- Carrier preferences
- Marketing relationships: online, code-share, interline

Itinerary characteristics

- Elapsed time
- Number of connections
- Direct flight indicator
- Equipment type
- Price
- Departure time of day

Price

- ARC database = **ticket-level price information** linked to specific itineraries and the time of purchase
- Average price by:
  - Product type (High yield, Low yield)
  - Advance purchase period (0-6 days, 7-20 days, 21 days or +)
  - Origin & Destination
  - Carrier
  - Level-of-service (NS, 1 CNX, 2 CNXS)
- Outbound (or inbound) price = total price/2
- Include taxes

Departure time of day

1. Departure time preferences vary by:
   - Length of haul
   - Direction of travel
   - Number of time zones
   - Day of week
   - Itinerary type (OW, OB, IB)
2. Continuous time of day preferences formulation is preferred over discrete formulation to avoid counter-intuitive forecasts

10 time of day classifications

Same time zone, $\leq$ 600 miles

Same time zone, $>$ 600 miles

1 time zone WB, $\leq$ 600 miles

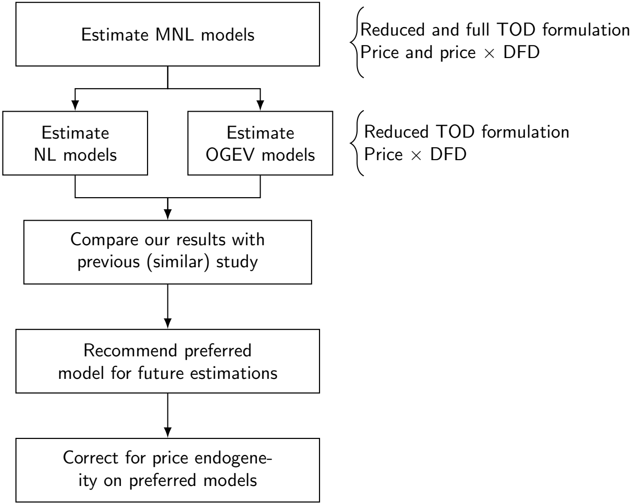
1 time zone WB, $>$ 600 miles

For each classification, we estimate separate time of day preferences for **outbound**, **inbound** and **one-way** itineraries and **day of week**

Define choice sets

- Construct **choice sets** for **each OD city pair that departs on day of week** d
- Create a **representative weekly schedule** as the Monday after the 9^{th} of the month [May 13 - May 19]
- Define a unique itinerary by org_1,$dest_1$,$opcarr_1$,$opfltnum_1$,$deptdow_1$ for legs $l = 1, 2, 3$
- Map all demand to representative schedule/unique itinerary
- Eliminate choice sets with demand $<$ 30 pax/month

Overview of modeling approach



Two-stage control-function (2SCF) method

- Stage 1: Estimate price by ordinary-least-square (OLS)

$$p_{ni} = \alpha_1 IV_{ni}^1 + ... + \alpha_k IV_{ni}^k + \boldsymbol{\gamma}_i' \mathbf{x}_{ni} + \mu_{ni} \quad (1)$$

- Stage 2: Estimate the choice model using the residuals δ from first stage

$$U_{ni} = \beta_{\widehat{\delta}} \widehat{\delta}_{ni} + \beta_p p_{ni} + \boldsymbol{\beta}_i' \mathbf{x}_{ni} + \varepsilon_{ni} \quad (2)$$

- Test: Estimate the choice model using the residuals δ from first stage and one instrument

$$U_{ni} = \beta_{\widehat{\delta}} \widehat{\delta}_{ni} + \beta_p p_{ni} + \boldsymbol{\beta}_i' \mathbf{x}_{ni} + \varepsilon_{ni} + \alpha_1 IV_{ni}^1 \quad (3)$$

Overview

Part I: Modeling Competition among Airline Itineraries

1. Motivation and research objective
2. Data and methodology
3. **Results**
4. Conclusions and limitation
5. New research objective
6. Preliminary results
7. Next steps

Departure time preferences

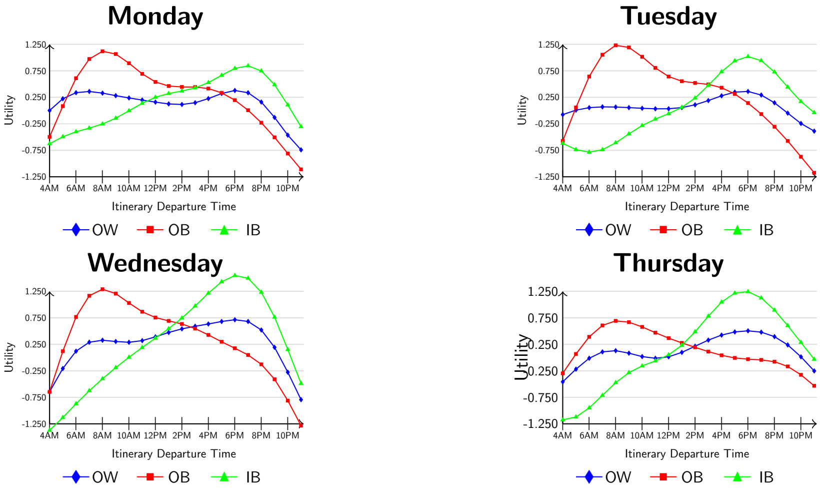


Figure: One TZ WB, distances $\leq$ 600 mi.

Departure time preferences

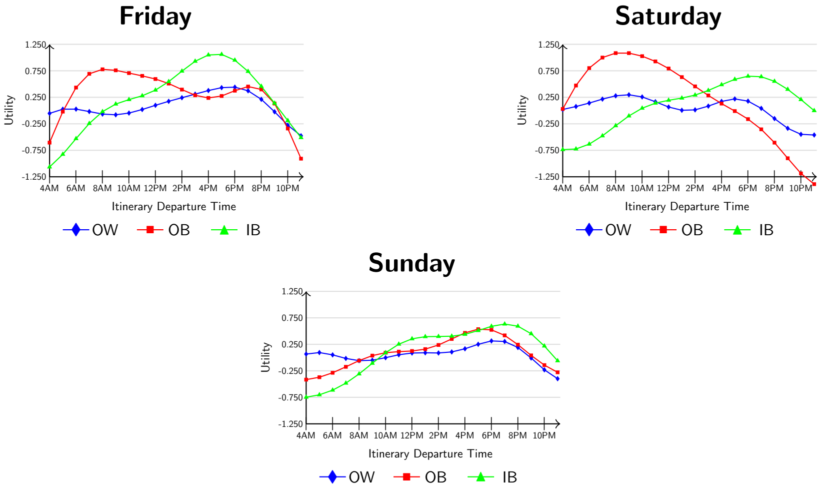


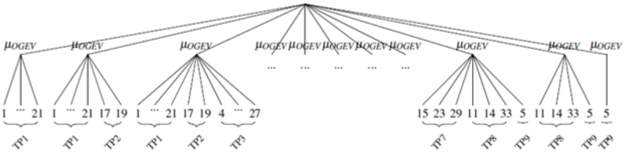
Figure: One TZ WB, distances $\leq$ 600 mi.

Main findings from MNL models

- NS $\succ$ 1 CNX $\succ$ 2 CNXS
- Cheap itineraries $\succ$ More expensive itineraries
- Short itineraries $\succ$ Longer itineraries
- Large aircraft $\succ$ Smaller aircraft
- Stability of MNL model results across time of day and price specifications

Competition among airline itineraries

- Models show strong inter-alternative competition by carrier and departure time.
  - Best NL model: by TOD and Carrier
  - Best OGEV model: Three Allocation (hourly)
- Results based on 2013 data are strikingly similar to those based on 2000 data.



Price elasticities

	Mean fare	Uncorrected	Corrected
Low yield products	240.20	-0.8976	-1.2567
High yield products	343.60	-0.5877	-0.8307

Value of Time Results

	Best MNL 6 TOD, price x DFD		Best NL TOD and Carrier 6 TOD, price x DFD		Best OGEV Three Alloc. hourly 6 TOD, price x DFD	
	uncorrected	*corrected*	*uncorrected*	*corrected*	*uncorrected*	*corrected*
VOT LY						
0-6 DFD	$73	**$43**	$73	**$42**	$71	**$42**
7-20 DFD	$57	**$36**	$56	**$35**	$55	**$35**
21+ DFD	$46	**$31**	$45	**$30**	$44	**$30**
VOT HY						
0-20 DFD	$161	**$84**	$157	**$81**	$157	**$84**
21+ DFD	$72	**$50**	$71	**$49**	$73	**$51**

Overview

1. Motivation and research objective
2. Data and methodology
3. Results
4. **Conclusions and limitation**
5. New research objective
6. Preliminary results
7. Next steps

Conclusions

- Importance to correct for price endogeneity
  - Over-estimation of customer's value of time
  - Biased price elasticities
  - Sub-optimal business decisions
- Highly refined departure time of day preferences
- Strong correlation across itineraries that share similar departure times
- Similar value of time estimates for MNL, NL, and OGEV models
- Similar results for models based on 2013 data and on 2000 data

Limitation

Choice set generation

- Each OD city pair that departs on day of week d

Limitation

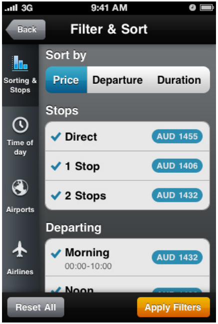
Choice set generation

- Each OD city pair that departs on day of week d
- **Up to 156 itineraries**

Limitation

Choice set generation

- Each OD city pair that departs on day of week d
- **Up to 156 itineraries**

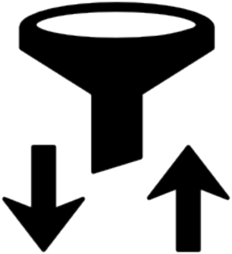


Overview

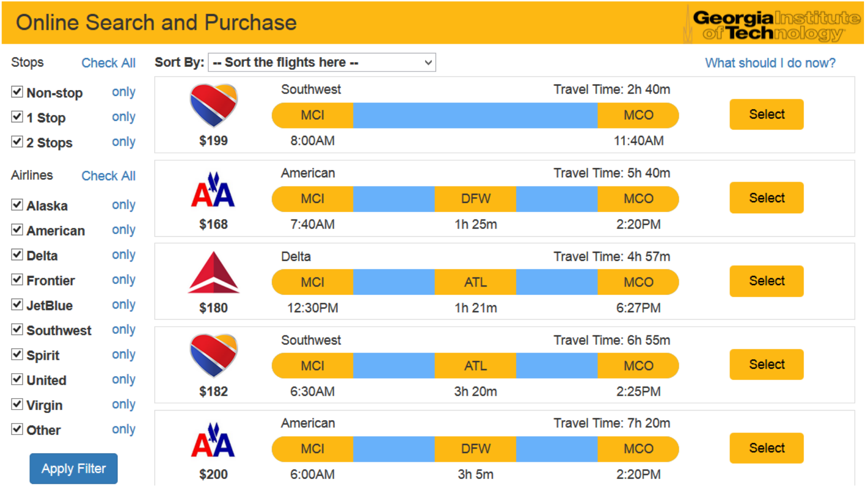
1. Motivation and research objective
2. Data and methodology
3. Results
4. Conclusions and limitation
5. **New research objective**
6. Preliminary results
7. Next steps

Ultimate objective

Develop a choice set generation model for itinerary choice models that incorporates sorting and filtering actions using an interactive online survey



Sorting and filtering actions



Intermediate objective

Determine if lower-cost crowdsourcing worksites provide similar results as more traditional survey panels

1. Amazon Mechanical Turk (AMT) is an online outsourcing platform with more than 500,000 workers in 190+ countries that perform microtasks , typically for $0.10 USD or less

amazon mechanical turk™
Artificial Artificial Intelligence

2. Qualtrics is a more traditional marketing firm that maintains a panel of respondents that complete surveys for a variety of clients

qualtrics

Overview

Comparison of AMT and Qualtrics

	AMT	Qualtrics
Number of respondents	690	553
-High yield respondents	62	62
-Low yield respondents	628	491
Data collection period	Oct-Nov 2016	March 2017
Total survey cost	$305.25	$3,835
Participant reimbursement	$0.25 regular workers $1.00 master workers	$0.65

Comparison of AMT and Qualtrics

Prior trip characteristics	Result
How often do you make air trips?	Same
When did you make this trip?	Same
Who paid for your ticket?	Same
How long before your trip did you purchase your ticket?	Same
What was the primary reason you flew?	Same
What day of the week did you depart?	Same
How many nights were you away?	Same
How many people travelled together?	Different
Traveler itinerary preferences	
I only fly certain airlines	Different
I generally shop for the cheapest flight	Different
I avoid small propeller and regional jet aircraft	Different
Travel times are more important to me than price	Different
Travel times are more important to me than carrier	Different
Price is more important to me than carrier	Different
Socio demographics	
What is your gender?	Different
What is your age?	Different
How many people are in your household?	Different
What was your annual household income last year?	Different

MNL models

- Similar preferences regarding:
  - travel times
  - number of connections
  - elapsed time

- Different preferences regarding price

	All	MTurk	Qualtrics
Ages 18-24	11.54	10.63	13.11
Ages 25-64	16.17	14.41	19.73
Ages 65+	23.61	46.09	22.57

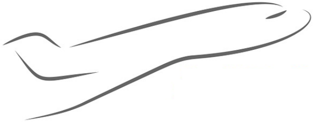
Overview

1. Motivation and research objective
2. Data and methodology
3. Results
4. Conclusions and limitation
5. New research objective
6. Preliminary results
7. **Next steps**

Next steps

- Analyze if AMT and Qualtrics respondents have the same behavior regarding the use of search and filter tools
- Develop a choice set generation model and compare results for AMT and Qualtrics
- Incorporate the developed choice set generation model into our state-of-the-art itinerary choice models
- Compare results from stated preferences data to the ones obtained using revealed preference data
- Estimate models using mixed data

Q&As



"Running an airline is like having a baby:
fun to conceive, but hell to deliver."

- C. E. Woolman

Thank you!

virginie.lurkin@pefl.ch