

Demand Estimation Report

Airlines Merger Simulation

August 24, 2025

1 Introduction

This report presents the results of the demand estimation analysis conducted as part of the Airlines Merger Simulation project. The tables and figures included here summarize the key findings.

We estimate logit and two-level nested-logit demand. In our application, the *top (group) level* partitions alternatives into **Outside** vs. **Inside** the air-travel market (the outside option is no trip / other modes). Within the **Inside** group, the *subgroup level* partitions products into **Legacy** vs. **Low-Cost Carrier (LCC)**.

2 Model and Nesting Structure

Let t index markets (e.g., OD-quarter). Products j belong to a subgroup $h \in \{\text{Legacy}, \text{Non-Legacy}\}$ (or LCC) nested inside a group $g \in \{\text{Outside (not flying), Inside carrier}\}$; by convention the outside good $j = 0$ belongs to $g = \text{Outside}$.

Utility (two-level nested logit)

For consumer i in market t ,

$$u_{ijt} = \delta_{jt} + (1 - \sigma_1) \varepsilon_{ijt} + (\sigma_1 - \sigma_2) \zeta_{iht} + \sigma_2 \eta_{igt}, \quad 0 \leq \sigma_2 \leq \sigma_1 < 1, \quad (1)$$

$$\delta_{jt} = \mathbf{x}'_{jt} \beta + \alpha p_{jt} + \mu_t + \xi_{jt}. \quad (2)$$

Here $\varepsilon_{ijt}, \zeta_{iht}, \eta_{igt}$ are independent Type I EV shocks generating correlation within subgroups (h) and groups (g). We follow the “ $+\alpha p$ ” convention in (2); economically, $\alpha < 0$ is expected.

Define the scale parameters $\lambda_1 \equiv 1 - \sigma_1$ (within-subgroup scale) and $\lambda_2 \equiv 1 - \sigma_2$ (within-group scale), with $0 < \lambda_2 \leq \lambda_1 \leq 1$.

Estimating equation (two-level)

Using the standard nested-logit inversion,

$$\log s_{jt} - \log s_{0t} = \mathbf{x}'_{jt} \beta + \alpha p_{jt} + \sigma_2 \log s_{j|h,t} + \sigma_1 \log s_{h|g,t} + \mu_t + \xi_{jt}. \quad (3)$$

In the airline application:

- *Group g* : {Outside, Inside} (inside vs. outside).
- *Subgroups h within $g = \text{Inside}$* : {Legacy, Non-Legacy} (or LCC).
- For $g = \text{Outside}$, the outside good forms a degenerate subgroup.

Special cases

- **One-level nested logit** (collapse the subgroup level): set $\sigma_1 = 0$ (so $\lambda_1 = 1$). Then

$$\log s_{jt} - \log s_{0t} = \mathbf{x}'_{jt} \beta + \alpha p_{jt} + \sigma_2 \log s_{j|h,t} + \mu_t + \xi_{jt},$$

where now h is the single (within-inside) nest (e.g., carrier-based).

- **Multinomial logit** (no within-group correlation): set $\sigma_1 = \sigma_2 = 0$.

Regressors and instruments

We take p_{jt} as average fare and \mathbf{x}_{jt} to include share nonstop, (thousand-mile) distance and its square, and $\log(1 + \text{rival carriers})$. We use standard rival-based shifters as instruments for p_{jt} (and, if needed, for the nesting terms): average rivals' number of markets served, average rivals' route distance, and the number of rival carriers in the market-time cell.

3 DB1B Data & Cleaning Summary

- **Source:** DOT DB1B “Market” files (10% ticket survey). Auxiliaries: city–market lookup (vacation flag), airport metadata (T_MASTER_CORD), airport–year populations, slot-controlled airports, carrier–airport hub table, and annual CPI index (2008=100; `cpi_index.csv`).
- **Coverage:** Years 2005–2019; quarters 1–4; CONUS only (U.S. excluding PR, VI, TT, HI, AK), filtered via T_MASTER_CORD.
- **Market definition:** City-pair O–D level using `origincitymarketid` and `destcitymarketid`; time unit is a calendar quarter.
- **Product definition:** O–D–tkcarrier–quarter. Monopoly O–D–quarter cells dropped (require ≥ 2 carriers).
- **Core filters:**
 - No interline tickets: `tkcarrierchange == 0`.
 - Real fare bounds: keep if real market fare $\in [25, 2500]$ USD, where real fare `mktfare_real` is `mktfare` deflated using annual CPI (2008=100).
 - Demand size screen at O–D–quarter: keep if total passengers $\geq 20 \times 365 / (4 \times 0.10) \approx 183$ in DB1B (i.e., at least 20 pax/day in the full universe, adjusted for the 10% sample).
- **Other variables:** CONUS filter and airport names/cities from T_MASTER_CORD (origin/destination); slot-controlled airport indicators at origin and destination; destination “vacation” flag from city lookup; airport–year populations for origin and destination (nonpositive values set to missing); carrier–airport hub indicators (wide \rightarrow long reshape).
- **Aggregation (passenger-weighted) to O–D–carrier–quarter:**
 - Average distance, nonstop miles, and fare; share of nonstop; extra miles = distance – nonstop miles.
 - Carrier presence at origin (origin–carrier passenger share), # destinations from the origin, and # markets served.
 - Rival statistics at the O–D–quarter: average rival presence, # markets, # destinations, and average rival distance.
- **Carriers:** Flags for carrier types (e.g., major, low-cost, legacy) created in `carrier_flags.do`; carrier code is `tkcarrier` (IATA 2-letter).
- **Output:** Final panel collapsed to O–D–carrier–quarter, with string IDs for market (`origin-destination`) and time (`year-quarter`); written to `data/processed/combined/airline_data.csv`.

Variable Summary Statistics

Lower level refers to lower nest define as *legacy vs. non-legacy carriers* in the table below.

Table 1: Summary statistics: variables used in demand estimation

	mean	sd	min	p25	p50	p75	max
Market share	0.001	0.006	0.0000	0.0001	0.000	0.001	0.473
Outside share	0.992	0.016	0.3365	0.9908	0.996	0.998	1.000
Inside share sum	0.008	0.016	0.0002	0.0024	0.004	0.009	0.664
Nest share (Upper level)	0.008	0.016	0.0002	0.0024	0.004	0.009	0.664
Nest share (Lower level)	0.005	0.011	0.0000	0.0014	0.003	0.005	0.607
Number of rival carriers	5.235	2.250	1.0000	4.0000	5.000	7.000	22.000
Number of destinations served	3583.900	1461.539	1.0000	3038.0000	4067.000	4661.000	5751.000
Average fare (dollars)	224.884	92.992	25.0000	171.7022	212.650	260.195	2492.016
Share nonstop flights	0.202	0.362	0.0000	0.0000	0.000	0.167	1.000
Average Distance (000s miles)	1.494	0.860	0.0670	0.8738	1.297	1.998	10.345
Average Distance sqr (000s miles)	2.971	3.733	0.0045	0.7635	1.683	3.992	107.019
Log(1 + fringe carriers)	0.350	0.455	0.0000	0.0000	0.000	0.693	2.565
Average distance to rival markets	1493.539	830.698	67.0000	905.1562	1287.021	1982.586	5996.580
Average number of rival destinations	52.183	26.142	1.0000	33.0000	51.000	70.000	153.000
Number of rival carriers	5.235	2.250	1.0000	4.0000	5.000	7.000	22.000
Number of rival carriers lower nest	2.969	1.601	0.0000	2.0000	3.000	4.000	15.000
Observations	1636916						

4 Demand Estimation Results

Column (1) presents the results of the logit model without instruments, while column (2) shows the logit model results with price instruments. Column (3) shows the nested-logit model without instruments, and column (4) presents the nested-logit model with instruments for both prices and nest shares. The nests are defined as inside goods (carriers that are not the outside good) and outside good. All columns include fixed effects for origin-destination markets. Robust standard errors are reported in parentheses below the coefficients.

First, we present the basic demand estimation results. Column (1) presents the results of the logit model without instruments, while column (2) shows the logit model results with price instruments. Column (3) shows the nested-logit model without instruments, and column (4) presents the nested-logit model with one level and instruments for both prices and nest shares.

Table 2: Demand Estimates (Logit and Nested-Logit)

	(1) $\ln s_{jt}/s_{0t}$	(2) $\ln s_{jt}/s_{0t}$	(3) $\ln s_{jt}/s_{0t}$	(4) $\ln s_{jt}/s_{0t}$
Average fare (dollars)	-0.0017*** (0.0000)	0.0102*** (0.0002)	-0.0005*** (0.0000)	-0.0019*** (0.0001)
Share nonstop flights	1.9504*** (0.0047)	2.0237*** (0.0057)	0.1189*** (0.0009)	0.3048*** (0.0032)
Average Distance (000s miles)	-4.1437*** (0.0139)	-4.5780*** (0.0209)	-0.0469*** (0.0024)	-0.4330*** (0.0091)
Average Distance sqr (000s miles)	0.2762*** (0.0033)	0.2153*** (0.0049)	0.0043*** (0.0005)	0.0398*** (0.0012)
Log(1 + fringe carriers)	-0.5032*** (0.0033)	-0.4860*** (0.0041)	0.0243*** (0.0005)	-0.0335*** (0.0010)
$\ln s_{jt g}$			0.9776*** (0.0002)	0.8740*** (0.0014)
Observations	1636916	1636916	1636916	1636916
Adjusted R^2	0.503	0.192	0.983	0.972
F-statistic (IV)		5853.3607		1575.4110

4.1 Full model, two-level nesting as legacy vs. non-legacy carriers

Below, we present the full model results including two levels of nesting and instruments for both prices and nest shares. The first column shows only one level of nesting (inside vs. outside), while the second column shows two levels of nesting (inside vs. outside, and legacy vs. within inside). All columns include fixed effects for origin-destination markets. Robust standard errors are reported in parentheses below the coefficients.

The instruments used are average rival distance, average number of markets served by rivals, the number of rival carriers and the number of rival carriers for the nest definition in the market.

Table 3: Demand Estimates (Logit and Nested-Logit)

	(1) $\ln s_{jt}/s_{0t}$	(2) $\ln s_{jt}/s_{0t}$
Average fare (dollars)	-0.0019*** (0.0001)	-0.0021*** (0.0001)
Share nonstop flights	0.3048*** (0.0032)	0.2986*** (0.0050)
Average Distance (000s miles)	-0.4330*** (0.0091)	-0.4182*** (0.0135)
Average Distance sqr (000s miles)	0.0398*** (0.0012)	0.0398*** (0.0013)
Log(1 + fringe carriers)	-0.0335*** (0.0010)	-0.0322*** (0.0013)
$\ln s_{jt g}$	0.8740*** (0.0014)	
$\ln s_{ht g}$		0.8866*** (0.0047)
$\ln s_{jt h}$		0.8753*** (0.0019)
Observations	1636916	1636916
Adjusted R^2	0.972	0.972
F-statistic (IV)	1575.4110	361.5578

4.2 Full model, two-level nesting as low-cost vs. non low-cost carriers

Now I show results where the lower (second) nest is low-cost vs. non low-cost carriers.

Table 4: Demand Estimates (Logit and Nested-Logit)

	(1) $\ln s_{jt}/s_{0t}$	(2) $\ln s_{jt}/s_{0t}$
Average fare (dollars)	-0.0019*** (0.0001)	-0.0025*** (0.0001)
Share nonstop flights	0.3048*** (0.0032)	0.2799*** (0.0049)
Average Distance (000s miles)	-0.4330*** (0.0091)	-0.3651*** (0.0139)
Average Distance sqr (000s miles)	0.0398*** (0.0012)	0.0389*** (0.0014)
Log(1 + fringe carriers)	-0.0335*** (0.0010)	-0.0311*** (0.0011)
$\ln s_{jt g}$	0.8740*** (0.0014)	
$\ln s_{ht g}$		0.9074*** (0.0043)
$\ln s_{jt h}$		0.8832*** (0.0020)
Observations	1636916	1636916
Adjusted R^2	0.972	0.970
F-statistic (IV)	1575.4110	396.5416

5 Elasticities

These are observation level elasticities calculated using the parameters from the two-level nested logit model with legacy vs. non-legacy carriers as the second level nest and instruments for both prices and nest shares.

Table 5: Elasticities: Own-price (Logit and Nested-Logit)

	count	mean	sd	min	p25	p50	p75	max
Logit Elasticity OLS	1636916	-0.386	0.160	-4.283	-0.447	-0.365	-0.295	-0.043
Nested-Logit Elasticity (1 Level)	1636916	-2.534	1.640	-38.207	-3.282	-2.460	-1.509	-0.048
Nested-Logit Elasticity (2 Levels)	1636916	-3.400	1.758	-45.444	-4.074	-3.205	-2.428	-0.085
Observations	1636916							

When low cost vs. non low-cost carriers is used as the second level nest, the elasticities are as follows (third row).

Table 6: Elasticities: Own-price (Logit and Nested-Logit)

	count	mean	sd	min	p25	p50	p75	max
Logit Elasticity OLS	1636916	-0.386	0.160	-4.283	-0.447	-0.365	-0.295	-0.043
Nested-Logit Elasticity (1 Level)	1636916	-2.580	1.645	-38.210	-3.325	-2.510	-1.586	-0.051
Nested-Logit Elasticity (2 Levels)	1636916	-4.869	2.596	-66.267	-5.891	-4.576	-3.405	-0.103
Observations	1636916							

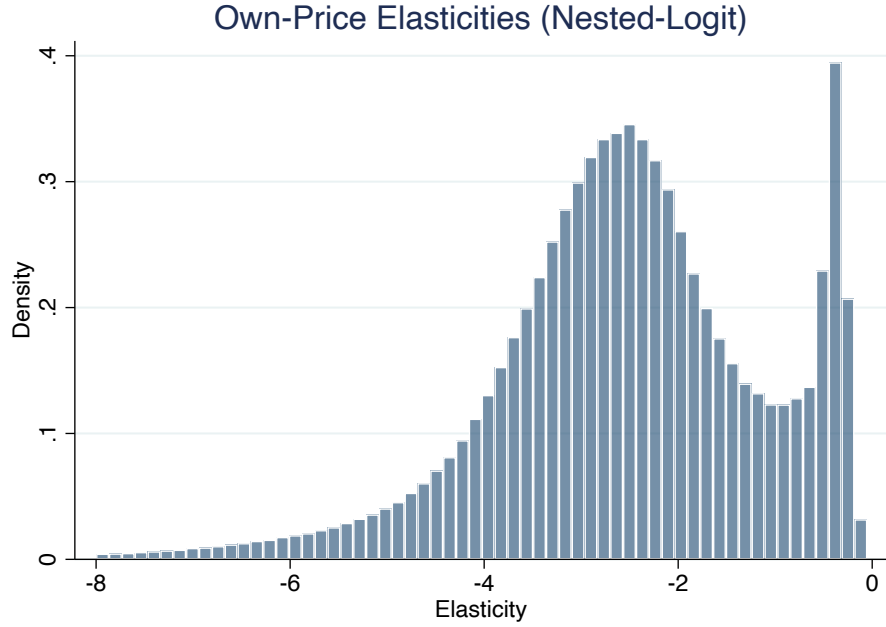


Figure 1: Histogram of Elasticities, one-level nest

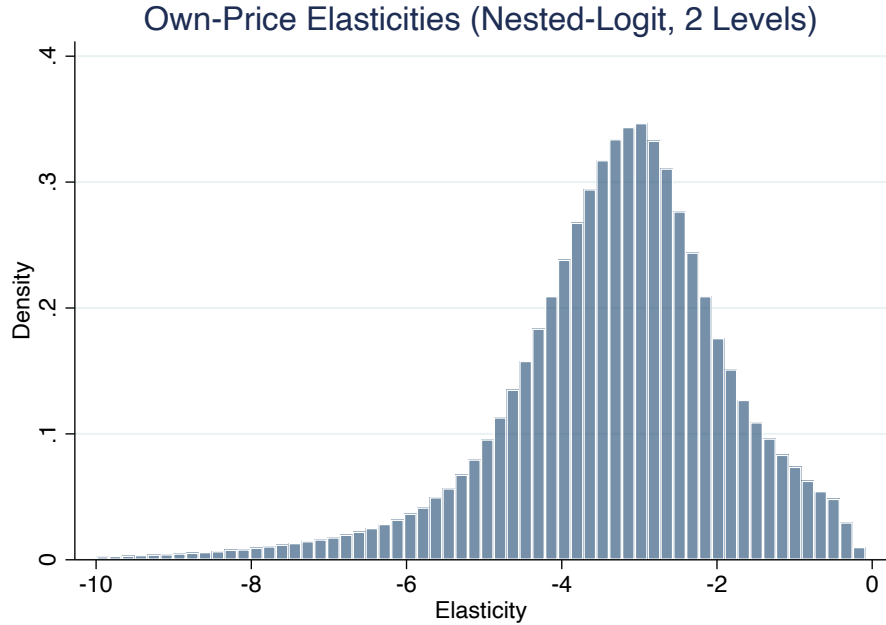


Figure 2: Histogram of Elasticities, two-level nest (legacy vs. non-legacy, inside vs. outside)

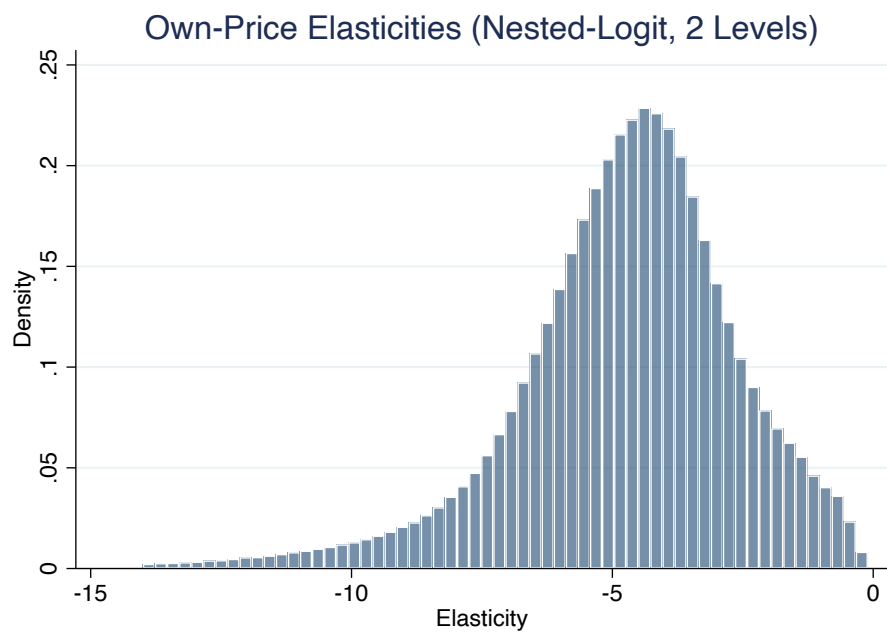


Figure 3: Histogram of Elasticities, two-level nest (low-cost vs. non low-cost, inside vs. outside)

6 Appendix: First Stage Results

Table 7: First Stage Results for Fare

	(1) Average fare (dollars)
Average distance to rival markets	0.0359*** (0.0008)
Average number of rival destinations	-0.8452*** (0.0088)
Number of rival carriers	-10.9626*** (0.0668)
Number of rival carriers lower nest	7.8044*** (0.0433)
Share nonstop flights	-1.9907*** (0.3057)
Average Distance (000s miles)	44.5059*** (1.8443)
Average Distance sqr (000s miles)	4.5607*** (0.4781)
Log(1 + fringe carriers)	3.0140*** (0.2175)
Observations	1636916
Standard errors in parentheses	
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$	

Table 8: First Stage Results for Legacy First Level Nest Share Variable

	(1) $\ln s_{ht g}$
Average distance to rival markets	0.0001*** (0.0000)
Average number of rival destinations	-0.0044*** (0.0001)
Number of rival carriers	-0.2495*** (0.0007)
Number of rival carriers lower nest	0.3341*** (0.0007)
Share nonstop flights	0.3309*** (0.0020)
Average Distance (000s miles)	-0.4301*** (0.0077)
Average Distance sqr (000s miles)	0.0518*** (0.0017)
Log(1 + fringe carriers)	0.0073*** (0.0017)
Observations	1636916
Standard errors in parentheses	
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$	

Below are the first stage for the second level nest share variable for legacy nests and low-cost nests.

Table 9: First Stage Results for Legacy Nest Share Variable (Second Level)

	(1) $\ln s_{jt h}$
Average distance to rival markets	0.0014*** (0.0000)
Average number of rival destinations	-0.0167*** (0.0002)
Number of rival carriers	-0.0430*** (0.0012)
Number of rival carriers lower nest	-0.3756*** (0.0009)
Share nonstop flights	1.4283*** (0.0045)
Average Distance (000s miles)	-3.4366*** (0.0126)
Average Distance sqr (000s miles)	0.1948*** (0.0027)
Log(1 + fringe carriers)	-0.1177*** (0.0039)
Observations	1636916

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: First Stage Results for Low-Cost Nest Share Variable (Second Level)

	(1) $\ln s_{jt h}$
Average distance to rival markets	0.0015*** (0.0000)
Average number of rival destinations	-0.0153*** (0.0002)
Number of rival carriers	-0.0293*** (0.0012)
Number of rival carriers lower nest	-0.3489*** (0.0009)
Share nonstop flights	1.5496*** (0.0046)
Average Distance (000s miles)	-3.4022*** (0.0136)
Average Distance sqr (000s miles)	0.1871*** (0.0031)
Log(1 + fringe carriers)	-0.0541*** (0.0039)
Observations	1636916

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$