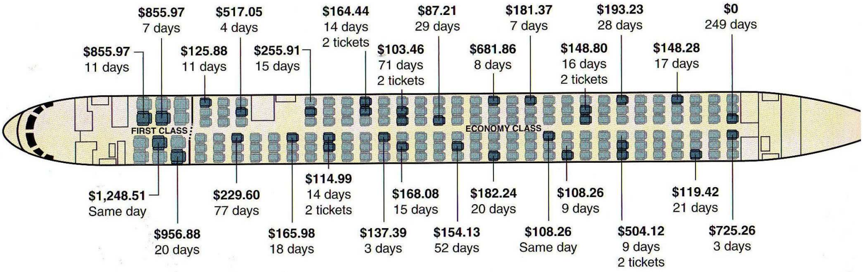
Demand Shifting across Flights and Airports in a Spatial Competition Model

Diego Escobari   Sang-Yeob Lee

November, 2010

Outline

Price dispersion in airlines



33 customers paid 27 different fares *(New York Times)*

- Borenstein and Rose (JPE, 1994): 36% difference.
- We focus on the effect of competition between flights and airports.

Contribution and Intuition

- Spatial competition similar to gas stations.
- Data collection and the econometrics control for other sources of price dispersion.
- Focus on the pricing interaction between flights and across airports.
- First empirical paper to analyze day-to-day pricing competition between flights.
- Estimates price reaction functions using spatial autoregressive models.
- The paper found that there is important demand shifting across flights.
- No demand shifting was found across New York City airports.

Some Related Literature

- Competition and price dispersion using DB1B:
  - Borenstein and Rose (1994)
  - Gerardi and Shapiro (2009)
- Day-to-day pricing using posted prices:
  - Stavins (2001)
  - Bilotkach (2006)
  - McAfee and te Velde (2007)
  - Escobari (2009)
  - Alderighi et al.(2011)

Dynamic pricing in airlines

Key characteristics:

- Perishable nature of airline tickets.
- Fixed capacity.
- Demand uncertainty.

Dynamic pricing in airlines

Airlines use dynamic pricing to:

- Price discriminate.
- Systematic and stochastic peak-load pricing.
- Cope with aggregate demand uncertainty and costly capacity.

Data

- Focus on one route: New York to Toronto.
- Three airports in New York: Newark Liberty International, John F. Kennedy and La Guardia.
- Lowest available fare and inventory levels from expedia.com
- A panel with 317 cross section observations and 15 time observations (collected every 3 days).
- One-way, non-stop, economy-class, departed between December 19 and December 24, 2008.
- American, Air Canada, Continental, Delta, Lan Chile and United.

Map of New York City airports

SAR and Price Reaction Functions

- SAR, across flights:

$$p_{it} = \lambda_1 W1 p_{it} + \mathbf{X}\beta + \eta_i + \varepsilon_{it}$$

- $W1$: Spatial weights of flights departing from the same airport.
- $\mathbf{X}$: Matrix of controls:
  - *DAYADV*: Days to departure.
  - *LOAD*: Load factor; capacity utilization.
- λ_1: SAR coefficient; demand shifting across flights.

SAR and Price Reaction Functions

- SAR, across airports:

$$p_{it} = \lambda_1 W1 p_{it} + \lambda_2 W2 p_{it} + \mathbf{X}\beta + \eta_i + \varepsilon_{it}$$

- $W2$: Spatial weights of flights departing from a different airport.
- λ_2: SAR coefficient; demand shifting across airports.

Market Definition and Spatial Weights

- Use matrices $W1$ and $W2$ to obtain weighted average prices of competing flights.
- Use the time difference between flights.
- Within a critical distance, each element in the matrix is defined as:

$$w_{ij} = \omega_{ij} / \sum_{j=1}^{N} \omega_{ij},$$

$$\omega_{ij} = 1/(1 + d_{ij}),$$

- d_{ij}: Distance between flight i and flight j.

Estimation and Interpretation

- IV following Lee (2003). Use instruments based on **X**.
- First interpretation: Usual.
- Second interpretation: Diversion ratio explained in Pinske et.al (2002) and used in Lee (2009).
- Assumptions:
  - Firms compete in a Bertrand-Nash simultaneous game.
- The price reaction function for each flight i:

$$p_i = \mathbf{X}_i \beta + \lambda \sum_{j \neq i}^{n} w_{ij} p_j + \mu_i$$

where, $w_{ij} = d_{ij} / \sum_{j \neq i}^{n} d_{ij}$ and $\lambda = 1/2 \sum_{j \neq i}^{n} d_{ij}$.
- $2 \times \lambda$: measures the fraction of sales lost by flight i when increasing its price.

Estimation

- Estimate via Kelejian and Prucha (1998) and Lee (2003).
- Two-step method.
- First step:
  - Estimate the SAR equation via 2SLS using $\mathrm{H} = [\mathrm{X}, \mathrm{WX}, \mathrm{W}^2\mathrm{X}]$ as instruments.
  - The estimator is $\hat{\theta}^{2SLS} = (\tilde{\mathrm{X}}'\mathrm{S}\tilde{\mathrm{X}})^{-1}\tilde{\mathrm{X}}'\mathrm{SP}$, where P is the vector of prices, $\tilde{\mathrm{X}} = [\mathrm{WP}, \mathrm{X}]$ is the matrix of explanatory variables, and $\mathrm{S} = \mathrm{H}(\mathrm{H}'\mathrm{H})^{-1}\mathrm{H}'$ is the weighting matrix.
- Second step:
  - Estimates an IV regression using instruments $\hat{\mathrm{Z}} = [\mathrm{X}, \mathrm{W}(\mathrm{I} - \hat{\lambda}^{2SLS})^{-1}\mathrm{X}\hat{\beta}^{2SLS}]$.
  - That is, $\hat{\theta}^{IV} = (\hat{\mathrm{Z}}'\tilde{\mathrm{X}})^{-1}\hat{\mathrm{Z}}'\mathrm{P}$.
  - This uses $\hat{\lambda}^{2SLS}$ and $\hat{\beta}^{2SLS}$ that are part of the vector $\hat{\theta}^{2SLS}$.

Flights by Carrier and Date

Table: Flights by Carrier and Date

	Fri Dec 19	Sat Dec 20	Sun Dec 21	Mon Dec 22	Tue Dec 23	Wed Dec 24	Total
American	12	8	11	12	13	11	67
Air Canada	19	9	13	20	20	16	97
Continental	8	5	7	8	8	6	42
Delta	4	4	4	4	4	3	23
Lan Chile	1	1	0	1	0	1	4
United	17	10	10	18	16	13	84
Total	61	37	45	63	61	50	317

Summary Statistics

Table: Summary statistics

Variables	Mean	Std. Dev.	Min.	Max.	Obs.
FARE	175.999	139.153	89.000	1265.000	4408
$W1p$	168.582	64.290	89.023	631.944	4408
$W2p$	170.826	59.690	35.277	479.224	4408
Location	3.462	2.241	0.267	5.861	4408
DAYADV	24.688	13.120	1.000	48.000	4408
LOAD	0.584	0.204	0.060	1.000	4408

Results, same airport and linear controls

Table: Results, same airport and linear controls

Variables	3 hours	6 hours	12 hours
W1p	0.309	0.407	0.557
	(13.53)	(15.52)	(18.00)
LOAD	92.916	89.553	88.352
	(4.08)	(3.91)	(3.90)
DAYADV	−2.847	−2.565	−2.346
	(−8.99)	(−8.51)	(−7.38)
Within R-squared	0.199	0.211	0.226
Observations	4398	4398	4398

Notes: Numbers in parentheses are t-statistics based on heteroskedasticity robust standard errors. All specifications include flight-specific effects.

Results, IV same airport and linear controls

Table: Results IV, same airport and linear controls

Variables	3 hours	6 hours	12 hours
$W1p$	0.546	0.580	0.934
	(6.81)	(7.60)	(15.93)
$LOAD$	157.336	170.642	176.890
	(8.92)	(10.42)	(14.06)
$DAYADV$	−2.304	−2.808	−3.650
	(−1.55)	(−1.98)	(−3.52)
Within R-squared	0.186	0.204	0.211
Observations	4398	4398	4398

Notes: Numbers in parentheses are t-statistics based on heteroskedasticity robust standard errors. All specifications control for flight-specific characteristics.

Results, IV same airport and nonlinear controls

Table: Results IV, same airport and nonlinear controls

Variables	3 hours	6 hours	12 hours
W1p	0.518	0.550	0.834
	(6.78)	(7.92)	(7.47)
LOAD	567.733	563.195	474.242
	(3.64)	(3.640	(3.00)
LOADSQ	−1431.52	−1414.62	−1174.66
	(−4.84)	(−4.80)	(−3.89)
LOADCU	1056.71	1158.10	972.18
	(6.59)	(6.62)	(5.34)
DAYADV	−9.257	−11.841	−12.085
	(−2.28)	(−2.95)	(−2.98)
DAYADVSQ	0.340	0.416	0.359
	(2.01)	(2.45)	(2.14)
DAYADVCU	−0.004	−0.005	−0.004
	(−1.86)	(−2.18)	(−1.75)
Within R-squared	0.229	0.249	0.250
Observations	4398	4398	4398

Results, across airports and nonlinear controls

Table: Results, across airports and nonlinear controls

Variables	3 hours	6 hours	12 hours
W1p	0.245	0.338	0.464
	(10.37)	(12.10)	(14.13)
W2p	0.018	0.005	−0.018
	(0.72)	(0.15)	(−0.63)
LOAD	300.562	288.778	234.743
	(1.75)	(1.69)	(1.38)
LOADSQ	−1292.45	−1288.24	−1143.76
	(−4.15)	(−4.12)	(−3.67)
LOADCU	1177.05	1177.98	1072.60
	(6.49)	(6.45)	(5.90)
DAYADV	−10.597	−13.043	−11.756
	(−2.93)	(−3.60)	(−3.27)
DAYADVSQ	0.348	0.437	0.381
	(2.14)	(2.68)	(2.35)
DAYADVCU	−0.004	−0.005	−0.004
	(−1.94)	(−2.31)	(−1.98)
Within R-squared	0.251	0.263	0.271
Observations	4398	4398	4398

Summary and Conclusions

- We investigate the nature of day-to-day competition between flights.
- We use instrumental variables methods and several spatial autoregressive models to estimate price reaction functions.
- The primary source of product differentiation is departure time.
- Controlling for other sources of price dispersion we find:
  - Flights respond to prices of competing flights.
  - Almost the entire demand shifting occurs within 3 hours of the departing flight.
  - No demand shifting was found across New York City airports.