Dynamic Price Discrimination in Airlines

Diego Escobari[†] Nicholas G. Rupp[‡] Joseph Meskey[‡]

[†]Department of Economics and Finance
The University of Texas - Pan American

[‡]Department of Economics
East Carolina University

EARIE
European Association for Research in Industrial Economics
Milan, Italy
August 2014

Outline

Motivation: Price dispersion

- Price dispersion (Borenstein and Rose, 1994).
- Price dispersion with panels (Gerardi and Shapiro, 2009)
- Role of costly capacity and demand uncertainty (with Gan, 2007); price competition (with Lee, 2012); price discrimination (with Jindapon, 2010); systematic peak-load pricing; demand learning; asymmetric pricing.
- Intertemporal price dispersion (Gaggero and Piga, 2011).
- Dynamic price dispersion (Mantin and Koo, 2009).
- Nonlinear pricing (Hernandez and Wiggins, 2010).
- Intertemporal price discrimination and dynamic adjustment to stochastic demand (Williams, 2013).

Motivation: Price discrimination

- Carriers exploit 'fences' such as:
  - Saturday-night-stayover.
  - Advance purchase discounts.
  - Minimum- and maximum-stay.
  - Refundable tickets.
  - Frequent flier miles.
  - Blackouts.
  - Volume discounts.
  - Fare classes (e.g. coach, first class)
  - Hour-of-day purchase.
- Airlines have the most sophisticated pricing systems in the world.

Contribution and Intuition

- Explains how sellers offer tickets at different prices at different hours of the day.
- Unique data on posted prices collected every hour.
- Conditions for price discrimination to exist:
  - Different valuation across consumer types.
  - Prevent arbitrage.
- Business travelers (high types) have higher valuations than tourists (low types).
- Incentive compatibility constraint: Prevent high types from behaving as low types.
- Airlines set higher prices during office hours.
- As the proportion of business travelers increases the difference increases.

Construction of the Data

- Lowest available fare (consistent with Deneckere and Peck, 2012).
- Pick a single day: Thursday, July 12, 2012.
- Initially follows Stavins (2001), but with two key improvements:
  - Panel.
  - Keep track of price changes by the hour.
- One-way, non-stop, economy-class.
  - Connecting passengers / sophisticated itineraries / legs.
  - Uncertainty in the return portion of the ticket.
  - Saturday-night-stayover / min- and max-stay.
  - Fare classes (e.g. coach, first class).
- 230 flights.
- 60 days prior to departure. (1422 observations per flight)
- American Airlines, Alaska, JetBlue, Delta, Frontier, AirTran Airways, United, US Airways, and Virgin America,

Summary Statistics

Table: Summary statistics

VARIABLES	(1) mean	(2) sd	(3) min	(4) max
FARE:				
Whole sample	375.3	223.0	58	1,288
$0 \leq$ ADV < 14.8	415.8	212.3	86.80	1,288
$14.8 \leq$ ADV < 29.6	358.2	226.0	58	968
$29.6 \leq$ ADV < 44.4	364.0	223.9	58	973
$44.4 \leq$ ADV < 59.3	370.1	223.7	58	968
ADV	29.91	16.64	0	59.3
DAYWEEK:				
MONDAY	0.184	0.387	0	1
TUESDAY	0.209	0.407	0	1
WEDNESDAY	0.206	0.405	0	1
THURSDAY	0.205	0.404	0	1
FRIDAY	0.196	0.397	0	1

Notes: The sample size is 145,458.

Static Pricing Equation

$$\text{LnFare}_{ijt} = \beta \text{Hour}_t + \sum_{k=0}^{59} \gamma_k \text{I}_{[\text{Adv}_t = k]} + v_{ij} + \varepsilon_{ijt},$$

i: flight; j: route; t: time.

- Day in advance fixed effects.
- v_{ij} controls for:
  - Time-invariant flight-, carrier-, and airport-pair-specific characteristics.
  - Flight-specific characteristics: Aircraft size, departure time, hub effect at the flight level, and systematic peak-load pricing.
  - Carrier-specific characteristics: Managerial capacity.
  - Airport-pair-specific characteristics: Herfindal index, distance between airport pairs, and hub effect.

Table: Static Pricing Equation

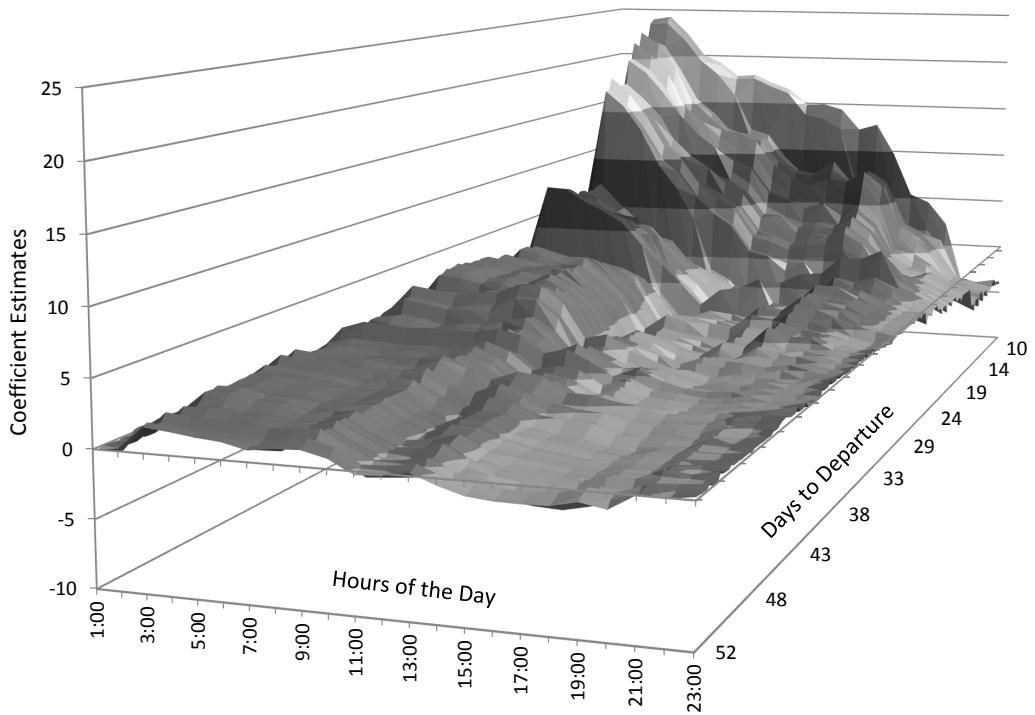
	(1)		(2)		(3)		(4)	
Variables	Coef	StErr	Coef	StErr	Coef	StErr	Coef	StErr
2:00	3.341*	(0.949)	3.319*	(0.971)	3.333*	(0.955)	3.589*	(1.008)
3:00	9.775*	(2.938)	9.819*	(2.937)	9.768†	(2.925)	10.06*	(2.874)
4:00	9.265*	(3.275)	9.360*	(3.274)	9.260†	(3.263)	9.792†	(3.278)
5:00	8.392†	(3.296)	8.443†	(3.301)	8.381†	(3.282)	8.555†	(3.223)
6:00	7.993*	(2.939)	8.054*	(2.921)	7.978†	(2.929)	7.803†	(2.847)
7:00	8.437*	(2.855)	8.478*	(2.849)	8.415†	(2.847)	7.825†	(2.938)
8:00	8.284*	(2.662)	8.307*	(2.653)	8.280†	(2.653)	7.223†	(2.788)
9:00	7.101*	(2.665)	7.173*	(2.653)	7.093†	(2.653)	5.964‡	(2.803)
10:00	5.428†	(2.118)	5.485*	(2.124)	5.424†	(2.106)	4.217‡	(2.142)
11:00	3.278†	(1.344)	3.338†	(1.349)	3.268†	(1.335)	1.890	(1.364)
12:00	3.376†	(1.474)	3.458†	(1.478)	3.361‡	(1.463)	1.706	(1.350)
13:00	2.935†	(1.396)	2.959†	(1.392)	2.889‡	(1.389)	1.137	(1.156)
14:00	1.749	(1.566)	1.748	(1.572)	1.750	(1.553)	0.130	(1.260)
15:00	0.118	(1.808)	0.147	(1.803)	0.126	(1.799)	-1.549	(1.454)
16:00	0.782	(1.688)	0.776	(1.692)	0.794	(1.677)	-1.060	(1.511)
17:00	0.885	(1.590)	0.907	(1.584)	0.883	(1.579)	-1.218	(1.362)
18:00	0.514	(1.461)	0.543	(1.466)	0.512	(1.453)	-2.136	(1.419)
19:00	0.534	(1.448)	0.562	(1.447)	0.530	(1.438)	-2.403	(1.381)
20:00	0.367	(1.198)	0.364	(1.202)	0.357	(1.191)	-2.992‡	(1.448)
21:00	0.0819	(0.787)	0.106	(0.805)	0.0952	(0.791)	-3.717†	(1.359)
22:00	-0.584	(0.957)	-0.578	(0.970)	-0.583	(0.961)	-3.997†	(1.670)
23:00	-1.583*	(0.520)	-1.581*	(0.535)	-1.578†	(0.519)	-3.797*	(1.124)
24:00	-0.901*	(0.279)	-0.891*	(0.276)	-0.893†	(0.279)	-1.183*	(0.264)
Monday_t	-9.826*	(3.655)	-9.815*	(3.667)	-9.809†	(3.652)		
Tuesday_t	-18.02*	(3.525)	-18.00*	(3.539)	-18.02*	(3.530)		
Wednesday_t	-34.31*	(6.441)	-34.32*	(6.471)	-34.31*	(6.451)		
Thursday_t	-34.51*	(8.410)	-34.51*	(8.426)	-34.55*	(8.381)		
Adv_t	-45.04*	(9.152)	-45.02*	(9.158)	-45.02*	(9.176)		
$\text{Adv}_t{}^2$	1.249*	(0.255)	1.248*	(0.255)	1.248*	(0.255)		
$\text{Adv}_t{}^3$	-0.0106*	(0.00229)	-0.0106*	(0.00229)	-0.0106*	(0.00229)		
Route FE	No		Yes		Yes		Yes	
Flight FE	No		No		Yes		Yes	
Adv FE	No		No		No		Yes	
Within R^2					0.308		0.321	

Table: Static Pricing Equation: Role of ADV_t.

Sample:	(1) $0 \leq \text{ADV} < 14.8$		(2) $14.8 \leq \text{ADV} < 29.6$		(3) $29.6 \leq \text{ADV} < 44.4$		(4) $44.4 \leq \text{ADV} < 59.3$	
Variables	Coef	StErr	Coef	StErr	Coef	StErr	Coef	StErr
2:00	7.518†	(2.930)	1.992	(1.503)	1.613†	(0.665)	0.205	(0.620)
3:00	32.78*	(9.450)	4.375	(3.147)	1.924†	(0.790)	2.181	(1.326)
4:00	33.31†	(10.19)	3.892	(3.589)	1.603	(0.924)	2.395	(1.328)
5:00	29.72†	(9.059)	3.309	(4.070)	0.748	(0.810)	2.382	(1.393)
6:00	26.11†	(8.007)	2.611	(4.255)	0.950	(0.767)	2.178‡	(1.102)
7:00	23.55†	(7.171)	2.984	(3.880)	1.123	(0.693)	1.873	(1.176)
8:00	20.28†	(6.258)	2.910	(3.372)	0.945	(0.861)	2.455‡	(1.162)
9:00	18.44†	(5.985)	1.638	(3.408)	1.734	(1.214)	2.874‡	(1.259)
10:00	17.82†	(6.034)	-2.831	(3.120)	0.673	(0.817)	1.108	(0.858)
11:00	15.45*	(3.922)	-5.406	(3.465)	0.329	(0.666)	-1.080	(0.979)
12:00	12.33*	(3.632)	-3.941	(4.121)	1.518	(0.887)	-0.634	(1.100)
13:00	11.72†	(4.423)	-4.271	(4.801)	-0.0909	(0.931)	-0.355	(1.190)
14:00	10.22*	(2.681)	-4.431	(5.171)	-1.043	(1.149)	-1.644	(1.767)
15:00	8.145‡	(3.739)	-5.465	(5.457)	-1.036	(1.179)	-1.482	(1.053)
16:00	7.923‡	(3.862)	-2.688	(4.284)	-1.160	(0.936)	-1.814‡	(0.889)
17:00	3.964	(3.608)	-2.359	(4.292)	-2.361†	(0.784)	-1.942‡	(1.015)
18:00	-0.879	(2.879)	-1.132	(3.705)	-2.534†	(0.816)	-2.057‡	(1.006)
19:00	-0.899	(1.984)	-0.753	(3.799)	-2.432†	(0.729)	-1.525	(0.825)
20:00	-7.183†	(2.650)	0.0696	(3.467)	-1.857†	(0.749)	-1.479	(1.016)
21:00	-12.72†	(4.631)	0.271	(1.921)	-1.762*	(0.448)	0.316	(0.574)
22:00	-13.11†	(4.372)	-0.934	(1.778)	-1.434†	(0.528)	0.0911	(0.905)
23:00	-12.24†	(4.789)	-2.666	(1.975)	-0.193	(0.565)	0.285	(0.334)
24:00	-4.231*	(0.959)	-0.635	(0.539)	-0.257	(0.546)	-0.0143	(0.0794)
Observations	31,037		36,377		39,405		38,550	
Within R^2	0.324		0.012		0.011		0.025	

Static Rolling Regression

Figure: Rolling Regression: Pricing Estimates



Dynamic Pricing Equation

$$\text{LnFare}_{jt} = \alpha \text{LnFare}_{j,t-1} + \beta \text{Hour}_t + \sum_{k=0}^{59} \gamma_k \text{I}_{[\text{Adv}=k]} + v_{ij} + \varepsilon_{ijt}.$$

- Buyers and sellers are allowed to behave dynamically.
  - $\text{LnFare}_{j,t-1}$ is treated as endogenous.
  - Forward looking.
  - Can form beliefs about future prices.
  - Estimation is consistent with rational expectations.

Dynamic Panels

- $\text{LnFare}_{j,t-1}$ as endogenous:
  - Allowed to be correlated with ε_{ijt} and earlier shocks.
  - Cannot be correlated with $\varepsilon_{ij,t+1}$ and future shocks.
- To estimate (α, β, γ), we use the methods described in Holtz-Eakin (1988) and Arellano and Bond (1991):
  - Take first differences to eliminate the fixed effect term v_{ij}.
  - Use $\Delta\varepsilon_{ijt}$ to construct the moments $E(\Delta\varepsilon_{ijt}\mathrm{M})$.
  - Assume ε_{ijt} is not serially correlated.
  - Lagged values of LnFare_{ijt} are valid instruments.
- Two specification tests:
  - Second order serial correlation test.
  - Sargan test for the validity of the instruments.

Table: Dynamic Pricing Equation: Role of ADV_t

Sample:	(1) $0 \leq \text{ADV} < 29.6$		(2) $29.6 \leq \text{ADV} < 59.3$		(3) $0 \leq \text{ADV} < 59.3$	
Variables	Coef	StErr	Coef	StErr	Coef	StErr
2:00	3.458*	(0.758)	0.693*	(0.253)	1.936*	(0.379)
3:00	10.46*	(0.770)	0.979*	(0.255)	5.095*	(0.383)
4:00	3.534*	(0.780)	0.689*	(0.256)	1.788*	(0.386)
5:00	2.680*	(0.783)	0.528†	(0.258)	1.311*	(0.389)
6:00	2.745*	(0.784)	0.571†	(0.261)	1.373*	(0.391)
7:00	2.531*	(0.791)	0.399	(0.264)	1.180*	(0.395)
8:00	2.101*	(0.796)	0.795*	(0.268)	1.208*	(0.399)
9:00	0.808	(0.798)	0.734*	(0.273)	0.529	(0.403)
10:00	0.328	(0.796)	-0.454‡	(0.274)	-0.348	(0.404)
11:00	-0.476	(0.798)	-1.155*	(0.275)	-0.937†	(0.405)
12:00	0.179	(0.800)	0.175	(0.275)	0.199	(0.406)
13:00	-0.481	(0.803)	-0.906*	(0.276)	-0.807†	(0.407)
14:00	-0.826	(0.802)	-1.237*	(0.277)	-1.069*	(0.407)
15:00	-0.925	(0.801)	-0.843*	(0.277)	-0.833†	(0.407)
16:00	-0.610	(0.800)	-1.202*	(0.276)	-0.901†	(0.406)
17:00	-0.841	(0.795)	-1.481*	(0.274)	-1.121*	(0.404)
18:00	-1.670†	(0.789)	-1.285*	(0.272)	-1.365*	(0.401)
19:00	-1.234	(0.785)	-1.099*	(0.272)	-1.083*	(0.400)
20:00	-0.797	(0.780)	-0.900*	(0.270)	-0.815†	(0.397)
21:00	-2.377*	(0.781)	-0.525‡	(0.270)	-1.423*	(0.398)
22:00	-1.589†	(0.767)	-0.326	(0.264)	-0.964†	(0.390)
23:00	-1.846†	(0.746)	-0.192	(0.257)	-1.034*	(0.379)
24:00	1.009	(0.752)	-0.00598	(0.255)	0.435	(0.380)
$\text{LNFARE}_{ij,t-1}$	0.630*	(0.00856)	0.392*	(0.0128)	0.625*	(0.00608)
Specification tests:						
Serial correlation	0.0211 [0.983]		0.0297 [0.976]		0.0900 [0.928]	
Sargan	370.5 [1]		263.2 [1]		443.8 [1]	

Regression Estimates

Table: Dynamic Pricing Equation. Role of ADV_t.

Sample:	$0 \leq \text{ADV} < 29.6$			$29.6 \leq \text{ADV} < 59.3$			$0 \leq \text{ADV} < 59.3$		
	Weekdays		Weekends	Weekdays		Weekends	Weekdays		Weekends
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variables:									
$\text{LNFARE}_{ij,t-1}$	0.629*	0.625*	0.743*	0.404*	0.404*	0.778*	0.625*	0.616*	0.794*
	(0.00840)	(0.00843)	(0.0372)	(0.0126)	(0.0126)	(0.0452)	(0.00602)	(0.00605)	(0.0231)
OFFICE_t [a]	-3.509*	-3.237*	0.0213	-1.001*	-0.836	0.138‡	-2.117*	-4.172*	0.189‡
	(0.340)	(0.708)	(0.220)	(0.120)	(0.639)	(0.0785)	(0.174)	(0.360)	(0.115)
EVENING_t	-4.439*	-7.054*	0.662*	-1.006*	-0.990‡	0.210*	-2.488*	-6.125*	0.474*
	(0.338)	(0.708)	(0.220)	(0.118)	(0.584)	(0.0789)	(0.172)	(0.359)	(0.116)
$\text{OFFICE}_t \cdot \text{ADV}_t$ [b]		-0.0190			-0.00387			0.0698*	
		(0.0421)			(0.0147)			(0.0108)	
$\text{EVENING}_t \cdot \text{ADV}_t$		0.169*			-0.000397			0.122*	
		(0.0409)			(0.0134)			(0.0106)	
Price discrimination:									
$\beta_{\text{OFFICE}} - \beta_{\text{EVENING}}$	0.931*		-0.640*	0.00477		-0.0719	0.371†		-0.284†
	[0.00519]		[0.00368]	[0.967]		[0.331]	[0.0290]		[0.0114]
At $\text{ADV}_t = 7$ [c]		2.502*						1.587*	
		[1.48e-07]						[6.65e-08]	
At $\text{ADV}_t = 21$		-0.128						0.857*	
		[0.753]						[9.11e-06]	
At $\text{ADV}_t = 35$					0.0329			0.126	
					[0.838]			[0.478]	
Specification tests:									
Serial correlation [d]	-0.0511	-0.0831	-0.868	0.0194	0.0198	-0.0687	-0.0438	-0.0976	-0.826
	[0.959]	[0.934]	[0.386]	[0.985]	[0.984]	[0.945]	[0.965]	[0.922]	[0.409]
Sargan [e]	331.4	325.2	232.8	280.5	270.9	174.2	380.4	387.5	328.4
	[1]	[1]	[0.903]	[1]	[1]	[0.998]	[1]	[1]	[1]

Conclusion

- Dynamic price discrimination for perishable good under uncertain demand is a challenging task.
- Fares variation over time cannot be labeled as price discriminatory.
- Estimate price differences within the same flight within the same day to departure.
- Airlines separate consumers in two types: Business and tourists.
- High valuation consumers pay higher prices during office hours.
- The difference is higher as the proportion of high types increases.