Price Discrimination in International Airline Markets*

Gaurab Aryal† Charles Murry‡ Jonathan W. Williams§

February 22, 2021

Abstract

We develop a model of inter-temporal and intra-temporal price discrimination by monopoly airlines to study the ability of different discriminatory pricing mechanisms to increase efficiency and the associated distributional implications. To estimate the model, we use unique data from international airline markets with flight-level variation in prices across time, cabins, and markets, as well as information on passengers' reasons for travel and time of purchase. We find that the ability to screen passengers across cabins every period increases total surplus by 35% relative to choosing only one price per period, with both the airline and passengers benefiting. However, further discrimination based on passenger's reason to traveling improve airline surplus at the expense of total efficiency. We also find that the current pricing practice yields approximately 89% of the first-best welfare. The source of this inefficiency arises mostly from dynamic uncertainty about demand, not private information about passenger valuations.

*We are very grateful to the U.S. Department of Commerce for providing the data for our analysis. We are thankful to research computing facilities at Boston College, UNC and UVA, with special thanks to Sandeep Sarangi at UNC research computing. Seminar (Arizona, BC, MIT, Penn State, Rice, Richmond Fed, Rochester, Texas A&M Toulouse, UC-Irvine, UNC-Chapel Hill, UWO, UWSTL, Vanderbilt, Virginia) and conference (SEA 2014, EARIE 2015, IIOC 2016, ASSA 2016, DSE 2019, SITE 2019) participants provided helpful comments. We thank Jan Brueckner, Ben Eden, Gautam Gowrisankaran, Jacob Gramlich, Qihong Liu, Brian McManus, Nancy Rose, Nick Rupp for insightful suggestions. All remaining errors are our own.

†Department of Economics, University of Virginia, aryalg@virginia.edu.

‡Department of Economics, Boston College, charles.murry@bc.edu.

§Department of Economics, University of North Carolina - Chapel Hill, jonwms@unc.edu.

1 Introduction

Firms with market power often use discriminatory prices to increase their profits. However, such price discrimination can have ambiguous implications on total welfare. Enhanced price discrimination may increase welfare by reducing allocative inefficiencies but may also reduce consumer welfare. So, an essential aspect of economic- and public-policy towards price discrimination is to understand how well various discriminatory prices perform in terms of the total welfare and its distribution, relative to each other and the first-best, (e.g., Pigou, 1920; Varian, 1985; Council of Economic Advisors, 2015).

We evaluate the welfare consequences of price discrimination and quantify sources of inefficiencies in a large and economically important setting, international air travel markets. To that end, we develop and estimate a model of inter-temporal and intra-temporal price discrimination by a monopoly airline and study the ability of different discriminatory mechanisms to increase welfare and the associated distributional implications. The model incorporates a rich specification of passenger valuations for two vertically differentiated seat classes on international flights, and a capacity-constrained airline that faces stochastic and time-varying demand. The airline screens passengers between two cabins while updating prices and seat offerings over time. Using the model estimates, we implement various counterfactuals in the spirit of Bergemann, Brooks, and Morris (2015), where we change the information the airline has about preferences and the timing of arrivals and measure the welfare under various discriminatory pricing strategies. Our counterfactual pricing strategies are motivated by recent airline practices intended to raise profits by reducing allocative inefficiencies, including attempts to solicit passengers' reason to travel and use of auctions (e.g., Nicas, 2013; Vora, 2014; Tully, 2015; McCartney, 2016).

We find that the ability to screen passengers across cabins increases the total surplus by 35% relative to choosing a single price each period (i.e., "shutting down" second-degree price discrimination across cabins), with both airline and passengers benefiting. However, further discriminatory practices based on passengers' reason to travel improve the airline's surplus but lower the total surplus. We also find that the current pricing practice yields approximately 89% of the first-best welfare, and that the source of this remaining inefficiency is mostly due to the dynamic uncertainty about demand, not the private information about passengers' valuations.[1] This suggests that airlines' attempts to improve dynamic allocations could provide efficiency improvements.

Our empirical strategy centers around a novel dataset of international air travel from the

[1] Here, the "first-best" welfare is defined in the context of our model. There may be other effects worth exploring, for example ticket resale in Lazarev (2013), that our model does not capture.

U.S. Department of Commerce's Survey of International Air Travelers. Compared to the extant literature, the novelty of these data is that we observe both the date of transactions and passenger characteristics for dozens of airlines in hundreds of markets. We document the late arrival of passengers traveling for business, who tend to have inelastic demand, and the associated changes in prices. Although business travelers' late arrival puts upward pressure on fares, fares do not increase monotonically for every flight. This pattern suggests that the underlying demand for air travel is stochastic and non-stationary.

To capture these salient data features, we propose a flexible but tractable demand system. Each period before a flight departs, a random number of potential passengers arrive and purchase a first-class ticket, an economy class ticket, or decide not to fly at all. Passengers' willingness-to-pay depends on the seat class and passenger's reason to travel. We allow passengers to have different willingness-to-pay for first-class, so, for some passengers, the two cabins are close substitutes but not for others. Furthermore, we allow the mix of the two *types* of passengers–business and leisure–to vary over time.

On the supply side, we model a monopoly airline's problem of selling a fixed number of economy and first-class seats. The airline knows the distribution of passengers' valuations and the expected number of arrivals each period but chooses prices and seats to release before it realizes actual demand.[2] At any time before the flight, the airline balances expected profit from selling a seat today against the forgone future expected profit. This inter-temporal trade-off results in a time-specific endogenous opportunity cost for each seat that varies with the expected future demand and number of unsold seats. Besides this temporal consideration, each period, the airline screens passengers between the two cabins. Thus, our model captures both the *inter*-temporal and *intra*-temporal aspects of price discrimination by airlines.

Estimation of our model presents numerous challenges. The richness of the demand, and our supply specification, result in a non-stationary dynamic programming problem that involves solving a mixed-integer nonlinear program for each state. We solve this problem to determine optimal prices and seat-release policies for each combination of unsold seats and days until departure. Moreover, our data include hundreds of flights across hundreds of routes, so not only do we allow for heterogeneity in preferences across passengers *within a flight*, we also allow different flights to have different distributions of passenger preferences. To estimate the model and recover the distribution of preferences across flights, we use a simulated method of moments approach based on Ackerberg (2009). Similar approaches to estimate a random coefficient specification has recently been used by Fox, Kim, and Yang (2016), Nevo, Turner, and Williams (2016), and Blundell, Gowrisankaran, and Langer

[2] We model the airline committing to a seat release policy to mimic the "fare bucket" strategy used by airlines in practice. See, for example, Alderighi, Nicolini, and Piga (2015) for more on these buckets.

(2020). Like these papers, we match empirical moments describing within-flight and across-flight variation in fares and purchases to a mixture of moments implied by our model.

Our estimates suggest that there is substantial heterogeneity across passengers within a flight and substantial heterogeneity across flights. The estimated marginal distributions of willingness-to-pay for business and leisure travelers are consistent with the observed distribution of fares. We estimate that the average willingness-to-pay for an economy seat in our data by leisure and business passengers is \$413 and \$506, respectively. Furthermore, on average, passengers value a first-class seat 23% more than an economy seat, which implies meaningful cross-cabin substitution. We also find declining arrivals of passengers overall, but with an increasing fraction of business travelers. Using the model estimates, we calculate the unobserved time-varying opportunity cost of selling a seat, which provides novel insight into airlines' dynamic incentives.

Using the estimates and the model, we characterize the level of efficiency and the associated distribution of surplus for alternative pricing mechanisms that provide new insights on the welfare consequences of price discrimination. In terms of efficiency, we find that airlines' current pricing practices increase total welfare by 35%, compared to a scenario where we prohibit airlines from charging multiple prices across cabins. We also find that the current pricing achieves 89% of the first-best welfare, and almost all of this inefficiency is due to the uncertainty about passengers' arrivals. However, greater discrimination based on the reason to travel –business versus leisure– improves airlines' surplus but slightly lowers total surplus.

In terms of the surplus distribution between airlines and passengers, we find that price discrimination skews the distribution of surplus in favor of the airlines. In particular, the gap between producer surplus and consumer surplus increases by approximately 37% when airlines price differently across cabins, compared to setting only one price per period. Additional price discrimination based on passengers' reasons to travel or willingness-to-pay leads to a higher airline surplus but a lower total surplus. We illustrate this "monopoly externality" by determining surplus division when the airline uses a Vickery-Clarke-Grove (VCG) auction, period-by-period. We find that using VCG more than doubles the consumer surplus, compared to the current pricing practices, while still achieving the same efficiency as eliminating all static informational frictions.

Contribution and Related Literature. Our paper relates to a vast research on the economics of price discrimination and research in the empirical industrial organization on estimating the efficiency and division of welfare under asymmetric information. Most of these empirical papers focus on either cross-sectional price discrimination (e.g., Ivaldi and Martimort, 1994; Leslie, 2004; Busse and Rysman, 2005; Crawford and Shum, 2006; McManus, 2007; Aryal and Gabrielli, 2019) or inter-temporal price discrimination dynamics (e.g., Nevo and Wolfram, 2002; Nair, 2007; Escobari, 2012; Jian, 2012; Hendel and Nevo, 2013; Lazarev, 2013; Cho et al., 2018; Kehoe, Larsen, and Pastorino, 2020).

There is also a literature that focuses on dynamic pricing (e.g., Graddy and Hall, 2011; Sweeting, 2010; Cho et al., 2018; Williams, 2020). However, none study intra-temporal price discrimination, inter-temporal price discrimination and dynamic pricing together, even though many industries involve all three. We contribute to this research by developing an empirical framework where both static discriminative pricing and dynamic pricing incentives are present and obtain results that characterize the welfare implications.[3]

Additionally, we complement recent research related specifically to airline pricing, particularly Lazarev (2013), Li, Granados, and Netessine (2014), and Williams (2020).[4] Lazarev (2013) considers a model of inter-temporal price discrimination with one service cabin and finds large potential gains from allowing reallocation among passengers arriving at different times before the flight (through ticket resale). Williams (2020) further allows for dynamic adjustment of prices in response to stochastic demand and finds that dynamic pricing (relative to a single price) increases total welfare at the expense of business passengers. Li, Granados, and Netessine (2014) use an instrumental variables strategy to study strategic behavior of passengers and infer that between 5% and 20% of passengers wait to purchase in a sample of domestic markets, with the share decreasing in market distance.[5] Incorporating this strategic consumer behavior into a structural model of pricing like Lazarev (2013), Williams (2020), and ours remains unexplored due to theoretical and computational difficulties, although Lazarev (2013) allows passengers to cancel their ticket.

Building on Lazarev (2013) and Williams (2020), we allow the airline to manage revenue by optimally choosing the number of seats to release and by screening passengers between two cabins. Through counterfactuals, our approach allows us to measure the importance of different channels through which inefficiencies arise in airline markets. In terms of estimation, we also allow a rich specification for unobserved heterogeneity across markets via a random coefficients approach to estimation. The richness of random coefficients allows us capture variation in demand parameters due to permanent and unobserved differences across markets.

[3] There is a long and active theoretical literature on static and inter-temporal price discrimination, see Stokey (1979); Gale and Holmes (1993); Wilson (1993); Dana (1999); Courty and Li (2000); Armstrong (2006) and references therein. Airline pricing has also been studied extensively from the perspective of *revenue management*; see van Ryzin and Talluri (2005).

[4] Relatedly, Borenstein and Rose (1994), Puller, Sengupta, and Wiggins (2012); Chandra and Lederman (2018) use a regression approach to study effects of oligopoly on price dispersion and price discrimination in airline markets.

[5] We expect the share to be small in our context because we study long-haul international markets.

2 Data

The Department of Commerce's Survey of International Air Travelers (SIAT) gathers information on international air passengers traveling to and from the U.S. Passengers are asked detailed questions about their flight itinerary, either during the flight or at the gate area before the flight. The SIAT targets both U.S. residents traveling abroad and non-residents visiting the U.S. Passengers in our sample are from randomly chosen flights from among more than 70 participating U.S. and international airlines, including some charter carriers. The survey contains ticket information, which includes the cabin class (first, business, or economy), date of purchase, total fare, and the trip's purpose (business or leisure). We combine fares that are reported as business class and first-class into a single cabin class that we label "first-class." This richness distinguishes the SIAT data from other data like the Origin and Destination Survey (DB1B) conducted by the Department of Transportation. In particular, the additional detail about passengers (e.g., time of purchase, individual ticket fares, and reason for travel) make the SIAT dataset ideal for studying price discrimination.

We create a dataset from the survey where a unit of observation is a single ticket purchased by a passenger flying a nonstop (or direct) route. We then use fares and purchase (calendar) dates associated with these tickets to estimate "price paths" for each flight in our data, where a flight is a single instance of a plane serving a particular route. For example, in our sample, we observe some nonstop passengers flying United Airlines from SEA to TPE on August 12, 2010, departing at 5:10 pm, then we say that this is one flight. From the data on fares and dates for this flight, we use kernel regression to estimate price paths for economy seats and first-class seats leading up to August 12, 2010. In this section, we detail how we selected the sample and display descriptive statistics that motivate our model and analysis.

2.1 Sample Selection

Our sample from the DOC includes 413,309 passenger responses for 2009-2011. We clean the data in order to remove contaminated and missing observations and to construct a sample of flights that will inform our model of airline pricing which we specify in the following section. We detail our sample selection procedure in Appendix A.1, but, for example, we exclude responses that do not report a fare, are part of a group travel package, or are non-revenue tickets. We supplement our data with schedule data from the Official Aviation Guide of the Airways (OAG) company, which reports cabin-specific capacities, by flight number. Using the flight date and flight number in SIAT we can merge the two data sets. We include flights for which we observe at least ten nonstop tickets after applying the sample selection criteria.

Nonstop Markets and Capacity. Like other studies that model discriminatory pricing by airlines (e.g., Lazarev (2013); Puller, Sengupta, and Wiggins (2012); Williams (2020)), we focus on nonstop travel in monopoly markets, where we define monopoly market criteria below. Although a physical flight will have both connecting and nonstop passengers, we assume that an airline devotes a specific portion of a plane to nonstop passengers before it starts selling tickets, and the airline does not change the plane's apportionment. We make this assumption to keep our model tractable because modeling airlines' pricing strategies for both nonstop and connecting passengers is a high-dimensional optimization problem that has to balance the cross-elasticities of passengers on all potential itineraries in the airline's network that could use one of the flights in our sample as a leg.

Specifically, we determine the ratio of nonstop travel on a flight in our survey by first determining the ratio of nonstop travel for the route in a quarter from the Department of Transportation's DB1B data. We then apply this ratio to the total capacity of the equipment used (which we observe and we describe below) in our sample to arrive at a nonstop capacity for each flight. As an illustration, consider the SEA to TPE flight on August 12, 2010. Suppose in this flight, United Airlines used a Boeing 757 with 235 economy seats, and according to the DB1B data in Q3:2010, only 45 percent of tickets for this market are nonstop tickets. Then, for us, this means that United Airlines reserved only 110 economy seats (out of 235) for nonstop travel, and so the initial economy class capacity for this flight market is 110. We repeat this exercise for the first-class seats.[6]

We observe the entire itinerary of a passenger, so to select ticket purchases for our final sample, we discard all passengers using a flight as part of a connecting itinerary. For example, if we see a passenger report flying from Cleveland to Taipei via Seattle, we drop them from our sample. The same is true if we observe a passenger flying from Toulouse to Charlotte via Paris.

"Monopoly" Markets. As mentioned earlier, we focus on monopoly markets. In international air travel, nonstop markets tend to be concentrated for all but the few busiest airport-pairs. We classify a market as a monopoly-market if it satisfies one of following two criteria: (i) one airline flies at least 95% of the total capacity on the route (where the capacity is measured using the OAG data); or (ii) there is a US carrier and foreign carrier that operate on the market with antitrust immunity from the U.S. Department Of Justice.

These antitrust exemptions come from market access treaties signed between the U.S. and the foreign country that specify a local foreign carrier (usually an alliance partner of the U.S. airline) that will share the route. For example, on July 20, 2010 antitrust exemption

[6] Few flights are not in the DB1B dataset, and for them, we use the nonstop ratio from the SIAT survey.

was granted to OneWorld alliance, which includes American Airlines, British Airways, Iberia, Finnair and Royal Jordanian, for 10 years subject to a slot remedy.[7]

In a few cases, we define markets at the city-pair level because we are concerned that within-city airports are substitutable. The airports that we aggregate up to a city-pairs market definition include airports in the New York, London, and Tokyo metropolitan. Thus, we treat a flight from New York JFK to London Heathrow to be in the same market as a flight from Newark EWR to London Gatwick.

Table 1: Top 20 Markets, ordered by sample representation

Market	Unique Flights	Obs.	Distance (miles)	Market	Unique Flights	Obs.	Distance (miles)
LAX-PVG	53	1,060	6,485	JFK-SCL	16	239	5,096
JFK-HEL	35	581	4,117	SFB-BHX	10	234	4,249
SFO-AKL	28	497	6,516	JFK-NCE	9	217	3,991
JFK-VIE	29	469	4,239	PHX-SJD	15	211	721
JFK-EZE	24	444	5,281	JFK-POS	13	208	2,204
JFK-WAW	34	434	4,267	JFK-PLS	8	205	1,303
JFK-JNB	26	368	7,967	SFO-HND	4	198	5,161
JFK-CCS	20	360	2,109	BOS-KEF	12	178	2,413
JFK-CMN	21	330	3,609	SFO-CDG	10	162	5,583
SEA-TPE	21	328	6,074	JFK-MTY	9	148	1,826

Note: The table displays the Top 25 markets, their representation in our sample, their distance (air miles) from U.S. Data from the Survey of International Air Travelers and sample described in the text.

Description of Markets and Carriers. After our initial selection and restriction on nonstop monopoly markets, we have 14,930 observations representing 224 markets and 64 carriers. We list the top 20 markets based on sample representation in Table 1, along with the number of unique flights and the total observations. The most common U.S. airports in our final sample are San Francisco (SFO), New York JFK (JFK), and Phoenix (PHX), and the three most common routes are SFO to Auckland, New Zealand (AKL), JFK to Helsinki, Finland (HEL), and JFK to Johannesburg, South Africa (JNB). The median flight has a distance of 4,229 miles, and the inter-quartile range of $2,287$ to $5,583$ miles. In Table 2 we display the top ten carriers from our final sample, which represent approximately 60% of our final observations.

2.2 Passenger Arrivals and Ticket Sale Process

Passengers differ in terms of time of purchases and reasons for travel and prices vary over time and across cabins. In this subsection we present key features in our data pertaining the

[7] To determine such markets, we use information from DOT and carriers' 10K reports filed with the SEC.

Table 2: Top Ten Carriers, ordered by sample representation

Carrier	**Unique Flights**	**Obs.**
American Airlines	88	1526
Continental	98	1483
US Airways	72	1200
Eastern China	53	1060
Finnair	35	581
Austrian	35	573
Lufthansa	40	558
Delta	35	550
Air New Zealand	30	520
ThomsonFly	20	470

Note: The table displays the top ten carriers, ordered by their frequency in the final sample. Data from the Survey of International Air Travelers and sample described in the text.

passengers and prices.

Timing of Purchase. Airlines typically start selling tickets a year before the flight date. Although passengers can buy their tickets throughout the year, in our sample most passengers buy in the last 120 days. To keep the model and estimation tractable, we classify the purchase day into a fixed number of bins. There are at least two factors that motivate our choice of bin sizes. First, it appears that airlines typically adjust fares frequently in the last few weeks before the flight date, but less often farther away from the flight date. Second, there is usually a spike in passengers buying tickets at focal points, like 30 days, 60 days etc. In Table 3 we present eight fixed bins, and the number of observations in each bin, and as we can see we use narrower bins when we are closer to the flight date. Each of these eight bins correspond to one period, giving us a total of eight periods.

Passenger Characteristics. We classify each passenger as either a business traveler or a leisure traveler based on the reason to travel. *Business* includes business, conference, and government/military, while *leisure* includes visiting family, vacation, religious purposes, study/teaching, health, and other, see Table A.1.1 for more.[8] We classify service cabins into economy class and first-class, where we classify every premium service cabin as latter.

In Table 4 we display some key statistics for relevant ticket characteristics in our sample. As is common in the literature, to make one-way and round-trip fares comparable we divide round-trip fares by two. Approximately 4.5% passengers report to have bought a one-way

[8] We also observe the channel through which the tickets were purchased: travel agents (36.47%), personal computer (39.76%), airlines (13.36%), company travel department (3.99%), and others (6.42%).

Table 3: Distribution of Advance Purchase

Days Until Flight	**Obs.**
0 – 3 days	597
4 – 7 days	753
8 – 14 days	1,038
15 – 29 days	1,555
30 – 44 days	2,485
45 – 60 days	2,561
61 – 100 days	2,438
101+ days	3,400

Note: The table displays the distribution of advance purchases. First column is the number of days before flight and the second column shows how many passengers bought their tickets in those days.

ticket, see Table A.1.1. From the top panel, in our sample, 92.5% of the passengers purchased economy class tickets, and the average fare was \$447 whereas 7.5% of passengers purchased first-class tickets and paid an average of \$897. The standard deviations in fares, which quantify both the across-market and within-flight variation in fares, are large, with the coefficient of variation 0.85 for economy and 1.06 for first-class.

In the second panel of Table 4, we display the same statistics by the number of days in advance of a flight's departure that the ticket was purchased (aggregated to eight "periods"). We see that 4% of the passengers bought their ticket in last three days before the flight; 5.08% bought 4-7 days; 7% bought 8-14 days; 10.4% bought 15-29 days; 16.7% bought 30-44 days; 17.2% bought 45-60 days, 16.4% bought 61-100 days and the rest 23% bought at least 100 days before the flight. While the average fare increases for tickets purchased closer to the departure date, so does the standard deviation.

Similarly, at the bottom panel of Table 4, we report price statistics by the passenger's trip purpose. About 14% of the passengers in our sample flew for business purposes, and these passengers paid an average price of \$684 for one direction of their itinerary. Leisure passengers paid an average of \$446. This price difference arises for at least three reasons: business travelers tend to buy their tickets much closer to the flight date, they prefer first-class seats, and they fly different types of markets.

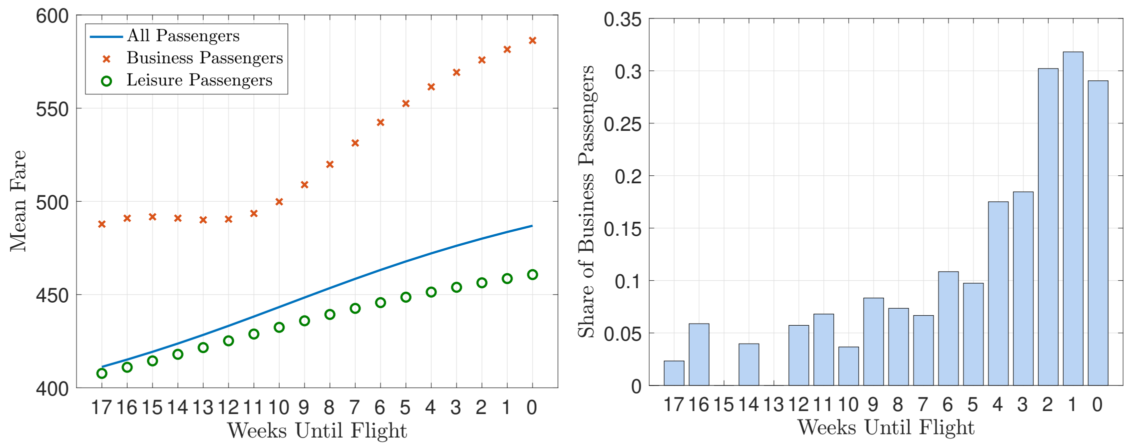
In Figure 1(a) we plot the average price for economy fares as a function when the ticket was purchased. Both business and leisure travelers pay more if they buy the ticket closer to the flight date, but the increase is more substantial for the business travelers. The solid line in Figure 1(a) reflects the average price across both reasons for travel. At earlier dates, the total average price is closer to the average price paid by leisure travelers, while it gets closer to the average price paid by the business travelers as the date of the flight nears. In

Table 4: Summary Statistics from SIAT, Ticket Characteristics

Ticket Class	Proportion of Sample	Fare Mean	Fare S.D.
Economy	92.50	447	382
First	7.50	897	956
Advance Purchase			
0-3 Days	4.03	617	636
4-7 Days	5.08	632	679
8-14 Days	7.00	571	599
15-29 Days	10.49	553	567
30-44 Days	16.76	478	429
45-60 Days	17.27	467	432
61-100 Days	16.44	414	315
101+ Days	22.93	419	387
Travel Purpose			
Leisure	85.57	446	400
Business	14.43	684	716

Note: Data from the Survey of International Air Travelers. Sample described in the text.

Figure 1: *Business versus Leisure Passengers before the Flight Date*

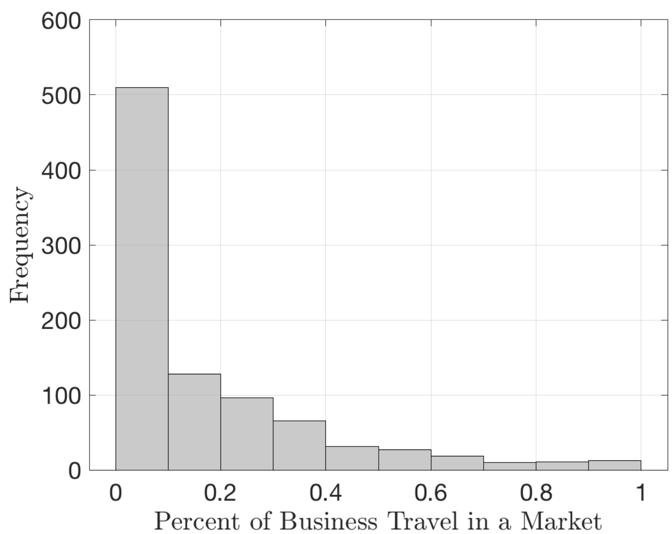


(a) Average Economy Fares Prior to Flight (b) Business and Leisure Travelers

Note: (a) Average price paths across all flights for tickets in economy class by week of purchase prior to the flight date, by self-reported business and leisure travelers. Individual transaction prices are smoothed using nearest neighbor with a Gaussian kernel with optimal bandwidth of 0.5198. (b) Proportion of business passengers across all flights, by advance purchase weeks.

Figure 1(b), we display the proportion of business to leisure travelers across all flights, by the advance purchase categories. In the last two months before flight, the share of passengers traveling for leisure is approximately 90%, which decreases to 65% a week before flight. Taken together, business travelers purchase closer to the flight date than leisure travelers, and markets with a greater proportion of business travelers have a steeper price gradient.

Figure 2: *Histogram of Percent of Business Passengers by Flight*



Note: Histogram of business-travel index (BTI). The business-traveler index is the flight-specific ratio of self-reported business travelers to leisure travelers. The mean is 0.154 and the standard deviation is 0.210.

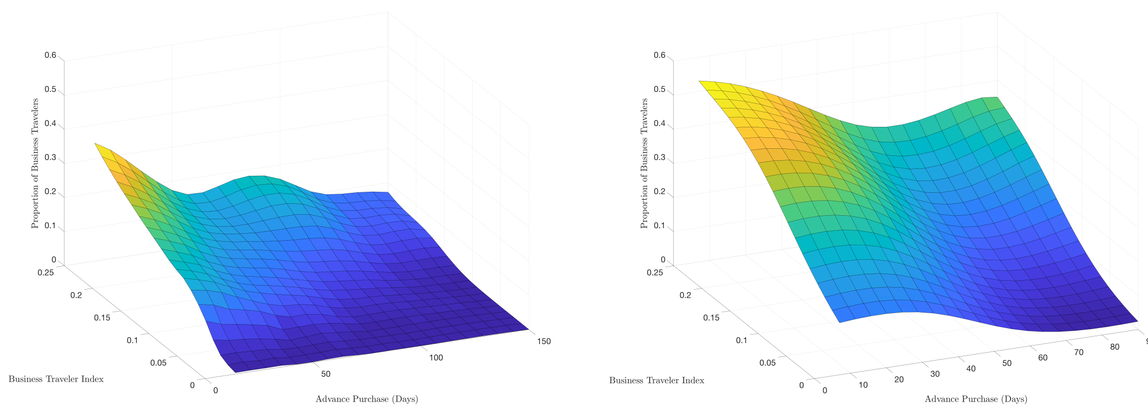
Observing the purpose of travel plays an important role in our empirical analysis, reflecting substantial differences in the behavior and preferences of business and leisure passengers. This passenger heterogeneity across markets drives variation in pricing, and this covariation permits us to estimate a model with richer consumer heterogeneity than the existing literature like Berry, Carnall, and Spiller (2006) and Ciliberto and Williams (2014). Further, a clean taxonomy of passenger types allows a straightforward exploration of the role of asymmetric information in determining inefficiencies and the distribution of surplus that arises from discriminatory pricing of different forms.[9]

To further explore the influence that this source of observable passenger heterogeneity has on fares, we present statistics on across-market variation in the dynamics of fares. Specifically, we first calculate the proportion of business travelers in each market, i.e., across all flights with the same origin and destination. Like Borenstein (2010), we call this market-

[9] In the raw data, 5.46% of passengers report that the main reason why they choose their airline is frequent flyer miles. In our final (estimation) sample, 1.7% passengers get upgraded, i.e., they buy economy but fly first-class. Our working hypothesis is that those upgrades materialized only on the day of travel.

specific ratio the business-traveler index (BTI). In Figure 2, we present the histogram of the BTI across markets in our data. If airlines know of this across-market heterogeneity and use it as a basis to discriminate both intra-temporally (across cabins) and inter-temporally (across time before a flight departs), different within-flight temporal patterns in fares should arise for different values of the BTI.

Figure 3: *Proportion of Business Travelers by Ticket Class*



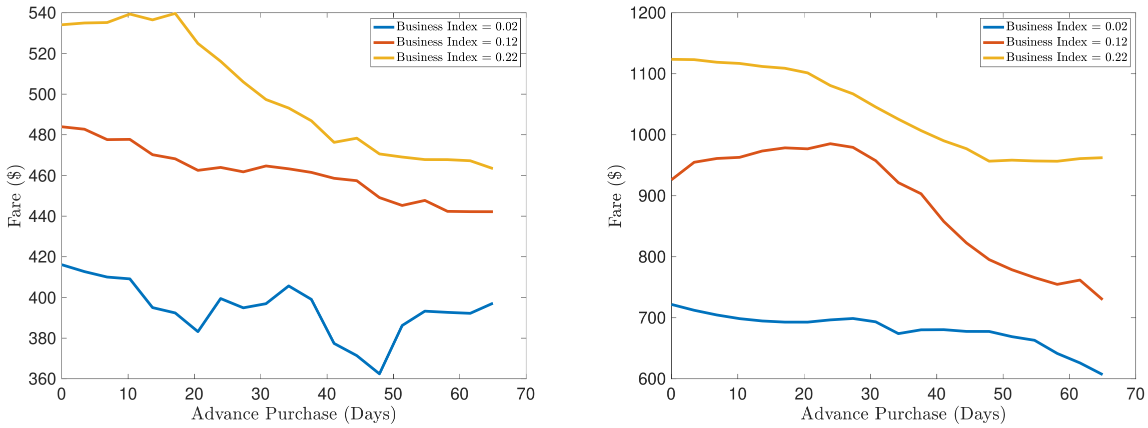
(a) Economy Class (b) First-Class

Note: The figure presents kernel regression of reason to travel on BTI and the purchase date. Panels (a) and (b) show the regression for economy and the first-class seats, respectively. The regression uses a Gaussian with optimal 'rule-of-thumb' bandwidth.

In Figure 3 we present the results of a bivariate kernel regression where we regress an indicator for whether a passenger is traveling for business on the BTI in that market and number of days the ticket was purchased in advance of the flight's departure. Figures 3(a) and 3(b) present the results for economy and first-class passengers, respectively. There are two important observations. First, across all values of the BTI, business passengers arrive later than leisure passengers. Second, business passengers disproportionately choose first-class seats. To capture this feature, in Section 3, we model the difference between business and leisure passengers in terms of the timing of purchases and the preference for quality by allowing the passenger mix to change as the flight date approaches, resulting in a non-stationary demand process.

The influence of business passengers is evident on prices. Like Figure 3, Figure 4(a) and Figure 4(b) present the results of a kernel regression with fare paid as the dependent variable for economy and first-class cabins, respectively. In both, we present cross-sections of these estimated surfaces for the 25^{th}, 50^{th}, and 75^{th} percentile values of the BTI. For both cabins, greater values of the BTI are associated with substantially higher fares. Further,

Figure 4: *Across-Market Variation in Fares*



(a) Economy (b) First-Class

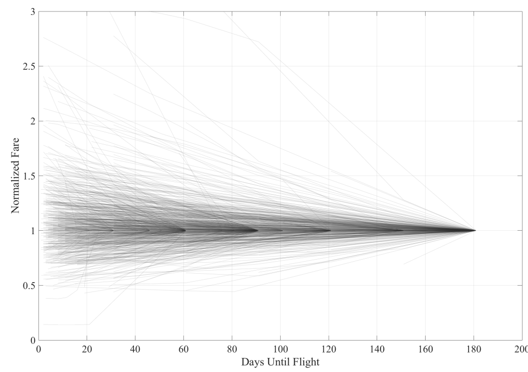
Note: The figure presents results from a kernel regression of fares paid on BTI and the purchase date. Panels (a) and (b) show the results from the regression evaluated at different values of BTI, for economy and the first-class seats, respectively. The regression uses a Gaussian with optimal 'rule-of-thumb' bandwidth.

there is a positive relationship between the rate of increase as the flight date approaches and the BTI, and this rate is positive as the flight date approaches only in markets with non-zero BTI. This pattern is most evident in first-class fares. Thus, the presence of business travelers is associated with both greater average fares and steeper increases in fares as the flight date approaches for both cabins. The larger increase in first-class fares as the flight date approaches, relative to economy fares, is consistent with the strong selection of business travelers into the first-class cabin.

While there are clear patterns in how the dynamics of *average* fares vary with the BTI, there is also substantial heterogeneity across flights in how fares change as the flight date approaches. To see the heterogeneity in temporal patterns for individual flights that Figure 3 masks, Figure 5 presents the time-paths of economy fares for all flights in our data. Specifically, for each flight, we estimate a smooth relationship between economy fares and time before departure using a kernel regression, and then normalize the path relative to the initial fare for that flight. Each line is a single flight from our data, and begins when we first observe a fare for that flight, and ends at 1, the day of the flight.

For most flights we observe little movement in fares until approximately 100 days before departure. Yet, for a small proportion of flights, there are substantial decreases and increases in fares as much as 5 months before departure. Further, by the date of departure, the interquartile range of the ratio of current fare to initial fare is 0.75 to 1.85. Thus, 25% of

Figure 5: *Flight-Level Dispersion in Fares*



Note: The figure presents the estimated fare paths from every flight we use in our estimation. Each line represents a flight. We normalize the fares to unity by dividing fares by the first fare we observe for the flight. The horizontal axis represents the number of days until the flight, where zero is the flight date.

flights experience a decrease of more than 25%, while 25% of flights experience an increase of greater than 85%. The variation in the temporal patterns in fares across flights is attributable to both the across-market heterogeneity in the mix of passengers, and how airlines respond to demand uncertainty.

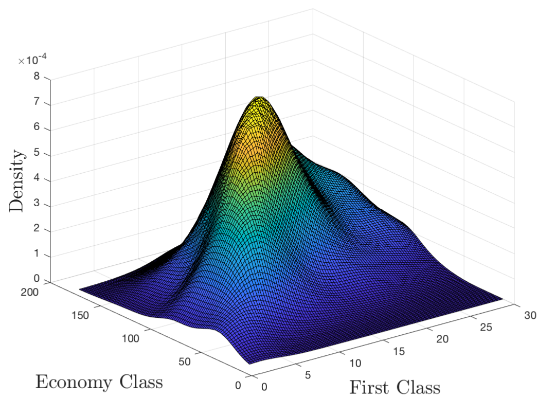
2.3 Aircraft Characteristics

Airlines' fares, and their responsiveness to realized demand, depend on the number of unsold seats. In Figure 6(a) we display the joint density of initial capacity of first and economy class in our sample. The median capacity of an aircraft in our sample is 116 economy seats and 15 first-class seats, and the mode is 138 economy and 16 first-class seats.

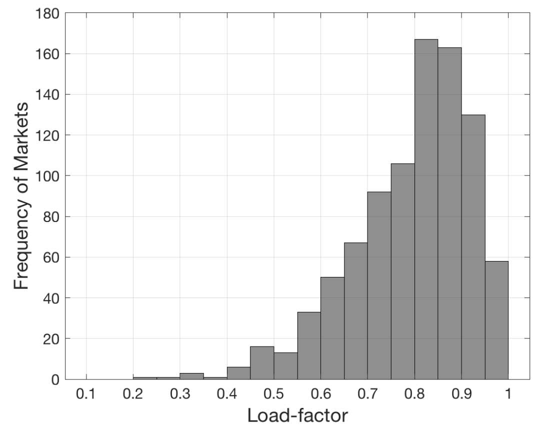
The three most common aircraft types in our sample are a Boeing 777, 747, and 737 (36% of flights in our sample). The 777 and 747 are wide-body jets. The 777 has a typical seating of around 350 seats and the 737 has a typical seating of around 160 seats (before adjusting for non-stop versus connecting passengers). The most common Airbus equipment is the A330, which makes up about 4% of the flights in our sample.

Across all flights, on average 88% of all seats are economy class. We merge the SIAT data with the Department of Transportation's T-100 segment data to get a measure of the load factor for our SIAT flights. From the T100, we know the average load factor across a *month* for a particular route flown by a particular type of equipment. In Figure 6(b) we display the density of load factor across flights in our sample. The median load factor is 82%, but there is substantial heterogeneity across flights.

Figure 6: *Initial Capacity and Load Factor*



(a) Density of Initial Capacities



(b) Histogram of Load Factor

Note: In part (a) this figure presents the Parzen-Rosenblatt Kernel density estimate of the joint-density of initial capacities available for nonstop travel. In part (b) this figure presents the histogram of the passenger load factor across our sample.

Overall, our descriptive analysis reveals a number of salient features that we capture in our model. We find that a business-leisure taxonomy of passenger types is useful to capture differences in the timing of purchase, willingness-to-pay, and preference for quality. Further, we find substantial heterogeneity in the business or leisure mix of passengers across markets, which airlines are aware of and responsive to, creating variation in both the level and temporal patterns of fares across markets. Finally, across flights we observe considerable heterogeneity in fare paths as the flight date approaches. Together, these features motivate our model of non-stationary and stochastic demand and dynamic pricing by airlines that we present in Section 3, as well as the estimation approach in Section 4.

3 Model

In this section, we present a model of dynamic pricing by a profit-maximizing multi-product monopoly airline that sells a fixed number of economy ($0 \leq K^e < \infty$) and first-class ($0 \leq K^f < \infty$) seats. We assume passengers with heterogeneous and privately known preferences arrive before the date of departure ($t \in \{0, \ldots, T\}$) for a nonstop flight. Every period the airline has to choose the ticket prices and the maximum number of unsold seats to sell at those prices before it realizes the demand (for that period).

Our data indicate essential sources of heterogeneity in preferences that differ by reason for travel: willingness-to-pay, valuation of quality, and purchase timing. Further, variability and non-monotonicity in fares suggest a role for uncertain demand. Our model's demand-side seeks to flexibly capture this multi-dimensional heterogeneity and uncertainty that serves as an input into the airline's dynamic-pricing problem. Furthermore, our model's supply-side seeks to capture the inter-temporal and intra-temporal tradeoffs faced by an airline in choosing its optimal policy.

3.1 Demand

Let N_t denote the number of individuals that *arrive* in period $t \in \{1, \ldots, T\}$ to consider buying a ticket. We model N_t as a Poisson random variable with parameter $\lambda_t \in \mathbb{N}$, i.e., $\mathbb{E}(N_t) = \lambda_t$. The airline knows λ_t for $t \in \{1, \ldots, T\}$, but must make pricing and seat-release decisions before the uncertainty over the number of *arrivals* is resolved each period. The arrivals are one of two types, for-business or for-leisure. The probability that a given individual is for-business varies across time before departure and denoted by $\theta_t \in [0, 1]$.

For a given individual, let $v \subset \mathbb{R}_+$ denote the value this person assigns to flying in economy cabin, and let the indirect utility of this individual from flying economy and first-class at price p, respectively, be

$$u^e(v, p, \xi) = v - p; \quad u^f(v, p, \xi) = v \times \xi - p, \qquad \xi \in [1, \infty).$$

Thus, ξ is the (utility) premium associated with flying in a first-class seat that captures the vertical quality differences between the two cabins. Arrivals are heterogeneous in terms of their v and ξ that are mutually independent and privately known to the individual.

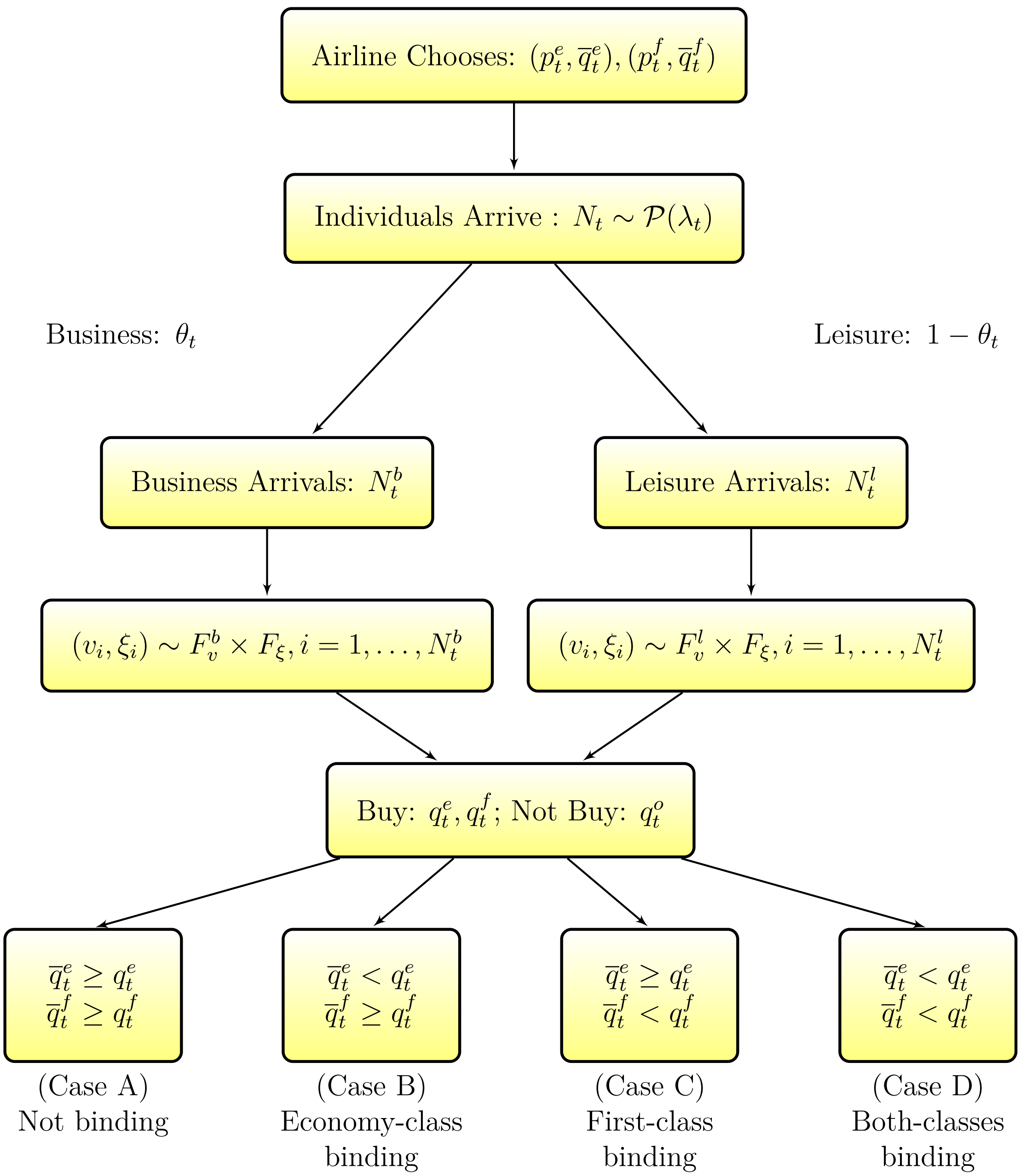
We assume that the distribution of these preferences across arrivals are realizations from type-specific distributions. Specifically, the v of for-business and for-leisure arrivals are drawn from $F_v^b(\cdot)$ and $F_v^l(\cdot)$, respectively, and ξ is drawn from $F_\xi(\cdot)$. Together with the arrival process, the type-specific distribution of valuations creates a stochastic and non-stationary demand process that we assume is known to the airline.

At given prices and a given number of seats available at those prices, Figure 7 summarizes a realization of the demand process for period t. Specifically, the realization of demand and timing of information known by the airline leading up to a flight's departure is as follows:

(i) Airline chooses a price and seat-release policy for economy cabin, $(p_t^e, \overline{q}_t^e)$, and the first-class cabin, $(p_t^f, \overline{q}_t^f)$, that determine the prices at which a maximum number of seats in the two cabins may be sold.

(ii) N_t many individuals *arrive*, the number being drawn from a Poisson distribution with parameter λ_t. Each *arrival* realizes their reason to fly from a Bernoulli

Figure 7: Realization of Demand



Note: A schematic representation of the timing of demand.

distribution with parameter θ_t (i.e., for-business equals one).

(iii) Each *arrival* observes their own (v, ξ), drawn from the respective distributions, $F_v^b(\cdot)$, $F_v^l(\cdot)$, and $F_\xi(\cdot)$.

(iv) If neither seat-release policy is binding (realized demand does not exceed the number of seats released in either cabin), *arrivals* select their most preferred cabin: first-class if $v \times \xi - p_t^f \geq \max\{0, v - p_t^e\}$, economy if $v - p_t^e \geq \max\{0, v \times \xi - p_t^f\}$, and no purchase if $0 \geq \max\{v \times \xi - p_t^f, v - p_t^e\}$. Those *arrivals* choosing the no-purchase option leave the market. If the seat-release policy is binding in either one or both cabins, we assume that *arrivals* make sequential decisions in a randomized order until either none remaining wishes to travel in the cabin with capacity remaining or all available seats are allocated.

(v) Steps (i)-(iv) repeat until the date of departure, $t = T$, or all of the seats are allocated.

In any given period (t), there are four possible outcomes given a demand realization: neither seat-release policy is binding, either one of the two seat-release policies is binding, or both are binding. If the seat-release policy is not binding for either of the two cabins, then the expected demand for the respective cabins in period t when the airline chooses policy $\chi_t := (p_t^e, \overline{q}_t^e, p_t^f, \overline{q}_t^f)$ is

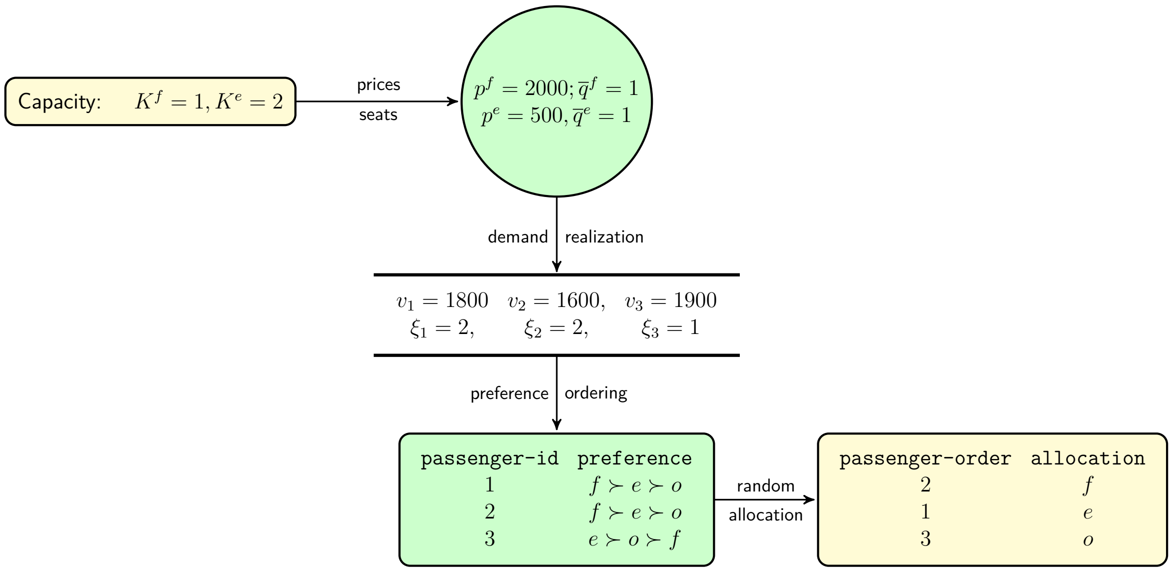
$$\mathbb{E}_t(q^e; \chi_t) := \sum_{n=0}^{\infty} \left\{ n \times \Pr(N_t = n) \underbrace{\Pr(v - p_t^e \geq \max\{0, v \times \xi - p_t^f\})}_{:= P_t^e(\chi_t)} \right\} = \lambda_t \times P_t^e(\chi_t);$$

$$\mathbb{E}_t(q^f; \chi_t) := \sum_{n=0}^{\infty} \left\{ n \times \Pr(N_t = n) \underbrace{\Pr(v \times \xi - p_t^f \geq \max\{0, v - p_t^e\})}_{:= P_t^f(\chi_t)} \right\} = \lambda_t \times P_t^f(\chi_t).$$

If one or both of the seat-release policies are binding, the rationing process creates the possibility for inefficiencies to arise both in terms of exclusion of passengers with a greater willingness-to-pay than those that are allocated a seat, as well as misallocations of passengers across cabins.

In Figure 8, we present a simple example to illustrate inefficiency arising from asymmetric information in this environment under random allocation. Assume the airline has one first-class and two economy seats remaining and chooses to release one seat in each cabin at $p^f = 2000$ and $p^e = 500$. Suppose three passengers arrive with values $v_1 = 2500$, $v_2 = 1600$, and $v_3 = 5000$, with $\xi_1 = \xi_2 = 2$ and $\xi_3 = 1$. Arrivals 1 and 2 are willing to pay twice

Figure 8: Illustration of Random Rationing Rule.



Note: Example to demonstrate how random-rationing rule can generate inefficiency in the model.

as much for a first-class seat as an economy seat, whereas arrival 3 values the two cabins equally. Suppose that under the random allocation rule, arrival 2 gets to choose first and arrival 3 is the last. As shown in Figure 8, the final allocation is inefficient because: a) arrival 2 gets first-class even though 1 values it more; and b) arrival 1 gets economy even though 3 values it more. This difference in arrival timing creates the possibility for multiple welfare-enhancing trades. Given the limited opportunity for coordination amongst arrivals to make such trades, and the legal/administrative barriers to doing so, we believe random rationing is a reasonable way to allocate seats within a period.

3.2 Supply

The airline has T periods, before the departure, to sell K^e and K^f economy and first-class seats, respectively. Each period, the airline chooses prices $\{p_t^e, p_t^f\}$ and commits to selling no more than $\{\overline{q}_t^e, \overline{q}_t^f\} \leq \omega_t$ seats at those prices, where $\omega_t := (K_t^e, K_t^f)$ is the number of unsold seats in each cabin. We model that airlines must commit to a seat release policy to mimic the "fare bucket" strategy that airlines use in practice (e.g., Alderighi, Nicolini, and Piga, 2015), which helps insure the airline against a "good" demand shock where too many seats are sold today at the expense of future higher willingness to pay passengers.[10] One of this

[10] Our modeling choices, in particular seat release policies and random assignments, imply that there may be instances when a passenger with a high willingness-to-pay shows up to the market and cannot be served.

market's defining characteristics is that the airline must commit to policies this period before realizing the current and future demand. The airline does not observe a passenger's reason to fly or valuations (v, ξ); however, the airline knows the underlying stochastic process that governs demand and uses the information to price discriminate, both within a period and across periods.[11]

Let c^e and c^f denote the constant cost of servicing a passenger in the respective cabins. These marginal costs, or so-called "peanut costs," capture variable costs like food and beverage service that do not vary with the timing of the purchase but may vary with the different levels of service in the two cabins. Let $\Psi := (\{F_v^b, F_v^l, F_\xi, c^f, c^e\}, \{\lambda_t, \theta_t\}_{t=1}^T)$ denote the vector of demand and cost primitives.

The airline maximizes the sum of discounted expected profits by choosing price and seat-release policies for each cabin, $\chi_t = \left(p_t^e, p_t^f, \overline{q}_t^e, \overline{q}_t^f\right)$, in each period $t = 1, \ldots, T$ given ω_t. The optimal policy is a vector $\{\chi_t : t = 1, \ldots, T\}$ that maximizes expected profit

$$\sum_{t=1}^{T} \mathbb{E}_t \left\{\pi(\chi_t, \omega_t; \Psi_t)\right\},$$

where $\pi(\chi_t, \omega_t; \Psi_t) = (p_t^f - c^f)q_t^f + (p_t^e - c^e)q_t^e$ is the per-period profit after the demand for each cabin is realized (q_t^e and q_t^f) and $\Psi_t = (\{F_v^b, F_v^l, F_\xi, c^f, c^e\}, \{\lambda_t, \theta_t\})$. The airline observes the unsold capacity (ω_t) at the time of choosing its policy, but not the particular realization of passenger valuations that determine the realized demand. The optimal seat-release policy must satisfy $\overline{q}_t^e \leq K_t^e$ and $\overline{q}_t^f \leq K_t^f$ and take on integer values.

The stochastic process for demand, capacity-rationing algorithm, and optimally chosen seat-release and pricing policies induce a non-stationary transition process between states, $Q_t(\omega_{t+1}|\chi_t, \omega_t, \Psi_t)$. The optimal policy in periods $t \in \{1, \ldots, T-1\}$ is characterized by the solution to the Bellman equation,

$$V_t(\omega_t, \Psi) = \max_{\chi_t} \mathbb{E}_t \left\{ \pi(\chi_t, \omega_t; \Psi_t) + \sum_{\omega \in \Omega_{t+1}} V_{t+1}(\omega_{t+1}, \Psi) \times Q_t(\omega_{t+1}|\chi_t, \omega_t, \Psi_t) \right\}, \tag{1}$$

where Ω_{t+1} represents the set of reachable states in period $t+1$ given ω_t and χ_t. The expectation, $\mathbb{E}_t$, is over realizations from the demand process (Ψ_t) from period t to the date

If, instead, we had used an optimal rationing, where seats are assigned in the order willingness-to-pay, it would lead to higher baseline efficiency. There still may be other modeling choices, but overall, we believe that our model is a reasonable given our sample and our final goal.

[11] See Barnhart, Belobaba, and Odoni (2003) for an overview of forecasting airline demand.

of departure T. In period T, optimal prices maximize

$$V_T(\omega_T, \Psi_T) = \max_{\chi_T} \mathbb{E}_T \pi(\chi_T, \omega_T; \Psi_T),$$

because the firm no longer faces any inter-temporal tradeoffs.[12] The dynamic programming that characterizes an airline's problem is useful for identifying the airline's tradeoffs and identifying useful sources of variation in our data.[13]

The optimal pricing strategy includes both inter-temporal and intra-temporal price discrimination. First, given the limited capacity, the airline must weigh allocating a seat to a passenger today versus a passenger tomorrow, who may have a higher mean willingness-to-pay because the fraction of for-business passengers increases as it gets closer to the flight date. This decision is difficult because both the volume (λ_t) and composition (θ_t) of demand changes as the date of departure nears. Thus, the good's perishable nature does not necessarily generate declining price paths like Sweeting (2010). Simultaneously, every period, the airline must allocate passengers across the two cabins by choosing χ_t such that the price and supply restriction-induced selection into cabins is optimal.

To illustrate the problem further, consider the trade-off faced by an airline from increasing the price for economy seats today: (i) decreases the expected number of economy seat purchases but increases the revenue associated with each purchase; (ii) increases the expected number of first-class seat purchases but no change to revenue associated with each purchase; (iii) increases the expected number of economy seats and decreases the expected number of first-class seats available to sell in future periods. Effects (i) and (ii) capture the multi-product tradeoff faced by the firm, while (iii) captures the inter-temporal tradeoff. More generally, differentiating Equation 1 with respect to the two prices gives two first-order conditions that characterize optimal prices given a particular seat-release policy:

$$\begin{pmatrix} \mathbb{E}_t(q^e; \chi_t) \\ \mathbb{E}_t(q^f; \chi_t) \end{pmatrix} + \begin{bmatrix} \frac{\partial \mathbb{E}_t(q^e;\chi_t)}{\partial p_t^e} & -\frac{\partial \mathbb{E}_t(q^f;\chi_t)}{\partial p_t^e} \\ -\frac{\partial \mathbb{E}_t(q^e;\chi_t)}{\partial p_t^f} & \frac{\partial \mathbb{E}_t(q^f;\chi_t)}{\partial p_t^f} \end{bmatrix} \begin{pmatrix} p_t^e - c^e \\ p_t^f - c^f \end{pmatrix} = \begin{pmatrix} \frac{\partial \mathbb{E}_t V_{t+1}}{\partial p_t^e} \\ \frac{\partial \mathbb{E}_t V_{t+1}}{\partial p_t^f} \end{pmatrix}. \tag{2}$$

[12] For model tractability, we assume that passengers cannot strategically time their purchases. So, their arrival times and their purchase times are the same and do not depend on the price path. This assumption is also used by Williams (2020) to model dynamic pricing. Board and Skrzypacz (2016) allow consumers to be strategic under an additional assumption that the seller has full commitment and chooses its supply function only once, in the first period. In the airline industry, however, the assumption that airlines choose their fares only once at the beginning is too strong.

[13] Although we focus only on one flight, an airline may also consider future flights. In the latter, fares across different flights can be interlinked. We conjecture that we can then approximate an airline's pricing problem with a non-zero continuation value in the last period. However, to estimate such a model, the SIAT survey data is insufficient because we would need to sample every flight sufficiently many times.

The left side is the contemporaneous marginal benefit net of static costs, while the right side is the discounted future benefit.

Equation 2 makes clear the two components of marginal cost: (i) the constant variable cost, or "peanut" cost, associated with servicing seats occupied by passengers; (ii) the opportunity cost of selling additional seats in the current period rather than in future periods. We refer to (iii), the vector on the ride side of the Equation 2, as the *shadow cost* of a seat in the respective cabins. These *shadow costs* depend on the firm's expectation regarding future demand (i.e., variation in volume of passengers and business/leisure mix as flight date nears), and the number of seats remaining in each cabin (i.e., K_t^f and K_t^e). The stochastic nature of demand drives variation in the *shadow costs*, which can lead to equilibrium price paths that are non-monotonic in time. This flexibility is crucial given the variation observed in our data (see Figure 5).[14]

The airline can use its seat-release policy to dampen both *intra*-temporal and *inter*-temporal tradeoffs associated with altering prices. For example, the airline can force everyone to buy economy by not releasing first-class seats in a period and then appropriately adjust prices to capture rents from consumers. Consider the problem of choosing the number of seats to release at each period $\overline{q}_t ;= (\overline{q}_t^e, \overline{q}_t^f) \leq \omega_t$. For a choice of $\overline{q}_t$ in period t, let $p_t(\overline{q}_t) := \{p_t^e(\overline{q}_t), p_t^f(\overline{q}_t)\}$ denote the optimal pricing functions as a function of the number of seats released. Then, the value function can be expressed recursively as

$$V_t(\omega_t, \Psi) = \max_{\overline{q}_t \leq \omega_t} \left\{ \pi_t((p_t(\overline{q}_t), \overline{q}_t), \omega_t; \Psi_t) + \sum_{\omega_{t+1} \in \Omega} V_{t+1}(\omega_{t+1}, \Psi) \times Q_t(\omega_{t+1} | (p_t(\overline{q}_t), \overline{q}_t), \omega_t, \Psi_t) \right\}.$$

The profit function is bounded, so this recursive formula is well defined, and under some regularity conditions, we can show that it is has a unique optimal policy. We present these regularity conditions and the proof of uniqueness in Appendix A.4.

4 Estimation and Identification

In this section, we discuss the parametrization of the model, (method of moments) estimation methodology, and the sources of identifying variation. The model's parametrization balances the dimensionality of the parameters and the desired richness of the demand structure, and the estimation algorithm seeks to limit the number of times we have to solve our model due to its computational burden. At the same time, we seek to avoid strong assumptions on

[14] The model implies a mapping between prices and the unobserved state (i.e., the number of remaining seats in each cabin). This rules out the inclusion of serially correlated unobservables (to the researcher) that shifts demand or costs that could otherwise explain variation in prices.

the relationship between model primitives and both observable (e.g., business travelers) and unobservable market-specific factors. Our identification discussion provides details of the moments we use in the estimation and how they identify each parameter.

4.1 Model Parametrization and Solution

Recall our model primitives, $\Psi = (\{F_b, F_l, F_\xi, c^f, c^e\}, \{\lambda_t, \theta_t\}_{t=1}^T)$, include distributions of valuations for business and leisure passengers, (F_b, F_l), distribution of valuations for 1st-class premium, F_ξ, marginal costs for economy and 1st-class, (c^f, c^e), and the time-varying Poisson arrival rate of passengers, λ_t, and the fraction of business passengers, θ_t.

Motivated by our data, we choose $T = 8$ to capture temporal trends in fares and passenger's reason for travel, where each period is defined as in Table 4. There are two demand primitives, λ_t and θ_t, that vary as the flight date approaches. To permit flexibility in the relationship between time before departure and these parameters, we use a linear parameterization,

$$\theta_t := \min\left\{\Delta^\theta \times (t-1), 1\right\}; \qquad \lambda_t := \lambda + \Delta^\lambda \times (t-1)$$

where Δ^θ, λ, and Δ^λ are scalar constants. This parametrization of the arrival process permits the volume (λ and Δ^λ) and composition (Δ^θ) of demand to change as the flight date approaches, while also limiting the number of parameters to estimate.

There are three distributions (F_b, F_l, F_ξ) that determine passenger preferences. We assume that business and leisure passenger valuations are truncated Normal random variables, F_b and F_l, respectively, left-truncated at zero. Given the disparity in average fares paid by business and leisure passengers, we assume $\mu^b \geq \mu^l$, which we model by letting $\mu^b = \mu^l \times (1 + \delta^b)$ with $\delta^b \geq 0$. The two cabins are vertically differentiated, and passengers weakly prefer first-class to the economy. To capture this product-differentiation, we assume that the quality premium, ξ, equals one plus an Exponential random variable with mean μ^ξ

Finally, we fix the marginal cost of supplying a first-class and economy seat, c^f and c^e, respectively, to equal industry estimates of marginal costs for servicing passengers. Specifically, we set $c^f = 40$ and $c^e = 14$ based on information from the International Civil Aviation Organization, Association of Asia Pacific Airlines, and Doganis (2002).[15] Our estimates and counterfactuals are robust to other values for these costs because the price variation is primarily due to inter-temporal and intra-temporal changes in the endogenous *shadow costs* of seats. In international travel, where the average fare is substantially greater than in domestic

[15] Doganis (2002) finds "peanut costs" in first-class are 2.9 times that of economy, and the average overall cost is \$17.8 (after adjusting for inflation). Given the relative sizes of the economy and first-class cabins for aircraft in our data, this implies costs for servicing one first-class and economy of \$40 and \$14, respectively.

travel, these *shadow costs* are more important than passenger-related services' direct costs.

Given this parametrization of the model, the demand process can be described by a vector of parameters, $\Psi = \left(\mu^l, \sigma^l, \delta^b, \sigma^b, \mu^\xi, \lambda, \Delta^\lambda, \Delta^\theta\right) \in [\underline{\Psi}, \overline{\Psi}] \subset \mathbb{R}^8$. The model is a finite period non-stationary dynamic program. We solve the model for state-dependent pricing and seat-release policies by working backward; computing expected values for every state in the state space, where the state is the number of seats remaining in each cabin. At each state, the optimal policy is the solution to a mixed-integer non-linear program (MINLP) because seats are discrete and prices are continuous controls.[16]

4.2 Estimation

Although the parameterization above in Section 4.1 is for a single flight, or a specific flight at a specific time between two airports, our data represent many diverse fights, and there may be many observed and unobserved factors that impact model primitives in an unknown way. For example the distance or commerce between cities may affect willingness-to-pay for a first-class seat. Instead of further parameterizing the model as a function of observables, we propose a flexible approach to estimate the distribution of flight-level heterogeneity. The approach has the added benefit of limiting the number of times the model is solved.

To illustrate our approach, consider the following example. We have many instances of the SEA-TPE route in our data, and we treat the prices and quantities sold on each instance of this route as a separate flight. Demand in such routes may vary across seasons; for example, there may be a higher willingness to pay for flights in the summer during the tourist season than in the winter. One approach would be to incorporate the observable characteristics of different flights (e.g., season, sporting events, college attendance) and allow them to affect the willingness-to-pay through some functional form.

Instead of relying on any such functional form assumption, we take a different approach and instead estimate a random coefficients model to estimate a distribution of demand primitives across flights. So, two different instances of the SEA-TPE route (two different flights) are allowed to differ in their demand primitives without us imposing any restriction, and the differences due to seasonality in demand will be captured by (parameters of) the distribution. We take this approach because (1) including enough observables to capture differences across flights would result in too many parameters to feasibly estimate the model, and (2) for our counterfactuals, our primary goal is to learn the distribution of demand, and not so

[16] We perform calculations in MATLAB R2020a, using MIDACO's MINLP solver (Schlüter, Gerdts, and Rückmann, 2012). For speed and accuracy, we use warm start points by (a) solving the no-price-discrimination problem first and (b) starting the MINLP problem for each state at the solution to an adjacent state.

much about the relationship between prices and flight-market observable characteristics.

Our approach combines the methodologies of Ackerberg (2009), Fox, Kim, and Yang (2016), Nevo, Turner, and Williams (2016), and Blundell, Gowrisankaran, and Langer (2020). We posit that empirical moments are a mixture of theoretical moments, with a mixing distribution known up to a finite-dimensional vector of parameters. To limit the computational burden of estimating these parameters that describe the mixing distribution, we rely on the importance sampling procedure of Ackerberg (2009). Our estimation proceeds in three steps. First, we calculate moments from the data to summarize the heterogeneity in equilibrium outcomes within and across flights. Second, we solve the model once, at S different parameter values that cover the parameter space $[\underline{\Psi}, \overline{\Psi}]$. Third, we optimize an objective function that matches the empirical moments to the analogous moments for a mixture of candidate data-generating processes. The mixing density that describes across-market heterogeneity in our data is the object of inference.

Specifically, for a given level of observed initial capacity, $\omega_1 := (K_1^f, K_1^e)$, our model produces a data-generating process characterized by parameters that describe demand and costs, $\Psi = \left(\mu^l, \sigma^l, \delta^b, \sigma^b, \mu^\xi, \lambda, \Delta^\lambda, \Delta^\theta\right)$. This data-generating process can be described by a set of N_ρ–many moment conditions that we denote by $\boldsymbol{\rho}(\omega_1; \Psi)$. We assume that the analogous empirical moment conditions, $\rho(\omega_1)$, can be written as a mixture of candidate moment conditions, i.e.,

$$\rho(\omega_1) = \int_{\underline{\Psi}}^{\overline{\Psi}} \boldsymbol{\rho}(\omega_1; \Psi) h(\Psi|\omega_1) d\Psi, \tag{3}$$

where $h(\Psi|\omega_1)$ is the conditional (on initial capacity ω_1) density of the parameters Ψ.[17]

The goal is to estimate the mixing density, $h(\Psi|\omega_1)$, that best matches the empirical moments (left side of Equation 3) to the expectation of the theoretical moments (right side of Equation 3). To identify the mixing density, we assume a particular parametric form for $h(\Psi|\omega_1)$ that reduces the matching of empirical and theoretical moments to a finite-dimensional nonlinear search. Specifically, we let the distribution of Ψ conditional on ω_1 be a truncated multivariate normal distribution, i.e.,

$$\Psi|\omega_1 \sim h(\Psi|\omega_1; \mu_\Psi, \Sigma_\Psi),$$

where μ_Ψ and Σ_Ψ are the vector of means and covariance matrix, respectively, of the non-

[17] Thus, we estimate a separate density for each initial capacity ω_1. This captures the possibility that the observed initial capacities may be correlated with the unobserved demand.

truncated distribution. We choose our estimates based on a least-squares criterion

$$\left(\hat{\mu}_\Psi(\omega_1), \hat{\Sigma}_\Psi(\omega_1)\right) = \arg\min_{(\mu_\Psi,\Sigma_\Psi)} \left(\hat{\rho}(\omega_1) - \widetilde{\mathbb{E}}(\boldsymbol{\rho}(\omega_1;\mu_\Psi,\Sigma_\Psi))\right)^\top \left(\hat{\rho}(\omega_1) - \widetilde{\mathbb{E}}(\boldsymbol{\rho}(\omega_1;\mu_\Psi,\Sigma_\Psi))\right) \quad (4)$$

where $\hat{\rho}(\omega_1)$ is an estimate of the $(M \times 1)$ vector of empirical moments and $\widetilde{\mathbb{E}}(\boldsymbol{\rho}(\omega_1;\mu_\Psi,\Sigma_\Psi))$ is a Monte Carlo simulation estimate of $\int_{\underline{\Psi}}^{\overline{\Psi}} \boldsymbol{\rho}(\omega_1;\Psi)h(\Psi|\omega_1;\mu_\Psi,\Sigma_\Psi)d\Psi$ equal to $\frac{1}{S}\sum_{j=1}^{S}\boldsymbol{\rho}(\omega_1;\Psi_j)$ with the S draws of Ψ taken from $h(\Psi|\omega_1;\mu_\Psi,\Sigma_\Psi)$.[18]

The dimensionality of the integral we approximate through simulation requires a large number of draws. After some experimentation to ensure simulation error is limited for a wide range of parameter values, we let $S = 10,000$. Thus, the most straightforward approach to optimization of Equation 4 would require solving the model $S = 10,000$ times for each value of (μ_Ψ, Σ_Ψ) until a minimum is found. Our model is complex, and the dimensionality of the parameter space to search over makes such an option prohibitive. For this reason, we appeal to the importance sampling methodology of Ackerberg (2009).

The integral in Equation 3 can be rewritten as

$$\int_{\underline{\Psi}}^{\overline{\Psi}} \boldsymbol{\rho}(\omega_1;\Psi)\frac{h(\Psi|\omega_1;\mu_\Psi,\Sigma_\Psi)}{g(\Psi)}g(\Psi)d\Psi,$$

where $g(\Psi)$ is a known well-defined probability density with strictly positive support for $\Psi \in \left[\underline{\Psi}, \overline{\Psi}\right]$ and zero elsewhere like $h(\Psi|\omega_1;\mu_\Psi,\Sigma_\Psi)$. Recognizing this, one can use importance sampling to approximate this integral with

$$\frac{1}{S}\sum_{j=1}^{S}\boldsymbol{\rho}(\omega_1;\Psi_j)\frac{h(\Psi_j|\omega_1;\mu_\Psi,\Sigma_\Psi)}{g(\Psi_j)}$$

where the S draws of Ψ are taken from $g(\Psi)$. Thus, the importance sampling serves to correct the sampling frequencies so that it is as though the sampling was done from $h(\Psi|\omega_1;\mu,\Sigma)$.

The crucial insight of Ackerberg (2009) is that this importance-sampling procedure serves to separate the problem of solving the model from the optimization of the econometric objective function. That is, we solve the model for a fixed number of S draws of Ψ from $g(\Psi)$, and then $\boldsymbol{\rho}(\omega_1;\Psi_j)$ is calculated once for each draw. After these calculations, optimization of the objective function to determine $(\hat{\mu}_\Psi(\omega_1), \hat{\Sigma}_\Psi(\omega_1))$ simply requires repeatedly calculating the ratio of two densities, $\frac{h(\Psi_j|\omega_1;\mu_\Psi,\Sigma_\Psi)}{g(\Psi)}$. To simplify the importance sampling process, we fix

[18] For the consistency of our estimator we assume that, for each initial capacity ω_1, the number of flights and the number of passengers in those flights are sufficiently large so that $\hat{\rho}(\omega_1)$ is a consistent estimator of the true moment $\rho(\omega_1)$, and our importance sampling procedure to determine $\widetilde{\mathbb{E}}(\boldsymbol{\rho}(\omega_1;\mu_\Psi,\Sigma_\Psi))$ is also consistent. For a formal analysis of the subject see Gourieroux, Monfort, and Renault (1993) Proposition 1.

the support of $g(\cdot)$ and $h(\cdot)$ to be the same, and let $g(\cdot)$ be a multivariate uniform distribution with the support $[\underline{\Psi},\overline{\Psi}]$ chosen after substantial experimentation to ensure it encompasses those patterns observed in our data.

To solve Equation 4, we use a combination of global search algorithms and multiple starting values. We repeat this optimization for each ω_1 which provides an estimate of the parameters of the distribution of market heterogeneity, $(\hat{\mu}_\Psi(\omega_1), \hat{\Sigma}_\Psi(\omega_1))$. To calculate the distribution of demand parameters across all flights, we then appropriately weight each estimate by the probability mass associated with that value of ω_1 (Figure 6). We calculate standard errors for the estimates and the counterfactuals by re-sampling the individual passenger observations in the SIAT data. This procedure accounts for error in survey responses as well as variation in our moments across flights. However, this procedure does not account for numerical error coming from the importance-sampling draws, but we argue that $S = 10,000$ is large enough for that to matter.

4.3 Identification

In this section, we introduce the moments we use in Equation 3 to estimate the market heterogeneity, $\Psi \sim h(\cdot|\omega_1; \mu_\Psi, \Sigma_\Psi)$, and present the identification argument that guides our choice. To that end, we present arguments that our moments vary uniquely with each element of Ψ, under the assumption that our data is generated from our model described in Section 3. In showing identification, we use several modeling and parametric assumptions, some of which are necessary, and some are to ease the computational burden and could be relaxed in principle.

Key to our identification is the *shadow* cost associated with each seat in the current period, which equals the expected revenue loss from selling the seat today instead of a future period. These *shadow*-costs depend on the demand and airline's capacity and can vary substantially across time for a flight due to the stochastic nature of demand.[19] Our model maps these *shadow*-costs to observables like prices of economy and first-class seats, price-paths, the timing of the purchase, passenger volumes, and business passengers' share. We use this mapping to construct flight-specific moments for each of these outcomes, which we then pool across flights with similar levels of capacity to construct aggregate moments.[20]

[19] Heuristically, our identification strategy is similar to that of Nevo, Turner, and Williams (2016), who study households that optimize their usage of telecommunications services when facing nonlinear pricing (fixed fee, allowance, and overage price) and uncertainty about their future usage. This uncertainty introduces a *shadow*-price for current usage that is a function of the overage price and probability of exceeding the usage allowance by the end of a billing cycle. If uncertainty is substantial and varies from month to month, it creates variation in shadow costs that provide useful variation to identify the household's preferences.

[20] To construct aggregate moments from flight-specific moments, we use Kernel density weights. These weights are higher (respectively, lower) for flights with similar (respectively, dissimilar) capacities.

This results in a set of empirical moments for each capacity, $\hat{\rho}(\omega_1)$, that we seek to match.

For a given initial capacity ω_1 and each period prior to the departure, we use the following moments conditions: (i) the fares for economy and first-class tickets, for various levels of BTI, which is shown in Figure 4; and (ii) the distribution of the maximum and minimum differences in first-class and economy fares over time, i.e., $\max_{t=1,\ldots,T}\{p_t^f - p_t^e\}$ and $\min_{t=1,\ldots,T}\{p_t^f - p_t^e\}$, respectively; (iii) the proportion of business traveler in each period and the economy/first-class fares, as shown in Figure 3; (iv) the joint distribution of flight-BTI and proportion of total arrivals for different periods; (v) the quantiles of passenger load factor which is shown in Figure 6(b); (vi) number of tickets, for each class, sold at various levels of BTI, which is similar to Figure 3 with the number of seats on the z-axis; and (vii) overall proportion of business travelers, see Figure 2.

Next, we explain why we chose these moments and determine conditions under which a unique set of model parameters rationalizes the data. In particular, we explain how the moments (i) and (ii) identify the willingness-to-pay parameters (i.e., μ^l, σ^l, δ^b, σ^b, μ^ξ) and how the remaining moments from (iii)-(vii) identify the arrival process and passenger mix parameters (i.e., $\lambda, \Delta^\lambda, \Delta^\theta$). For notational ease, we suppress the dependence on ω_1.

4.3.1 Willingness-to-Pay

Moments (i) describe the variation in prices, both within and across flights, and provide information that identifies the parameters that determine the distribution of willingness-to-pay. To see this, consider the decision of an individual of type (v, ξ) who arrives in period t and faces prices (p_t^f, p_t^e). If we assume that the seat-release policy is not binding, the passenger's optimal choice is given by

$$\begin{aligned}
\texttt{first-class, if} &\quad v_i \times \xi - p^f \geq \max\{0, v_i - p_t^e\}\\
\texttt{economy, if} &\quad v_i - p^e \geq \max\{0, v_i \times \xi - p_t^f\}\\
\texttt{do not buy, if} &\quad \max\{v_i \times \xi - p_t^f, v_i - p_t^e\} \leq 0.
\end{aligned}$$

Therefore, the probability of purchase is decreasing in prices, and the rate of decrease depends on the distribution of v. Conditional on purchase, the fraction of passengers buying first-class in a flight at time t is the probability that $v \geq (p_t^f - p_t^e)/(\xi - 1)$, and the fraction buying economy is the probability that $v \leq (p_t^f - p_t^e)/(\xi - 1)$. Because F_b and F_l are time invariant, conditional on knowing the distribution of ξ, the variation in fares and the resulting differences in these probabilities by reason for travel, which in turn vary with flight date, trace the distributions F_b and F_l and reveal (μ^l, σ^l, δ^b, σ^b). Note this implies that we treat both anticipated and unanticipated "demand shocks" as the same, and that any seasonality in our data will affect the variance of the estimate of parameter density.

Next we consider the identification of preference parameters (μ^l, σ^l, δ^b, σ^b) when the seat-release policies are binding. When seats bind in period t, we possibly only observe a subset of passengers. However, the fact that F_b and F_l are time-invariant means that variation in the fares over time is sufficient for the identification of the preference parameters, and thus rationing only affects the identification of the parameters that govern the arrival process (λ_t, θ_t). Moreover, conditional on identifying $(\lambda_t, \theta_t, F_\xi)$, we also have variation in prices across-markets with similar ω_1, which are informative about these parameters. For instance, if there is an increase in the demand for economy tickets relative to business (e.g., Christmas seasonal effect), making the change in fares greater for the economy class than for the first-class tickets, then as long as there is sufficient variation in fares this surge will affect the size of the market, not the willingness-to-pay.

Next, we consider the identification of the distribution $F_\xi(\cdot;\mu^\xi)$, under the assumption that the distribution is known up to the mean parameter, μ^ξ. The moments (ii) use the variation in the extreme differences of fares across cabins and help identify the mean of the quality premium $(\xi - 1)$. Note that for a passenger with (v, ξ) who buys first-class, ξ must be at least $(p_t^f - p_t^e)/v$, and for a passenger with (v, ξ) who buys economy ξ must be at most $(p_t^f - p_t^e)/v$. Comparing across all passengers and all times gives

$$\frac{\max_t (p_t^f - p_t^e)}{\min\{v : \texttt{bought first-class}\}} \leq (\xi - 1) \leq \frac{\min_t (p_t^f - p_t^e)}{\max\{v : \texttt{bought economy}\}},$$

where, for example, $\min\{v : \texttt{bought first-class}\}$ and $\max\{v : \texttt{bought economy}\}$ are the minimum and maximum value among those who buy first-class and economy, respectively. Thus, moments capturing the covariation between cabin-specific quantities and the price differential across cabins identify μ^ξ. Implicitly, we are assuming that for some t, the expected maximum and the expected minimum fare differences are the same, in which case the above inequalities become equal. Although we assume that ξ is an Exponential, this assumption is not necessary, we could have used other distribution, but we can only identify the mean μ^ξ.

4.3.2 Arrival Process and Passenger Mix

Moments (iii)-(v) capture within-flight price dispersion as the flight date approaches and identify the arrival rate. Recall our assumption that $\lambda_t = \lambda + \Delta^\lambda \times (t-1)$ so at the initial period $\lambda_1 = \lambda$, and given our Poisson assumption λ is an average number of passengers that arrive in the first period, where seat-release policy is less likely to be binding. Although we do not observe the number of passengers that arrive in $t = 1$, we can exploit the observed price-path, in particular, its value at $t = 1$ and its slope because, all else equal, the shadow cost of a seat is proportional to λ and Δ^{λ}. We also use the fact that while the arrival process is linear in t, prices are not.

Heuristically, higher demand (relative to the capacity) manifests itself in the form of a more significant increase in prices, over time, because it implies more significant variation in the opportunity cost of a seat due to more substantial "surprises" in the number of sales at given prices. Thus, the monotonic relationship between the size of the demand and variability in price paths for a given flight suggests that we can use the dispersion of price paths from their initial levels to identify $(\lambda, \Delta^{\lambda})$. For instance, suppose λ is small but Δ^{λ} is high.

Thus for the identification, we have relied on the assumption that the arrival process is Poisson, and passengers are myopic because the latter assumption implies that higher dispersion in paid fares as the departure gets closer is only due to the increasing share of business passengers and their preferences. This data feature suggests that while we can relax the Poisson assumption and use the Negative Binomial, say, to model the arrival process, it is essential that the rate of change in arrival, Δ^{λ}, is constant and passengers are myopic.

To identify the passenger mix, θ_t, we use both the reason to travel and the covariation between deviations of fares from their initial level, which is captured by the moment conditions (vi) and (vii). To see why we need the former, note that we have

$$\Pr(\texttt{business}) = \Pr(\texttt{business|buy}) \times \Pr(\texttt{buy}) + \Pr(\texttt{business|not-buy}) \times \Pr(\texttt{not-buy}),$$

where $\Pr(\texttt{business|buy})$ is estimable, and given λ we know $\Pr(\texttt{buy})$ and $\Pr(\texttt{not-buy})$, but $\Pr(\texttt{business|not-buy})$ is unknown. Business travelers have higher mean willingness-to-pay than leisure travelers, so $\Pr(\texttt{leisure|not-buy}) > \Pr(\texttt{business|not-buy})$. This selection is, however, the smallest at $t = 1$ (Figure 1(b)) and we assume that $\Pr(\texttt{business|not-buy}) = 0$ at $t = 1$. Then, using the assumption that $\theta_t = \min\{\Delta^{\theta}(t-1), 1\}$, which gives us $\theta_1 = 0$, and the number of sales in $t = 1$ we can determine $\Pr(\texttt{buy})$. In other words, we rely on the assumption that at the start $t = 1$ there are no business travelers – see Figure 1. This probability varies over time, and even though it is a function of both preferences and Δ^{θ}, the only reason for its time variation is Δ^{θ}, which in turn is captured in variation in price paths (Figures 3 and 5). So the distribution of changes in fares relative to the initial fare is informative about Δ^{θ}.

5 Results

In this section, we present our estimation results. First, we discuss how our estimates capture sources of across-market heterogeneity. Second, we calculate the distribution of opportunity costs for a seat and show how they vary across cabins and time until departure. We discuss model fit in Appendix A.3.

5.1 Market Heterogeneity

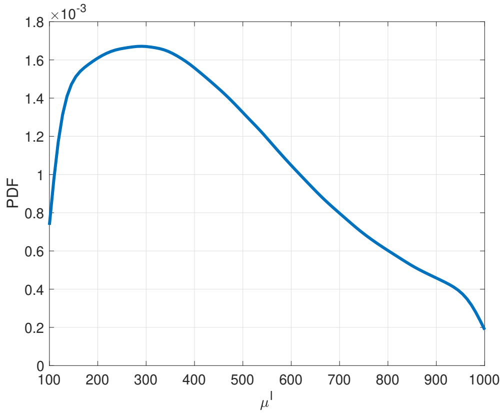
Recall that we estimate means and covariances of the parameters' distribution across all markets in our sample. In Table 5 we present the means (i.e., the first moments) of these distributions averaged across all of our markets, and means for the market with the modal initial-capacity, of 117 economy seats and 15 first-class seats. The mean willingness-to-pay for a (one-way) economy class seat across our entire sample is \$413, and the modal market is \$508. The mean coefficient of variation of willingness-to-pay is 0.608, and the mean business traveler values an economy class ticket 22.6% more than the mean leisure passenger. We estimate that, on average, about 23 passengers show up in the first period (110+ days before the flight) and, on average, arrivals decrease by 7% each period while the fraction of business passengers increases by 7% each period (across eight periods).

Table 5: Demand Estimates

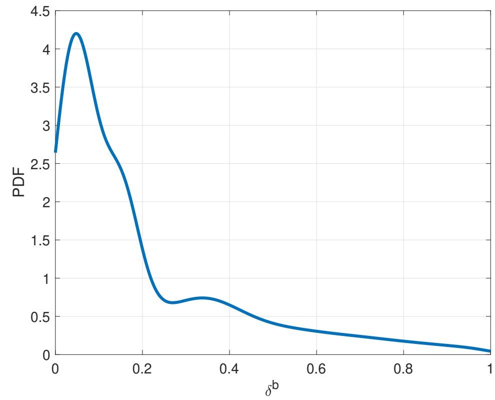
Demand Parameters	Mean All Markets (1)	Mean Markets with Modal Capacities (2)
μ^l	413.234	508.054
	(19.539)	(30.942)
$\frac{\sigma^l}{\mu^l}$	0.608	0.247
	(0.04)	(0.165)
δ^b	0.226	0.116
	(0.011)	(0.04)
$\frac{\sigma^b}{\mu^b}$	0.353	0.371
	(0.044)	(0.145)
μ^ξ	0.230	0.267
	(0.042)	(0.004)
λ	23.318	18.119
	(1.019)	(3.385)
Δ^λ	-0.071	-0.052
	(0.003)	(0.011)
Δ^θ	0.077	0.071
	(0.003)	(0.019)

Note: In this table we present the mean demand parameters from the marginal distribution $h(\Psi; \widehat{\mu}_\Psi, \widehat{\Sigma}_\Psi)$ (in column (1)) and the conditional density $h(\Psi|\omega_1^*; \widehat{\mu_\Psi}, \widehat{\Sigma_\Psi})$, where ω_1^* is the modal capacities (in column (2)). In our sample, a market with modal capacities has 115 economy class seats and 14 1st-class seats. Bootstrapped standard errors are in the parentheses.

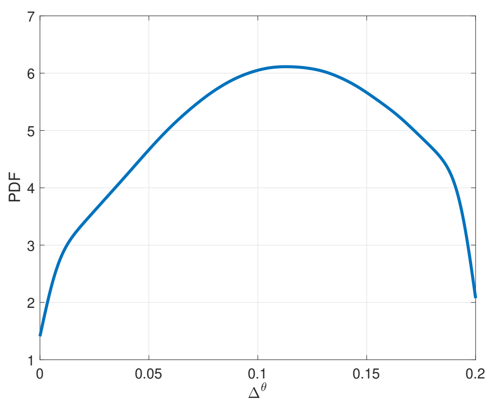
Figure 9: *Market Heterogeneity: Marginal Densities of Demand Parameters*



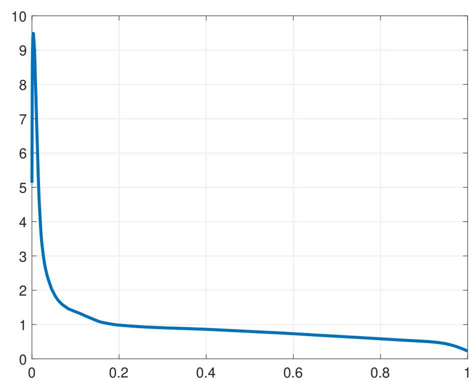
(a) PDF of μ^l



(b) PDF of δ^b



(c) PDF of Δ^θ



(d) PDF of μ^ξ

Note: This figure displays the marginal densities for (a) leisure travelers' average willingness-to-pay for an economy seat (μ^l); (b) difference in business travelers' average willingness-to-pay for an economy seat μ^b and μ^l expressed as a ratio of μ^l (δ^b), (d) change in the share of business travelers each period (Δ^θ), and (d) the average additional valuation for a first-class seat relative to an economy seat (μ^ξ).

However, we have many different markets in our sample. To get a sense of the heterogeneity across these markets, in Figure 9 we show the marginal densities of four parameters (out of the eight parameters in Ψ). The main takeaway is that there is, indeed, substantial heterogeneity in demand across markets. For example, from Figure 9-(a) we can see that the mode value of μ^ℓ is approximately \$300, which is close to the market average of \$413 as shown in Table 5(1). However, there is a long upper tail, with a positive mass at the upper bound value of \$1,000. In Figure 9(b), we present the density of δ^b that determines the difference in willingness-to-pay between leisure and business passengers as $\mu^b = \mu^l \times (1+\delta^b)$. We present the rate of change in the fraction of business passenger arrivals in panel (c). This distribution implies substantial heterogeneity in the fraction of business arrivals.

We present the density of the taste for first-class service (ξ) in panel (d). On average, passengers' willingness-to-pay for a first-class seat is 23% more than for an economy seat, with substantial heterogeneity across passengers. There is a mass close to zero, which implies thre is meaningful cross-cabin substitution, but also a long tail, which implies there is substantial potential surplus for airlines to capture.

In Appendix A.2, Table A.2.1 we display the full variance-covariance estimates for the modal market. And in Tables A.2.3 and A.2.2 we display mean of the demand parameters across all twenty initial capacities for which we estimated the model.

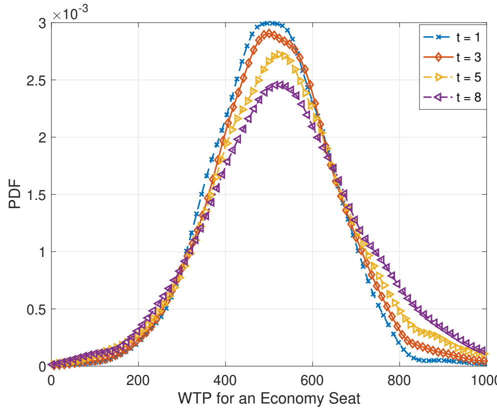
5.2 Implications of Demand Estimates

Using our estimates, we can determine the implied densities of the willingness-to-pay for an economy seat and a first-class seat and how these densities change over time. Recall that the density of the willingness-to-pay for an economy seat in period t is the mixture $\theta_t \times f_v^b(\cdot) + (1-\theta_t) \times f_v^l(\cdot)$, and the density of the willingness-to-pay for an first-class seat in period t is similar to the economy seat augmented by $f_\xi(\cdot)$. These densities, for periods $t = 1, 3, 5, 8$ are displayed shown in Figure 10. As expected, densities of willingness-to-pay for a first-class seat is "shifted" to the right of the densities of willingness-to-pay for an economy seat by ξ. Using the estimates for λ_t and θ_t from column 2 of Table 5, we can also determine the average number of arrivals by their reason to travel. From the definition of $\lambda_t = \lambda + \Delta^\lambda \times (t-1)$ we find the average number of arrival is 18 in period $t = 1$, and every period decreases at the rate of $\Delta^\lambda = -0.052$. The share of business travelers increases at a rate of $\Delta^\theta = 0.071$.

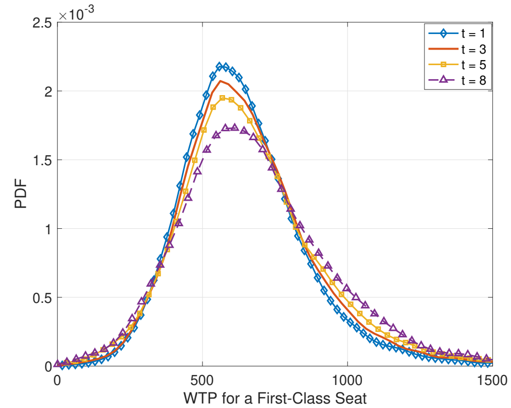
It is also illustrative to consider what these parameters imply about the total (opportunity or shadow) marginal cost of a seat. The total marginal cost of a seat comprises its "peanut" cost, which is constant, and the opportunity cost varies over time depending on the state's evolution, i.e., the number of unsold economy seats and first-class seats. The shadow cost is the right-hand side of Equation 2, the change in expected value for a change in today's price. In other words, the shadow cost is the cost of future revenues to the airline of selling an additional seat today.

In Figure 11 we present the state's evolution, in terms of the contours corresponding to the state's joint density, as implied by our model estimates. Consider ω_1, which is the initial capacity for this modal capacity market. So, when we move to the next few periods, we see that the uncertainty increases. However, as we get closer to the departure time, the contours move towards the origin, which denotes that fewer seats might remain unsold with time. The contour of the state at the time of departure ($\omega_{\texttt{dept}}$) denotes the distribution of

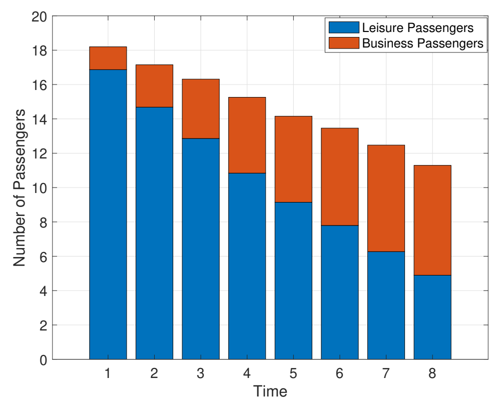
Figure 10: *Willingness-to-Pay for a Seat by Cabin and the Arrivals, for Modal Capacity*



(a) PDF of WTP for an Economy Seat



(b) PDF of WTP for a First-Class Seat



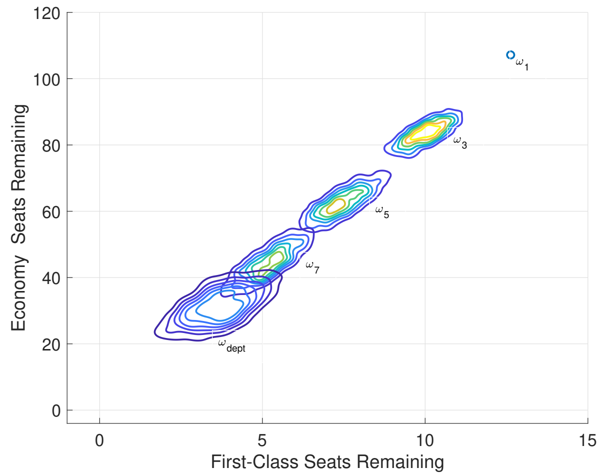
(c) Number of Arrivals, over time

Note: This figure displays, for the modal capacity and mean market, the implied densities of WTP for an economy seat and WTP for a first-class seat, and the number of arrivals by their purpose of travel

the state at the time of departure. The fact that even at the time of departure, the contours are not degenerate is consistent with the load factor observed in our sample, see Figure 6(b). Thus we can conclude that there is substantial uncertainty (or volatility) about demand.

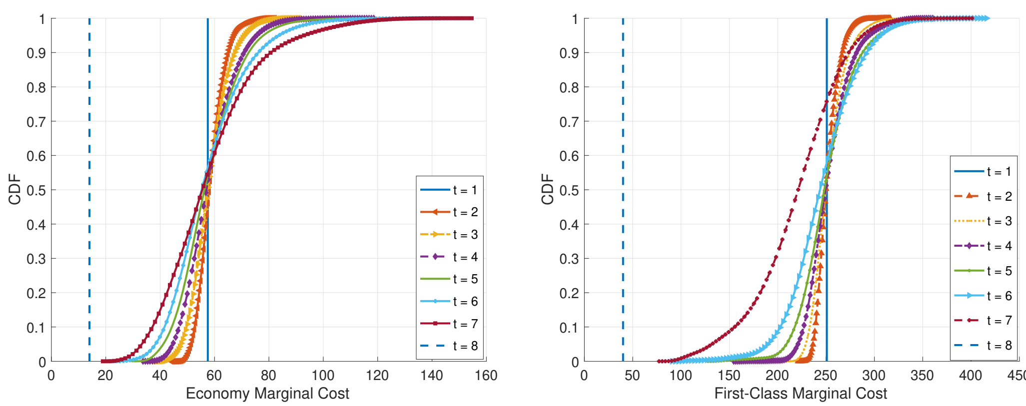
One of the implications of this demand volatility is the implied volatility in the value of a seat to the airline, i.e., the seat's opportunity cost. In Figure 12, we present the distributions of the marginal cost for an economy and first-class seat that are realized in equilibrium, averaged across all markets and all capacities. This feature graphically relates the state transitions to the shadow cost of a seat. In particular, in Figure 12, we take the distribution of states realized in a given period (Figure 11) and sample the total marginal costs (sum of the derivative of value functions with respect to price and the "peanut" cost) based on those frequencies and then plot the distributions. In panel (a), we present the

Figure 11: *Evolution of Seats Remaining*



Note: The figure displays the contours corresponding to the joint density of unsold seats for every period, over all capacities and markets.

Figure 12: *Distributions of marginal (opportunity) cost of a seat*



(a) Economy (b) First-Class

Note: Panels (a) and (b) display the distributions of opportunity cost distributions at each time, for an economy and a first-class seat, respectively.

distributions for an economy seat, and in panel (b), we present the distributions of a first-class seat. As can be seen, there is a significant variation in the costs. These variations are crucial for our identification as they are the underlying reason for dispersion in the observed fares.

In the first period, $t = 1$, there is no uncertainty about the state, which in turn means the marginal cost is degenerate at \$57.64 for an economy seat and \$217.74 for a first-class seat. With $t > 1$, the distributions become more dispersed but with little change in the mean.

For instance, the means of the marginal costs for an economy seat in periods $t = 3, 4, 5$ and 7 are \$84.54, \$84.84, \$84.98, and \$84.87, respectively. In contrast, the variances increase substantially from \$68 to \$100 to \$132 to \$298, during the same periods. We observe similar pattern for a first-class seat; the mean marginal costs are, approximately, \$245, \$239, \$243 and \$238 in periods $t \in \{3, 4, 5, 7\}$, respectively. And in the same period the variance increases from \$570 in $t = 3$ to \$851 in $t = 7$. Finally, in the last period, $t = 8$, the opportunity cost of a seat is zero as the marginal cost is only the peanut cost.

6 Inefficiency and Welfare

There are two sources of informational frictions in this market that contribute to inefficiency: asymmetric information about passengers' valuations and uncertainty about future demand. A passenger's valuation may be due to idiosyncratic preferences and may also be associated with their reason for travel. Airlines' inability to price based on a passenger's reason for travel or even the idiosyncratic valuation can distort the seats' final allocation.[21] The second source leads to inefficient allocations of limited capacity because the airline chooses its prices and seat-release policies before the demand is realized. Intra-temporal and inter-temporal misallocation introduced by these frictions represent opportunities for welfare-improving trade. Using counterfactual pricing and allocation mechanisms, we quantify the inefficiencies attributable to these sources. We first show how to visualize these sources of inefficiencies using a schematic representation of a welfare triangle. We then present and discuss our results.

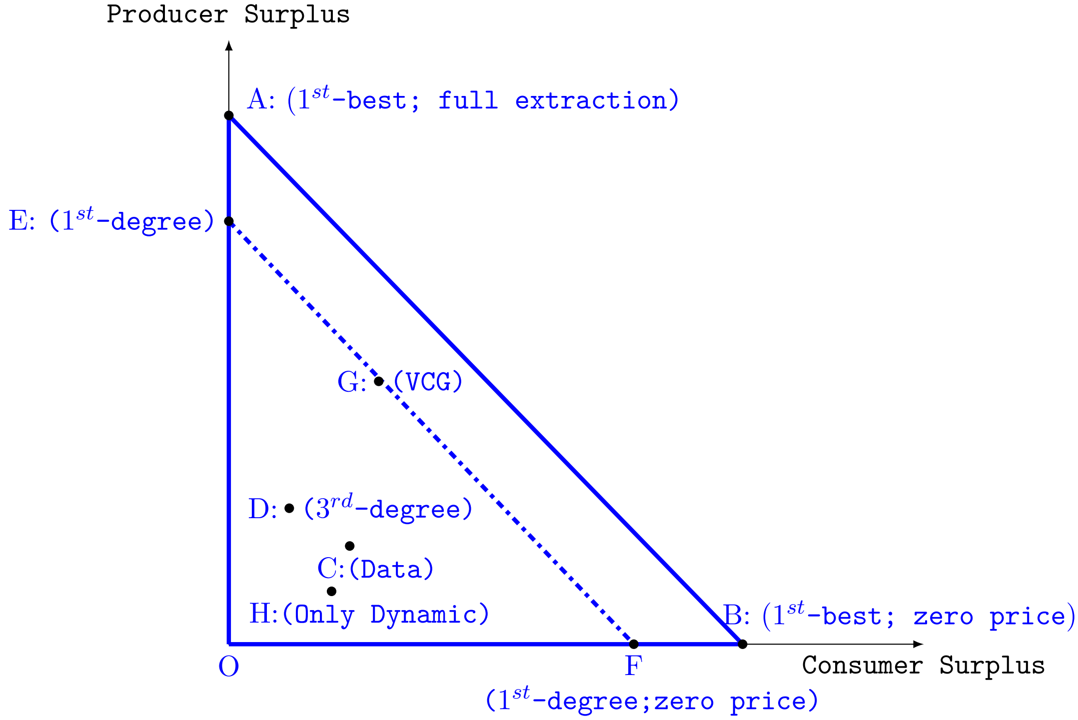
6.1 Welfare Triangle

Consider the first-best allocation: seats are allocated to the highest valuation passengers (v, ξ) regardless of the timing of their arrival to the market. Under this allocation, the division of surplus would depend on the prices. Figure 13 shows the line between A and B associated with this efficient benchmark and forms the welfare triangle (OAB). Point A represents the full extraction of consumer surplus (i.e., price equals valuation), and point B represents maximum consumer surplus (i.e., prices equal the peanut costs).

Point C ("Data") in Figure 13, which is in the interior of the triangle OAB, denotes the division of surplus resulting from current pricing practices by airlines that we observe in our data and that we use to estimate our model. This outcome is preferable to no trade

[21] This source of inefficiency is present in a static monopoly setting when the monopolist chooses one price. However, in our setting, the monopolist must consider the unknown valuations of future arriving passengers.

Figure 13: *Welfare Triangle*



Note: A schematic representation of the welfare triangle under a different pricing regime.

for both the airline and consumers but is strictly inside the welfare frontier due to the two inefficiency sources discussed above. The distance of C from A–B illustrates the magnitude of welfare-improving opportunities relative to current practice.

Suppose the airline did not employ second-degree price discrimination each period before the flight, and only changed prices across time. Or, in other words, the airline did not exploit the difference in quality between an economy seat and a first-class seat when choosing its prices, and each period chose one price for both cabins.[22] The airline would still adjust the price each period depending on the opportunity cost of a seat. This counterfactual of choosing one price across cabins corresponds to point H in Figure 13 ("Only Dynamic"). While the producer surplus under H will be lower than the producer surplus under C, the effect on consumer surplus is theoretically ambiguous. Choosing one price across two cabins should improve welfare for those who buy first-class under the current prices, but it should lower welfare because economy class seats become expensive and total sales will adjust.

Airlines do not observe arrivals' reasons for travel, limiting their ability to price based on

[22] Formally, airline solves $\max_{\{p_t,q_t\}_{t=1}^T} \sum_{t=1}^T \mathbb{E}_t \{\pi(\chi_t, \omega_t; \Psi_t)\}$, where the flow profit is $\pi(\chi_t, \omega_t; \Psi_t) = (p_t - c^f) \times \mathbb{E}(q_t^f | (p_t, q_t)) + (p_t - c^e) \times \mathbb{E}(q_t^e | (p_t, q_t))$, and $\mathbb{E}(q_t^f | (p_t, q_t))$ and $\mathbb{E}(q_t^e | (p_t, q_t))$ are expected demands.

the difference between business and leisure arrivals' willingness-to-pay, possibly resulting in the exclusion of leisure arrivals on account of expectations of greater demand from business arrivals. Permitting the airline to price based on the reason for travel, i.e., third-degree price discrimination, can increase profits for the airline, but the implication for passengers is ambiguous. Leisure passengers may benefit, but it may come at the cost of business passengers. Since leisure and business travelers arrive at different times, and the airline faces capacity concerns, the change in consumer surplus depends on the entire demand process. Furthermore, the number of total seats sold may increase or decrease, affecting the total welfare. Point D ("3^{rd}-degree") in Figure 13 represents a division of welfare when the airline can charge different prices based on passenger's reason for travel and one seat-release policy for each cabin.

Even with the airline's ability to price based on the reason for travel, asymmetric information about idiosyncratic valuations can create inefficiencies. For example, some leisure passengers may have unusually high valuations, and some business travelers may have unusually low valuations. To ascertain the importance of this information asymmetry, we consider a setting where the airline practices first-degree price discrimination. The airline observes valuations each period and decides which arrivals to accommodate, charging each arrival its valuation. However, the airline is still uncertain about future demand realizations. This outcome corresponds to point E in Figure 13. Likewise, point F in Figure 13 corresponds to the first-degree allocation of seats but with the price equal to the peanut cost.

The (dotted) line that joins E and F is informative about the extent of dynamic inefficiency in the market. In particular, the line E–F represents the frontier of the welfare triangle (OEF) when the airline knows (v, ξ) for passengers in a given period but cannot foresee future realizations of the demand process. One way to divide the surplus along the E-F frontier is by implementing Vickery-Clarke-Groves auctions every period. Such a division of surplus is denoted by point G ("VCG") in Figure 13. Thus, the set of potential outcomes in OAB but not in OEF represents lost surplus due to inter-temporal demand uncertainty. One could envision a secondary-market run by the airline that could resolve these dynamic inefficiencies, and our estimates provide the value that could be created by such an exchange.

6.2 Counterfactual Results

Table 6 presents the welfare estimates under all of these alternative pricing strategies, averaged across all markets and capacities observed in our data. Columns of Table 6 are indexed by a letter (e.g., C) that corresponds to the point in the welfare triangle of Figure 13.

Current Pricing. Recall that point C in Figure 13 denotes the division of surplus resulting from the airlines' current pricing practices. The surplus associated with this pricing strategy is in the first column of Table 6. Total surplus is $56,050 for the average flight, with 69% of the surplus going to the airline. This outcome is preferable to no trade for both the airline and consumers but is strictly inside the welfare frontier due to the two inefficiency sources discussed above. Comparing the total surpluses under columns C and A of Table 6 we find that the surplus associated with current pricing represents 89% of the average market's potential attainable surplus.

Table 6: Price Discrimination Counterfactuals

	Data C	Only Dynamic H	3^{rd}-Degree D	1^{st}-Degree E	VCG G	1^{st}-Best $\overline{\text{AB}}$
Producer Surplus	38,654	28,475	38,903	56,109	14,987	-
	(2,297)	(1,571)	(2,315)	(4,031)	(1,564)	
Consumer Surplus	17,396	12,977	17,042	-	41,122	-
	(1,714)	(1,127)	(1,697)		(2,886)	
– **Business**	5,416	4,621	5,192	-	10,516	-
	(474)	(417)	(438)		(587)	
– **Leisure**	11,980	8,355	11,850	-	30,605	-
	(1,510)	(1,021)	(1,388)		(2,724)	
Total Surplus	56,050	41,452	55,946	56,109	56,109	63,226
	(3,917)	(2,619)	(3,904)	(4,031)	(4,031)	(4,728)

Note: In this table we present measures of welfare for six different outcomes, corresponding to points A-H in Figure 13, where Column C is the data. These calculations are performed for all market types receiving positive weight for a given capacity, then averaged across types and capacities. Bootstrapped standard errors are in parentheses.

Only Dynamic Pricing. When we restrict the airline to choose only one price each period for both cabins, as expected, it lowers the total welfare. Note that the airline can still change prices over time. Producer surplus is 74% of the baseline (Column C), and the total surplus is 74% of the baseline. Although the airline's ability to use second-degree price discrimination to screen passengers between cabins (from H to C) increases the total surplus, airlines capture roughly two-thirds of the additional surplus.

Group Pricing: Business versus Leisure. Column D of Table 6 provides surplus estimates when the airlines are permitted to price based on the reason for travel, i.e., third-degree price discrimination. Relative to current pricing practice (i.e., Column C), we find airline surplus increases by about one percent and consumer surplus falls by about the same, leaving total surplus nearly unchanged. Thus, group pricing based on reason to travel slightly increases revenue but lowers the total surplus.

Static versus Dynamic Inefficiencies. We begin with the surplus under the first-degree price discrimination in column E of Table 6, where the airline can price equal to the arrivals' willingness-to-pay but still faces uncertain future demand. By construction, the airline can capture the entire surplus. However, the total surplus is only slightly higher than the second and third-degree price discrimination (C and D of Table 6, respectively). We find that a VCG auction would result in relatively low prices, or that consumers would capture 73% of the surplus in the presence of a period-by-period VCG auction. Thus, these results suggest that airlines' increased effort to learn passenger information to group price discrimination does not increase total surplus in the market and instead only transfers surplus from passengers to the airline. Comparing total surpluses under C and A, we find that stochastic demand and asymmetric information lead to approximately 11% loss of welfare. Comparing E and A, we find that almost all of this inefficiency is due to stochastic demand.

7 Conclusion

We develop a model of intra-temporal and inter-temporal price discrimination by airlines that sell a fixed number of seats of different quality to heterogenous consumers arriving before a flight. We specify demand as non-stationary and stochastic, which accommodates the salient features of airline pricing. Using unique data from international airline markets, we flexibly estimate the distribution of preferences for flights. The estimation exploits the relationship between a passenger's seat chosen, timing of purchases, reasons for travel, and the fare paid to identify how effectively airlines discriminate using sources of passenger heterogeneity. We find that the flexibility of the model and estimation algorithm are successful in capturing key features of our data.

Next, through several counterfactual exercises, we use the estimates to explore the role that stochastic demand and asymmetric information have on efficiency and the distribution of surplus. We find that current pricing practices result in substantial inefficiencies relative to the first-best outcome. In particular, total welfare is only 89% of the welfare without demand uncertainty and asymmetric information. To isolate the role of different sources of asymmetric information in determining welfare, we solve for optimal seat-release and prices when the airline can discriminate based on passengers' reason to travel, and also when the airline can observe their preferences. The first case (i.e., third-degree price discrimination) achieves 88% of the first-best welfare, representing a 1% decrease from current practices. Business passengers' and leisure passengers' surpluses decrease due to the loss of informational rent and reduction in seats sold. The second case (i.e., first-degree price discrimination) where the only remaining source of inefficiency is inter-temporal demand uncertainty, has an in-significant effect on welfare compared to the first case. Thus, demand uncertainty accounts for almost all of the total welfare loss, while asymmetric information accounts for none.

There are many avenues for future research on related topics. First, like other studies of dynamic pricing, we model a monopolistic market structure that accurately reflects our data. This limits our ability to examine the impact of competition on discriminatory-pricing practices. Another interesting path for future research is to consider the possibility that consumers are strategic in their purchasing decisions. While this is difficult to conclude with our data, purchases of numerous goods are increasingly made online, which allows firms to track search behavior and adapt pricing accordingly. Given the growing theoretical literature on this topic (e.g., Board and Skrzypacz (2016) and Dilme and Li (Forthcoming)) that yield testable implications from strategic behavior by consumers, empirical studies like ours and Sweeting (2010) represent an opportunity to offer insight to future modeling efforts.

Relatedly, as firms gather more information about the preferences and purchasing habits of consumers, exploitation of this information becomes an important concern. For more on the role of privacy and efficiency, see Hirshleifer (1971) and Posner (1981). Although there are few papers that study the role of privacy (Taylor, 2004; Calzolari and Pavan, 2006) in price discrimination, more empirical research in this topic is needed. Future research should seek to understand the trade-off between efficiency and privacy, especially in industries where firms have greater access to such information.

References

Ackerberg, Daniel. 2009. "A New Use of Importance Sampling to Reduce Computational Burden in Simulation Estimation." *Quantitative Marketing and Economics* 7 (4):343–376.

Alderighi, Marco, Marcella Nicolini, and Claudio A. Piga. 2015. "Combined Effects of Capacity and Time on Fares: Insights from the Yield Management of a Low-Cost Airline." *The Review of Economics and Statistics* 97 (4):900–915.

Armstrong, Mark. 2006. "Recent Developments in the Economics of Price Discrimination." In *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, edited by Richard Blundell, Whitney K. Newey, and Torsten Persson. 97–141.

Aryal, Gaurab and Maria. F. Gabrielli. 2019. "An Empirical Analysis of Competitive Nonlinear Pricing." *International Journal of Industrial Organization* 68:102538.

Barnhart, Cynthia, Peter Belobaba, and Amedeo R. Odoni. 2003. "Application of Operations Research in the Air Transport Industry." *Transportation Science* 37 (4):368–391.

Bergemann, Dirk, Benjamin Brooks, and Stephen Morris. 2015. "The Limits of Price Discrimination." *American Economic Review* 105 (3):921–957.

Berry, Steven, Michael Carnall, and Pablo T. Spiller. 2006. "Airline Hubs: Costs, Markups and the Implications of Customer Heterogeneity." *Competition policy and antitrust* :183–213.

Blundell, Wesley, Gautam Gowrisankaran, and Ashley Langer. 2020. "Escalation of Scrutiny: The Gains from Dynamic Enforcement of Envrionmental Regulations." *American Economic Review* 110:2558–25585.

Board, Simon and Andrzej Skrzypacz. 2016. "Revenue Management with Forward-Looking Buyers." *Journal of Political Economy* 124 (4):1046–1087.

Borenstein, S. 2010. "An Index of Inter-City Business Travel for Use in Domestic Airline Competition Analysis." *NBER Working Paper* .

Borenstein, Severin and Nancy L. Rose. 1994. "Competition and Price Dispersion in the US Airline Industry." *Journal of Political Economy* 102 (4):653–683.

Busse, Meghan and Marc Rysman. 2005. "Competition and price discrimination in yellow pages advertising." *RAND Journal of Economics* :378–390.

Calzolari, Giacomo and Alessandro Pavan. 2006. “On the Optimality of Privacy in Sequential Contracting.” *Journal of Economic Theory* 130 (1):168–204.

Chandra, Ambarish and Mara Lederman. 2018. “Revisiting the Relationship between Competition and Price Discrimination.” *American Economic Journal: Microeconomics* 10 (2):190–224.

Cho, Sungjin, Gong Lee, John Rust, and Mengkai Yu. 2018. “Optimal Dynamic Hotel Pricing.” *Mimeo* .

Ciliberto, Federico and Jonathan W. Williams. 2014. “Does Multimarket Contact Facilitate Tacit Collusion? Inference on Conduct Parameters in the Airline Industry.” *RAND Journal of Economics* 45 (4):764–791.

Council of Economic Advisors. 2015. “Big Data and Differential Pricing.” Tech. rep., United States, The White House.

Courty, Pascal and Hao Li. 2000. “Sequential Screening.” *Review of Economic Studies* 67 (4):697–717.

Crawford, G.S. and Matthew Shum. 2006. “Monopoly Quality Degradation in Cable Television.” *Journal of Law and Economics* 49 (1):181–209.

Dana, James D. 1999. “Using Yield Management to Shift Demand when the Peak TIme is Unknown.” *RAND Journal of Economics* 30 (3):456–474.

Dilme, F. and F. Li. Forthcoming. “Revenue Management without Commitment: Dynamic Pricing and Periodic Fire Sales.” *Review of Economic Studies* .

Doganis, Rigas. 2002. *Flying Off Course: The Economics of International Airlines*. Routledge, 3rd ed.

Escobari, Diego. 2012. “Dynamic Pricing, Advance Sales and Aggregate Demand Learning in Airlines.” *The Journal of Industrial Economics* 60 (4):697–724.

Fox, Jeremy, K. Kim, and Chenyu Yang. 2016. “A Simple Nonparametric Approach to Estimating the Distribution of Random Coefficients in Structural Models.” *Journal of Econometrics* 195 (2):236–254.

Gale, David and Hukukane Nikaido. 1965. “The Jacobian Matrix and Global Univalence of Mappings.” *Mathematische Annalen* 159 (2):81–93.

Gale, Ian L and Thomas J Holmes. 1993. "Advance-purchase discounts and monopoly allocation of capacity." *The American Economic Review* :135–146.

Gourieroux, Christian, A. Monfort, and Eric Renault. 1993. "Indirect Inference." *Journal of Applied Econometrics* 8:S85–S118.

Graddy, Kathryn and George Hall. 2011. "A Dynamic Model of Price Discrimination and Inventory Management at the Fulton Fish Market." *Journal of Economic Behavior & Organization* 80 (1):6–19.

Hendel, Igal and Aviv Nevo. 2013. "Intertemporal Price Discrimination in Storable Goods Markets." *American Economic Review* 103 (7):2722–2751.

Hirshleifer, J. 1971. "The Private and Social Value of Information and the Reward to Incentive Activity." *American Economic Review* 61 (4):561–574.

Ivaldi, Marc and David Martimort. 1994. "Competition under Nonlinear Pricing." *Annales d'Economie et de Statistique* 34:71–114.

Jian, Kang. 2012. *When Are the Tickets Cheapest? A Structural Model of an Airline's Dynamic Pricing under Fixed Capacity.* Ph.D. thesis, University of Virginia.

Johnson, Oliver. 2007. "Log-concavity and the Maximum Entropy Property of the Poisson Distribution." *Stochastic Processes and their Applications* 117 (6):791–802.

Kehoe, Patrick J., Bradley J. Larsen, and Elena Pastorino. 2020. "Dynamic Competition in the Era of Big Data." *Working Paper* .

Lazarev, John. 2013. "The Welfare Effects of Intertemporal Price Discrimination: An Empirical Analysis of Airline Pricing in U.S. Monopoly Markets." *Working Paper* .

Leslie, Phillip. 2004. "Price Discrimination in Broadway Theater." *RAND Journal of Economics* 35 (3):520–541.

Li, Jun, Nelson Granados, and Serguei Netessine. 2014. "Are Consumers Strategic? Structural Estimation from the Air-Travel Industry." *Management Science* 60 (9):2114–2137.

McCartney, Scott. 2016. "Going Once, Going Twice: Airlines Auction Seat Upgrades." *The Wall Street Journal* .

McManus, Brian. 2007. "Nonlinear Pricing in an Oligopoly Market: The Case of Specialty Coffee." *Rand Journal of Economics* 38 (2):512–532.

Nair, Harikesh. 2007. “Intertemporal Price Discrimination with Forward-Looking Consumers: Application to the US Market for Console Video-Games.” *Quantitative Marketing and Economics* 5 (3):239–292.

Nevo, Aviv, J. Turner, and Jonathan W. Williams. 2016. “Usage-Based Pricing and Demand for Residential Broadband.” *Econometrica* 84 (2):411–443.

Nevo, Aviv and Catherine Wolfram. 2002. “Why Do Manufacturers Issue Coupons? An Empirical Analysis of Breakfast Cereals.” *RAND Journal of Economics* 33 (2):319–339.

Nicas, Jack. 2013. “How Airlines Mine Personal Data In-Flight.” *The Wall Street Journal* .

Pigou, A. C. 1920. *The Economics of Welfare*. London: McMillan, 1 ed.

Posner, Richard A. 1981. “The Economics of Privacy.” *American Economic Review* 71 (2):405–409.

Puller, Steven L., Anirban Sengupta, and Steven N. Wiggins. 2012. “Does Scarcity Drive Intra-route Price Dispersion in Airlines?” *NBER Working Paper* 15555.

Schlüter, Martin, Matthias Gerdts, and Jan-J. Rückmann. 2012. “A Numerical Study of MIDACO on 100 MINLP Benchmarks.” *Optimization* 61 (7):873–900.

Stokey, Nancy. 1979. “Intertemporal Price Discrimination.” *Quarterly Journal of Economics* 93 (3):355–371.

Sweeting, Andrew. 2010. “Dynamic Pricing Behavior in Perishable Goods Markets: Evidence from Secondary Markets for Major League Baseball Tickets.” *Journal of Political Economy* 120 (6):1133–1172.

Taylor, C. 2004. “Consumer Privacy and the Market for Customer Information.” *RAND Journal of Economics* 35 (4):631–650.

Tully, Shawn. 2015. “Southwest Bets Big on Business Travelers.” *Fortune* .

van Ryzin, Garrett and Kalyan Talluri. 2005. “An Introduction to Revenue Management.” In *Emerging Theory, Methods, and Applications*, chap. 6. INFORMS, 142–194.

Varian, Hal R. 1985. “Price Discrimination and Social Welfare.” *American Economic Review* :870–875.

Vora, Shivani. 2014. “For Lufthansa Flights, Bidding for Upgrades.” *The New York Times* .

Williams, Kevin R. 2020. "Dynamic Airline Pricing and Seat Availability." *Cowles Foundation Discussion Paper* .

Wilson, Robert. 1993. *Nonlinear Pricing*. Oxford University Press.

Appendix

A.1 Survey of International Air Travel

We present additional details about the SIAT data. As we mentioned in Section 2, the data are collected by the Department of Commerce. The DOC contracts with a private survey firm, CIC Research Inc. We use data from the surveys conducted in 2009, 2010, and 2011. There are two data collection methods: (1) Direct participation of the airlines, which arrange for their flight crews to distribute and collect surveys on-board; (2) Use of sub-contractors to distribute and collect the questionnaires in the airport departure gate area. According to the SIAT, in 2009, these two methods accounted for 60% and 40% of all collections, respectively. The dataset can be purchased at `https://rb.gy/fop8cc`. A copy of the survey questionnaire is available at `http://charliemurry.github.io/files/SIAT_Data_Doc_2009.pdf`.[23]

There are 413,309 survey responses in the data we receive from the Department of Commerce. We impose many restrictions to arrive at our final sample. In Table A.1.1, we display summary statistics at four stages of the sample selection process: (1) the original data, (2) after we drop responses that do not report a price, (3) after we make additional selection criteria, like dropping flights with less than 10 responses, responses with other partial information, non-revenue and other exotic tickets, and connecting tickets and (4) our final sample after we select monopoly markets and merge with auxiliary data on capacities. Approximately 38% of the original survey responses do not have information about fares, and we drop those. Out of the remaining 62% who report fares, approximately 53% report traveling with at least one companion. If there are multiple people traveling together, e.g., a family, the survey is intended to be administered to one person in the group. When a respondent reports flying with other passengers, we duplicate the ticket data for each passenger they report flying with. We exclude respondents who report buying their tickets as a part of a tour package, or using airlines miles, or through any other discounted fare. We also restrict our sample to responses that report traveling with at most 10 people in their group (which is 98.23% of the original sample) to minimize the chances that the tickets were bought as part of some tour package.

[23] DOC processes the surveys and implements a quality control process to ensures data integrity. Any foreign language responses are translated into English.

Table A.1.1: Summary Statistics

Variable	Mean	Min	Median	Max	Non-missing	Mean	Min	Median	Max	Non-missing
	(1) Raw Data					(2) Non-missing Fares				
Business Passenger	0.17	–	–	–	413,309	0.19	–	–	–	256,843
– busines	0.04	–	–	–	413,309	0.04	–	–	–	256,843
– convention	0.02	–	–	–	413,309	0.02	–	–	–	256,843
– government	0.00	–	–	–	413,309	0.00	–	–	–	256,843
– teaching	0.02	–	–	–	413,309	0.02	–	–	–	256,843
– vacation	0.17	–	–	–	413,309	0.20	–	–	–	256,843
– visit friends/family	0.12	–	–	–	413,309	0.15	–	–	–	256,843
– religion	0.01	–	–	–	413,309	0.01	–	–	–	256,843
– health	0.01	–	–	–	413,309	0.01	–	–	–	256,843
Nonstop	0.49	–	–	–	413,309	0.46	–	–	–	256,843
Economy Class	0.90	–	–	–	413,309	0.90	–	–	–	256,843
Advanced Purchase	72.37	0.00	45.00	365.00	413,309	67.39	0.00	45.00	365.00	256,843
Fare (on-way adjusted)	567.97	0.00	400.00	24,258.00	256,843	567.97	0.00	400.00	24,258.00	256,843
Total travelers	2.98	1.00	2.00	10.00	413,309	2.80	1.00	2.00	10.00	256,843
Trav. with Family* .	0.36	–	–	–	413,309	0.34	–	–	–	256,843
US Resident*	0.48	–	–	–	413,309	0.51	–	–	–	256,843
Male*	1.54	1.00	2.00	2.00	363,586	1.56	1.00	2.00	2.00	236,658
Age*	43.01	18.00	42.00	99.00	234,546	43.10	18.00	42.00	92.00	153,092
Income Bin*	5.68	1.00	5.00	11.00	305,647	5.77	1.00	5.00	11.00	213,284
	(3) Selection On Additional Variables					(4) Final Sample				
Business Passenger	0.16	–	–	–	25,945	0.14	–	–	–	14,930
– busines	0.04	–	–	–	25,945	0.03	–	–	–	14,930
– convention	0.02	–	–	–	25,945	0.02	–	–	–	14,930
– government	0.00	–	–	–	25,945	0.00	–	–	–	14,930
– teaching	0.02	–	–	–	25,945	0.01	–	–	–	14,930
– vacation	0.19	–	–	–	25,945	0.17	–	–	–	14,930
– visit friends/family	0.14	–	–	–	25,945	0.13	–	–	–	14,930
– religion	0.01	–	–	–	25,945	0.00	–	–	–	14,930
– health	0.01	–	–	–	25,945	0.01	–	–	–	14,930
Nonstop	0.58	–	–	–	25,945	1.00	–	–	–	14,930
Economy Class	0.93	–	–	–	25,945	0.92	–	–	–	14,930
Advanced Purchase	73.65	0.00	60.00	365.00	25,945	75.25	0.00	60.00	365.00	14,930
Fare (on-way adjusted)	528.43	70.00	400.00	5000.00	25,945	480.37	70.00	375.00	5000.00	14,930
Total travelers	2.90	1.00	2.00	10.00	25,945	3.06	1.00	2.00	10.00	14,930
Travel with Family*	0.38	–	–	–	25,945	0.42	–	–	–	14,930
US Resident*	0.51	–	–	–	25,945	0.47	–	–	–	14,930
Male*	1.54	1.00	2.00	2.00	23,698	1.54	1.00	2.00	2.00	13,604
Age*	43.67	18.00	43.00	88.00	16,209	43.34	18.00	42.00	88.00	9,434
Income Bin*	5.69	1.00	5.00	11.00	21,247	5.70	1.00	5.00	11.00	12,011

Note: In this table we present the summary statistics of variables that we observe from the SIAT survey. Variables with an asterisk (*) denote the variables that are not used in our empirical analysis. The table is separated into four sub-tables, where each sub-table displays the summary statistics after each round of sample selection. Sub-table (1) denotes the original data; Sub-table (2) denotes the sample after we have dropped any observation that does not report fares; Sub-table (3) denotes the sample after we drop flights with less than 10 responses, responses with partial information for our purposes, and non-revenue tickets, and connecting tickets; and sub-table (4) denotes the sample after we restrict to monopoly markets.

A.2 Additional Estimation Results

Table A.2.1: Mean and Variance-Covariance of joint truncated Normal distribution for Markets with Modal Capacity

Parameter	Mean (μ_Ψ)	Variance-Covariance Matrix (Σ_Ψ)							
		μ^l	$\frac{\sigma^l}{\mu^l}$	δ^b	$\frac{\sigma^b}{\mu^b}$	μ^ξ	λ	Δ^λ	Δ^θ
μ^l	427.8 (70.344)	61,315 (107.083)	•	•	•	•	•	•	•
$\frac{\sigma^l}{\mu^l}$	0.523 (0.244)		0.428 (0.199)	•	•	•	•	•	•
δ^b	0.074 (0.037)	•	•	5.9×10^{-3} (0.0364)	•	•	•	•	•
$\frac{\sigma^b}{\mu^b}$	0.409 (0.172)	•	•	•	0.015 (0.156)	•	•	•	•
μ^ξ	0.431 (0.006)	•	•	•	•	0.271 (0.019)	•	0.039 (0.022)	0.04 (0.017)
λ	18.838 (6.795)	•	•	•	•	•	4.288 (9.293)	•	•
Δ^λ	-0.0287 (0.006)	•	•	•	•	0.039 (0.022)	•	0.012 (0.01)	0.011 (0.014)
Δ^θ	0.0948 (0.019)	•	•	•	•	0.04 (0.017)	•	0.011 (0.014)	0.011 (0.017)

Note: In this table we present the estimates of mean and variance-covariance matrix of the modal demand parameter Ψ given at the modal initial capacity ω_1^*. Our modal capacity has 115 economy class seats and 14 1st-class seats. Bootstrapped standard errors are in the parentheses.

Table A.2.2: Estimated Mean Values of Ψ, by Capacities

Parameters\Capacity Index	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
μ^l	411.53	431.14	455.91	511.26	298.82	455.92	405.24	354.38	429.78	411.84	427.50	436.63	354.30	330.03	444.48	478.27	504.97	422.39	498.49	438.26
$\frac{\sigma^l}{\mu^l}$	0.53	0.60	0.48	0.39	0.96	0.53	0.50	0.55	0.61	0.55	0.43	0.50	0.97	0.95	0.32	0.39	0.48	0.44	0.40	0.68
δ^b	0.23	0.16	0.04	0.04	0.66	0.15	0.09	0.36	0.31	0.09	0.53	0.26	0.37	0.40	0.05	0.06	0.03	0.07	0.04	0.15
$\frac{\sigma^b}{\mu^b}$	0.12	0.16	0.53	0.51	0.16	0.29	0.29	0.45	0.14	0.34	0.59	0.14	0.56	0.40	0.56	0.40	0.55	0.55	0.48	0.33
μ^ξ	0.01	0.24	0.22	0.25	0.24	0.32	0.01	0.01	0.44	0.01	0.02	0.29	0.15	0.04	0.35	0.26	0.29	0.32	0.29	0.12
λ	24.23	26.29	20.97	34.38	14.13	18.12	27.65	21.97	22.44	27.40	12.26	23.21	18.82	17.51	28.62	23.86	33.20	27.59	30.52	26.46
Δ^λ	-0.08	-0.08	-0.04	-0.08	-0.05	-0.09	-0.07	-0.08	-0.09	-0.07	-0.09	-0.09	-0.09	-0.09	-0.09	-0.07	-0.08	-0.09	-0.08	-0.06
Δ^θ	0.07	0.07	0.09	0.05	0.08	0.07	0.08	0.09	0.05	0.08	0.06	0.08	0.07	0.08	0.06	0.09	0.07	0.08	0.06	0.10

Note: In this table we present the mean demand parameters from the conditional distribution $h(\Psi|\omega_1; \widehat{\mu}_\Psi, \widehat{\Sigma}_\Psi)$, given the initial-capacity ω_1. We estimated the model for 20 different capacities, and each capacity index corresponds to one such capacity.

Table A.2.3: Estimated Mode Values of Ψ, by Capacities

Parameters\ Capacity Index	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
μ^l	100.35	433.05	128.12	444.19	119.05	224.42	381.73	107.06	114.21	106.32	287.67	476.89	101.62	308.74	258.56	381.76	615.27	210.74	376.61	587.11
$\frac{\sigma^l}{\mu^l}$	0.48	0.00	0.50	0.03	0.96	0.56	0.96	0.91	0.91	0.95	0.53	0.99	1.00	1.00	0.08	0.17	0.17	0.00	0.00	0.93
δ^b	0.02	0.16	0.04	0.04	0.63	0.14	0.09	0.34	0.32	0.07	0.58	0.03	0.34	0.10	0.06	0.06	0.00	0.00	0.00	0.15
$\frac{\sigma^b}{\mu^b}$	0.01	0.17	0.53	0.06	0.18	0.20	0.84	0.27	0.17	0.90	0.44	0.14	0.63	0.24	0.52	0.38	0.32	0.60	0.26	0.22
μ^ξ	0.00	0.05	0.08	0.68	0.42	0.26	0.01	0.00	0.95	0.00	0.00	0.02	0.16	0.04	0.22	0.01	0.64	0.15	0.63	0.07
λ	13.27	27.04	7.82	30.92	5.48	8.29	26.95	9.28	26.41	39.54	5.00	22.60	16.65	13.13	24.73	6.72	34.12	40.85	12.24	16.09
Δ^λ	-0.10	-0.09	0.00	-0.05	-0.03	-0.10	-0.10	-0.10	-0.08	-0.09	-0.10	-0.09	-0.09	-0.10	-0.09	-0.04	-0.05	-0.08	-0.05	-0.10
Δ^θ	0.07	0.02	0.12	0.12	0.12	0.03	0.06	0.11	0.11	0.09	0.00	0.06	0.07	0.10	0.10	0.12	0.12	0.11	0.12	0.05

Note: In this table we present the mode of the demand parameters from the conditional distribution $h(\Psi|\omega_1; \widehat{\mu}_\Psi, \widehat{\Sigma}_\Psi)$, given the initial-capacity ω_1. We estimated the model for 20 different capacities, and each capacity index corresponds to one such capacity.

A.3 Model Fit

In Figure A.3.1 we display the empirical moments and the model-implied moments evaluated at the estimated parameters. The moments take the form of deciles of the cumulative density functions of the data and the model predictions.[24] The deciles from our data are shown in red and the deciles predicted by the model are shown in blue. In our estimation step, for each initial capacity, we seek to match 620 moments to determine weights, i.e., the conditional density $h(\Psi|\omega_1)$ in Equation 3 for each period.[25] Here, we display the fit for the modal capacity, which is 115 economy seats and 14 first-class seats.

A.4 Uniqueness of the Optimal Policy

In this section we show that the optimal policy is unique under some regularity conditions. These regularity conditions are widely used in the literature and ensure that demand is decreasing in its own and cross price, and that the demand for each seat class is concave. We begin by presenting these conditions below, but for notational ease suppress the time index.

Assumption 1. *1. (**Downward Demand**): $\left(\frac{\partial \mathbb{E} q^e(p^e,p^f)}{\partial p^e}\right) \leq 0$, and $\left(\frac{\partial \mathbb{E} q^f(p^e,p^f)}{\partial p^f}\right) \leq 0$.*

*2. (**Concave Demand**): $\frac{\partial}{\partial p^e}\left(\frac{\partial \mathbb{E} q^e(p^e,p^f)}{\partial p^e}\right) < \frac{\partial}{\partial p^f}\left(\frac{\partial \mathbb{E} q^e(p^e,p^f)}{\partial p^e}\right) \leq 0$, and $\frac{\partial}{\partial p^f}\left(\frac{\partial \mathbb{E} q^f(p^e,p^f)}{\partial p^f}\right) < \frac{\partial}{\partial p^e}\left(\frac{\partial \mathbb{E} q^f(p^e,p^f)}{\partial p^f}\right) \leq 0$.*

*3. (**Cross Price Curvature**): $\frac{\partial}{\partial p^f}\left(\frac{\partial \mathbb{E} q^e(p^e,p^f)}{\partial p^f}\right) < 0$, and $\frac{\partial}{\partial p^e}\left(\frac{\partial \mathbb{E} q^f(p^e,p^f)}{\partial p^e}\right) < 0$.*

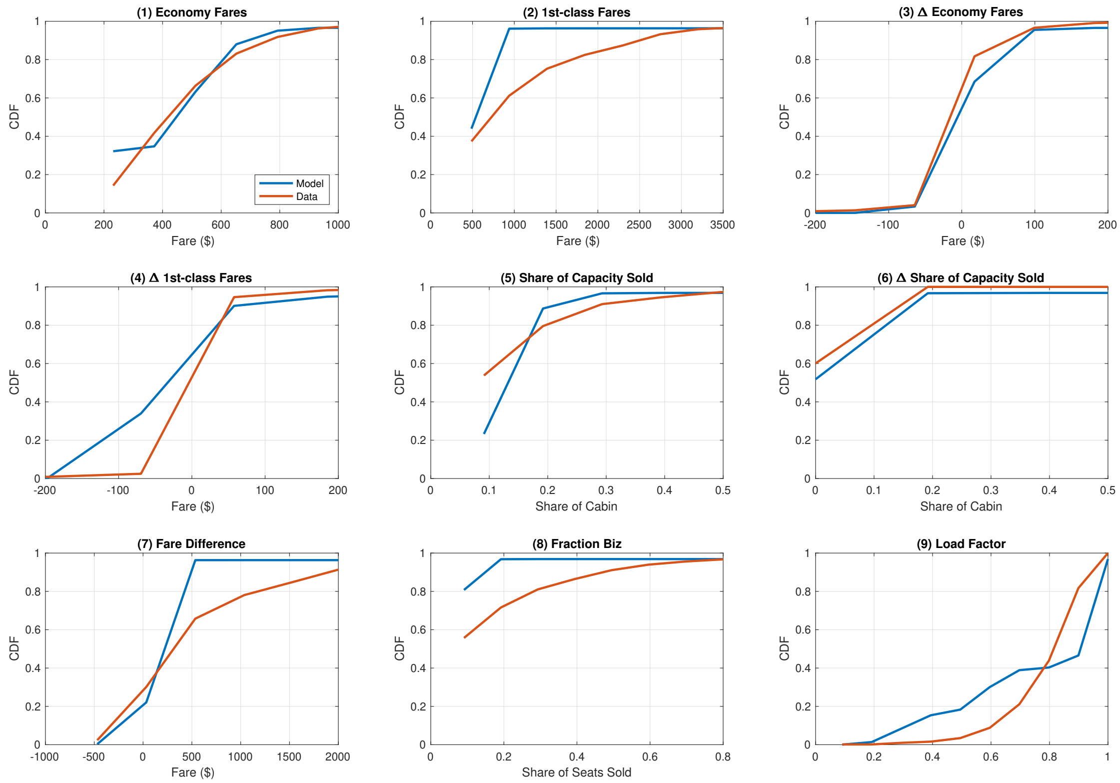
The first assumption says that the demand for either of the cabins must be weakly decreasing in its own price. The second assumption says that the demand is concave in its own price, which ensures the revenue is well defined. It also says that the decrease in demand of economy seat decreases more with respect to economy fare than with respect to business fare. The third assumption says that the change in demand for economy seats with respect to first-class price decreases with the first-class price and vice versa. Although these assumptions are not on the primitives of the model, we present them in these forms because they are more intuitive, self-explanatory and thus easier to understand than the equivalent assumptions on the primitives.

Lemma 1. *Under Assumption 1 there is a unique policy function $\{\sigma_t : t =, 1 \ldots, T\}$.*

[24] In practice, we match moments period by period, but here we display the densities aggregated across periods.

[25] We use these 20 conditional densities to generate average welfare for our counterfactuals in Table 6.

Figure A.3.1: Model Fit



Note: This figure displays the deciles of the variables that we use in estimation, aggregated across all the periods. The empirical moments are in red and include all markets in the SIAT data, and the model implied moments are in blue. Panels (1) and (2) display the moments of the economy fares and first-class fares, respectively; panels (3) and (4) display the moments of the change in economy fares and first-class fares, respectively, across two adjacent periods; panel (5) displays the moments of the share of initial capacity sold; panel (6) displays the moments of the change in the share of initial capacity sold across two adjacent periods; panel (7) displays the moments of the difference in fares between first-class and economy class; panel (8) displays the moments of business passengers's share; and panel (9) displays the moments of the load factor.

Proof. To prove this result we use induction on T:

1. Suppose $T = 1$ and K^e and K^f denote the cabin specific capacities. There are $N := (K^e+1) \times (K^f+1)$ possible seats combinations (state-variables) that could be realized. We show that for each $n \in \{1, \ldots, N\}$ there is a unique optimal pair $\{p^e(n), p^f(n)\}$.

2. Suppose uniqueness is true for $T = \tilde{t}$, then we show that the uniqueness holds even when $T = \tilde{t} + 1$.[26]

[26] With time we also have to keep track of the remaining seats, we will use the short-hand notation of n_t to mean $n_t \in N$.

Step 1

Here, $T = 1$ and for notational ease suppress the time index. The airline solves:

$$
\begin{aligned}
V(\sigma^*) &= \max_{p^e,p^f}\left\{\sum_{k=e,f}(p_t^k - c^k)\int q^k(p^e,p^f)g^k(q^k(p^e,p^f);p^e,p^f)dq^k\right\} \\
&= \max_{p^e,p^f}\sum_{k=e,f}(p_t^k - c^k)\mathbb{E}q^k(p^e,p^f) \qquad \text{(A.1)}
\end{aligned}
$$

Then the equilibrium prices (p^e, p^f) solve the following system of equations:

$$
\begin{bmatrix} \mathbb{E}q^e(p^e,p^f) + (p^e - c^e)\frac{\partial \mathbb{E}q^e(p^e,p^f)}{\partial p^e} + (p^f - c^f)\frac{\partial \mathbb{E}q^f(p^e,p^f)}{\partial p^e} \\ \mathbb{E}q^f(p^e,p^f) + (p^f - c^f)\frac{\partial \mathbb{E}q^f(p^e,p^f)}{\partial p^f} + (p^e - c^e)\frac{\partial \mathbb{E}q^e(p^e,p^f)}{\partial p^f} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \qquad \text{(A.2)}
$$

The above system has a unique solution (p^e, p^f) if the negative of the Jacobian corresponding to the above system is a P-matrix (Gale and Nikaido, 1965). In other words, all principal minors of the Jacobian matrix are non-positive, which follows from Assumption 1.

Step 2

Suppose we have a unique solution when $T = \tilde{t}$ and all finite pair $\{K^e, K^f\}$. Now we want to show that the solution is still unique if we have one additional period, i.e., $T = \tilde{t} + 1$. Consider the value function

$$
V(\sigma^{*\tilde{t}}) := \max_{\{\sigma_t\}_{t=0}^{\tilde{t}}} \mathbb{E}_0\left[\sum_{t=0}^{\tilde{t}}\sum_{k=e,f}(p_t^k - c^k)q_t^k(\sigma_t)\big|K_0^f, K_0^e\right] \qquad \text{(A.3)}
$$

where $\sigma^{*\tilde{t}} := (\sigma_1^*, \ldots, \sigma_{\tilde{t}}^*)$ is the unique optimal policy. Now, suppose we have $\tilde{t} + 1$ periods to consider. So the maximization problem faced by the airline becomes

$$
\begin{aligned}
&\max_{\{\sigma_t\}_{t=0}^{\tilde{t}+1}} \mathbb{E}_0\left[\sum_{t=0}^{\tilde{t}+1}\sum_{k=e,f}(p_t^k - c^k)q_t^k(\sigma_t)\big|K_0^f, K_0^e\right] = \max_{\{\sigma_t\}_{t=0}^{\tilde{t}}} \mathbb{E}_0\left[\sum_{k=e,f}(p_t^k - c^k)q_t^k(\sigma_t)\big|K_0^f, K_0^e\right] \\
&+ \max_{\sigma_{\tilde{t}+1}} \sum_{\omega_{\tilde{t}+1}\Omega_{\tilde{t}+1}} \mathbb{E}_{\tilde{t}+1}\left[\sum_{k=e,f}(p_t^k - c^k)q_t^k(\sigma_t)\big|\omega_{\tilde{t}+1}\right]\Pr(\omega_{\tilde{t}+1}|\sigma^{\tilde{t}})
\end{aligned}
$$

Consider the last period. We have shown that for any realized state space $\omega_{\tilde{t}+1}$ there is a unique optimal policy that solves the second term. The question is if the uniqueness is preserved when we take an expectation with respect to the state variable $\omega_{\tilde{t}+1}$.

For the solution to be unique it is sufficient that the transition probability $\Pr(\omega_{\tilde{t}+1}|\sigma^t)$ is log-concave, which guarantees the expected profit is quasi-concave, hence the solution is unique.[27] Then, the fact that the uniqueness extends from $\tilde{t}$ to $\tilde{t}+1$ follows from the usual backward induction argument of finite-periods maximization problem. Therefore it is enough to show that the transition probability is a (generalized) Poisson distribution, which is log-concave (see Johnson, 2007).

For simplicity, and to provide some intuition as to why transition probability is a (generalized) Poisson, we present the derivation of the transition probability when there is only one cabin and without censoring. Extending the argument to two cabins and incorporate rationing is straightforward, albeit tedious, once we recognize that the Poisson structure is preserved under truncation. Suppose there is only one cabin and no seat-release policy and hence no censoring. And let $\tilde{K}_t = m$ is the number of seats remaining at time t. Then, the probability of reaching $\tilde{K}_{t+1} = m'$ in $t+1$ from m in period t is

$$\Pr(m'|m, \text{Price} = p) = \Pr(Sales_t = \underbrace{m - m'}_{=d}|p) = \sum_{n=0}^{\infty} \Pr(Sales_t = d, N_t = n|p)$$

$$\begin{aligned}
&= \sum_{n=0}^{\infty} \Pr(Sales_t = d|N_t = n, p) \times \Pr(N_t = n|p) = \sum_{n=d}^{\infty} \binom{n}{d} (1 - \tilde{F}_t(p))^d (\tilde{F}_t(p))^{n-d} \frac{e^{-\lambda_t}\lambda_t^n}{n!} \\
&= e^{-\lambda_t}(1 - \tilde{F}_t(p))^d \lambda_t^d \sum_{n=d}^{\infty} \binom{n}{d} (\tilde{F}_t(p))^{n-d} \frac{\lambda_t^{n-d}}{n!} = e^{-\lambda_t} \frac{[\lambda_t(1 - \tilde{F}_t(p))]^d}{d!} \sum_{n=d}^{\infty} \frac{[\lambda_t \tilde{F}_t(p)]^{n-d}}{(n-d)!} \\
&= e^{-\lambda_t} \frac{[\lambda_t(1 - \tilde{F}_t(p))]^d}{d!} e^{\lambda_t \tilde{F}_t(p)} = e^{-\lambda_t(1 - \tilde{F}_t(p))} \frac{[\lambda_t(1 - \tilde{F}_t(p))]^d}{d!} \\
&= e^{-\lambda_t(1 - \tilde{F}_t(p))} \frac{[\lambda_t(1 - \tilde{F}_t(p))]^{(m-m')}}{(m - m')!},
\end{aligned}$$

where

$$\tilde{F}(t) = \begin{cases} \int_0^t f_t(v)dv, & \texttt{economy cabin} \\ \int_0^t \int_0^\infty f_t(v) f_\xi(w/v) \frac{1}{v} dv dw, & \texttt{first-class cabin} \end{cases}$$

and in the fourth equality the sum starts from $n = d$ as the probability of d sales when $n < d$ is zero. Therefore, the transition probability is a Poisson with parameter $\lambda_t(1 - \tilde{F}_t(p))$.

$\square$

[27] A positive and discrete random variable V is log-concave if its p.m.f. $\Pr(V = i)$ forms a log-concave sequence. A non-negative sequence $\{r_i : i \geq 0\}$ is log-concave if for all $i \geq 1 : (r_i)^2 \geq (r_{i-1}) \times (r_{i+1})$.