Dynamic Pricing, Advance Sales, and Aggregate Demand Learning in Airlines

Escobari, Diego

The University of Texas - Pan American

17 December 2011

Dynamic Pricing, Advance Sales, and Aggregate Demand Learning in Airlines[*]

(Journal of Industrial Economics, Forthcoming)

Diego Escobari[†]

December 17, 2011

Abstract

This paper uses a unique U.S. airlines panel data set to empirically study the dynamic pricing of inventories with uncertain demand over a finite horizon. I estimate a dynamic pricing equation and a dynamic demand equation that jointly characterize the adjustment process between prices and sales as the flight date nears. I find that the price increases as the inventory decreases, and decreases as there is less time to sell. Consistent with aggregate demand learning and price adjustment, demand shocks have a positive and much larger effect on prices than the positive effect of anticipated sales.

Keywords: Pricing, demand uncertainty, demand learning, airlines

JEL Classifications: C23, L93

[*]This paper is an extension of my doctoral thesis (Escobari (2008)). I am grateful to my committee members Li Gan, Hae-shin Hwang, Steven Puller, and Ximing Wu. I appreciate useful comments from Damian Damianov, Paan Jindapon, Qi Li, Joon Park, Thomas R. Saving, John K. Wald, and Steven Wiggins. Two anonymous referees and the editor provided insightful comments that helped improve the paper. Stephanie C. Reynolds provided outstanding assistance with the data. Financial support from the Private Enterprise Research Center at Texas A&M and the Bradley Foundation is gratefully appreciated.

[†]Department of Economics & Finance, The University of Texas - Pan American, Edinburg, TX 78539, Phone: (956) 665-2104, Fax: (956) 665-5020, Email: escobarida@utpa.edu., URL: http://faculty.utpa.edu/escobarida

1 Introduction

Despite the large theoretical literature on airline pricing in economics, management, marketing, and operations research, there exists little empirical understanding of the dynamics of prices as the flight date nears. The goal in this paper is to empirically investigate three closely related questions about the dynamics of prices and advance sales. First, for a given inventory of seats, do fares rise as the departure date nears? Second, at a given point prior to the departure date, do fares increase as inventory decreases? Finally, do airlines learn about the aggregate demand and adjust their prices as new information about the pattern of sales is revealed? Most neoclassical economists might argue that this learning and price adjustment is an inherent feature of all markets. However, this learning is about advance sales dynamics, not spot market dynamics. The importance of this question arises due to the large theoretical work (e.g., Prescott (1975), Eden (1990), Deneckere et al. (1996), Dana (1999b), to most recently Deneckere and Peck (2010, section 3)) that shows how prices can respond to aggregate demand uncertainty *without* exploiting learning. In addition, by controlling for price dispersion across flights, this paper contributes to the growing literature on price dispersion in the airline industry (e.g., Gerardi and Shapiro (2009)) by showing the importance of aggregate demand uncertainty and advance sales as a source of price dispersion within flights.

There are three features that make the dynamics of prices and inventories as the flight date nears particularly interesting. First, airlines offer tickets in advance, and unsold tickets expire at departure. Second, capacity is also set in advance and can only be modified at a relatively large marginal cost. Finally, there is uncertainty about the aggregate demand. Hence, airline ticket sales represent an example of dynamic pricing of inventories with uncertain demand over a finite horizon. This problem arises in a variety of good and services, such as hotel rooms, cabins on cruise liners, car rentals, and entertainment and sporting events. To some extent, it is also present in goods that are not necessarily perishable, but where production decisions are made in advance and demand is uncertain and concentrated during a selling season (e.g., the Christmas shopping season).

The paper takes advantage of a unique panel data set collected from the online travel agency Expedia.com, which contains prices and seat inventories at the ticket level for 103 days prior to the departure of 228 U.S. domestic flights. This is different from most of the empirical research on airlines that uses aggregate data from the Bureau of Transportation and Statistics (e.g., Borenstein and Rose (1994), and Gerardi and Shapiro (2009)) and research with posted prices without inventories. Stavins (2001) uses prices from the Official Airline Guide, and more recently Bilotkach (2006), McAfee and te Velde (2007), and Bilotkack and Rupp (2011) use posted prices from online travel agencies. To focus on the dynamics of prices as the flight date nears, the construction of the data set controls for product heterogeneities and "fences" that segment consumers (e.g., Saturday-night stayover, different connections/legs, refundability, and fare class).

To answer how prices depend on days to departure and inventories, I estimate a pricing equation that is consistent with the theoretical models in Gallego and van Ryzin (1994), Zhao and Zheng (2000), and Deneckere and Peck (2010). The estimation controls for time-invariant flight-, carrier-, and route-specific characteristics and takes into account the dynamic feedback between sales and prices. This means that sellers and buyers can behave dynamically: The decisions to price and to buy today can be affected by previous realizations of fares and sales. Moreover, sellers and buyers can adopt a forward-looking perspective and form beliefs about the future evolution of prices and inventories. Supporting the theoretical prediction in Gallego and van Ryzin (1994), the results show that the price is lower if there is less time to sell: For every day that passes without sales, the price falls 57.1 cents. Furthermore, prices increase 7 and 14 days before departure, consistent with the arrival of higher valuation travelers (see Zhao and Zheng (2000) and Su (2007)). Finally, consistent with Gallego and van Ryzin (1994), Zhao and Zheng (2000), and models where price dispersion arises from the combination of costly capacity and aggregate demand uncertainty (e.g., Prescott (1975)), the estimates indicate that one fewer available seat increases fares by 1.53 dollars.

I follow a two-step approach to answer whether carriers learn about the aggregate demand. In the first step I estimate a dynamic demand equation that formalizes the feedback from prices to cumulative bookings, and use the estimates to separate bookings into expected bookings (booking curve) and unexpected bookings (demand shocks). In the second step I estimate a dynamic pricing equation that differentiates the effects of these two booking components on pricing decisions. The results show that prices respond to new information about the pattern of sales (demand shocks). Moreover, the estimates indicate that the positive response of prices to unexpected sales (demand shocks) is statistically and economically greater than the positive response of prices to anticipated sales. This pattern is consistent with aggregate demand learning and price adjustment, as opposed to the hypothesis that prices adjust mechanically to sales, whether or not they are expected. Aggregate demand learning combined with the estimated downward sloping dynamic demand means that airlines can partially control the evolution of bookings via prices.

The paper is related to the literature on peak-load pricing.[1] Borenstein and Rose (1994) provide a distinction between two types of peak-load pricing in airlines. The first is systematic peak-load pricing, which reflects variation in the shadow cost of capacity due to demand fluctuations known at the time the flight is scheduled. The second is stochastic peak-load pricing, which reflects demand uncertainty resolved only as sales progress.[2] In this paper, I control for systematic peak-load pricing and find empirical support to specific predictions from some stochastic peak-load pricing models. When prices are set before sales begin Prescott's (1975) static stochastic peak-load pricing model predicts an upward schedule of fares. Eden (1990) formalizes Prescott's model and Dana (1999b) extends it to monopoly and imperfect competition.[3] Eden (1990) and Lucas and Woodford (1993) point out that Prescott's static model has an interesting time-consistency property. Even if prices are allowed to change during the selling season, learning about the final state of the demand is needed to deviate from the original price schedule. Deneckere and Peck (2010) present a generalized multiple period version of Prescott's model, where demand is gradually learned over time and prices are allowed to change each period. While advance-purchase sales are important for learning, they can serve other purposes as well. Gale and Holmes (1993) show that advance-purchase discounts can be used to divide uncertain peak demand more evenly between two departures, while Dana (1998) shows that they can be used to screen consumers and price discriminate.[4]

[1] For peak-load pricing under certainty, see Boiteaux (1980) and Williamson (1966), and under uncertainty, see Visscher (1973) and Carlton (1977).

[2] Using ex-ante known demand intensities, Escobari (2009) estimates a congestion premia associated with systematic peak-load pricing.

[3] Escobari and Gan (2007) find empirical support for these models.

[4] Escobari (2009) provides empirical support for the main empirical prediction in Gale and Holmes (1993)

The paper is also related to a strand of literature that comes from operations research. A number of papers consider the problem of dynamically pricing a stock of perishable product over a finite time horizon (e.g., Gallego and van Ryzin (1994), Zhao and Zheng (2000), and Su (2007)).[5] The closest to my paper is Lin (2006), who presents a model where the seller uses realized demand to infer about the arrival rate, update the future demand distribution, and set the price. Consistent with my results, Lin finds that higher prices should be set when demand is expected to be larger.

The organization of the paper is as follows. Section 2 explains the data, provides preliminary evidence of learning, and presents an overview of the estimation. The motivation and estimation of the dynamic pricing equation is presented in Section 3. Section 4 deals with the motivation and estimation of the dynamic demand and discusses the beliefs. Section 5 separates the evolution of bookings into booking curve and demand shocks, and presents the dynamic pricing with learning. Section 6 concludes.

2 Data, Preliminary Evidence, and Overview of the Estimation

2.1 Data

The data for this paper were collected between March and June 2006 from the online travel agency Expedia.com. It is a panel with 228 cross-sectional observations and 35 observations over time. Each cross-sectional unit is a specific non-stop one-way flight from a carrier on a route. There are 81 routes in the sample, and each route is defined as a pair of departure and destination airports. The observations in time start 103 days prior to the departure date and were collected every three days until the day before departure. All flights depart on Thursday, June 22, 2006. The carriers considered are American, Alaska, Continental, Delta, United, and US Airways, with the proportion of flights of each carrier chosen to be close to its share in the U.S. market. This data set is similar to the one used in Stavins

— less discount seats on ex-ante known peak flights. Puller et al. (2008) find modest support for Dana (1999b) and Gale and Holmes (1993), and strong support for models of second-degree price discrimination.

[5] See McGill and van Ryzin (1999) and Elmaghraby and Keskinocak (2003) for references to and descriptions of various theoretical models.

(2001), but with two important improvements. First, it is a panel, which allows to control for unobserved time-invariant flight-, carrier-, and route-specific characteristics. Second, it has information about seat availability at each fare, obtained from the seat-availability map, where the available preferred or prime seats are counted as available seats.

The construction of the data set controls for important sources of price dispersion in the industry. By picking one-way flights, the paper controls for fare differences associated with round-trip tickets (e.g., Saturday-night stayover, minimum stay and maximum stay). Selecting non-stop flights controls for price variation that arises in more sophisticated itineraries (e.g., different connections/legs). Economy-class tickets control for the fare class, and by selecting the least expensive price, I control for the existence of more expensive refundable tickets. Moreover, tickets obtained through frequent-flyer programs are excluded from the sample. Additionally, having only one-way non-stop flights is helpful to define a single inventory level at each posted price. Including complex itineraries, additional fare classes, round-trip tickets, or international destinations would impose an important burden on the empirical section. For the inventory levels, it would mean having more than one inventory at each posted fare. For the fares, it would involve including ticket characteristics that the carrier can use to screen consumers and price discriminate.

[Table 1, here.]

Table 1 displays the summary statistics of the variables used in the analysis. Fare is the one reported by Expedia.com and corresponds to the least expensive economy-class fare for a particular flight. DayAdv is the number of days in advance, and Load $\in [0, 1]$ is the ratio of unavailable seats to total seats in the aircraft. I refer to this ratio as the load factor, which is a ticket-level load factor that can change at each posted fare, ranging from zero if the plane is empty to one if it is full.[6] Figure 1 displays the average and the standard deviation of fares across the 228 flights at different days prior to the departure date. Two things are worth noting in this figure. First, average fares appear to increase over time. Second, the dispersion of fares across flights is fairly constant, with only a slight increase close to the departure date.

[6] The literature on airlines defines load factor only once the plane has departed, and it is the percentage of seats filled with paying passengers.

[Figure 1, here.]

2.2 Preliminary Evidence of Demand Learning

Preliminary evidence and the intuition behind aggregate demand learning and price adjustment is illustrated in Figure 2 for Delta flight 1588, which goes from from Atlanta, GA (ATL) to San Jose, CA (SJC). The figure plots the dynamics of fares, load factors (actual bookings), and demand shocks, defined as the deviations of actual bookings from the expected evolution of bookings. The key point in this figure is that it illustrates the different responses of fares to expected and unexpected changes in load factors. Between 85 and 82 days in advance, the load factor increased by 0.14 (14% of the aircraft's capacity). It is reasonable to believe that this jump in sales came as a surprise to Delta, meaning actual bookings 82 days in advance were above expected bookings. Demand learning and price adjustment means that Delta is able to find out about this new information regarding the pattern of bookings (demand shock) and respond by increasing fares. That appears to have happened 82 days in advance, when the price increased from $469 to $664.[7] Fares are nearly unresponsive to the 0.22 increase in load factor between 58 and 19 days before departure, because this increase appears to be within the expected evolution of cumulative bookings.

[Figure 2, here.]

2.3 Overview of the Estimation

To answer how fares respond to days to departure and inventories, I estimate a dynamic pricing equation. I follow a two-step procedure to answer whether airlines learn about the aggregate demand and adjust their fares in response to new information. In the first step I estimate a dynamic demand to characterize the evolution of bookings and to separate bookings into the booking curve and demand shocks. The second step consists of estimating a dynamic pricing equation with learning, where anticipated and unanticipated bookings are allowed to have a different effect on fares.

[7] A similar (negative) demand shock appears to have occurred at 64 days in advance. The decrease in the load factor may reflect reserved tickets that were never bought, or tickets that were canceled/returned.

3 Dynamic Pricing

3.1 Dynamic Pricing Equation

I estimate the following dynamic specification to answer whether fares rise as the departure date nears and whether fares increase as inventory decreases:

$$\ln(\textsc{Fare})_{ijt} = \alpha \ln(\textsc{Fare})_{ij,t-1} + \gamma \textsc{DayAdv}_t + \beta \textsc{Load}_{ij,t-1} + \nu_{ij} + \varepsilon_{ijt}. \quad (1)$$

The subscript i refers to the flight, j to the route, and t to time. The variable $\ln(\textsc{Fare})_{ijt}$ is the logarithm of fare and $\textsc{DayAdv}_t$ is the number of days prior to the departure date, both measured at time t.[8] $\textsc{Load}_{ijt}$ is the load factor at the end of period t; therefore, Equation 1 is consistent with the price posting theoretical model in Deneckere and Peck (2010), where carriers post prices based on the beginning-of-period cumulative bookings, $\textsc{Load}_{ij,t-1}$. This specification is also consistent with Gallego and van Ryzin (1994) and Zhao and Zheng (2000), where prices depend on inventories and on time to departure. ν_{ij} captures the time-invariant flight-, carrier-, and route-specific effects, and ε_{ijt} denotes the remaining disturbance. Notice that ν_{ij} captures the time-invariant carrier- and route-specific characteristics in a flexible way, allowing them to vary across flights within the same carrier and across flights within the same route. Time-invariant characteristics comprise most of the controls included in Stavins (2001) (e.g., Herfindahl index, distance, and hub) and unobservables, such as managerial capacity and systematic peak-load pricing, which arises due to congestion known at the time the flight is scheduled.[9]

Even though the coefficient on the lagged dependent variable is not of direct interest, allowing for dynamics in the underlying process may be crucial for recovering consistent estimates of the effect of DayAdv and Load on fares. The correlation between prices and cumulative bookings may reflect a common driving force that arises from a dynamic adjustment process. Because cumulative bookings at time t come from the aggregation of previous single-period demands that depend on their contemporaneous posted prices,

[8] $\textsc{DayAdv}_t$ can be calculated as $1 + 3 \times (35 - t)$ for $t = 1, 2, \ldots, 35$.

[9] Borenstein and Rose (1994) control for systematic peak-load pricing under the assumption that it is correlated with airlines' fleet utilization rates and airports' operation rates, both being time-invariant and part of ν_{ij} in Equation 1. Flights at more congested departure times that are associated with a larger shadow cost of capacity will have a larger ν_{ij}, all else equal.

$\textsc{Load}_{ij,t-1}$ will be treated as weakly exogenous. Maintaining that the disturbances are serially uncorrelated, $\textsc{Load}_{ij,t-1}$ is predetermined in the sense that it is uncorrelated with ε_{ijt}, but $\textsc{Load}_{ij,t-1}$ may be correlated with $\varepsilon_{ij,t-1}$ and earlier shocks,

$$\left.\begin{array}{ll} E(\textsc{Load}_{ij,s-1}\varepsilon_{ijt}) = 0, & s \leq t \\ E(\textsc{Load}_{ij,s-1}\varepsilon_{ijt}) \neq 0, & s > t \end{array}\right\}, \qquad \forall ij. \tag{2}$$

Serially uncorrelated disturbances mean that ε_{ijt} corresponds to an unexpected change in prices, and that previous unexpected changes cannot be used to predict future unexpected changes. ε_{ijt} represents the random part of prices that may include a "last-minute deal" that consumers cannot predict from past variables or the state of sales. Moreover, weak exogeneity does not restrict consumers or sellers from adopting a forward-looking perspective, and it is consistent with rational expectations models.

To allow for this dynamic feedback between $\textsc{Load}$ and fares and to obtain consistent estimates of the coefficients of interest γ and β, I will initially use the difference GMM estimator for dynamic panel data models proposed by Holtz-Eakin et al. (1988) and Arellano and Bond (1991). They suggest estimating Equation 1 by taking first differences to eliminate the unobserved time-invariant flight-, route-, and carrier-specific characteristics ν_{ij} to obtain

$$\Delta \ln(\textsc{Fare})_{ijt} = \alpha \Delta \ln(\textsc{Fare})_{ij,t-1} + \gamma \Delta \textsc{DayAdv}_t + \beta \Delta \textsc{Load}_{ij,t-1} + \Delta\varepsilon_{ijt}. \tag{3}$$

Then there is the need of a vector Z of instruments to construct the moments $E(\Delta\varepsilon_{ijt}\mathrm{Z})$ and to estimate Equation 3 via GMM. Under the assumptions that the error term ε_{ijt} is not serially correlated and that $\textsc{Load}_{ij,t-1}$ is weakly exogenous, lagged values of $\textsc{Load}_{ij,t-1}$ are valid instruments for $\Delta\textsc{Load}_{ij,t-1}$. By construction the new error term $\Delta\varepsilon_{ijt}$ is correlated with the lagged dependent variable; hence, $\ln(\textsc{Fare})_{ij,t-2}$ and earlier lags are used as instruments for $\Delta \ln(\textsc{Fare})_{ij,t-1}$. Because $\textsc{DayAdv}_t$ is treated as strictly exogenous, I simply use $\Delta\textsc{DayAdv}_t$ as its own instrument in the difference equation.

Blundell and Bond (1998) point out a statistical shortcoming with this difference GMM estimator. If $\ln(\textsc{Fare})$ and $\textsc{Load}$ are persistent over time, lagged levels of these variables are weak instruments for the regression equation in differences. Hence, I combine Equations 1 and 3 and use the system GMM estimator that Blundell and Bond proposed. The additional moments for the equation in levels are $E[(\nu_{ij} + \varepsilon_{ijt})\mathrm{W}] = 0$. Then the instruments for the regression in differences are the same as above and the instruments W for the regression in levels are the lagged differences of $\ln(\textsc{Fare})_{ijt}$ and $\textsc{Load}_{ijt}$. While the levels of $\ln(\textsc{Fare})_{ij,t-1}$ and $\textsc{Load}_{ij,t-1}$ may be correlated with ν_{ij}, for the instruments W to be valid, W is assumed to be uncorrelated with ν_{ij}. The assumption is realistic because Load measures sold inventories relative to the aircraft size. Thus, Load will not be affected by any ν_{ij} that impact the total number of seats sold, but leave the ratio of seats sold to total seats unchanged. Moreover, because ν_{ij} does not change during the selling season, if observed by the carrier it will be before sales begin and when the aircraft size can still be modified at a relatively low marginal cost. Even if scheduling a different-sized aircraft is not possible, the carrier will likely absorb any (observed) ν_{ij} by setting higher/lower prices across all tickets. The idea is that the carrier will not want any (observed) ν_{ij} to affect $\textsc{Load}_{ijt} - \textsc{Load}_{ij,t-1}$ throughout the selling season because Load should evolve between zero and one on every flight, regardless of ν_{ij}. This means that ν_{ij} is almost certainly correlated with the levels of $\ln(\textsc{Fare})_{ij,t-1}$. Finally, there may be reasons to believe that ν_{ij} may be correlated with $\textsc{Fare}_{ijt} - \textsc{Fare}_{ij,t-1}$ (no logs) if the change in the dollar amount of fares is greater in more expensive flights; however, that is less likely to be the case for the first differences of $\ln(\textsc{Fare})_{ijt}$, $\ln(\textsc{Fare}_{ijt}/\textsc{Fare}_{ij,t-1})$.

3.2 Dynamic Pricing Estimates

The results from the estimation of the pricing equation are reported in Table 2. For comparison purposes, the first four columns report four sets of estimates that assume strict exogeneity of $\ln(\textsc{Fare})_{ij,t-1}$ and $\textsc{Load}_{ij,t-1}$, while the GMM estimates that relax this assumption are reported in the last four columns. In column 1, the negative coefficient on the DayAdv variable indicates that fares are higher closer to the departure date, but this estimate appears to be downwards-biased mainly because of the omitted variable $\textsc{Load}_{ij,t-1}$. A negative omitted-variable bias is consistent with the negative correlation between $\textsc{Load}_{ij,t-1}$ and $\textsc{DayAdv}_t$ (fewer available seats closer to the departure date) and the positive correlation between $\ln(\textsc{Fare})_{ijt}$ and $\textsc{Load}_{ij,t-1}$ (higher fares with fewer available seats). The second column presents a (misspecified) static pricing equation, where the estimate on the DayAdv variable appears downwards-biased largely because of the omitted variable $\ln(\text{Fare})_{ij,t-1}$ (when compared to column 4) and because it ignores the dynamic feedback between current inventories and previous prices (when compared to the GMM specifications).[10]

The estimates in columns 3 and 4 behave as expected in the presence of flight-specific effects. Consistent with the Monte Carlo simulation results in Blundell et al. (2000), the Pooled OLS appears to give an upwards-biased estimate of the coefficient on the lagged dependent variable, while the Within appears to give a downwards-biased estimate of this coefficient. Also consistent with Blundell et al. (2000), the estimates on DayAdv and Load appear to have a large negative bias in the Pooled OLS and a smaller negative bias in the Within specification. Blundell et al. (2000) find that the bias is larger when the regressor is persistent, which is the case with DayAdv and to a lesser extent with Load. What is puzzling is the difference in the estimated DayAdv coefficient between the Within specification in column 4 and the GMM specifications, especially because DayAdv_t is always treated as strictly exogenous and the GMM are specifically utilized to deal with the weakly exogeneity and the endogeneity of $\text{Load}_{ij,t-1}$ and $\ln(\text{Fare})_{ij,t-1}$ respectively. This difference is consistent with the asymptotic ($N \to \infty$) bias of strictly exogenous variables for the Within estimator derived in Nickell (1981, p. 1424). He finds that the Within estimator of γ is downwards biased if the time-demeaned exogenous variable is negatively related (in the regression sense) with the lagged time-demeaned dependent variable, which is the case here.[11]

[Table 2, here.]

Columns 5 and 6 present the two-step first-differenced GMM panel estimates.[12] Column

[10]A negative omitted-variable bias would be consistent with the negative correlation between DayAdv_t and $\ln(\text{Fare})_{ij,t-1}$, and the positive correlation between $\ln(\text{Fare})_{ijt}$ and $\ln(\text{Fare})_{ij,t-1}$. The dynamic demand estimates will show why $\text{Load}_{ij,t-1}$ needs to be treated as weakly exogenous in Equation 1.

[11]Specifically, the asymptotic bias derived in Nickell (1981) is

$$\operatorname*{plim}_{N\to\infty}(\hat{\theta}-\theta) = -\operatorname*{plim}_{N\to\infty}\left[(\ddot{\mathrm{X}}'\ddot{\mathrm{X}})^{-1}\ddot{\mathrm{X}}'\ddot{y}_{-1}\right]\operatorname*{plim}_{N\to\infty}(\hat{\alpha}-\alpha),$$

where $\theta' = (\gamma,\beta)$, $\hat{\theta}$ and $\hat{\alpha}$ are from the Within estimator of Equation 1, y_{-1} is ln(Fare) lagged once, $\mathrm{X} = [\text{DayAdv} \vdots \text{Load}]$, and $\ddot{y}$ and $\ddot{\mathrm{X}}$ are time-demeaned transformations. For γ, both of the terms on the right-hand side are negative, making its Within estimator downwards biased.

[12]Given the apparent persistence of fares and load, I follow Blundell and Bond (2000) by estimating simple

5 uses $\ln(\textsc{Fare})_{ij,t-2}$ and the second lag of $\textsc{Load}$ as instruments, while column 6 uses $\ln(\textsc{Fare})_{ij,t-2}$ and the second and third lags of $\textsc{Load}$ as instruments. The validity of these specifications is addressed with two tests. To assess the assumption that the error term ε_{ijt} is not serially correlated, I test whether the differenced error term is second-order serially correlated. The large p-values provide strong support for a valid specification. The Sargan test of over-identifying restrictions to test the overall validity of the instruments shows that the null hypothesis — that the lagged levels dated $t-2$ (column 5) as instruments are not correlated with the residuals — is rejected. However, at a 5% significance level, the null in the Sargan test for the lagged levels dated $t-3$ (and earlier) as instruments in column 6 is not rejected. The two-step GMM system estimates are reported in columns 7 and 8. The additional instruments for the levels equations are $\Delta\ln(\textsc{Fare})_{ij,t-1}$ and $\Delta\textsc{Load}_{ij,t-1}$. The second-order serial correlation test strongly supports the assumption of no serial correlation. Furthermore, the Sargan test of over-identifying restrictions, which analyzes the sample analogs of the moment conditions used in the GMM estimation, validates the instrument list. The Difference Sargan test validates the additional instruments used in the levels equations.

The positive and highly significant effect of cumulative bookings on fares across all specifications of Table 2 indicate that, all else equal, fares increase as the aircraft's remaining capacity becomes scarcer. This positive sign is consistent with the theoretical models in Gallego and van Ryzin (1994) and Zhao and Zheng (2000). This is also consistent with models where price dispersion arises from the combination of costly capacity and aggregate demand uncertainty (e.g., Prescott (1975), Eden (1990), Dana (1999a), and Dana (1999b)). The estimated coefficient in column 8 indicates that in a 100-seat aircraft, fares increase by \$1.53 (\$291.1×0.527/100) for each fewer available seat. A standard deviation increase in utilized capacity increases fares by \$38.65, which corresponds to a 49.09% within flight standard deviations of fares.

The positive and significant coefficients on $\textsc{DayAdv}$ indicate that, holding inventories constant, fares decrease as there is less time to sell. The point estimate in column 8 reads that on average when a day passes without sales, the price falls 57.1 cents

AR(1) specifications for $\ln(\textsc{Fare})$ and $\textsc{Load}$. The results confirm $\textsc{Load}$ to be persistent, but the estimates of the autoregressive terms were smaller and more precise than those in Blundell and Bond (2000).

($\$291.1 \times 1.961/10^3$). In their review of airline pricing, McAfee and te Velde (2007) explain that falling prices as takeoff approaches is a remarkably robust prediction of theories. This includes the theoretical predictions in Kincaid and Darling (1963) and Gallego and van Ryzin (1994), where the intuitive explanation of falling prices follows from the perishable nature of airline seats. A given inventory of seats will be more difficult to sell if there is less time; hence, there is the incentive to lower the price.[13]

[Table 3, here.]

The key assumption behind decreasing prices over time in Gallego and van Ryzin (1994) is that the reservation price distribution is the same across all consumers. This assumption is unlikely to hold in airlines because consumers who purchase closer to the departure date tend to have higher valuations. Zhao and Zheng (2000) find that prices can increase as time to expire decreases if the reservation price distribution shifts to the right. Moreover, Su (2007, p. 735) shows that higher prices closer to departure can exist because business travelers who have lower waiting costs do not mind committing to travel schedules later. The predictions in these two papers are in line with the common observation of higher prices closer to the departure date, as Figure 1 suggests. While lower inventories and the positive LOAD coefficient can explain higher prices closer to takeoff, Figure 1 shows that prices increase much faster in the last two weeks. To capture potential non-linearities, Table 3 offers additional results that include the indicator variables $1_{[\text{DAYADV}<k]}$ for $k = 7, 14$, which equal one if DAYADV is less than k, zero otherwise. The main findings from Table 2 hold; moreover, the indicator variables capture jumps in prices. The point estimates in column 6 indicate that there are jumps of \$21.25 and \$10.19 at 7 and 14 days to departure, respectively. The results from this table indicate that for a given inventory, prices decrease as time to departure decreases, but increase 7 and 14 days before departure, consistent with the time where most business travelers decide to buy.

[13] Lazear (1986) also finds decreasing prices over time in a simple two-period model. Moreover, Sweeting (2010) documents the existence of decreasing prices in online resale markets for Major League Baseball tickets.

4 Dynamic Demand and the Evolution of Sales

While the dynamic pricing equation did not require models for Load to be specified to estimate the parameters (α, γ, β), modeling Load and formalizing the feedback mechanism from prices to cumulative bookings serves two purposes. First, the feedback mechanism has a dynamic demand interpretation and shows how bookings depend on previous prices. Second, the feedback mechanism can be used to characterize the evolution of cumulative bookings and to separate bookings into expected and unexpected bookings. This section focuses on the estimation and discussion of the dynamic demand, whereas the next section discusses how to separate bookings into these two components.

The demand is dynamic because at each point prior to departure newly arrived and existing consumers form expectations about future prices and future product availability, which affect their decision to buy a ticket at the going price, wait to purchase later, or exit (see, for example, Su 2007). Moreover, the current decision affects future utility because, even though consumers buy a ticket today, the good is consumed in the future. Demand dynamics in airlines is related to demand dynamics for storable and durable products because in all these cases demand anticipation is important.[14]

4.1 Dynamic Demand Equation

The stock variable Load comes from the aggregation of sales that occurred during previous periods. Sales during period t can be obtained as the difference between beginning-of-period and end-of-period cumulative bookings, $\Delta\text{Load}_{ijt} = \text{Load}_{ijt} - \text{Load}_{ij,t-1}$. I model period t demand, ΔLoad_{ijt}, to depend on the demand last period (which brings on the dynamics), contemporaneous posted fares, and the number of days to departure:

$$\Delta\text{Load}_{ijt} = \rho\Delta\text{Load}_{ij,t-1} + \phi \ln(\text{Fare})_{ijt} + \delta\text{DayAdv}_t + \eta_{ij} + u_{ijt}. \qquad (4)$$

η_{ij} captures the time-invariant fight-, carrier-, and route-specific effects, and u_{ijt} are random error terms independent of all random variables introduced earlier. This feedback mechanism between prices and sales is consistent with the theoretical model in Deneckere

[14]Pesendorfer (2002) and Hendel and Nevo (2006) show that consumers anticipate their demand for storable products. When prices are low consumers increase their purchases to store for future consumption.

and Peck (2010), where within each period firms start posting prices for that period (Equation 1), then consumers arrive in random order, observe posted prices, and decide whether to purchase (Equation 4). A downward-sloping demand curve, $\phi < 0$, will let sellers control the intensity of the demand via prices, as in Gallego and van Ryzin (1994).

An issue in the estimation of Equation 4 is the potential endogeneity of fare. Endogeneity arises if there is correlation between $\ln(\text{Fare})_{ijt}$ and the unobserved $\eta_{ij} + u_{ijt}$. The most common cause of this correlation is if the carrier sets prices knowing more about the error term than the econometrician. Taking first differences eliminates the time-invariant effect η_{ij}:

$$\Delta^2\text{Load}_{ijt} = \rho\Delta^2\text{Load}_{ij,t-1} + \phi\Delta\ln(\text{Fare})_{ijt} + \delta\Delta\text{DayAdv}_t + \Delta u_{ijt}. \tag{5}$$

Then the GMM dynamic panel estimators allow for different assumptions on the contemporaneous correlation between $\ln(\text{Fare})_{ijt}$ and u_{ijt}. Notice that under the assumption of serially uncorrelated u_{ijt}, these errors can be interpreted as the random part of sales (demand shocks) that cannot be predicted based on previous realizations of the variables or previous realizations of the error term.

A first approach will assume $\ln(\text{Fare})_{ijt}$ is predetermined; hence, u_{ijt} represents a true demand shock for both the carrier and the econometrician. This means that the carrier sets prices after observing previous realizations of the demand shocks but does not observe the contemporaneous or future demand shocks. Arellano and Bond (1991) propose estimating the equation in differences using the moments $E(\Delta u_{ijt}\text{M}) = 0$, where M is the vector of instruments that contains the lags of $\ln(\text{Fare})_{ijt}$, and because Δu_{ijt} is correlated with $\Delta^2\text{Load}_{ij,t-1}$, M also includes the lags of $\text{Load}_{ij,t-1}$. For the system GMM estimator — which combines Equations 4 and 5 — the additional moment conditions for the equation in levels are $E[(\eta_{ij} + u_{ijt})\text{H}] = 0$. The vector of valid instruments H includes $\Delta^2\text{Load}_{ij,t-1}$, $\Delta\ln(\text{Fare})_{ijt}$, and the lags of both. The assumption for the validity of these additional instruments is that they should be uncorrelated with η_{ij}; however, there could still be correlation between $\Delta\text{Load}_{ij,t-1}$ or $\ln(\text{Fare})_{ijt}$ and η_{ij}. To see why this is a reasonable assumption, keep in mind that η_{ij} captures the time-invariant characteristics that affect sales relative to capacity (ΔLoad_{ijt}) on every period t prior to departure. Correlation between $\Delta^2\text{Load}_{ij,t-1}$ and η_{ij} can be interpreted as flight-, carrier-, or route-specific characteristics that affect the rate at which sales increase (or decrease) throughout the selling period. This kind of correlation is very unlikely because $\textsc{Load} \in [0, 1]$ and airlines will adjust capacity before sales begin to take into account any observed η_{ij} that can affect the demand for a flight. For example, a flight with a particularly large demand will be assigned larger capacity, making the correlation between $\Delta\textsc{Load}_{ij,t-1}$ and η_{ij} unlikely (while still allowed) and the correlation between $\Delta^2\textsc{Load}_{ij,t-1}$ and η_{ij} even less likely. Similar to the discussion of the instruments for Equation 1, it is also unlikely to have time-invariant characteristics that affect sales to also change with $\Delta \ln(\textsc{Fare})_{ijt}$. Treating $\ln(\textsc{Fare})_{ijt}$ as predetermined when it is endogenous will yield a biased estimate of ϕ. In particular, if u_{ijt} is positively correlated with price, the estimate of ϕ will be upwards-biased.

A second approach treats $\ln(\textsc{Fare})_{ijt}$ as endogenous by allowing for contemporaneous correlation between u_{ijt} and $\ln(\textsc{Fare})_{ijt}$. Treating $\ln(\textsc{Fare})_{ijt}$ as potentially endogenous invalidates $\ln(\textsc{Fare})_{ij,t-1}$ and $\Delta \ln(\textsc{Fare})_{ijt}$ as instruments. Hence, the vector M of instruments for the difference equation can only include $\ln(\textsc{Fare})_{ij,t-2}$ and its lags, while the vector H of instruments for the levels equation can only include $\Delta \ln(\textsc{Fare})_{ij,t-1}$ and its lags. Modeling $\ln(\textsc{Fare})_{ijt}$ as endogenous when it is predetermined still yields consistent estimates; however, it does not use all the available instruments. Treating $\ln(\textsc{Fare})_{ijt}$ as endogenous affects the interpretation of u_{ijt}; while it is still a shock for the econometrician, it may not be an unobserved demand shock for the carrier.

The dynamic adjustment process between prices and sales, as characterized by Equations 1 and 4, imply that consumers and sellers can behave dynamically. Equation 1 suggests that previous sales affect the current posted price, while Equation 4 suggests the decision to buy a ticket today at the current posted price can be affected by previous realizations of fares. Equations 2 only imply that the consumer's decision to buy a ticket today must be uncorrelated with future price shocks ε_{ijt}. Moreover, weak exogeneity or endogeneity of $\ln(\textsc{Fare})_{ijt}$ means that the price today is uncorrelated with future demand shocks u_{ijt}. This do not restrict consumers or sellers from adopting forward-looking perspectives. Weak exogeneity and endogeneity are consistent with rational expectations models, in which sellers' and buyers' beliefs would be equal to the true data-generating process. However, sellers and consumers can have their own subjective beliefs about the evolution of prices and sales, not necessarily following Equations 1 and 4. As explained in Arellano and Bond (1991), short-run dynamics will compound influences from expectations formations and decision processes. Even if all travelers have rational expectations about the evolution of fares and sales, the presence of private information and arrival rates implies variance in who actually buys in any period. Private information arises because consumers are heterogeneous and they privately know their own individual demand and valuation, in addition to potential heterogeneity in the formation of beliefs.

In airlines, agents' beliefs about the evolution of sales and prices are formed not only based on current and past realizations of sales and prices for one particular flight. Airlines use historical data, and buyers are likely to be familiar with price patterns based on previous trips. Beliefs are important because consumers can purchase at the ongoing price or delay their purchase decisions. If a consumer is optimistic about future prices, he might expect the possibility of a "last-minute deal" and prefer to postpone his purchase. Forward-looking behavior combined with the existence of higher expected fares closer to the departure — which does not rule out occasional last-minute deals — can lead consumers to buy tickets as soon as they solve their individual demand uncertainty. For concern about its reputation, it is not optimal for the carrier to have frequent last-minute deals because this can result in a large fraction of buyers delaying their purchases. At some point, it may be more valuable for the airline to fly with idle capacity than to set lower fares close to departure.

4.2 Dynamic Demand Estimates

Table 4 reports the results from the estimation of the dynamic demand in Equation 4. DayAdv_t is treated as strictly exogenous in all the specifications. $\ln(\text{Fare})_{ijt}$ is treated as strictly exogenous in columns 1 and 2, as weakly exogenous in columns 3 through 6, and as endogenous in columns 7 and 8.[15] The four system GMM estimates pass all the specification tests. There is strong evidence that u_{ijt} is not serially correlated, and the

[15] Instruments for the first-differenced equations are lags 1 and 2 of $\Delta\text{Load}_{ij,t-1}$ and $\ln(\text{Fare})_{ijt}$ in columns 3 and 5, and additionally lags 3 in columns 4 and 6. The instruments used in the levels equations of columns 5 and 6 are $\Delta^2\text{Load}_{ij,t-1}$ and $\Delta\ln(\text{Fare})_{ijt}$. Treating $\ln(\text{Fare})_{ijt}$ as potentially endogenous invalidates $\ln(\text{Fare})_{ij,t-1}$ and $\Delta\ln(\text{Fare})_{ijt}$ as instruments. Hence, the first-differenced equations in column 7 use lags 1 and 2 of $\Delta\text{Load}_{ij,t-1}$ and $\ln(\text{Fare})_{ij,t-1}$, and in column 8 additionally use lag 3. The levels equations in columns 7 and 8 use $\Delta^2\text{Load}_{ij,t-1}$ and $\Delta\ln(\text{Fare})_{ij,t-1}$.

instruments and the additional instruments are validated by the Sargan and the Difference Sargan respectively.

[Table 4, here.]

All GMM specifications find a statistically significant negative coefficient for ln(Fare) — a downward sloping period t demand. Moreover, there is almost no difference between the estimates that treat prices as weakly exogenous and those that treat prices as endogenous. This is evidence that u_{ijt} represents a shock not only for the econometrician but also for the carrier. The coefficient on ln(Fare) in column 6 indicates that when the price of a ticket in period t increases by 10%, contemporaneous sales decrease by 0.28 seats in a 100-seat aircraft.[16] This estimate is reasonable because if a carrier expects to sell 2 seats on a given period, with a 10% higher price expected sales drop to 1.72 seats. All four system GMM specifications also agree on the sign and provide a similar magnitude for the effect of days in advance. The estimates suggest that sales increase as departure nears: the estimate in column 6 reads that, for each day closer to the departure date, sales increase by 0.0476 seats for a 100-seat aircraft. Additional results presented in the Appendix show very similar estimates for specifications that include a second-order autoregressive term and use different sets of instruments.

5 Dynamic Pricing with Learning

5.1 Separating the Booking Curve and the Demand Shocks

An alternative way to write Equation 4 to illustrate the feedback mechanism from previous prices to cumulative bookings is

$$\textsc{Load}_{ijt} = (1+\rho)\textsc{Load}_{ij,t-1} - \rho\textsc{Load}_{ij,t-2} + \phi\ln(\textsc{Fare})_{ijt} + \delta\textsc{DayAdv}_t + \eta_{ij} + u_{ijt}. \quad (6)$$

When $\phi = 0$, $\textsc{Load}_{ij,t-1}$ in Equation 1 is strictly exogenous. When $\phi \neq 0$, $\textsc{Load}_{ij,t-1}$ is weakly exogenous and depends via $\ln(\textsc{Fare})_{ij,t-1}$ on all past disturbances, not just on $\varepsilon_{ij,t-1}$. Without $\textsc{Load}_{ij,t-2}$ and $\textsc{DayAdv}_t$, Equation 6 is one of the characterizations of

[16] This is obtained using $\Delta\Delta\textsc{Load} \approx (-0.028/100)(\%\Delta\textsc{Fare})$. A 10% increase in price ($\%\Delta\textsc{Fare} = 10$) decreases sales ($\Delta\textsc{Load}$) by 0.0028, which in a 100-seat aircraft is 0.28.

the feedback mechanism presented in Bun and Kiviet (2006).[17] Equation 6 can be written to emphasize the existence of two different components:

$$\text{Load}_{ijt} = E[\text{Load}_{ijt}|\text{Load}_{ij,t-1}, \text{Load}_{ij,t-2}, \ln(\text{Fare})_{ij,t-1}, \text{DayAdv}_t, \rho, \phi, \delta, \eta_{ij}] + u_{ijt}. \tag{7}$$

The first term on the right-hand side is the "anticipated" component, which depends on all previous sales and prices, captured by the lags of Load_{ijt} and $\ln(\text{Fare})_{ijt}$, and the second term is the "unanticipated" component or demand shock at period t.

The anticipated component describes the expected evolution of sales as the departure date nears under normal or average conditions. It is known in airlines and in operations research as the booking curve, BC_{ijt}. Then the difference between actual bookings and the booking curve, $\text{S}_{ijt} \equiv \text{Load}_{ijt} - \text{BC}_{ijt}$, contains the necessary information to know whether capacity in flight i at time t is above or below expectations. The booking curve can be obtained as the fitted values, $\text{BC}_{ijt} = \widehat{\text{Load}}_{ijt}$, and the unexpected component of the demand as the residuals, $\text{S}_{ijt} = \text{Load}_{ijt} - \widehat{\text{Load}}_{ijt}$.

5.2 Dynamic Pricing with Learning Equation

After separating the evolution of cumulative booking into the booking curve and demand shocks, I replace $\text{Load}_{ij,t-1}$ in Equation 1 with $\text{BC}_{ij,t-1}$ and $\text{S}_{ij,t-1}$ to estimate the effect of each of these components on price,

$$\ln(\text{Fare})_{ijt} = \alpha \ln(\text{Fare})_{ij,t-1} + \gamma \text{DayAdv}_t + \beta_{\text{BC}} \text{BC}_{ij,t-1} + \beta_{\text{S}} \text{S}_{ij,t-1} + \nu_{ij} + \varepsilon_{ijt}, \tag{8}$$

with both $\text{BC}_{ij,t-1}$ and $\text{S}_{ij,t-1}$ treated as weakly exogenous. $\text{S}_{ij,t-1}$ is the new piece of information about the pattern of sales that was not available when the carrier set the price last period. A positive β_S coefficient means that if there is a positive demand shock — signaling that actual bookings are above expectations — the carrier sets a higher price, evidence that the airline responds to new information about the pattern of sales. Moreover, aggregate demand learning and price adjustment occurs if the response to a demand shock

[17] Bun and Kiviet (2006) formalize the feedback mechanism to analyze the finite sample behavior of particular least-squares and method-of-moments estimators. A similar characterization is used in Blundell et al. (2000) in Monte Carlo simulations.

is greater than the response to an anticipated sale. There is learning in the sense that the carrier can distinguish between expected sales and demand shocks.

Following the estimation of Equation 1, after first differencing Equation 8 I use the moments $E(\Delta\varepsilon_{ijt}\mathrm{Z}) = 0$. The vector of instruments Z will have lags of $\ln(\textsc{Fare})_{ij,t-1}$, $\mathrm{BC}_{ij,t-1}$, and $\mathrm{S}_{ij,t-1}$ to instrument for $\Delta\ln(\textsc{Fare})_{ij,t-1}$, $\Delta\mathrm{BC}_{ij,t-1}$, and $\Delta\mathrm{S}_{ij,t-1}$, respectively. Moreover, $\Delta\textsc{DayAdv}_t$ serves as its own instrument. For the system estimator, the additional moments $E[(\nu_{ij} + \varepsilon_{ijt})\mathrm{W}] = 0$ for the levels equation use lags of $\Delta\ln(\textsc{Fare})_{ijt}$, $\Delta\mathrm{BC}_{ijt}$, and $\Delta\mathrm{S}_{ijt}$ for the vector of instruments W. Notice that the booking curve and the demand shock that appear in the pricing equation are estimated regressors derived from a first-stage estimation of the dynamic demand. Including BC alone in the estimation of Equation 8 would yield incorrect standard errors because of the additional variation that arises from the estimation error in BC. Equation 8 follows one of the models of Pagan (1984) and accounts for the estimation error associated with the first-step estimation, $\mathrm{S}_{ijt} = \textsc{Load}_{ijt} - \widehat{\textsc{Load}}_{ijt}$, explicitly by including it in the estimated equation.[18]

5.3 Identification

The identification of the pricing and demand equations comes from variation over time in prices and sales. As previously discussed, the identification in the first-difference estimator requires limited serial correlation in the error term and instruments that are exogenous. These assumptions are tested using a serial-correlation test and the Sargan over-identifying restrictions test, respectively. When the variables are persistent the first-difference estimator may suffer from weak instruments, hence the use of the system estimator. The validity of the additional moment conditions in the system estimator is tested with the Difference Sargan test.

Notice that the identification of the demand shocks comes from the specific flight's demand pattern relative to other contemporaneous flights. This is because the estimation of the booking curve uses the same data as the estimation of the pricing equation, with all the flights sharing the same departure date. Airlines have more information than what is used in this paper, and it is reasonable to believe that they use historical data within flights to estimate the booking curve and not necessarily data across flights.

[18] Abowd et al. (1999) have a more recent implementation of this model.

One limitation in this strategy to identify demand shocks is that even though the estimation controls time-invariant characteristics, there may be some time-variant characteristics not captured by the BC that would be captured if I were using historical data. For example, airlines know whether some flights should fill earlier than others (e.g., flights with more tourists fill earlier). In this case, a higher LOAD that is not captured by the BC will appear to be a demand shock. Hence, there is some additional variation in S that corresponds to the BC, and this should bias the estimates against finding learning and price adjustment. One benefit of using data across flights is that all flights in the sample share the same departure date and the same dates prior to departure. This is helpful in controlling for time-variant characteristics common to all flights. If sales are particularly different during specific dates prior to departure (e.g., higher closer to departure), this is known ex-ante by the carrier and will also be part of the estimated BC. Hence, it should not mislead the carrier into thinking that it is a flight-specific demand shock.

5.4 Dynamic Pricing with Learning Estimates

Table 5 reports the parameter estimates of the dynamic pricing with learning equation. To separate cumulative bookings into BC and S, columns 1 and 2 use the estimates from the dynamic demand in column 8 of Table 4, which treats price as endogenous. Columns 3 through 10 use the estimates from column 6 of Table 4, which treats price as predetermined.[19] To illustrate some of the dynamics of S, Figure 2 plots the demand shocks obtained using the estimates in column 6 of Table 4. The fast increase in the load factor at 82 days to departure is captured as a positive demand shock.

[Table 5, here.]

All the columns in Table 5 pass the three specification tests. The two main coefficients of interest — the marginal effects of BC and S — are both positive and highly significant across the first four specifications. Evaluated at the sample mean of fares, the coefficient on BC in column 4 indicates that in an aircraft with 100 seats, every additional expected sale increases fares by \$0.79 (\$291.1×0.273/100). The positive and highly significant coefficient

[19] Estimates of ν_{ij} are not needed because the estimation of Equation 8 controls for time-invariant characteristics.

on S shows strong evidence that prices respond to new information about the pattern of sales. The coefficient on S from column 4 indicates that if there is a positive demand shock in period t, every additional unanticipated sale will increase fares by \$1.92. Across these first four specifications, at at least 1% significance level, the response of prices to an unanticipated sale is greater than the response to an anticipated sale. This is evidence of aggregate demand learning and price adjustment as opposed to prices mechanically adjusting to sales, whether or not they are expected. Columns 5 and 6 show that when allowing for breaks at 7 and 14 days prior to departure the coefficient on BC is no longer significant, but the main result holds: an unanticipated sale has a larger effect on prices than an anticipated sale.

The combination of the estimated coefficients on BC and S can predict price drops for a sufficiently large negative demand shock. For example, using the estimates in column 4 and for a 100-seat aircraft, if the carrier expects to sell two seats but only sells one, then the overall effect is a drop in fares by \$ 0.34 (\$291.1×(2×0.273−1×0.664)/100). The combined effect of BC and S on fares is consistent with the positive DayAdv coefficient found in the estimation of Equation 1. DayAdv considers the particular case in which time to departure changes and a negative demand shock exactly offsets expected sales; hence, no actual sales occur. This can explain why DayAdv is not significant in these specifications.

An obvious implication from aggregate demand learning, price adjustment, and the downward-sloping dynamic demand estimated in Equation 4 is that airlines can use prices to partially control the path of cumulative bookings and affect the final state of the aggregate demand. This explains the nature of the dynamic interaction between prices and sales. Previous prices affect cumulative bookings — Load is weakly exogenous in Equation 1 — and realized demand affects pricing decisions — ln(Fare) has to be treated as endogenous or weakly exogenous in Equation 4.

Columns 7 through 10 provide two additional results. Columns 7 and 8 illustrate how the effect of new information changes with the identity of the carrier. AA is a dummy variable equal to one if the carrier is American Airlines, zero otherwise. The positive and significant coefficient on the interaction term S × AA is slightly bigger that the coefficient on S. This indicates that the effect of a demand shock has a little more than twice the impact on prices in an American Airlines flight than in a flight of a different carrier.[20] Columns 9 and 10 show how learning and price adjustment changes as the departure date nears. Intuitively, being one seat above expectations long before departure should have a smaller impact on fares than being one seat above expectations when there is less time to sell. The coefficients in column 10 evaluated at the sample average of fares illustrate that in a 100-seat aircraft, a demand shock of one additional unexpected sale increases fares by \$2.52 at four weeks before departure, but increases fares by \$3.12 at one week before departure.[21] I present various robustness results in the Appendix.

These aggregate demand learning and price adjustment results are consistent with serial nesting of booking classes. In an expected peak flight, when a higher booking class sells more quickly than expected, serial nesting of booking classes allows sales from a lower booking class to be available. Therefore, for booking classes within the same cabin, inventories for a lower class are always counted as available for higher booking classes. This occurs in such a way that higher booking classes can never be sold out before a lower booking class. On the other hand, seats from a higher booking class might be released into a lower booking class in an expected off-peak flight. This serial nesting changes the lowest available fare in the same way it is observed in the data.

Finally, while focusing on one-way non-stop tickets helps to control for various sources of price dispersion and helps to define the Load variable, the inventory of seats is also sold as part of round-trips and longer itineraries. Even if one-way tickets are a small fraction of overall tickets sold, this should not affect the estimation of the pricing equation as long as the carrier adjusts the observed one-way price based on the current inventory. In the dynamic demand equation, ϕ measures the response of ΔLoad to a change in the logarithm of the one-way (ow) price, $\ln(\text{Fare})^{\text{ow}}$. It is worth noting that this marginal effect may be channeled through the prices of other tickets for the same flight, for example, a round-trip (rt) ticket, $\frac{\partial \Delta \text{Load}}{\partial \ln(\text{Fare})^{\text{rt}}} \cdot \frac{\partial \ln(\text{Fare})^{\text{rt}}}{\partial \ln(\text{Fare})^{\text{ow}}} = \phi$. The estimation is also capturing $\frac{\partial \Delta \text{Load}}{\partial \ln(\text{Fare})^{\text{rt}}} = \phi$ if $\frac{\partial \ln(\text{Fare})^{\text{rt}}}{\partial \ln(\text{Fare})^{\text{ow}}} = 1$, which is the case when the round-trip price is always two times the

[20]Similar specifications with interactions of S with dummies of other carriers found that the interactions were not statistically significant.

[21]Notice that these specifications no longer follow Pagan (1984); hence, I bootstrap the two-stage procedure to obtain bootstrap standard errors, clustered by flight.

one-way price.[22]

6 Conclusion

This paper uses a unique panel data set of U.S. domestic flights to empirically study the dynamics of prices and inventories as the departure date nears. The construction of the data set controls for important sources of price dispersion in the industry (e.g., Saturday-night stayover, fare class, refundability, different connections/legs, minimum- and maximum-stay), while the panel structure is key to control for unobserved time-invariant flight-, route-, and carrier-specific characteristics. The results showed that, for a given inventory, fares decrease as there is less time to sell, with breaks at 7 and 14 days to departure when price increases. Moreover, at a fixed point prior to the departure date, fares rise as the inventory decreases. These findings are consistent with various theoretical models of optimal pricing under uncertain demand and perishable inventories (e.g., Gallego and van Ryzin (1994), Zhao and Zheng (2000), Su (2007) and various references therein). In addition, the higher fares with lower inventories is also consistent with models that predict dispersed prices when capacity is costly, demand is uncertain, and when prices are set ex-ante (e.g., Prescott (1975)).

To assess whether carriers learn about the aggregate demand during the sales season, I estimate a dynamic demand equation and a dynamic pricing equation that jointly characterize the adjustment process between prices and sales as the flight date nears. These two equations are consistent with rational expectations models and allow sellers and buyers to behave dynamically: Current decisions to price and buy can be affected by prior realizations of fares and sales. In addition, agents can form their own beliefs and adopt forward-looking perspectives. The results show that demand shocks have a positive and much larger effect on prices than the positive effect of anticipated sales. This is evidence that carriers differentiate between expected and unexpected sales, and adjust their prices as new information about the pattern of sales is revealed. Aggregate demand learning and price adjustment, combined with a downward-sloping dynamic demand, mean that carriers

[22] This is a standard assumption to obtain the one-way price based on the round-trip price, see for example Borenstein and Rose (1994, p. 677), and Gerardi and Shapiro (2009, p. 5).

can partially control the evolution of cumulative bookings using prices.

Appendix

Table 6 provides robustness results for the estimation of the dynamic demand equation. The first three columns treat ln(Fare) as weakly exogenous, while columns 4 through 6 treat it as endogenous. Instruments for the first-differenced equations in the first three columns are lags of $\Delta\text{Load}_{ij,t-1}$ and $\ln(\text{Fare})_{ijt}$, with column 1 using the first two lags, column 2 the first three lags, and column 4 the first four lags. The instruments used in the levels equations in the first three columns are all the same: $\Delta^2\text{Load}_{ij,t-1}$ and $\Delta\ln(\text{Fare})_{ijt}$. When $\ln(\text{Fare})_{ijt}$ is treated as potentially endogenous, the instruments in the first-differenced equations are lags of $\Delta\text{Load}_{ij,t-1}$ and $\ln(\text{Fare})_{ij,t-1}$. Columns 4, 5, and 6 use the first 2, 3, and 4 lags respectively. For all the levels equations the instruments are $\Delta^2\text{Load}_{ij,t-1}$ and $\Delta\ln(\text{Fare})_{ij,t-1}$. ΔDayAdv_t acts as its own instrument in the difference equation across all specifications.

[Table 6, here.]

The results show that the coefficients on ln(Fare) and DayAdv from Table 4 are robust to the selection of the instrument list and the addition of the second-order autoregressive term. I use the results in columns 3 and 6 — which pass all three specification tests — to obtain alternative estimates of the booking curve and the demand shocks. Table 7 provides robustness results for the dynamic pricing with learning equation using these alternative BC and S. The instruments across different columns follow the same criteria as previous pricing equations. All the specifications show that the response of prices to demand shocks is greater than the response to anticipated sales, which is evidence of aggregate demand learning and price adjustment. Note that in Table 7 the estimates of the autoregressive term suggest that ln(Fare) could be a random walk. Following Blundell and Bond (2000) and Bond et al. (2005), I estimated a simple OLS AR(1) specification for ln(Fare), which rejected the unit-root null. This is important because if ln(Fare) followed a random walk, lagged values of ln(Fare) would be uncorrelated with its first differences and the difference GMM estimator would not identify α. In that case Bond et al. (2005) and Binder et al. (2005) explain that identification comes from the levels equation under the additional assumption that var(ln(FARE)$_{ij1}$)$< \infty$. That is, the variance of the initial logarithm of prices is finite.

[Table 7, here.]

References

Abowd, J.M., F. Kramarz and D.N. Margolis, 1999. High wage workers and high wage firms. *Econometrica*, 67, pp. 251-333.

Arellano, M. and S. Bond, 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, 58, pp. 277-297.

Arellano, M. and O. Bover, 1995. Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68, pp. 29-51.

Bilotkach, V., 2006. Understanding price dispersion in the airline industry: Capacity constraints and consumer heterogeneity. In *Advances in Airline Economics*, Vol. 1, edited by Darin Lee. New York: Elsevier Sci.

Bilotkack, V. and N. Rupp, 2011. A guide to booking airline tickets online. Mimeo, University of California, Irvine.

Binder, M., C. Hsiao and M.H. Pesaran, 2005. Estimation and inference in short panel vector autoregressions with unit roots and cointegration. *Econometric Theory*, 21, pp. 795-837.

Blundell, R.W. and S. Bond, 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87, pp. 115-143.

Blundell, R.W. and S. Bond, 2000. GMM Estimation with persistent panel data: An application to production functions. *Econometric Reviews*, 19, pp. 321-340.

Blundell, R.W., S. Bond and F. Windmeijer, 2000. Estimation in dynamic panel data models: Improving on the performance of the standard GMM estimator. In: Baltagi,

B. (Ed.) Advances in Econometrics, Vol. 15. Nonstationary Panels, Panel Cointegration and Dynamic Panels. JAI Press, Elsevier Science, Amsterdam, pp. 53-91.

Bond, S., C. Nauges and F. Windmeijer, 2005. Unit roots: Identification and testing in micro panels. Institute of Fiscal Studies working paper CWP07/05, London.

Borenstein, S. and N.L. Rose, 1994. Competition and price dispersion in the U.S. airline industry. *Journal of Political Economy*, 102, pp. 653-683.

Boiteaux, M., 1980. Peak load pricing. *Journal of Business*, 33, pp. 157-179.

Bun, M.J.G. and J.F. Kiviet, 2006. The effects of dynamic feedbacks on LS and MM estimator accuracy in panel data models. *Journal of Econometrics*, 132, pp. 409-444.

Carlton, D.W., 1977. Peak load pricing with stochastic demand. *American Economic Review*, 67, pp. 1006-1010.

Dana, Jr. J., 1998. Advance-purchase discounts and price discrimination in competitive markets. *Journal of Political Economy*, 106, pp. 395-422.

Dana, Jr. J., 1999a. Using yield management to shift demand when the peak time is unknown. *Rand Journal of Economics*, 30, pp. 456-474.

Dana, Jr. J., 1999b. Equilibrium price dispersion under demand uncertainty: The roles of costly capacity and market structure. *Rand Journal of Economics*, 30, pp. 632-660.

Deneckere, R., H.P. Marvel and J. Peck, 1996. Demand uncertainty, inventories, and resale price maintenance. *Quarterly Journal of Economics*, 111, pp. 885-913.

Deneckere, R. and J. Peck, 2010. Dynamic competition with random demand and costless search: A theory of price posting. Mimeo, University of Wisconsin.

Eden, B., 1990. Marginal cost pricing when spot markets are complete. *Journal of Political Economy*, 98, pp. 1293-1306.

Elmaghraby, W. and P. Keskinocak, 2003. Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management Science*, 49, pp. 1287-1309.

Escobari, D., 2008. Essays on pricing under uncertainty. Ph.D. Thesis, Texas A&M University.

Escobari, D., 2009. Systematic peak-load pricing, congestion premia and demand diverting: Empirical evidence. *Economics Letters*, 103, pp. 59-61.

Escobari, D. and L. Gan, 2007. Price dispersion under costly capacity and demand uncertainty. NBER working paper 13075, Cambridge, MA.

Gale, I. and T. Holmes, 1993. Advance-purchase discounts and monopoly allocation of capacity. *American Economic Review*, 83, pp. 135-146.

Gallego, G. and G.J. van Ryzin, 1994. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science*, 40, pp. 999-1020.

Gerardi, K.S. and A.H. Shapiro, 2009. Does competition reduce price dispersion? New evidence from the airline industry. *Journal of Political Economy*, 117, pp. 1-37.

Hendel, I. and A. Nevo, 2006. Sales and consumer inventory. *Rand Journal of Economics*, 37, pp. 543-561.

Holtz-Eakin, D., W. Newey and H. Rosen, 1988. Estimating vector autoregressions with panel data. *Econometrica*, 56, pp. 1371-1395.

Kincaid, W.M. and D.A. Darling, 1963. An inventory pricing problem. *Journal of Mathematical Analysis and Applications*, 7, pp. 183-208.

Lazear, E.P., 1986. Retail pricing and clearance sales. *American Economic Review*, 76, pp. 1432.

Lin, K.Y., 2006. Dynamic pricing with real-time demand learning. *European Journal of Operational Research*, 174, pp. 522-538.

Lucas, R.E. and M. Woodford, 1993. Real effects of monetary shocks in an economy with sequential purchases. NBER working paper 4250, Cambridge, MA.

McAfee, R.P. and V. te Velde, 2007. Dynamic pricing in the airline industry. In *Handbook of Economics and Information Systems*, vol. 1, edited by T.J. Hendershott. New York: Elsevier Sci.

McGill, J.I. and G.J. van Ryzin, 1999. Revenue management: Research overview and prospects. *Transportation Science*, 33, pp. 233-256.

Nickell, S., 1981. Biases in dynamic models with fixed effects. *Econometrica*, 49, pp. 1417-1426.

Pagan, A., 1984. Econometric issues in the analysis of regressions with generated regressors. *International Economic Review*, 25, pp. 221-247.

Pesendorfer, M., 2002. Retail sales: A study of pricing behavior in supermarkets. *Journal of Business*, 75, pp. 33-66.

Prescott, E., 1975. Efficiency of the natural rate. *Journal of Political Economy*, 83, pp. 1229-1236.

Puller, S.L., A. Sengupta and S.N. Wiggins, 2008. Testing theories of price dispersion and scarcity pricing in the airline industry. Mimeo, Texas A&M University.

Stavins, J., 2001. Price discrimination in the airline market: The effect of market concentration. *The Review of Economics and Statistics*, 83, pp. 200-202.

Su, X, 2007. Intertemporal pricing with strategic consumer behavior. *Management Science*, 53, pp. 726-741.

Sweeting, A., 2010. Dynamic Pricing Behavior in Perishable Goods Markets: Evidence from Secondary Markets for Major League Baseball Tickets. Mimeo, Duke University.

Visscher, M.L., 1973. Welfare-maximizing price and output with stochastic demand: Comment. *American Economic Review*, 63, pp. 224-229.

Williamson, O.E., 1966. Peak load pricing and optimal capacity under constraints. *American Economic Review*, 56, pp. 810-827.

Zhao, W. and Y.S. Zheng, 2000. Optimal dynamic pricing for perishable assets with nonhomogeneous demand. *Management Science*, 46, pp. 375-388.

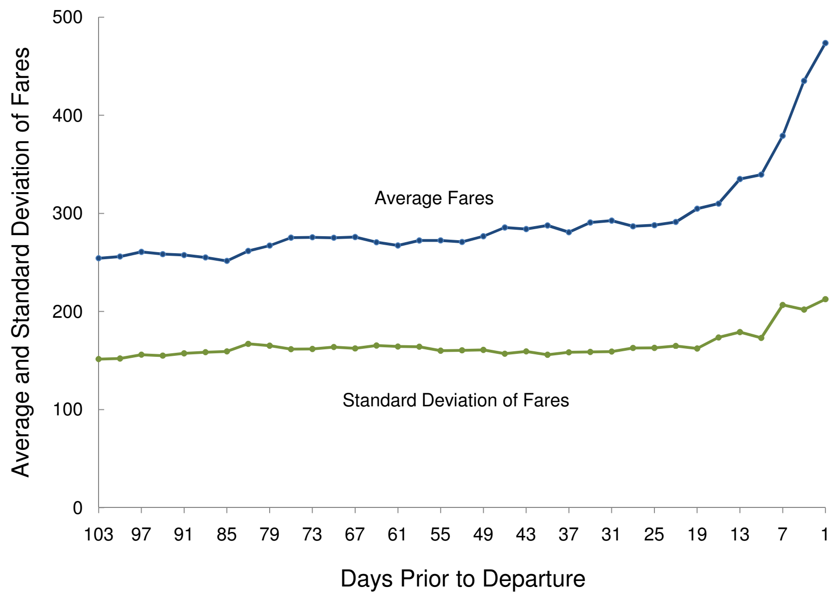

Figure 1: Average and standard deviation of fares

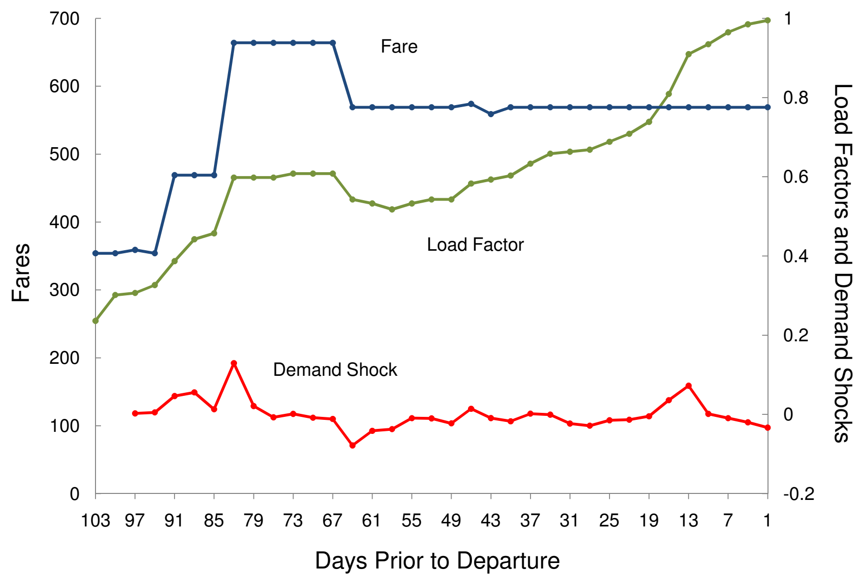

Figure 2: Fares, load factors, and demand shocks (Delta 1588 ATL-SJC)

Table 1: Summary Statistics

VARIABLES	Mean (1)	Std.Dev. (2)	Min. (3)	Max. (4)
FARE [a]	291.087	171.879	54.000	1,224.000
DAYADV	52.289	30.154	1.000	103.000
LOAD	0.509	0.252	0.012	1.000
BC [b]	0.518	0.246	0.017	1.049
S [b]	0.000	0.043	-0.436	0.448
BC [c]	0.518	0.247	0.016	1.051
S [c]	0.001	0.043	-0.438	0.479

Notes: The sample size is 7,933. [a] The standard deviation *between* flights is 152.933 and *within* is 78.751. [b] Based on column 6, Table 4. [c] Based on column 8, Table 4.

Table 2: Dynamic Pricing Estimates

LOAD treated as:	Strictly exogenous				Weakly exogenous			
Estimator:	Within		Pooled	Within	Difference		System	
Instruments:					$t-2$	$t-3$	$t-2$	$t-3$
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\ln(\text{FARE})_{ij,t-1}$	0.723*		0.948*	0.706*	0.827*	0.788*	0.938*	0.932*
	(0.030)		(0.004)	(0.030)	(0.045)	(0.047)	(0.008)	(0.020)
$\text{DAYADV}_t/10^3$	-1.802*	-1.280*	-0.380*	-0.107	1.197*	1.248*	1.763*	1.961*
	(0.115)	(0.349)	(0.093)	(0.171)	(0.257)	(0.242)	(0.278)	(0.671)
$\text{LOAD}_{ij,t-1}$		0.516*	0.106*	0.300*	0.449*	0.482*	0.492*	0.527*
		(0.066)	(0.013)	(0.032)	(0.052)	(0.051)	(0.052)	(0.134)
Serial correlation^a (p-value)					0.874	0.860	0.900	0.902
Sargan^b (p-value)					0.005	0.086	0.492	0.919
Difference Sargan^c (p-value)							1.000	1.000

Notes: The dependent variable is $\ln(\text{FARE})_{ijt}$. Figures in parentheses for the OLS are White heteroskedasticity-consistent estimates of the asymptotic standard errors, $N \to \infty$. For the GMM are the Windmeijer finite-sample corrected standard errors of the GMM two-step estimates. ‡ significant at 10%; † significant at 5%; * significant at 1%. ^{a} For the GMM, the null hypothesis is that the errors in the first-difference regression exhibit no second-order serial correlation (valid specification). ^{b} The null hypothesis is that the instruments are not correlated with the residuals (valid specification). ^{c} The null hypothesis is that the additional instruments used in the levels equations are not correlated with the residuals (valid specification).

Table 3: Dynamic Pricing GMM System Estimates

Instruments:	$t-2$	$t-3$	$t-2$	$t-3$	$t-2$	$t-3$
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(\text{Fare})_{ij,t-1}$	0.943*	0.938*	0.953*	0.947*	0.951*	0.946*
	(0.007)	(0.018)	(0.008)	(0.019)	(0.008)	(0.022)
$\text{DayAdv}_t/10^3$	1.728*	1.923*	1.461*	1.666*	1.547*	1.753*
	(0.274)	(0.618)	(0.286)	(0.626)	(0.294)	(0.661)
$1_{[\text{DayAdv}<7]}$	0.091*	0.090*			0.073*	0.073*
	(0.019)	(0.020)			(0.020)	(0.020)
$1_{[\text{DayAdv}<14]}$			0.060*	0.059*	0.037*	0.035†
			(0.010)	(0.014)	(0.011)	(0.015)
$\text{Load}_{ij,t-1}$	0.434*	0.469*	0.351*	0.390*	0.359*	0.398†
	(0.052)	(0.120)	(0.059)	(0.138)	(0.061)	(0.158)
Serial correlation[a] (p-value)	0.948	0.949	0.840	0.844	0.901	0.905
Sargan[b] (p-value)	0.495	0.920	0.492	0.920	0.500	0.924
Difference Sargan[c] (p-value)	1.000	1.000	1.000	1.000	1.000	1.000

Notes: The dependent variable is $\ln(\text{Fare})_{ijt}$. Figures in parentheses are the Windmeijer finite-sample corrected standard errors of the GMM two-step estimates. ‡ significant at 10%; † significant at 5%; * significant at 1%. [a] [b] [c] See notes on Table 2.

Table 4: Dynamic Demand Estimates

ln(Fare) treated as:	Strictly exogenous		Weakly exogenous				Endogenous	
Estimator:	Pooled	Within	Difference		System		System	
Instruments:			$t-2$	$t-3$	$t-2$	$t-3$	$t-2$	$t-3$
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta\text{Load}_{ij,t-1}$	-0.111*	-0.140*	-0.186*	-0.184*	-0.193*	-0.192*	-0.196*	-0.193*
	(0.032)	(0.027)	(0.025)	(0.026)	(0.025)	(0.025)	(0.025)	(0.025)
$\ln(\text{Fare})_{ijt}$	-0.004*	-0.030*	-0.025*	-0.027*	-0.027*	-0.028*	-0.029*	-0.030*
	(0.001)	(0.003)	(0.007)	(0.007)	(0.005)	(0.005)	(0.006)	(0.006)
$\text{DayAdv}_t/10^3$	-0.340*	-0.446*	-0.424*	-0.446*	-0.461*	-0.476*	-0.474*	-0.483*
	(0.020)	(0.023)	(0.033)	(0.030)	(0.031)	(0.029)	(0.032)	(0.031)
Serial correlation^a (p-value)			0.624	0.681	0.494	0.513	0.436	0.488
Sargan^b (p-value)			0.000	0.004	0.091	0.389	0.084	0.372
Difference Sargan^c (p-value)					0.999	1.000	0.999	1.000

Notes: The dependent variable is ΔLoad_{ijt}. Figures in parentheses for the OLS are White heteroskedasticity-consistent estimates of the asymptotic standard errors, $N \to \infty$. For the GMM are the Windmeijer finite-sample corrected standard errors of the GMM two-step estimates. ‡ significant at 10%; † significant at 5%; * significant at 1%. $^{a\ b\ c}$ See notes on Table 2.

Table 5: Dynamic Pricing with Learning GMM System Estimates

BC and S from:	Table 4, column 8		Table 4, column 6							
Instruments:	$t-2$	$t-3$	$t-2$	$t-3$	$t-2$	$t-3$	$t-2$	$t-3$	$t-2$	$t-3$
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\ln(\text{FARE})_{ij,t-1}$	0.973*	0.966*	0.971*	0.964*	0.995*	0.987*	0.992*	0.985*	0.987*	0.979*
	(0.006)	(0.010)	(0.006)	(0.011)	(0.007)	(0.011)	(0.007)	(0.010)	(0.009)	(0.013)
$\text{DAYADV}_t/10^3$	0.256	0.513	0.325	0.583	-0.180	0.097	-0.082	0.159	0.141	0.415
	(0.254)	(0.361)	(0.251)	(0.423)	(0.266)	(0.409)	(0.290)	(0.369)	(0.366)	(0.455)
$1_{[\text{DAYADV}<7]}$					0.068*	0.068*	0.065*	0.065*	0.074*	0.075*
					(0.020)	(0.018)	(0.018)	(0.018)	(0.018)	(0.019)
$1_{[\text{DAYADV}<14]}$					0.056*	0.053*	0.056*	0.054*	0.048*	0.046*
					(0.010)	(0.010)	(0.010)	(0.010)	(0.009)	(0.012)
$\text{BC}_{ij,t-1}$	0.214*	0.260*	0.227*	0.273*	0.012	0.066	0.032	0.0790	0.064	0.116
	(0.043)	(0.067)	(0.043)	(0.079)	(0.053)	(0.076)	(0.047)	(0.068)	(0.060)	(0.093)
$\text{S}_{ij,t-1}$	0.636*	0.668*	0.629*	0.664*	0.537*	0.573*	0.462*	0.488*	1.074*	1.140*
	(0.102)	(0.130)	(0.101)	(0.136)	(0.106)	(0.141)	(0.150)	(0.165)	(0.219)	(0.284)
$\text{S}_{ij,t-1}\times$ AA							0.562†	0.599†		
							(0.275)	(0.247)		
$\text{S}_{ij,t-1} \times \text{DAYADV}_t/10^3$									-9.202†	-9.837†
									(3.948)	(4.868)
Serial correlation[a] (p-value)	0.894	0.896	0.895	0.897	0.872	0.877	0.888	0.894	0.871	0.875
Sargan[b] (p-value)	0.381	0.874	0.380	0.876	0.382	0.877	0.995	1.000	0.383	0.876
Difference Sargan[c] (p-value)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$H_0: \beta_{BC} = \beta_S$[d] (p-value)	0.000	0.002	0.000	0.000	0.000	0.000				

Notes: The dependent variable is $\ln(\text{FARE})_{ijt}$. The figures in parentheses in columns 1 through 4 are the Windmeijer finite-sample corrected standard errors of the GMM two-step estimates. For columns 7 through 10 are bootstrap standard errors based on the two-step procedure, 500 replications and clustered by flight. ‡ significant at 10%; † significant at 5%; * significant at 1%. [a] [b] [c] See notes on Table 2. [d] The null hypothesis is that coefficients on expected and on unexpected sales (demand shocks) are the same.

Table 6: Dynamic Demand GMM System Estimates

ln(Fare) treated as:	Weakly exogenous			Endogenous		
Instruments:	$t-2$	$t-3$	$t-4$	$t-2$	$t-3$	$t-4$
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta\text{Load}_{ij,t-1}$	-0.257*	-0.236*	-0.225*	-0.262*	-0.240*	-0.227*
	(0.038)	(0.038)	(0.029)	(0.038)	(0.038)	(0.037)
$\Delta\text{Load}_{ij,t-2}$	-0.097*	-0.074*	-0.061*	-0.097*	-0.076*	-0.063*
	(0.024)	(0.023)	(0.020)	(0.025)	(0.024)	(0.022)
$\ln(\text{Fare})_{ijt}$	-0.027*	-0.028*	-0.029*	-0.030*	-0.031*	-0.033*
	(0.005)	(0.005)	(0.009)	(0.007)	(0.006)	(0.010)
$\text{DayAdv}_t/10^3$	-0.530*	-0.530*	-0.523*	-0.548*	-0.544*	-0.540*
	(0.039)	(0.037)	(0.047)	(0.041)	(0.040)	(0.052)
Serial correlation[a] (p-value)	0.033	0.095	0.172	0.049	0.099	0.162
Sargan[b] (p-value)	0.067	0.343	0.787	0.059	0.329	0.814
Difference Sargan[c] (p-value)	1.000	1.000	1.000	1.000	1.000	1.000

Notes: The dependent variable is ΔLoad_{ijt}. Figures in parentheses for the OLS are White heteroskedasticity-consistent estimates of the asymptotic standard errors, $N \to \infty$. For the GMM are the Windmeijer finite-sample corrected standard errors of the GMM two-step estimates. ‡ significant at 10%; † significant at 5%; * significant at 1%. [a] [b] [c] See notes on Table 2.

Table 7: Dynamic Pricing with Learning GMM Estimates

BC and S from:	Table 6, column 3				Table 6, column 6			
Instruments:	$t-2$	$t-3$	$t-2$	$t-3$	$t-2$	$t-3$	$t-2$	$t-3$
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\ln(\text{Fare})_{ij,t-1}$	0.981*	0.974*	1.005*	0.997*	0.978*	0.970*	1.001*	0.993*
	(0.006)	(0.012)	(0.007)	(0.013)	(0.006)	(0.010)	(0.007)	(0.012)
$\text{BC}_{ij,t-1}$	0.152*	0.201†	-0.063	-0.007	0.183*	0.232*	-0.032	0.024
	(0.044)	(0.081)	(0.054)	(0.086)	(0.044)	(0.067)	(0.053)	(0.084)
$\text{S}_{ij,t-1}$	0.634*	0.665*	0.538*	0.578*	0.614*	0.644*	0.523*	0.561*
	(0.103)	(0.136)	(0.107)	(0.131)	(0.101)	(0.129)	(0.105)	(0.129)
$\text{DayAdv}_t/10^3$	-0.121	0.159	-0.601†	-0.304	0.048	0.328	-0.433	-0.137
	(0.269)	(0.492)	(0.282)	(0.473)	(0.261)	(0.369)	(0.275)	(0.460)
$1_{[\text{DayAdv}<7]}$			0.067*	0.067*			0.068*	0.068*
			(0.020)	(0.018)			(0.020)	(0.018)
$1_{[\text{DayAdv}<14]}$			0.058*	0.056*			0.057*	0.055*
			(0.010)	(0.010)			(0.010)	(0.009)
Serial correlation^a (p-value)	0.883	0.885	0.860	0.864	0.884	0.886	0.862	0.866
Sargan^b (p-value)	0.265	0.792	0.277	0.795	0.264	0.794	0.273	0.797
Difference Sargan^c (p-value)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$H_0 : \beta_{BC} = \beta_S{}^d$ (p-value)	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000

Notes: The dependent variable is $\ln(\text{Fare})_{ijt}$. Figures in parentheses are the Windmeijer finite-sample corrected standard errors of the GMM two-step estimates. ‡ significant at 10%; † significant at 5%; * significant at 1%. $^{a\ b\ c}$ See notes on Table 2. ^{d} The null hypothesis is that coefficients on expected and on unexpected sales (demand shocks) are the same.