Incorporating Search and Sales Information in Airline Demand Estimation

Ali Hortaçsu

University of Chicago and NBER

Olivia R. Natan

University of Chicago–Booth School of Business

Hayden Parsley

Tobin Center for Economic Policy

Timothy Schwieg

Becker Friedman Institute

Kevin R. Williams*

Yale School of Management and NBER

June 2020

<u>Preliminary; Comments welcome; Please do not cite</u>

Abstract

We propose a novel methodology to estimate demand in advance purchase markets where demand is uncertain and the willingness to pay of arriving consumers changes over time. We estimate the distribution of stochastic consumer arrivals as well as consumer preferences by combining information on consumer search activity with product-level sales. The discrete choice model allows for flexible discrete unobserved heterogeneity, endogenous prices, and accommodates zero sale observations at the product level. We apply the method to rich data provided by a large international air carrier based in the United States. The model fits the data well, even in periods in which sales are infrequently observed. We find that there is a significant increase in willingness to pay toward the departure date as an increasing number of price insensitive consumers look for tickets.

*The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research. We thank the anonymous airline for giving us access to the data used in this study. Under the agreement with the authors, the airline had "the right to delete any trade secret, proprietary, or Confidential Information" supplied by the airline. We agreed to take comments in good faith regarding statements that would lead a reader to identify the airline and damage the airline's reputation. All authors have no material financial relationships with entities related to this research.

Emails: hortacsu@gmail.com, olivia.natan@chicagobooth.edu, timothyschwieg@uchicago.edu, hayden.parsley@yale.edu, kevin.williams@yale.edu

1 Introduction

Granular data often pose challenges for existing demand modeling techniques due to the presence of zero sales observations. Logit-style demand systems are popular and easy to compute, but widely used estimation methodologies typically require zero sale observations to be dropped. Simply removing the zeros creates a selections bias in the demand estimates, and aggregating the data may not be feasible in many applications. In advance purchase markets—such as in the airline, hotel, and entertainment industries—these zero-sales observations may arise for a particular reason—demand is stochastic. In the extreme case, no consumers may consider purchasing on a given day. On the other hand, consumers who do in fact consider making a purchase may end up opting not to do so if the current price is sufficiently high.

Distinguishing between low demand (few consumers entering the market) and elastic demand (few consumers opting to purchase) is essential for optimizing prices and product availability. For example, understating the amount of uncertainty in demand in a market will lead to an overstatement in estimates of consumer price insensitivity. Because observed demand is an outcome that depends on both preferences and uncertainty, it is difficult to ascertain the role either plays in markets based on particular realizations of demand alone.

In this paper, we propose a new demand methodology to investigate product-level demand that simultaneously estimates the magnitude of stochastic demand along with consumer preferences. Although we tailor our specification to the airline industry, the method is appropriate for many settings in which zero sale observations are observed and the researcher has access to both search and sales data. The method addresses sparse sales at granular levels without requiring data to be aggregated along common dimensions, such as over products or time. Importantly, the method recovers the stochastic process that causes variation in observed product sales, controlling for consumer preferences. This is especially important in the airline context, because estimates of the arrival process can be used as inputs for forecasting models and can be incorporated into demand learning frameworks.

The classic problem of missing "no purchase" data is well known in operations research. Historical sales data is likely insufficient to estimate demand when sales are sufficiently sparse, since these data do not contain the finite number of consumers who opted to not purchase. In such contexts, empirical shares (sales ÷ market size) are random variables and we cannot appeal to the law of large numbers when the number of purchase opportunities is small. In order to estimate consumer preferences when demand is stochastic, it is essential to control for the variation in sales that are driven by stochastic arrivals. Much of the previous work that models stochastic demand—for example Gallego and Van Ryzin (1994), Talluri and Van Ryzin (2004), and Vulcano, Van Ryzin, and Ratliff (2012), either assumes the arrival rate or notes the difficulty in separating elastic demand from low arrivals. Moreover, recent advancements in the estimation of nonparametric demand systems assume that observed product shares represent true purchase probabilities (Compiani, 2019; Tebaldi, Torgovitsky, and Yang, 2019). Although these methods produce estimates of very flexible demand systems, they would necessarily lead to incorrect inferences of consumer preferences when demand is stochastic. For example, few sales would be rationalized through elastic demand, when, in fact, few sales may occur because few consumers are interested in buying.

Whereas demand is typically estimated solely with historical sales data, our approach incorporates consumer search activity to inform the magnitude of stochastic demand. By using search data to measure the number of "no purchases", we estimate how demand uncertainty varies over time. This has been shown to be an important consideration for optimal pricing (Bitran and Mondschein, 1997; Bitran, Caldentey, and Mondschein, 1998; Zhao and Zheng, 2000). Moreover, our approach also allows for a rich demand specification. The model scales well and is able to produce granular demand estimates—for example, we estimate demand for flights by day-of-week and time-of-day. Practitioners may find the method appropriate for estimating differences in willingness-to-pay across product lines.

The demand model combines elements of stochastic demand models in operations research with random coefficients demand specifications widely used in empirical industrial organization and quantitative marketing research. Following work in operations (Talluri and Van Ryzin, 2004; Vulcano, van Ryzin, and Chaar, 2010), we consider a discrete time demand model in which consumer arrivals are modeled via a Poisson distribution. We allow for time-varying arrival parameters and estimate them using search data. We allow for unobserved preference heterogeneity through a discrete random coefficients specification, following Berry, Carnall, and Spiller (1996). Incorporating time-varying consumer heterogeneity is important both for accurately characterizing demand and for appropriately applying revenue management methodologies. Arriving consumers belong to one of two types, which we label as "business" or "leisure" travelers. We allow the fraction on consumer types to change over time in a flexible and unrestricted way. This allows us to measure changes in the composition of arriving consumers toward the departure date. We incorporate a number of restrictions common to the literature, including that arriving consumers do not delay purchase for potential fare decreases. In addition, we assume consumers solve straightforward discrete choice utility maximization problems. The structure of the demand model is similar to Williams (2020). However, our method does not use supply-side conditions for estimation and we incorporate a flight-level unobservable that is potentially correlated with price.

Information on consumer search behavior is used to address the presence of zero sale observations in daily flight demand. As noted in Berry, Linton, and Pakes (2004), zero sales are problematic because these observations must be dropped from analysis, and this leads to a selection bias in the demand estimates. Recent methodologies have been proposed to address this issue, including adjusting zero sales off the bound (Gandhi, Lu, and Shi, 2017) and modeling the occurrence as stemming from incomplete product consideration (Dubé, Hortaçsu, and Joo, 2020). In Quan and Williams (2018), zeros are rationalized by modeling the finite multinomial distribution that generates observed purchases; however, the presence of zeros in this application are driven by purchase opportunities being small relative to the size of the choice set. In airline markets, zeros are driven by the extensive margin of search—whether any product is considered—rather than variation in which subsets of products are considered. We explain sampling variation in sales as being driven by the presence of stochastic consumer arrivals. Importantly, our approach differs from the prior literature in that we explicitly model the underlying process that generates sparsity in observed sales.

We use a Bayesian Markov Chain Monte Carlo approach for estimation. We extend the methods of Jiang, Manchanda, and Rossi (2009) and Rossi, Allenby, and McCulloch (2012) by developing a Gibbs sampling routine for Poisson demand models with random-effects. A major advantage of this approach is that it allows us to consider randomness in sales from two sources—arrivals and random-effects in utility. This allows us to estimate a rich demand specification with price endogeneity while simultaneously rationalizing the observed sparsity of sales. Estimation via MCMC provides computational advantages. By searching for parameter values via simulation instead of optimization, we sidestep solving a complex non-linear system.[1]

We first demonstrate the performance of the method using simulation studies. We then apply

[1]The existence of priors helps identification by adding concavity to posteriors. This improves convergence of the algorithm relative to non-linear optimization techniques which are known to get “stuck” in flat regions of the objective function.

our proposed methodology to granular data provided by a large international air carrier based in the United States. In addition to detailed bookings, prices, and product characteristics, we derive a measure of aggregate search activity by counting the number of consumers who initiate bookings requests for relevant itineraries on the air carrier's website and mobile application. Although we observe all bookings, do we not observe all relevant searches. In particular, we do not observe searches for tickets purchased through certain travel agencies. We propose a modification to our empirical procedure to account for searches made via indirect booking channels.

In our empirical application, we estimate demand for select origin-destination pairs where the carrier offers several flight frequencies a day. We estimate time-of-day and departure date preferences. In addition, we allow the rate of arrivals to depend on the departure date and time until departure. We find that consumers who arrive close to departure are less price sensitive than the consumers who arrive well in advance of the departure date. Our reported demand elasticities range between -3.0 and -1.1, depending on the departure date, departure time, and the number of days before departure. Our flexible fixed effects specification captures systematic features of the markets studied, including higher demand for specific days of the week and a distaste for early morning flights.

The rest of the paper is organized as follows. In Section 2, we introduce the consumer demand model. In Section 3, we outline the Bayesian estimation approach. In Section 4, we present a Monte Carlo study to demonstrate the performance of the estimator. In Section 5, we introduce the data and estimate demand in select flight markets. The conclusion follows.

2 Model of Demand

We consider the demand for flights for a particular origin-destination pair departing on a particular departure date d. We focus on consumers booking nonstop itineraries from a single air carrier and abstract from potential correlations in demand across alternative connecting flight options and alternative departure dates. Time is discrete. The definition of a market is an origin-destination (r), departure date (d), and day before departure (t) tuple. The booking horizon for each flight j leaving on date d is $t \in \{0, ..., T\}$. The time index counts down towards zero; the first period of sale is $t = T$, and the flight departs when the time index equals zero. In each of the sequential markets t, all consumers face a single price for each flight j.[2] Arriving consumers simultaneously choose the flights that maximize their individual utilities. If demand exceeds remaining capacity for any flight offered, we incorporate the rationing assumptions noted below.

An integer number of consumers arrive according to a stochastic process ($A_{r,t,d}$) that is independent of consumer demand. We formalize the arrival process in Section 2.2. Each consumer solves a discrete choice utility maximization problem, described in detail in Section 2.1. Consumers choose from the set of nonstop flights leaving on the focal departure date, which we denote by $J(r,t,d)$, and the outside option. Each consumer purchases a single ticket on one of flight options or chooses the outside option, which corresponds to not traveling (via this airline). We denote this option by $j = 0$. Flights are imperfect substitutes, and we allow for product-specific preferences. We assume arriving consumers do not substitute across alternative departure dates d' or wait to purchase on an alternative day before departure t'. For notational parsimony, we suppress the r subscript in the notation; all parameters are origin-destination specific.

Because capacity is limited, it may be the case that more consumers wish to travel than there are seats remaining. We incorporate two assumptions that greatly simplify the demand system. Note that these assumptions are only binding in the period in which a flight actually sells out (only 5.4 percent of flights sell out in advance of the departure date). Denote $C_{j,t,d}$ to be the remaining capacity for flight j leaving on date d, t days before departure. First, we assume that consumers are not forward looking and do not strategically choose flights based on remaining capacity, $C_{j,t,d}$. This avoids the complication that consumers may choose a less preferred option in order to increase the chances of securing a seat. Second, we assume that when demand exceeds remaining capacity for a particular flight, consumers who chose that flight are randomly shuffled; the first $C_{j,t,d}$ are selected and the rest receive the outside option. We highlight the importance of this assumption in the following subsection.

2.1 Utility Specification

Arriving consumers have heterogeneous preferences, and we adopt a random coefficients demand model that is commonly applied to the airline setting. Following Berry, Carnall, and Spiller (1996),

[2] It is possible for a fare class to sell out mid-day. If that occurs, late-arriving consumers would face a higher price. In our data, this occurs very rarely; we abstract away from within-day price changes.

we assume consumers are one of two types, corresponding to leisure (L) travelers and business (B) travelers. An individual consumer is denoted as i and her consumer type is denoted by $\ell \in \{B, L\}$. These types need not correspond exactly to the lay definition of leisure and business travel; instead they capture discrete unobserved heterogeneity. The probability that an arriving consumer is a business traveler is equal to γ_t.

We assume the indirect utilities are linear in product characteristics and given by

$$u_{i,j,t,d} = \begin{cases} X_{j,t,d}\beta - p_{j,t,d}\alpha_{\ell(i)} + \xi_{j,t,d} + \varepsilon_{i,j,t,d}, & j \in J(t,d) \\ \varepsilon_{i,0,t,d}, & j = 0 \end{cases}.$$

In the specification, $X_{j,t,d}$ denote product characteristics other than price $p_{j,t,d}$. Consumer preferences over product characteristics and price are denoted by $\left(\beta, \alpha_\ell\right)_{\ell \in \{B,L\}}$. For notational parsimony, we commonly refer to the collection $\{\alpha_B, \alpha_L\}$ as α. Our specification allows for heterogeneity in preferences on price, but we assume that consumers have common preferences for other product characteristics. The term $\xi_{j,t,d}$ denotes an unobserved (to the econometrician) product quality that is allowed to be correlated with price. The term $\varepsilon_{i,j,t,d}$ is an unobserved random component of utility and is assumed to be distributed according to a type-1 extreme value distribution. All consumers solve a straightforward utility maximization problem; consumer i chooses flight j if, and only if,

$$u_{i,j,t,d} \geq u_{i,j',d,t}, \forall j' \in J \cup \{0\}.$$

The distributional assumption on the idiosyncratic error term leads to analytical expressions for the individual choice probabilities of consumers (Berry, Carnall, and Spiller, 1996). In particular, the probability that consumer i wants to purchase a ticket on flight j is equal to

$$s^i_{j,t,d} = \frac{\exp\left(X_{j,t,d}\beta - p_{j,t,d}\alpha_{\ell(i)} + \xi_{j,t,d}\right)}{1 + \sum_{k \in J(t,d)} \exp\left(X_{k,t,d}\beta - p_{k,t,d}\alpha_{\ell(i)} + \xi_{k,t,d}\right)}.$$

Since consumers are one of two types, we define $s^L_{j,t,d}$ be the conditional choice probability for a leisure consumer (and $s^B_{j,t,d}$ for a business consumer). Integrating over consumer types, we have

$$s_{j,t,d} = \gamma_t s^B_{j,t,d} + (1-\gamma_t) s^L_{j,t,d}.$$

Although $s(\cdot)$ has an analytical form, there is no analytic inversion to recover ξ. However, since we can write the analytic form of $s(\cdot)$ as a function of ξ, we write this function conditional on the demand parameters to emphasize that the ξs are the unobserved random effects in the model. These random effects ξ function as residuals, explaining the components of demand unobserved to the econometrician that still influence the purchasing decision. These may include advertisements by the airline or conferences and events that yield higher demand for particular departure dates. Any distribution assumed on ξ implies a distribution of the shares. This will later be used to determine the likelihood of sampled shares. We denote this relationship by $f(\cdot)$,

$$s_{j,t,d} = f\Big(\xi_{j,t,d}, \beta, \alpha, \gamma \,|\, X_{j,t,d}\Big).$$

The market shares determine the probability of an arriving customer purchasing, but the sparsity of observed sales forces us to model shares as unobserved. That is, we cannot simply average observed sales and equate to market shares because zero market shares can be due to zero consumer arrivals or no arriving consumers wanting to purchase. Observed sales in the data are $q_{j,t,d}$, which are not only a function of the product shares, but also the number of people that arrive in each time period. This is important because in periods with low arrivals, the probability of $q_{j,t,d} = 0$ is quite high, but the probability of $s_{j,t,d} = 0$ is never equal to zero. Thus, when we observe few arrivals, we must properly account for the sampling variation to be expected in sales quantities.

Finally, we allow ξ, the unobserved component of demand, to be correlated with price. We control for this endogeneity with traditional price instruments. Instrumental variables are correlated with price but unrelated to per-period demand shocks. Our treatment of ξ allows for a rich level of price endogeneity and maintains the interpretation of unobservable demand determinants as in previous demand estimation methodologies (Berry, 1994; Berry, Levinsohn, and Pakes, 1995).

2.2 Stochastic Arrivals

We assume consumer arrivals follow a Poisson distribution with rate $\lambda_{t,d}$. Note that these rates are unobserved to the econometrician, as each rate corresponds to a single observation (the data are at the t,d level and one cannot estimate the rate of a Poisson distribution with a single observation). We will place additional parametric assumptions on the Poisson rates that allow for identification in the following section. Search and purchase data inform consumer arrivals, however, it may be that some searches are unobserved. For example, a consumer who searches and purchases through a travel agency will result in a sale, but no search activity will be recorded on the air carrier's website. To account for a fraction of observed searches, we introduce an exogenous, fixed-proportion scaling parameter ζ. Implicitly, we assume that consumers who search and purchase via other channels have the same distribution of preferences.[3] It follows that

$$A_{t,d} \sim \text{Poisson}\Big(\lambda_{t,d} \cdot \zeta\Big).$$

Three assumptions allow us to construct analytic expressions for demand: arrivals are independent of preferences; consumers have no knowledge of remaining capacity; consumers solve the above utility maximization problems. In situations in which a flight is observed to sell out, capacity is randomly rationed: consumers are randomly ordered, the first $C_{j,t,d}$ are selected to travel, and the remaining consumers receive the outside option. With these assumptions, conditional on prices and product characteristics, demand for each flight j is distributed according to a Poisson distribution. Specifically, demand is distributed

$$\tilde{q}_{j,t,d} \sim \text{Poisson}\Big(\lambda_{t,d} \cdot s_{j,t,d}\Big).$$

Importantly, the demand for flight j is independent of the demand for flight j', conditional on $X_{t,d}$ and $p_{t,d}$. We denote $q_{j,t,d}$ to be the realized demand for flight j; it is equal to

$$q_{j,t,d} = \min\Big\{\tilde{q}_{j,t,d}\,,\, C_{j,t,d}\Big\}.$$

3 Estimation Approach

In this section, we discuss how we use search data to separate consumer arrivals from preferences as well as a Bayesian procedure to estimate the demand model. We adopt a Bayesian approach to demand estimation that builds upon existing discrete choice demand methodologies (Rossi and Allenby, 1993; Jiang, Manchanda, and Rossi, 2009; Rossi, Allenby, and McCulloch, 2012).

[3] We conduct robustness to our assignment of ζ.

3.1 Using search information to separate arrivals from preferences

In our model, consumers may arrive infrequently and in small numbers—leading to many zero shares. Thus, we cannot treat shares as fixed and appeal to a law of large numbers in the number of purchasing opportunities. If we treated shares as fixed, observing few sales would lead us to believe that demand is elastic. However, if the number of consumer arrivals is also small, we would expect to see few, if any, sales in a given period. We solve the issue of missing "no purchase" data by directly measuring the magnitude of stochastic demand. We do so by using search data to inform the arrival process.

Although the estimation procedure described below jointly estimates arrivals and preferences, identification of the model parameters come from distinct data series. The arrival process is estimated from consumer search data. This is similar to a Poisson regression model in that the arrival rates correspond to regression coefficients of search counts on a vector of control variables. With these estimates, we explicitly model the probability of observing various sales quantities. For example, consider a search series of mostly zeros and ones. Our model would estimate a low arrival rate. Also suppose that sales also are mostly zeros and ones, with more ones than zeros. The model then estimates preferences, given the arrival rates, based on the probability distribution of sales. Because the arrival rates are low and the shares are relatively high in this example, we would infer that consumers have inelastic demands.

Estimation of the full model differs in two key ways. First, the full model contains an unobservable component of demand that is potentially correlated with price. We assume a particular correlation structure between this demand shock and price. We address price endogeneity with instrumental variables; in our empirical application, we use cost shifters and variables that proxy for the opportunity cost of capacity. Second, the full model also has random coefficients that capture unobserved heterogeneity in consumer preferences. We estimate the price coefficient and the probability of each consumer type based on the distribution of sales, conditional on prices and the arrival rates. For example, suppose the arrival rates are constant over time. If prices rise substantially and the purchasing rate stays constant, the model will infer that later arrivals are less price sensitive. Our empirical approach matches the distribution of sales to predicted model sales, given the estimates of the arrival process.

3.2 Estimation Procedure

We use a Bayesian approach to demand estimation. The Bayesian approach has a some benefits relative to traditional approaches to demand estimation, specifically, relative to using linear generalized method of moments (GMM) in Berry (1994) and Berry, Levinsohn, and Pakes (1995). First, our model is nonlinear due to modeling demand as a Poisson distribution. The Bayesian approach removes the need to solve a series of nonlinear optimality conditions. Second, we have found that the Bayesian approach scales well. This allows us to accommodate rich arrival process covariates as well as many fixed effects in the utility specification.

In the Bayesian approach, we start with a prior distribution, $p(\theta)$, on consumer arrivals and preferences. We then build up posterior distributions, $p(\theta|\text{data})$, by constructing the likelihood of the model—that is,

$$p(\theta\,|\,\text{data}) \propto p(\text{data}\,|\,\theta)p(\theta).$$

The output of the demand estimation procedure are the posterior distributions, conditional on the historical data observed.

To sample from the posterior distributions, we use a hybrid-Gibbs sampler. This iterates through all the parameters of the model. In each step, we find the marginal distribution of the parameter of interest holding all others constant. We use conjugate prior distributions whenever possible to simplify drawing from the posterior distribution. When conjugate priors are not available, we use the Metropolis-Hastings algorithm to draw from the posterior.

Estimation is split into several distinct parts: arrivals, shares, demand parameters, and the price endogeneity parameters. Arrival parameters allow us to rationalize zero sale observations; share draws allow us to use traditional demand analysis for the preference parameters; and the price endogeneity parameters allow for price to be correlated with the unobserved components of demand.[4] In the following subsections, we discuss each step of the sampler.

[4] We start estimation by initializing the first draws from parameter and share distributions. We initialize the arrival rates using a Poisson regression of the search data against both departure dates and day before departure. Shares are initialized as purchases divided by the observed searches. The remaining demand parameters are initialized following Berry (1994). Given shares, we assign $\gamma_t = .5$ and then invert the demand system to recover mean utilities. We set the initial β based on the output of the instrumental variables regression of mean utilities on demand covariates.

3.2.1 Updating arrivals, λ

To update consumer arrival rates, λ, we use both the search and the quantity data. The likelihood is the joint probability of seeing these outcomes ($A_{t,d}$ and $q_{j,t,d}$), taking $s_{j,t,d}$ as given. Note that we only have a single arrival observation, $A_{t,d}$, to make inference about each arrival rate $\lambda_{t,d}$. We cannot estimate this level of arrival rates from a single observation. Thus, we place additional restrictions on the arrival process. We allow λ to vary between day of departure and days before departure without imposing a covariance structure between different $\lambda_{t,d}$s. We assume

$$\lambda_{t,d} = \exp(\theta_t + \theta_d).$$

This allows for the search data from the same departure day to be relevant for other days of departure and reduces the complexity of the state space.[5]

For each pair t, d where capacity does not bind, we can construct the likelihood of arrivals based on

$$\begin{aligned} q_{j,t,d} &\sim \text{Poisson}(\lambda_{t,d} \cdot s_{j,t,d}), \\ A_{t,d} &\sim \text{Poisson}(\lambda_{t,d} \cdot \zeta). \end{aligned} \tag{1}$$

However, it may be the case that capacity is binding for a particular flight j at time (t, d). When this occurs, we observe a right-censored estimate of the true number of individuals that wished to purchase. That is, for a given capacity $C_{j,t,d}$,

$$\begin{aligned} q_{j,t,d} &= \min\left\{\tilde{q}_{j,t,d}\, , \, C_{j,t,d}\right\}, \\ \tilde{q}_{j,t,d} &\sim \text{Poisson}(\lambda_{t,d} \cdot s_{j,t,d}), \\ A_{t,d} &\sim \text{Poisson}(\lambda_{t,d} \cdot \zeta). \end{aligned} \tag{2}$$

We draw the parameters θ_t and θ_d sequentially according to a Metropolis-Hastings step using a candidate normal distribution to sample. To construct the likelihood given candidate values of θ_t and θ_d, we account for the censoring in quantity. When $q_{j,t,d}$ is not equal to the remaining capacity of the plane, the likelihood contribution of from quantities is the probability mass function of the

[5] Within each route, this amounts to restricting the shape of arrivals towards departure to evolve the same way for all departure dates. Each departure date shifts the shape by the multiplier $\exp(\theta_d)$. This is not restrictive in our context. In other contexts, alternative parameterizations may be more appropriate.

Poisson distribution evaluated at $q_{j,t,d}$: $f_x(q_{j,t,d}|\theta_t, \theta_d)$, based on Equation 1. When the capacity constraint is observed to bind, the contribution is instead $1 - F_x(q_{j,t,d}|\theta_t, \theta_d)$, based on Equation 2.

3.2.2 Updating shares, $s(\cdot)$

We treat product shares as unobserved because we do not appeal to a law of large numbers on market sizes. The posterior distribution of shares is shaped by both the quantity data and the demand parameters (including the random effect ξ). Randomness factors into product shares two ways. First, quantities are impacted by the sampling of the arrivals. Second, shares are influenced by the the random effect of the unobserved components of demand, ξ.

The likelihood of the shares is constructed from data on quantities (and product characteristics), given our estimates of the arrival process. We also incorporate capacity constraints, which impact quantities, into the likelihood for shares. For flights that are not observed to sell out, we appeal to Equation 1 and construct the likelihood of seeing a particular quantity q sold based on

$$q_{j,t,d} \sim \text{Poisson}\Big(\lambda_{t,d} \cdot s_{j,t,d}\Big)$$

For flights in which the capacity constraint binds, we again evaluate the expressions in Equation 2:

$$q_{j,t,d} = \min\Big\{\tilde{q}_{j,t,d}\,,\, C_{j,t,d}\Big\}, \text{ where } \tilde{q}_{j,t,d} \sim \text{Poisson}(\lambda_{t,d} \cdot s_{j,t,d})$$

We treat the demand shock ξ as a random effect, which induces a distribution on shares, regardless of the data observed. Conditional on α, β, γ, and the data, the shares are a function of the random effect. This unobservable ξ influences market shares and the distribution we impose on it affects how the shares are distributed. However, since ξ is assumed to be correlated with price, we allow for the distribution of ξ to be correlated with the firm's pricing decision. We follow the standard Bayesian framework for simultaneity with discrete choice models (Rossi and Allenby, 1993; Jiang, Manchanda, and Rossi, 2009; Rossi, Allenby, and McCulloch, 2012). Using a set of price instruments $Z_{j,t,d}$, we assign

$$\left.\begin{aligned} \xi_{j,t,d} &= f^{-1}\Big(s_{j,t,d}\,|\,\beta, \alpha, \gamma, X_{t,d}\Big) \\ \upsilon_{j,t,d} &= p_{j,t,d} - Z'_{j,t,d}\eta \end{aligned}\right\} \sim \mathcal{N}_{\text{iid}}(0, \Sigma) \quad \text{that such} \quad \Sigma = \begin{pmatrix} \tau & \rho \\ \rho & \kappa \end{pmatrix}.$$

This treatment of ξ allows for a rich level of price endogeneity. However, the approach also rationalizes zero sale observations. Correlation between the unobserved components of demand and price is captured by ρ.

Randomness enters into this model in two locations, both in $\xi_{j,t,d}$ as well as in $q_{j,t,d}$, so both must be accounted for when drawing shares. This part of the estimation is more computationally taxing. Our imposed distributional assumption on ξ carries through to induce a distribution on $s_{j,t,d}$. This is the distribution of shares if we only considered the demand parameters for data augmentation. For notational parsimony, we omit the conditional of α, γ, β, υ, Σ, but note that each function is implicitly conditioned on each. We refer to the share equation as

$$s_{j,t,d} = f(\xi_{j,t,d}),$$

and since f is injective, the density of $s_{j,t,d}$ is given by

$$f_{s_{j,t,d}}(x) = f_{\xi_{j,t,d}}\left(f^{-1}(x)\right) \cdot \left| J_{\xi_{j,t,d} \to s_{j,t,d}} \right|^{-1}.$$

With this notation, $J_{\xi_{j,t,d} \to s_{j,t,d}}$ represents the Jacobian matrix of model shares with represent to ξ,

$$\left[J_{\xi_{j,t,d} \to s_{j,t,d}} \right]_{mn} = \left[\frac{\partial s}{\partial \xi} \right]_{mn} = \begin{cases} \gamma_t s^B_{m,t,d} s^B_{n,t,d} + (1-\gamma_t) s^L_{m,t,d} s^L_{n,t,d} & \text{if } m \neq n \\ -\gamma_t s^B_{m,t,d}(1 - s^B_{m,t,d}) - (1-\gamma_t) s^L_{m,t,d}(1 - s^L_{m,t,d}) & \text{if } m = n, \end{cases}$$

Thus, $\left| J_{\xi_{j,t,d} \to s_{j,t,d}} \right|^{-1}$ denotes the inverse of the determinant of the Jacobian.

Importantly, the distribution assumed on ξ defines a distribution for the shares. Using this, we can determine the likelihood of shares by finding the ξ values these shares imply. The Jacobian term ensures that the shares are a proper probability distribution. For every share draw, we must invert the share system to get ξ. Given this inversion, we compute the likelihood by determining the distribution of ξ conditional on Σ and υ.

Since υ and ξ are assumed to be jointly normal, knowing υ provides information about the magnitude of the demand shock. This joint normality does not factor into the Jacobian of the shares distribution, since neither $s_{j,t,d}$ nor $\xi_{j,t,d}$ appear in the first-stage regression and the first-stage is a linear system. However, we must use the correct conditional distribution for ξ. Conditioning on both η and Σ is enough to pin down the the correlation structure between ξ and υ, and to "observe" υ as well. Drawing on the structure of the bivariate normal distribution, we have

$$\xi|\upsilon \sim \mathcal{N}\left(\frac{\rho\upsilon}{\kappa}, \tau - \frac{\rho^2}{\kappa}\right),$$

where

$$\begin{pmatrix} \xi \\ \upsilon \end{pmatrix} \sim \mathcal{N}(0, \Sigma), \quad \Sigma = \begin{pmatrix} \tau & \rho \\ \rho & \kappa \end{pmatrix}.$$

One interpretation of this treatment of simultaneity is that price gives information about the draw of ξ and so the conditional distribution of ξ is higher or lower depending on the unobservable that influences price. With our modeling assumptions, the conditional distribution of shares of a particular market (t,d) is then given by

$$\prod_{j=1}^{J(t,d)} \left[\phi\left(\frac{f^{-1}(s_{j,t,d}) - \frac{\rho\upsilon}{\kappa}}{\sqrt{\tau - \frac{\rho^2}{\kappa}}} \right) \right] \cdot \left| J_{\xi_{t,d} \to s_{t,d}} \right|^{-1},$$

where $\phi(\cdot)$ is the standard normal density function.

The posterior likelihood is constructed by taking the product of the each market's inversion multiplied by the likelihood contribution of each good sold in the market, following the capacity constraint logic above. Since this distribution is not conjugate-prior, we sample from the posterior using a Metropolis-Hastings step.[6]

3.2.3 Updating price coefficients, α_L, α_B

We use a two-type random coefficient model following Berry, Carnall, and Spiller (1996), which prohibits us from using the technique in Jiang, Manchanda, and Rossi (2009) to sample the distribution of price coefficients. Here, we extend their framework to accommodate discrete unobserved heterogeneity in the random coefficients model.

For a fixed set of product shares, randomness only enters the model through the unobservable

[6]Candidates are accepted and rejected according to a standard Metropolis-Hastings algorithm. We sample candidates using a draw from the Dirichlet distribution centered around the previous shares, scaling every element by a constant search parameter that can be tuned for efficiency. We use the Dirichlet distribution since it samples on a simplex, which is always imposed on shares. We restrict results to be within the interval $(0,1)$ with a threshold of 10^{-5} for numerical efficiency.

ξ. For these constant shares, any choice of (α,β) implies a residual ξ. Therefore, we can use the assumed distribution on ξ to compute the likelihood of a particular (α,β).

The likelihood of $\alpha = (\alpha_B, \alpha_L)$ can be constructed using the same logic as in determining the prior for the share draws. Instead of sampling different shares, α is sampled. Different price sensitivities change the residual ξ, for which we can invert the demand system and evaluate the likelihood. Conditional on η and Σ, we can compute the distribution of ξ and determine the likelihood of a particular draw of α. The likelihood is given by

$$\prod_{(t,d)} \prod_{j=1}^{J(t,d)} \left[\phi\left(\frac{f^{-1}(s_{j,t,d}) - \frac{\rho \upsilon}{\kappa}}{\sqrt{\tau - \frac{\rho^2}{\kappa}}} \right) \right] \cdot \left| J_{\xi_{j,t,d} \to s_{j,t,d}} \right|^{-1},$$

where $\phi(\cdot)$ is again the standard normal density function.

Due to the lack of availability of conjugate priors, we can use any prior distribution on the price coefficients. We impose a multivariate normal prior on α,

$$\alpha \sim \mathcal{N}(\alpha_0, \Sigma_\alpha).$$

which implies the posterior distribution is given by

$$\prod_{(t,d)} \prod_{j=1}^{J(t,d)} \left[\phi\left(\frac{f^{-1}(s_{j,t,d}) - \frac{\rho \upsilon}{\kappa}}{\sqrt{\tau - \frac{\rho^2}{\kappa}}} \right) \right] \cdot \left| J_{\xi_{j,t,d} \to s_{j,t,d}} \right|^{-1} \cdot \phi\left(\Sigma_\alpha^{-\frac{1}{2}} (\alpha - \alpha_0) \right). \tag{3}$$

Sampling from this distribution is done with a Metropolis-Hastings step as well.[7]

3.2.4 Updating probabilities on consumer types, γ

The final demand parameter required in the model is γ_t, which are the probabilities on consumer types. The likelihood for these parameters follow the same function as for α, β.

[7] We select candidates from a normal distribution centered around the previous accepted value. The variance of the sampling distribution is then tweaked for optimal mixing.

The likelihood of gamma given the shares drawn is equal to

$$\prod_{(t,d)} \prod_{j=1}^{J(t,d)} \left[\phi \left(\frac{f^{-1}(s_{j,t,d}) - \frac{\rho \upsilon}{\kappa}}{\sqrt{\tau - \frac{\rho^2}{\kappa}}} \right) \right] \cdot \left| J_{\xi_{j,t,d} \to s_{j,t,d}} \right|^{-1}. \tag{4}$$

Unlike α, where we use a normal prior, we impose a Beta prior on γ,

$$\gamma_t \sim \text{Beta}\Big(\alpha_\gamma, \beta_\gamma\Big).$$

This ensures that all posterior draws are contained within the interval $(0,1)$, but still allows for informative priors. We sample particular values of gamma using a Metropolis-Hastings.[8]

3.2.5 Updating remaining preferences, β

The draw for product preferences (β) is substantially simpler than the update for α because these preferences are assumed to be constant across consumer types. If these preferences were also type specific, a similar method to the draw of α would be required. Our distributional assumptions that allow us to use a Bayesian regression technique to draw β. This method is efficient, scales to high dimensions easily, and does not require estimation tuning.

We define $\Xi_{j,t,d} = X_{j,t,d}\beta + \xi_{j,t,d}$ where $\xi_{j,t,d}$ is taken as the random effects draw from the previous step. As before, we still need to deal with the correlation structure between ξ and price. Note that

$$\xi | \upsilon \sim \mathcal{N}\left(\frac{\rho \upsilon}{\kappa}, \tau - \frac{\rho^2}{\kappa} \right)$$

and after subtracting the mean and dividing by the standard deviation, we obtain a standard normal distribution. This is given by

$$\frac{\xi - \frac{\rho}{\kappa}\upsilon}{\sqrt{\tau - \frac{\rho^2}{\kappa}}} \sim \mathcal{N}(0,1).$$

[8]The candidate is drawn using a normal distribution centered around the previous accepted value using the above likelihood.

The linear system can then be written as

$$\Xi_{j,t,d} - \frac{\rho}{\kappa}\upsilon_{j,t,d} = X_{j,t,d}\beta + U_{j,t,d},$$

where $U_{j,t,d} \sim_{iid} \mathcal{N}\left(0, \sqrt{\tau - \frac{\rho^2}{\kappa}}\right)$. This equation is now a standard linear regression. By imposing a normal prior on β, we can use the conjugacy of a standard Bayesian linear regression model to obtain the posterior distribution of β. We assign the prior distribution to be

$$\beta \sim \mathcal{N}(\mu_\beta, V_\beta),$$

which implies the posterior distribution for β is given by

$$\beta \sim \mathcal{N}\left[\left(X'X + V_\beta{}^{-1}\right)^{-1}(X'X\left((X'X)^{-1}X'\left(\Xi - \frac{\rho}{\kappa}\upsilon\right)\right) + V_\beta{}^{-1}\mu_\beta), (X'X + V_\beta{}^{-1})^{-1}\left(\tau - \frac{\rho^2}{\kappa}\right)\right]. \quad (5)$$

With this notation, X is the stacked matrix of all $X_{j,t,d}$, and Ξ is a vector.

3.2.6 Updating the Price Equation: η

Updating η is essential for maintaining the correlation structure between the unobserved components of demand (ξ) and price. Conditional on shares, α, β, γ, and therefore ξ are constant. We can treat this as "observing" ξ. This gives information about the distribution of υ based on the bi-variate normal structure.

Recall that the pricing equation and the conditional distribution of υ given ξ are

$$p_{j,t,d} = Z_{j,t,d}\eta + \upsilon_{j,t,d},$$
$$\upsilon|\xi \sim \mathcal{N}\left(\frac{\rho}{\tau}\xi, \kappa - \frac{\rho^2}{\tau}\right).$$

Subtracting the mean and dividing by the standard deviation produces a standard normal distribution,

$$\frac{\upsilon - \frac{\rho}{\tau}\xi}{\sqrt{\kappa - \frac{\rho^2}{\tau}}} \sim \mathcal{N}(0,1).$$

The first-stage pricing regression can then be written as

$$p_{j,t,d} - \frac{\rho}{\tau}\xi_{j,t,d} = Z_{j,t,d}\eta + U_{j,t,d},$$

where $U_{j,t,d} \sim iid \ \mathcal{N}\left(0, \sqrt{\kappa - \frac{\rho^2}{\tau}}\right)$. Drawing η also uses standard Bayesian methodology. We impose a normal prior on η, specifically

$$\eta \sim \mathcal{N}(\mu_\eta, V_\eta).$$

Thus, the posterior distribution of η is equal to

$$\eta \sim \mathcal{N}\left(\left(Z'Z + V_\eta{}^{-1}\right)^{-1}(Z'Z\left((Z'Z)^{-1}Z'\left(P - \frac{\rho}{\tau}\xi\right)\right) + V_\eta{}^{-1}\mu_\eta), (Z'Z + V_\eta{}^{-1})^{-1}\left(\kappa - \frac{\rho^2}{\tau}\right)\right). \quad (6)$$

Where Z, P are matrices containing all the distinct elements of $Z_{j,t,d}$ and $p_{j,t,d}$.

3.2.7 Updating the Correlation Structure: Σ

The final step of the estimation process is to sample the parameters of the joint-normal distribution of ξ and η. We impose an Inverse-Wishart prior for a conjugate-prior distribution. Since we have conditioned on all other parameters, we "observe" both ξ and η by inverting the share equation and the first-stage price equation, respectively. Given the prior

$$\Sigma \sim IW(\nu, V),$$

the posterior distribution is then given by

$$\Sigma \sim IW(\nu + N, V + S) \quad \text{such that} \quad S = \sum_{(t,d)} \sum_{j=1}^{J(t,d)} \begin{pmatrix} \upsilon \\ \xi \end{pmatrix} \begin{pmatrix} \upsilon \\ \xi \end{pmatrix}'. \quad (7)$$

4 Monte Carlo Study

We demonstrate our empirical approach with a Monte Carlo study. We create a synthetic data set representing 100 different departure dates, with a booking horizon of 120 days. This provides a total of 22,000 observations for the arrivals process. This mimics conducting analysis on just more than an annual quarter of data.

We specify the arrival process to be

$$A_{t,d} \sim \begin{cases} \text{Poisson}(1) & \text{for } t \geq 60 \\ \text{Poisson}(2) & \text{for } 30 \leq t < 60 \\ \text{Poisson}(4) & \text{for } 7 \leq t < 30 \\ \text{Poisson}(10) & \text{for } t < 7. \end{cases}$$

For each departure date, we select the number of potential flight options from a discrete uniform distribution. The minimum number of flight options is one; the maximum number of flights is six. This is meant to create a scenario similar to observed flight frequencies. We allow capacity to be sufficiently large such that the choice set does not change over time—this additional variation would help identify substitution patterns. Next, we assume product characteristics are drawn from a random uniform distribution. We draw the price disturbances $\upsilon_{j,t,d}$ and the unobserved product characteristics $\xi_{j,t,d}$ from the following multivariate normal distribution,

$$\begin{pmatrix} \xi_{j,t,d} \\ \upsilon_{j,t,d} \end{pmatrix} \sim \mathcal{N}(0, \Sigma), \quad \text{where } \Sigma = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 1.0 \end{pmatrix}.$$

Two instruments are drawn from a standard normal distribution, and we define $\eta = [-4, 8]$. Prices are then calculated to be

$$p_{j,t,d} = Z\eta + \nu_{j,t,d}.$$

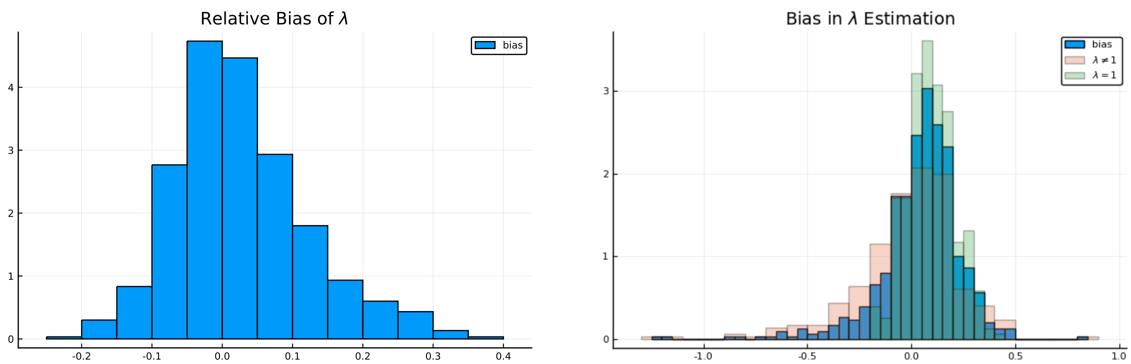
We specify γ_t to vary over three or seven day windows across the booking period. The window size is set to seven for $t > 30$, and three for rest of the booking horizon. For the first time window, we set $\gamma_t = .2$ and the final time window $\gamma_t = .8$. We increase γ_t by a constant factor for each window until the terminal interval. This creates an increasing sequence of γ_t across the booking horizon.

Our choices on the arrival process, structural errors, and γ_t , create an environment in which product-level demand is low, there is a high degree of endogeneity between price and the demand shocks, and the inelastic consumers arrive towards the end of the booking horizon. All these features are present in our data, which is discussed in detail in section 5, and together motivate our proposed estimation method.

With all the key data generating process parameters defined, we first calculate the conditional choice probabilities (defined in Section 2). Next, we multiply these values by the Poisson arrival rates to determine the distribution of demand. We draw sales quantities to complete the construction of the data set. Finally, we apply the demand methodology presented in Section 3. After an initial burn in period of 5,000 interactions, we sample chains for another 5,000 iterations.

Figure 1: Monte Carlo Estimates of the Arrival Process

(a) (b)



Note: Density plot of parameters for 5,000 sampled chains. (a) (b)

Figure 1 presents a summary of the estimates of the arrival process. Because the number of parameters estimated is large, we present summaries over all of the parameters. Generally, we find the arrival rates are precisely estimated, even though the specification contains over 200 parameters. Panel (a) plots the distribution of the relative bias, $(\hat{\lambda}_t - \lambda_t)/\lambda_t$, where $\hat{\lambda}_t$ is the predicted value. We find that 93.4 percent of the parameters fall within the 95th percentile credible sets. Panel (b) plots the distribution of the relative bias for periods in which $\lambda_t = 1$ and $\lambda_t \neq 1$. We see that the distribution of relative bias for the $\lambda_t = 1$ periods are skewed to the right, while for the other periods we have a distribution centered around 0. As the rate of arrivals increase, our estimation accuracy improves.

Mechanically, the procedure works to ensure that searches can rationalize observed sales. That is, the procedure increases any search rate in instances where the sum of inside good sales makes the necessary arrival draw very unlikely. In addition, arrival rates can be adjusted downward depending on restrictions imposed by other parameters of the model. This is rare and we have only observed very small decreases in Monte Carlo studies.

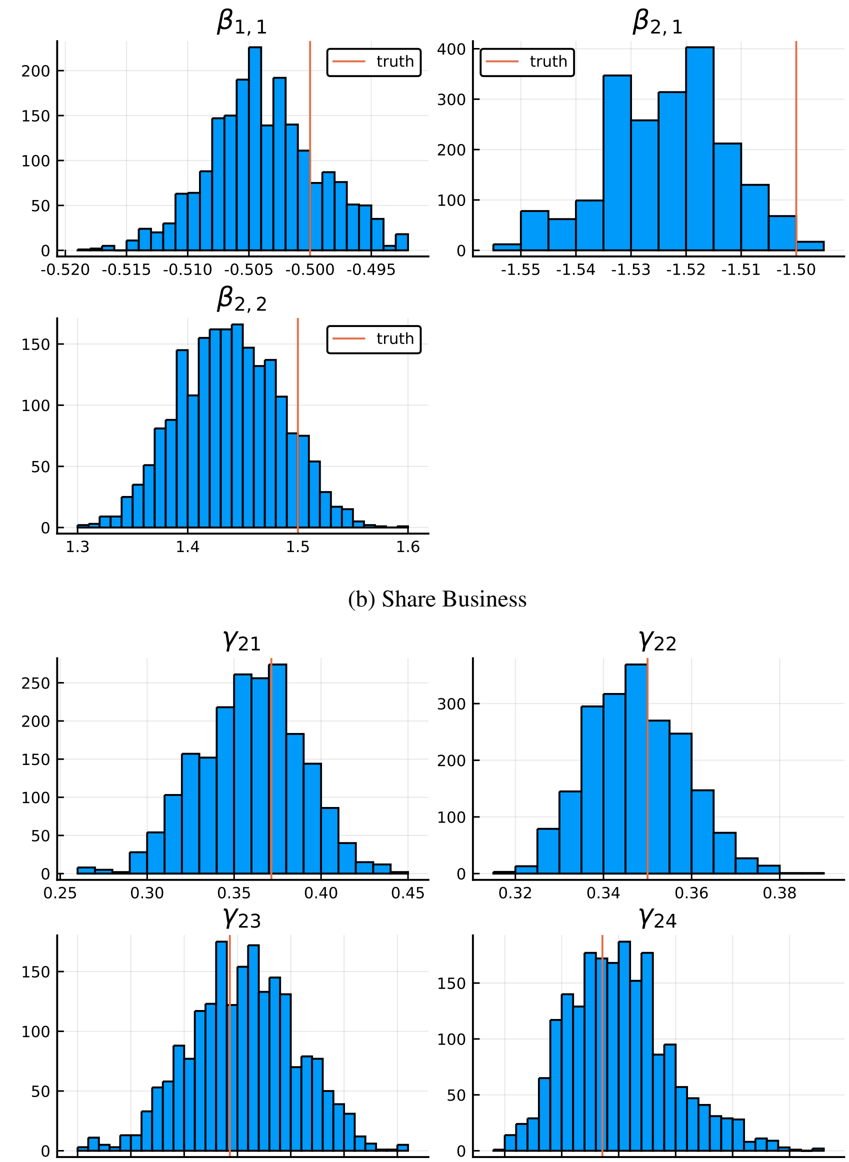
Figure 2 shows the posterior distributions for selected preference parameters. Panel (a) shows the posterior distributions for the constant and price parameters for both consumer types. Despite having high endogeneity, the true value to the pricing parameters appear in our 95th percentile credible region and have relatively low variance. Panel (b) shows the posterior distributions for the first couple of γ_t parameters, where arrivals are sparse. With only 300 or 700 observations corresponding to each window, each of the true values lies within the credible region. Even with a sparse arrival rate, our estimator is able to produce posteriors that contain the true value.

We find that with small volumes of data, instruments, and hundreds of parameters, our estimation methodology is able to capture the true value for many of the parameters in the 95th percentile credible region of the sampled posteriors. Importantly, the method also scales well computationally because estimating λ and β (the parameters with the largest dimension) take relatively little estimation time within the sampler. For example, adjusting the arrival process so that it contains hundreds more parameters does not change total computation time substantially.

Computation burden can increase with larger samples for two reasons. First, if the choice set increases, inverting market shares—which is required in several steps of the Gibbs procedure—can take more time. If shares are small, the inversion can take longer to compute. Second, additional markets (more departure dates or days before departure) necessarily means inverting shares in more markets. However, the additional time requirement can be addressed by conducting the inversion separately for each market and parallelizing the computation. We also recommend a jacobian or accelerated fixed point method when implementing the contraction, as these methods reduce the number of iterative steps it takes to solve for the structural errors.

Figure 2: Monte Carlo Preference Parameters

(a) Price Parameters



Note: Density plot of parameters for 5,000 sampled chains. (a) $\beta_{1,1}$ is the price coefficient for the leisure type, $\beta_{2,1}$ is the price coefficient for the business type, $\beta_{2,2}$ is the intercept. (b) The first four γ parameters which correspond to when arrivals are sparse.

5 Empirical Application

We apply our proposed methodology to granular data provided by a large international air carrier based in the United States. In Section 5.1, we introduce the different data series used. In Section 5.2, we document key data patterns that must be rationalized by the model. In Section 5.3, we detail our empirical selection. In Section 5.4, we discuss model identification. Finally, in Section 5.5, we estimate the model using selected markets.

5.1 Data

We use data provided by the air carrier for domestic flights over a thirteen month period. We use the publicly available DB1B and T100 tables provided by the Bureau of Transportation Statistics to select airline routes for analysis. We select markets in which the air carrier maintains a high percentage of all nonstop traffic. We also select routes that minimize the percentage of connecting flight traffic.

The first data set used contains bookings information. Each entry in the table includes the fare paid, the number of passengers involved, the particular flights included in the itinerary, the booking channel (direct, travel agency, etc.), and the purchase date. These data allow us to form a complete picture of quantities sold for nonstop bookings. The fare data contain the lowest available fare leading up to departure, which we use to construct prices of all products offered by the carrier (we supplement this with filed fare information). We simplify our analysis a number of ways. First, we focus on economy class tickets. Second, we assume consumers purchase the lowest available fare—although customers may purchase a more expensive economy fare class, this is rare. Third, we assume all consumers are offered the same price for a particular flight within a day—in our sample, less than one percent of observed inventory decisions prohibit all observed buyers from purchasing at the same price. Finally, we assume consumers purchase a single ticket only—average bookings are between one and two seats.

The second data set used are inventory data. These data contain inventory allocations for each flight in the sample.[9] In addition, the inventory data contain a daily, flight-level measure of the air carrier's assigned opportunity cost of a seat sold. These values are an output of the air carrier's revenue management software. We will use this data column as an instrument when estimating demand.

The final data set used contains search information. The search data measure the aggregate

[9]The economy availability, combined with the booking data, allow us to capture quantity sold that are not present in our bookings data. Unobserved (to us) sales are from connecting traffic, rewards bookings, or consumers altering existing reservations.

searches for each market, which we define as a origin-destination-departure date-purchase date tuple. We construct these data using consumer clickstream data. Consumers arrive at the air carrier's website or mobile app. Their activity in a browsing session is tracked, even if they are not logged into the site or app. We count the unique number of consumers who execute search requests across all markets. As an example, if a consumer searches for "market A," then "market B," then "market A" again, we count this as a single search for "market A" and a single search for "market B."

Note that we measure the market size for the direct channel, which is one of three booking channels. Consumers can also purchase via an Online Travel Agency (OTA), such as Expedia, or via a travel agent. In Section 5.2, we summarize how bookings vary across channels. We cannot account for searches made from other channels, however, because we observe all bookings, we account for searches made via the unobserved channels through a scaling factor (this is based on the fraction of sales through the direct channel).

We use the sum of unique users for each market to define the search count, however, this does not account for multiple passenger itineraries. Thus, we scale searches by the observed number of passengers per booking.

5.2 Descriptive Evidence

We document several facts to motivate our approach to modeling demand in this context. The timing of the market—both in terms of sales and pricing decisions—highlights the need to leverage search data to disentangle changes in market size from changes in preferences (composition of consumers in the market).

Airline tickets are typically available up to one year in advance of purchase. Figure 3 shows how all markets in our sample evolve toward the departure date. All lines are scaled by their series maximum so that they all lie between zero and one. The graph shows three important findings: (i) searches and sales are low well in advance of the departure date, (ii) searches and bookings increase closer to the departure date, and (iii) fares are U-shaped over the booking horizon but increase greatly close to the departure date. In addition to what is shown in these scaled series, we note that the average bookings over the entire horizon are less than one per day, and the median rate of bookings is zero up until less than ninety days before departure. This pattern is mirrored in consumer searches.

Figure 3: Search, Bookings, and Prices Over Time - Full Booking Horizon

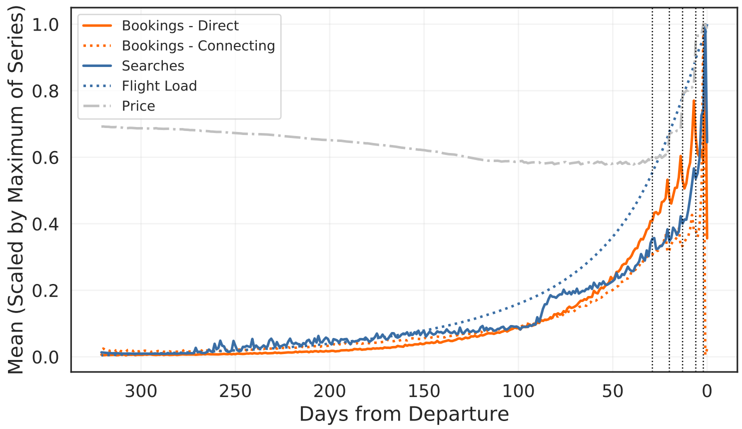
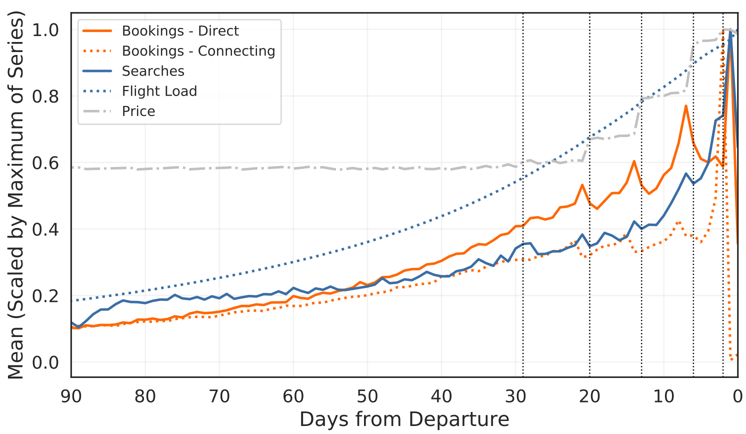


Figure 4: Search, Bookings, and Prices Over Time - Final 90 Days



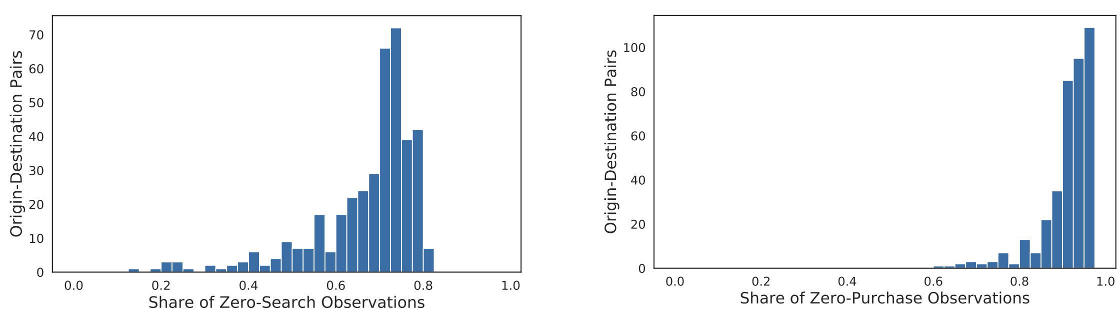
Although we observe the entire booking horizon for all flights, a majority of seats are sold within a few months of the departure date. For this reason, we focus on the final ninety days before departure. Only seventeen percent of bookings occur outside this window. Figure 4 repeats Figure 3 but zooms in on the final ninety days before departure. The total size of the market, as measured by consumer search, is increasing until the day prior to departure. Moreover, the booking rate also increases during this period. The dips in the booking rate coincide with the expiration of fares with advance purchase discount requirements. There is relatively little variation in the average fare outside of these regular price changes.

As mentioned in Section 5.1, we observe bookings from all channels, but we do not observe search from indirect purchase channels. Our estimation approach accounts for these unobserved arrivals by incorporating the inverse of direct channel share into the exogenous scaling factor. The extent to which this adjustment alters the market size considerably varies across routes, but it does not vary over the booking horizon.

For many routes, searches are sparse over the entire booking horizon. We frequently observe either no searches or a single search request in many markets. Figure 5 plots the distribution of route-level search and sales sparsity over the entire booking horizon. Figure 6 plots the distributions of route-level sparsity using only the last ninety days before departure. Both searches and sales become less sparse towards the departure date. However, we measure sales sparsity by aggregating over all flights within the day—the frequency of zero sales at the flight level is much higher. Together, these empirical findings highlight the need for a demand model that (i) accounts for across-route heterogeneity in arrivals and preferences, (ii) captures the possibility of a market size of zero, and (iii) accounts for frequent zero-purchase observations.

Figure 5: Market Sparsity Across Routes

(a) Search Sparsity (b) Sales Sparsity



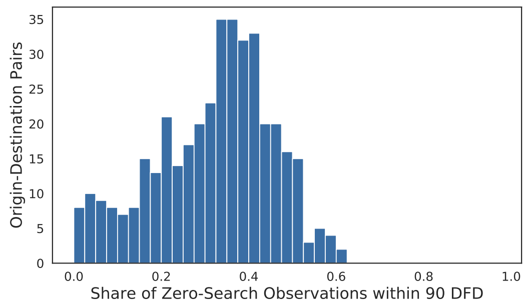
Note: (b) Zero sales are measured here as zero inside share - aggregating across different flight times in the same day.

5.3 Empirical Specification

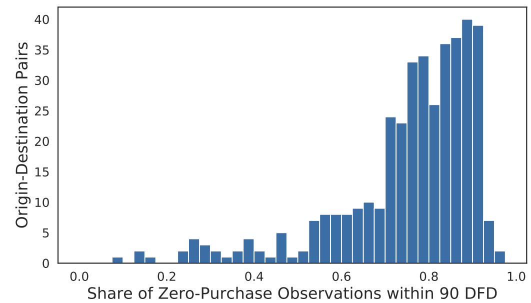
To demonstrate our methodology, we select two origin-destination pairs where the air carrier offers several daily flights. We select these routes for two reasons. First, we can exploit variation in the number of products to estimate time-of-day departure parameters. This allows us to investigate heterogeneity in demand within a departure date. Second, these markets also offer substantial variation in search activity over time. This allows us to demonstrates how the methodology accommodates variation in consumer behavior as the departure date approaches.

Figure 6: Market Sparsity Across Routes Close to Departure

(a) Search Sparsity



(b) Sales Sparsity



Note: (b) Zero sales are measured here as zero inside share - aggregating across different flight times in the same day.

For each market, we use the following utility specification,

$$u_{i,j,t,d} = \beta_0 - \alpha_{\ell(i)} p_{j,t,d} + \text{FE(Time of Day } j) + \text{FE}(d) + \xi_{j,t,d} + \varepsilon_{i,j,t,d}.$$

We allow the share of business travelers to vary over time, but we put some structure onto the problem to improve identification by pooling over days before departure. We employ a blocking function that maps days before departure into different γ parameters. We choose a blocking function that becomes more granular (pools fewer days) as the departure date approaches. The specification is shown graphically in the model fit figures below. Finally, recall that we estimate the arrival rates with additive day before departure and departure date parameters.

5.4 Identification

The difficulty in estimating a model with stochastic demand is handling the problem of separably identifying shocks to arrivals from shocks to preferences. Conditional on observing information about the arrival process, identification of our model follows from standard arguments from the literature on demand in differentiated products using market level data. This is discussed in Berry and Haile (2014, 2016). The time of day taste parameters are identified from the variation between the number of flights and times that they are offered across departure dates. The parameters that determine the mix of business and leisure consumers are identified from a unique feature of the airline data—prices increase greatly but the booking rate does not fall substantially under higher prices. This observation suggests that preferences of arriving consumers evolve over the booking horizon. Our specification picks this up through adjustments to γ_t.

Because we allow the unobserved product level demand shocks to be correlated with price, we require instruments to provide sufficient exogenous variation to identify the parameters that measure price sensitivity between the two consumer types. We use jet fuel prices, the carrier's shadow price of capacity (an output of the revenue management software), plane capacity, and total number of outbound bookings from a route's destination airport as our instruments.[10]

Fuel prices act as a cost shifter. The shadow price informs the opportunity cost of capacity, but might be correlated with the contemporaneous unobserved demand shock. We have assumed the unobserved demand shocks are not correlated over time, so we use a one period lag on the shadow value to ensure exclusion. The plane capacity captures the fact that the opportunity cost of selling one seat is lower for a larger plane serving the same route. The total number of outbound bookings at the destination airport captures the change in opportunity cost for flights that are driven by demand shocks in other markets. For example, for a flight from A to B, where B potentially provides service elsewhere, we use all traffic from B onward to other destinations C or D. We assume demand shocks are independent across markets, so shocks to $B \to C$ and $B \to D$ are unrelated to demand for $A \to B$. Thus, a positive shock to onward traffic, out of hub B, will raise the opportunity cost of serving $A \to B \to C$ or $A \to B \to D$. This propagates to price set on the $A \to B$ leg.

5.5 Estimation Results

Figure 7 and Figure 8 contain summary plots of the estimated demand systems. Panel (a) compares the time series of the estimated search process by day before departure with its empirical counterpart,

$$\frac{1}{|D|}\sum_{d\in D}\widehat{A_{t,d}}(\widehat{\theta}_t,\widehat{\theta}_d), \quad \text{and} \quad \sum_{d\in D}\frac{1}{|D|}\sum_{d\in D}A^{\text{data}}_{t,d}.$$

The search process is computed at the posterior mean. Panel (b) compares the sum of inside model shares (probability of purchasing a ticket on any flight) model to empirical shares in a similar man-

[10]For a route with origin O and destination D, the total number of outbound bookings from the route's destination airport is can be defined as the following; $\sum_{i=1}^{K} Q_{D,O'}$. Where $Q_{D,O'}$ is the sum of the total number of bookings in period t, across all flights, for markets where the origin is the original route's destination. In this definition, we assume that there are K many routes that have original route's destination as an origin.

ner.

$$\frac{1}{|D|}\sum_{d\in D}\sum_{j\in J(d,t)}\widehat{s_{j,t,d}}(\widehat{\theta}_t,\widehat{\theta}_d),\quad\text{and}\quad\frac{1}{|D|}\sum_{d\in D}\sum_{j\in J(d,t)}\frac{q_{j,t,d}^{\text{data}}}{A_{t,d}^{\text{data}}}.$$

Panel (c) shows how the probability on consumer types changes over time in solid (orange). In addition to plotting $\widehat{\gamma}_t$, we plot the probability of purchase conditional on an arriving consumer being a business customer. These time series are also calculated at the posterior means. Finally, panel (d) depicts demand elasticities. The solid (orange) line shows the average product elasticity (over flights and departure dates), for each day before departure. The two dashed lines show the least and most elastic flights. These lines demonstrate the heterogeneity in estimates across flights.

In both markets, our specification on arrivals is very flexible and the estimated arrival processes closely match the data. The vertical axes on the plots are normalized, but they show arrivals increase over time. There is a substantial increase in searches close to the departure date. Fitting the market size closely is only possible through our flexible fixed effects specification, particularly over days from departure.

The model fits shares reasonably well. The sparsity in sales well in the advance of the departure date is the most difficult feature of the data to match. This difficulty is shown in panel (b) for Route 2, where model shares are higher and less dispersed than data shares. Because searches and bookings increase closer to the departure date, the model closely matches the data here. There is a noticeable drop in shares in the last week before departure. For Route 1, the increase in searches, the lack of an increase in quantity sold, and non-rising prices close to the departure date, cause γ_t to drop immediately before departure. That is, although we identify most arrivals as business travelers very close to the departure date, the fact that shares decrease with prices mostly constant is rationalized by additional leisure travelers entering the market.

Our estimates suggest that a systematic change in the types of consumers interested in traveling occurs over time. For Route 1, we estimate that γ_t ranges from close to zero all the way to one. For Route 2, we estimate γ_t ranges from 0.4 to one. Close to departure, the customers arriving are nearly all business-type. We do not interpret these segments literally; the business segment is less price sensitive and arrives later, on average. These characteristics need not only apply to customers traveling for business—business is the label we give to the discrete, unobserved types corresponding to these demand characteristics.

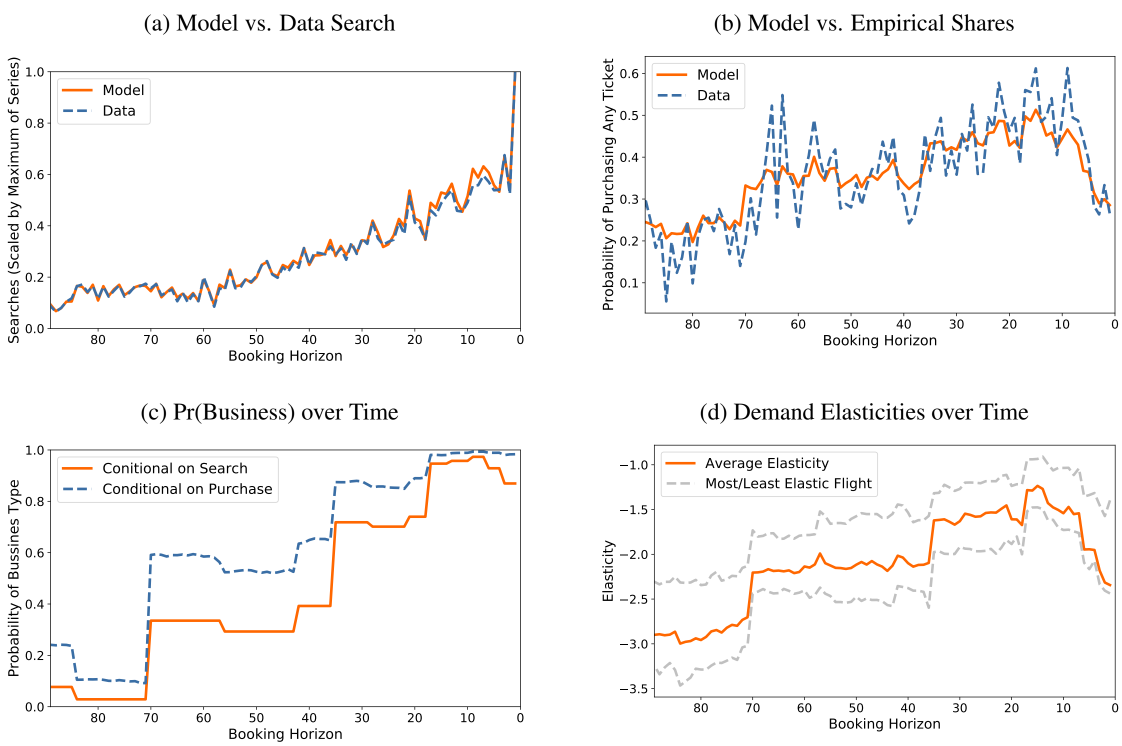
We estimate flight-level demand elasticities to range from -3.0 to -1.3 for Route 1 and -2.1 to -1.1 for Route 2. Although we estimate that the demand elasticity falls close to the departure date, willingness to pay is high—both markets produce γ_ts close to one within the last week of departure. We also find significant variation in demand across flights and departure dates. The difference in demand elasticities within any given t typically varies by close to one, except for Route 2, close to the departure date.

In addition to elasticity variation, the market size variation driven by θ_d—the departure date fixed effect—is substantial. For example, the highest-arrival departure date has around three times as many arrivals as the lowest-arrival date for Route 2. For Route 1, this is less stark, with the highest arrival departure date about 75 percent higher than the lowest. Demand also varies based on departure date, conditional on arrivals and customer type composition.

The variation in arrivals and demand across departure dates highlights the importance of pricing that is responsive to these differences. Some of the departure date variation is systematic. All else equal, arrivals are higher for weekend travel than weekday. The same is true for demand, conditional on arrivals. Much of the variation in the departure date fixed effects in the utility specification move systematically with the day of week.

Our estimates also suggest that, within a departure date, flight departure time alters willingness to pay. For Route 1, which is a westbound long-haul flight, consumers are willing to pay \$150 more for a late morning or early evening one-way flight than the earliest departure. For Route 2, which is a shorter flight, consumers are willing to pay \$125 more for a late morning or afternoon one-way flight than the earliest departure times. These time of day preferences are likely highly variable across routes; for example, the eastbound flights on Route 1 may have additional willingness to pay for early morning departures.

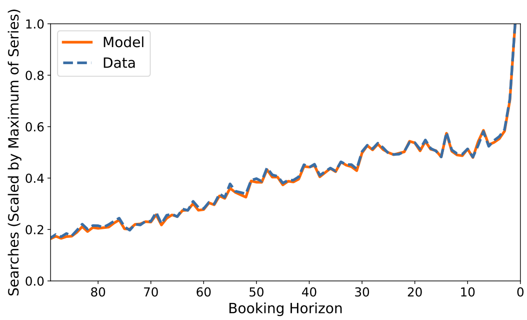
Figure 7: Model Estimates for Route 1



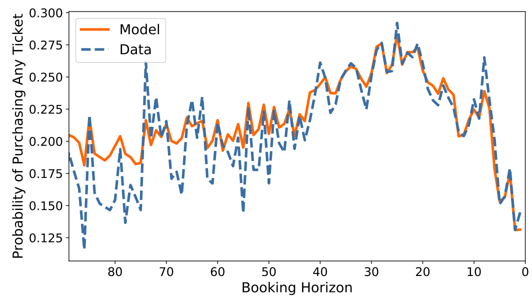
Note: The horizontal axis of all plots denotes the negative time index, e.g. zero corresponds to the last day before departure. (a) Normalized model fit of searches with data searches. (b) Model fit of product shares with empirical shares. (c) Fitted values of γ_t over time, along with the probability a consumer is business conditional on purchase. (d) Mean product elasticities over time, along with the least and most elastic flights.

Figure 8: Model Estimates for Route 2

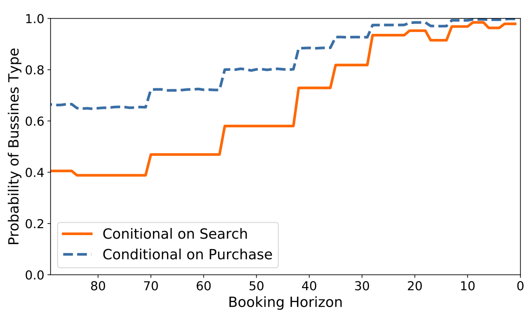
(a) Model vs. Data Search



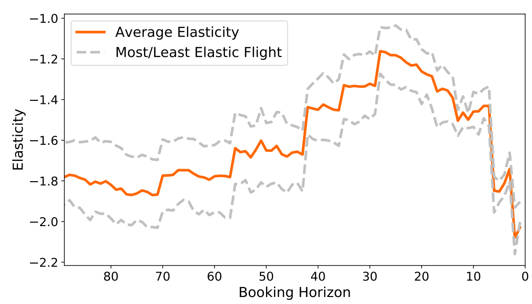
(b) Model vs. Empirical Shares



(c) Pr(Business) over Time



(d) Demand Elasticities over Time



Note: The horizontal axis of all plots denotes the negative time index, e.g. zero corresponds to the last day before departure. (a) Normalized model fit of searches with data searches. (b) Model fit of product shares with empirical shares. (c) Fitted values of γ_t over time, along with the probability a consumer is business conditional on purchase. (d) Mean product elasticities over time, along with the least and most elastic flights.

6 Conclusion

In this paper, we introduce a tractable demand methodology that can be used to simultaneously estimate the arrival process of customers and their associated preferences. The demand model features stochastic consumer arrivals with time-varying parameters and a mass-point random coefficients utility specification where the mixture of consumer types is allowed to change over time. We propose using an aggregate measure of search activity—the total number of consumers who execute search requests—to inform the size of the market and to estimate the magnitude of stochastic demand. We develop and implement a Bayesian estimation procedure that accommodates zero sale observations and prices to be correlated with a product-level unobservable demand shock. We apply our method to granular data provided by a large international air carrier based in the United States. We find that demand becomes more inelastic and the average number of interested travelers increases toward the departure date.

Given the degree of departure date heterogeneity in arrivals, it may be difficult to anticipate the potential market size for each flight. Output from our demand procedure can be used to investigate learning about the market's size in real time (Lin, 2006). In addition, if demand is correlated over time, past arrivals may be useful to inform current and future pricing decisions. We leave exploring the supply-side implications of empirical procedure for future work.

References

Berry, S., and P. Haile (2014): "Identification in Differentiated Products Markets Using Market Level Data," *Econometrica*, 82(5), 1749–1798.

——— (2016): "Identification in Differentiated Products Markets," *Annual Review of Economics*, 8, 27–52.

Berry, S. T. (1994): "Estimating Discrete-Choice Models of Product Differentiation," *The RAND Journal of Economics*, pp. 242–262.

Berry, S. T., M. Carnall, and P. T. Spiller (1996): "Airline Hubs: Costs, Markups and the Implications of Customer Heterogeneity," Working Paper.

Berry, S. T., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, pp. 841–890.

Berry, S. T., O. B. Linton, and A. Pakes (2004): "Limit Theorems for Estimating the Parameters of Differentiated Product Demand Systems," *The Review of Economic Studies*, 71(3), 613–654.

Bitran, G., R. Caldentey, and S. Mondschein (1998): "Coordinating clearance markdown sales of seasonal products in retail chains," *Operations research*, 46(5), 609–624.

Bitran, G. R., and S. V. Mondschein (1997): "Periodic pricing of seasonal products in retailing," *Management science*, 43(1), 64–79.

Compiani, G. (2019): "Market Counterfactuals and the Specification of Multi-Product Demand: A Nonparametric Approach," Working Paper.

Dubé, J.-P. H., A. Hortaçsu, and J. Joo (2020): "Random-Coefficients Logit Demand Estimation with Zero-Valued Market Shares," Working paper, National Bureau of Economic Research.

Gallego, G., and G. Van Ryzin (1994): "Optimal Dynamic Pricing of Inventories with Stochastic Demand over Finite Horizons," *Management Science*, 40(8), 999–1020.

Gandhi, A., Z. Lu, and X. Shi (2017): "Estimating Demand for Differentiated Products with Zeroes in Market Share Data," Working Paper.

Jiang, R., P. Manchanda, and P. E. Rossi (2009): "Bayesian Analysis of Random Coefficient Logit Models using Aggregate Data," *Journal of Econometrics*, 149(2), 136–148.

Lin, K. Y. (2006): "Dynamic pricing with real-time demand learning," *European Journal of Operational Research*, 174(1), 522–538.

Quan, T. W., and K. R. Williams (2018): "Product Variety, Across-Market Demand Heterogeneity, and the Value of Online Retail," *The RAND Journal of Economics*, 49(4), 877–913.

Rossi, P. E., and G. M. Allenby (1993): "A Bayesian Approach to Estimating Household Parameters," *Journal of Marketing Research*, 30(2), 171–182.

Rossi, P. E., G. M. Allenby, and R. McCulloch (2012): *Bayesian Statistics and Marketing*. John Wiley & Sons.

Talluri, K., and G. Van Ryzin (2004): "Revenue Management Under a General Discrete Choice Model of Consumer Behavior," *Management Science*, 50(1), 15–33.

Tebaldi, P., A. Torgovitsky, and H. Yang (2019): "Nonparametric Estimates of Demand in the California Health Insurance Exchange," Working Paper.

Vulcano, G., G. van Ryzin, and W. Chaar (2010): "Choice-based revenue management: An empirical study of estimation and optimization," *Manufacturing & Service Operations Management*, 12(3), 371–392.

Vulcano, G., G. Van Ryzin, and R. Ratliff (2012): "Estimating primary demand for substitutable products from sales transaction data," *Operations Research*, 60(2), 313–334.

Williams, K. R. (2020): "Dynamic Airline Pricing and Seat Availability," Working Paper.

Zhao, W., and Y.-S. Zheng (2000): "Optimal Dynamic Pricing for Perishable Assets with Nonhomogeneous Demand," *Management Science*, 46(3), 375–388.