Submitted to *Operations Research*
manuscript

Network Revenue Management under a Spiked Multinomial Logit Choice Model

Yufeng Cao
Antai College of Economics and Management, Shanghai Jiao Tong University,
yufeng.cao@sjtu.edu.cn

Anton J. Kleywegt, He Wang
H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology,
anton@isye.gatech.edu, he.wang@isye.gatech.edu

Airline booking data have shown that the fraction of customers who choose the cheapest available fare class often is much greater than that predicted by the multinomial logit choice model calibrated with the data. For example, the fraction of customers who choose the cheapest available fare class is much greater than the fraction of customers who choose the next cheapest available one, even if the price difference is small. To model this spike in demand for the cheapest available fare class, a choice model called the *spiked multinomial logit* (spiked-MNL) model was proposed. We study a network revenue management problem under the spiked-MNL choice model. We show that efficient sets, i.e., assortments that offer a Pareto-optimal trade-off between revenue and resource usage, are *nested-by-revenue* when the spike effect is nonnegative. We use this result to show how a deterministic approximation of the stochastic dynamic program can be solved efficiently by solving a small linear program. The solution of the small linear program is used to construct a booking limit policy, and we prove that the policy is *asymptotically optimal*. This is the first such result for a booking limit policy under a choice model, and our proof uses an approach that is different from those used for previous asymptotic optimality results. Finally, we evaluate different revenue management policies in numerical experiments using both synthetic and airline data.

Key words: network revenue management; assortment optimization; discrete choice model; context-dependent utility

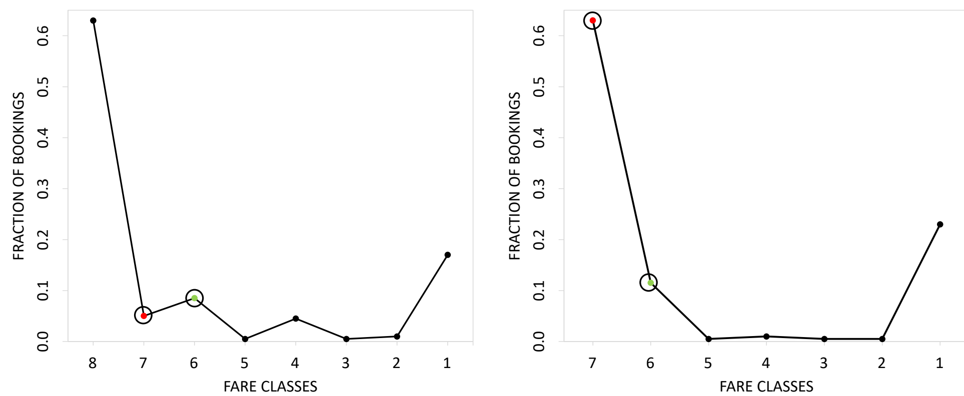
1. Introduction

Revenue management (RM) is widely used by airlines to maximize revenues through inventory control and pricing. Most airlines have a number of fare classes for each itinerary, where an itinerary refers to a timed sequence of flights. Each fare class has certain booking rules (e.g., refundable/nonrefundable, change fees and frequent flyer credits) and a price that for RM purposes is regarded as predetermined. Airlines then control prices by opening or closing fare classes. We refer to a combination of an itinerary and a fare class as a *product*. Thus, a basic decision in airline RM is to select which subset of products to offer to customers at each point in time. The subset of products made available to customers is called an assortment. Two RM problems studied in the literature are the static *assortment optimization* problem, which selects a fixed assortment to maximize expected revenue, and the *network revenue management* problem, which dynamically adjusts the assortment to maximize expected revenue over a horizon subject to resource constraints. Airlines dynamically adjust assortments because the remaining seats on flights and the time until departure change over time. The assortment decisions should be considered jointly for the flights in an airline network, because some itineraries use capacity on multiple flights and because customers may substitute among different itineraries. The scale of airline networks makes solving assortment optimization problems and revenue management problems challenging.

Unlike the independent demand model, which assume that each customer requests one specific product, choice-based demand models assume that customers have heterogeneous preferences over products and that each customer chooses the product that she prefers most from the set of available products (including the no-purchase alternative). One of the most popular choice models is the multinomial logit (MNL) model. The MNL model has several desirable properties: (1) It has a parsimonious and easily interpretable structure. (2) Many problems under the MNL model are tractable, including parameter estimation, assortment optimization, and pricing (Luce 1959, McFadden 2001, Train 2009, Talluri and Van Ryzin 2004a, Keller et al. 2014). However, the MNL model also has shortcomings. For example, the MNL model has the independence from irrelevant alternatives (IIA) property, that is, the relative choice probabilities of two alternatives in the choice set do not depend on the presence of other alternatives in the choice set. The IIA property often is too restrictive for modeling and explaining observed choice behavior. In this paper we address the following choice behavior observed in airline data. The fraction of customers who choose the cheapest fare class (among a considered set of available products) often significantly exceeds the fraction predicted by the MNL model (Boyd and Kallesen 2004, Dai et al. 2014).

As an example, consider Figure 1, which shows an airline's booking data for a specific flight. The fare classes are ordered such that Class 1 has the highest price and Class 8 has the lowest price. In the left figure, we show the fraction of bookings in each fare class for the flight when Classes 1–8 are open. In the right figure, we show the fraction of bookings in each fare class for the flight when only Classes 1–7 are open. Note that in both cases, the cheapest available fare class (Class 8 on the left and Class 7 on the right) receives more than 60% of bookings. Moreover, the fraction of bookings in Class 7 is much more than that in Class 6 when Class 7 is the cheapest available fare class (on the right), but the fraction of bookings in Class 7 is less than that in Class 6 when Class 7 is not the cheapest available fare class (on the left). Thus, the IIA property is violated, because the ratio between the fractions of bookings in Class 7 and Class 6 is affected by the inclusion of another alternative (i.e., Class 8). (We provide a more thorough analysis of this airline dataset in Section 8.)

Figure 1 Historical booking data for a flight when Classes 1 to 8 are open (left) and when only Classes 1 to 7 are spectively.



The phenomenon that the alternative in an assortment with the most attractive value of a specific attribute gets a large fraction of bookings is called the *spike effect*. More specifically, the phenomenon that the *cheapest* available alternative in a consideration set gets a large fraction of the bookings for that consideration set is called the *cheapest fare spike effect*. In general, a consideration set can be either the entire assortment or a subset, and its definition depends on the application and the modeler. In this paper, the set of available products for each itinerary forms a consideration set, so the demand for the cheapest available alternative for each itinerary is spiked. The *spiked multinomial logit (spiked-MNL)* choice model was first introduced by Dai et al. (2014) to capture the spike effect for both the cheapest available fare class as well as the cheapest fully refundable fare class for each flight. This model has the advantages that it captures spike effects while retaining the desirable properties of the MNL model such as tractable estimation and tractable assortment optimization. While Dai et al. (2014) focuses on general insights that were acquired during a revenue management project with an airline, the current paper focuses on the theoretical properties of the spiked-MNL model. The main contributions in the paper are the following:

- Motivated by airline data, we consider a revenue management problem under the spiked-MNL model, and we show that when the spike effect is nonnegative, the efficient sets (i.e., assortments that offer a Pareto-optimal trade-off between revenue and resource usage) are nested-by-revenue.
- We consider a deterministic (fluid) approximation of the network revenue management problem under the spiked-MNL model, known as the choice-based deterministic linear program (CDLP). Even though the number of decision variables of this linear program is exponential in the number

of products, we show that it can be solved in time that is polynomial in the number of products by solving a much smaller sales-based linear program (SBLP) and by exploiting the nested-by-revenue structure of efficient sets.

- We show how the CDLP solution can be used to construct a nested booking limit policy of the form widely used in practice, and we prove the asymptotic optimality of such a booking limit policy. To the best of our knowledge, this is the first such result for a booking limit policy. (Previous literature has established asymptotic optimality results for time limit policies, that is, policies that offer a deterministic sequence of assortments, with each assortment offered for a deterministic fraction of time. Airlines prefer to use booking limit policies, that is, policies that offer a random sequence of assortments, with each assortment offered until the first time that the sales of a product in the assortment reaches its booking limit.) To deal with the random sequence of assortments resulting from applying the booking limit policy, our proof uses an approach that is very different from those used for previous asymptotic optimality results.
- We perform extensive numerical experiments using both synthetic data and real airline data to compare the spiked-MNL model with other discrete choice models. We also compare the performance of different RM policies under the spiked-MNL model.

Notation

Let $\mathbb{R}$ and $\mathbb{R}_+$ denote the set of real numbers and the set of nonnegative real numbers, respectively. Let $\mathbb{Z}$ and $\mathbb{Z}_+$ denote the set of integers and the set of nonnegative integers. We use boldface lower-case and upper-case letters to represent vectors and matrices, respectively. For a vector $\mathbf{x}$, let x_j denote its j-th component. The abbreviation i.i.d. stands for independent and identically distributed.

2. Literature Review

There is an extensive literature on RM and assortment optimization. The surveys by Hübner and Kuhn (2012) and Kök et al. (2015) contain comprehensive overviews of this literature. The idea of airline RM can be traced back to Littlewood's proposal for controlling the availability of two fare classes (Littlewood 1972). Traditional RM demand models assume that each customer requests a specific product (i.e., itinerary and fare class combination). The seller then decides whether to accept or reject the customer's request. This modelling assumption is known as the *independent demand model*. McGill and Van Ryzin (1999) surveys RM literature under this demand model.

The independent demand model does not account for customer choice behavior and may lead to cascading deterioration of revenue performance (Cooper et al. 2006). Some partial modeling remedies such as buy-downs and buy-ups, or spill-and-recapture, have been proposed to incorporate demand substitution (see, e.g., Gallego et al. 2009, Walczak et al. 2010, Cooper and Li 2012). The use of choice-based demand models in RM has been studied more recently (Strauss et al. 2018). Talluri and Van Ryzin (2004a) considered the problem of RM under a general discrete choice model for a single flight. They formulated the problem as a dynamic program (DP), introduced the concept of efficient sets, and showed that optimal assortments are efficient sets. They also showed for single flight RM under the MNL model that optimal assortments are nested by fare order. Due to the curse of dimensionality, the computational burden of solving the DP increases exponentially for airline networks. Therefore, Gallego et al. (2004) proposed a choice-based deterministic linear program (CDLP) as an approximation of the DP. Zhang and Cooper (2005) considered RM for parallel flights and developed a simulation-based heuristic. An important assumption in their paper was that customers would only switch between flights, but not between fare classes within a flight. Later, van Ryzin and Vulcano (2008) studied a network RM problem using virtual nesting controls. Liu and Van Ryzin (2008) extended the concept of efficient sets from Talluri and Van Ryzin (2004a) and proved that the solution of the CDLP is asymptotically optimal for the DP. Even though the number of efficient sets usually is much less than the number of subsets, the number of efficient sets could still be exponential in the number of flights, and thus even if one restricted the decision variables of the CDLP to efficient sets, the number of decision variables of the CDLP could still be very large. Liu and Van Ryzin (2008) suggested solving the CDLP using column generation. Zhang and Adelman (2009) approximated the DP value functions with affine functions, and proposed a column generation algorithm to solve the resulting approximate DP problem under the MNL choice model with disjoint consideration sets. Talluri (2014) proposed a new approach called segment-based deterministic concave program, which is a concise relaxation of the CDLP. Recently, Gallego et al. (2015) proposed a sales-based linear program (SBLP) for general attraction demand models, including the MNL model. The SBLP has a polynomial number of variables under the MNL model, and an optimal solution of the SBLP can be converted in polynomial time to an optimal solution of the CDLP.

In addition to RM under general choice models, many researchers have considered RM under specific choice models. Several choice models have been considered in the RM literature, including MNL (Talluri and Van Ryzin 2004a, Liu and Van Ryzin 2008, Gallego et al. 2015), robust MNL (Rusmevichientong et al. 2010, Rusmevichientong and Topaloglu 2012), nested logit (Davis et al. 2014, Feldman and Topaloglu 2015), mixed MNL (Bront et al. 2009, Rusmevichientong et al. 2014), Markov chain choice model (Feldman and Topaloglu 2017), and rank-based choice models (Farias et al. 2013, Bertsimas and Mišic 2015). Among these models, the MNL model is widely used in the literature as a benchmark, because of its desirable properties mentioned above.

Typical airline reservation control systems use either booking limits (e.g., Talluri and Van Ryzin 1998, Bertsimas and De Boer 2005) or bid-prices (e.g., Bertsimas and Popescu 2003) to control the availability of fare classes. Originally, these controls were motivated by the structure of optimal solutions for single flight RM problems under the independent demand model, as well as their simplicity. Although optimal policies cannot be implemented in general with booking limits or bid prices, it nevertheless is of practical importance to find good booking controls that can be implemented with an airline's given reservation control system.

The spike effect has been noticed in the airline industry before. Boyd and Kallesen (2004) considered a mixture of two customer segments: One segment of customers are primarily concerned with price and will always buy the cheapest available product, and the other segment of customers are interested in specific products as in an independent demand model. They discussed how mixing the two customer segments affects revenue management practice and provided simulation-based illustrations. Their two-segment model exhibits the cheapest spike effect but ignores product substitution. Also, the two-segment model does not have the desirable tractability properties of the MNL model. As shown in the example in Figure 1, the spike effect violates the IIA property, and therefore the spike effect cannot be represented by the MNL choice model. Therefore, Dai et al. (2014) proposed the spiked-MNL model to incorporate the spike effect observed in airline data. In their formulation, the spiked-MNL model can have either positive or negative cheapest fare spikes, and they proposed an SBLP formulation for such a model. Ding (2017) provided more detailed results related to some parts of Dai et al. (2014), such as the identifiability of the spiked-MNL model. The spiked-MNL model has similar tractability properties as the MNL model. Specifically, spiked-MNL model parameters are estimated in the same way as MNL model parameters. Also, as shown in this paper, assortment optimization and revenue management problems under the spiked-MNL model are as tractable as under the MNL model. Compared to Dai et al. (2014) and Ding (2017), the current paper focuses on the theoretical properties of the spiked-MNL model. Given the airline context, we constrain the cheapest fare spikes to be nonnegative and we characterize the structure of efficient assortments under this assumption. The structural result provides a more concise SBLP formulation than that of Dai et al. (2014). Further, we propose booking limit policies based on the SBLP solution and we prove asymptotic optimality of the booking limit policy. We are not aware of any research besides Dai et al. (2014) and Ding (2017) that considers the spike effect in choice models.

3. Model Formulation

We consider an RM problem for a network of flights that are marketed by a single airline and that have the same departure date. Let $\mathcal{F}$ denote the set of flights, and let $m := |\mathcal{F}|$ denote the number of flights. For each flight $f \in \mathcal{F}$, let c_f denote the seat capacity of flight f, and let $\mathbf{c} := (c_f, f \in \mathcal{F}) \in \mathbb{Z}_+^m$. An itinerary consists of a subset of flights, such as a single flight or a sequence of connecting flights. Let $\mathcal{G}$ denote the set of itineraries. For each itinerary, the airline offers multiple fare classes. Each fare class has its own price and set of rules (e.g., cancellation fee, eligibility for upgrade, and frequent flyer miles earned). A *product* is an itinerary and fare class combination. Let $\mathcal{J}$ denote the set of products, and let $n := |\mathcal{J}|$ denote the number of products. For each flight $f \in \mathcal{F}$ and product $j \in \mathcal{J}$, let $a_f^j \in \{0,1\}$ denote the number of seats on flight f used by product j, and let $\mathbf{a}^j := (a_f^j, f \in \mathcal{F}) \in \{0,1\}^m$. Let r_j denote the net revenue of product j, and let $\mathbf{r} := (r_j, j \in \mathcal{J}) \in \mathbb{R}^n$. For each itinerary $g \in \mathcal{G}$, let $\mathcal{J}^g \subset \mathcal{J}$ denote the set of products for itinerary g, and let $n(g) := |\mathcal{J}^g|$ denote the number of fare classes for itinerary g.

The selling horizon is partitioned into discrete periods indexed by $t = 0, 1, \ldots, T$. We assume that the time periods are sufficiently short so that there is at most one customer arrival in each period. In each period t, the airline selects an assortment $A(t) \subset \mathcal{J}$ to offer to customers. Each customer considers only a subset of the products in $\mathcal{J}$ for purchase; we call this the customer's *consideration set*. For example, a customer who travels from origin O to destination D will consider only those products with itineraries that start at O and end at D. We assume that the collection of possible consideration sets form a partition of $\mathcal{J}$. Let $\{\mathcal{J}(h) : h \in \mathcal{H}\}$ denote such a partition of $\mathcal{J}$ into consideration sets. Thus $\mathcal{H}$ is the index set of consideration sets. (Each $h \in \mathcal{H}$ will also be called a market.) If a customer arrives in period t, let $C(t) \subset \mathcal{J}$ denote the customer's consideration set. Thus, for each t, $C(t) = \mathcal{J}(h)$ for some $h \in \mathcal{H}$. For period t, let $S(t) := A(t) \cap C(t)$ denote the products that are currently available in the customer's consideration set. Let $j = 0$ denote the no-purchase alternative (or the null alternative), which is always available to each customer. Thus, if a customer arrives in period t, then the customer chooses an alternative in the customer's choice set $S(t) \cup \{0\}$. With a slight abuse of notation, let $g \in \mathcal{J}(h)$ denote an itinerary g for which the itinerary-fare class combinations belong to market h. In summary, the set of products $\mathcal{J}$ is partitioned into a subset of products $\mathcal{J}(h)$ for each market $h \in \mathcal{H}$, and for each market h, $\mathcal{J}(h)$ is partitioned into a subset of products $\mathcal{J}^g$ for each itinerary $g \in \mathcal{J}(h)$.

In each period t, a customer arrives with consideration set $C(t) = \mathcal{J}(h)$ with probability λ_h; with probability $1 - \sum_{h \in \mathcal{H}} \lambda_h$, no customer arrives. If the customer has consideration set $C(t) = \mathcal{J}(h)$, then she chooses product $j \in S(t)$ with probability $P^h_{j:S(t)}$ or chooses the no-purchase alternative with probability $P^h_{0:S(t)} = 1 - \sum_{j \in S(t)} P^h_{j:S(t)}$.

For an assortment $A \subset \mathcal{J}$ and an alternative $j \in A \cup \{0\}$, the probability that a customer who arrives in period t chooses j if A is offered in period t is given by

$$P_{j:A} \quad := \quad \frac{\lambda_{h(j)}}{\sum_{h' \in \mathcal{H}} \lambda_{h'}} P^{h(j)}_{j:A \cap \mathcal{J}(h(j))},$$

where $h(j)$ is the market that product j belongs to, that is, $j \in \mathcal{J}(h(j))$. Let the expected seat capacity on flight $f \in \mathcal{F}$ consumed by a customer if A is offered be denoted by $Q_f(A) := \sum_{j \in A} a_f^j P_{j:A}$. Also, let the expected revenue per customer arrival if A is offered be denoted by $R(A) := \sum_{j\in A} r_j P_{j:A}$. The customer choices in different time periods are independent conditional on the assortments offered.

Next we present a dynamic programming formulation of the RM problem. Given initial capacities of the flights $\mathbf{c}(0)$, the airline dynamically selects an assortment for each period t in order to maximize the expected total revenue. Let $c_f(t)$ denote the remaining capacity of flight f at the beginning of period t, and let $\mathbf{c}(t) := (c_f(t), f \in \mathcal{F}) \in \mathbb{Z}_+^m$. Let $\mathcal{J}_{\mathbf{c}(t)} := \{j \in \mathcal{J} : c_f(t) \geq a_f^j \ \forall f \in \mathcal{F}\}$ denote the set of products that can be offered with remaining capacities $\mathbf{c}(t)$. Let $V_t : \mathbb{Z}_+^m \mapsto \mathbb{R}$ denote the optimal revenue-to-go function at time t. The optimality equation is given by

$$\begin{aligned} V_t(\mathbf{c}(t)) &= \max_{A\subset \mathcal{J}_{\mathbf{c}(t)}} \left\{ \sum_{h\in\mathcal{H}} \sum_{j\in A\cap\mathcal{J}(h)} \lambda_h P^h_{j:A\cap\mathcal{J}(h)} \left[r_j + V_{t+1}(\mathbf{c}(t) - \mathbf{a}^j)\right] \right. \\ &\qquad \left. + \left[\sum_{h\in\mathcal{H}} \lambda_h P^h_{0:A\cap\mathcal{J}(h)} + 1 - \sum_{h\in\mathcal{H}} \lambda_h\right] V_{t+1}(\mathbf{c}(t)) \right\} \qquad (1) \\ &= \max_{A\subset \mathcal{J}_{\mathbf{c}(t)}} \left\{ \sum_{h\in\mathcal{H}} \sum_{j\in A\cap\mathcal{J}(h)} \lambda_h P^h_{j:A\cap\mathcal{J}(h)} \Big[r_j - \big(V_{t+1}(\mathbf{c}(t)) - V_{t+1}(\mathbf{c}(t) - \mathbf{a}^j)\big)\Big] \right\} + V_{t+1}(\mathbf{c}(t)) \end{aligned}$$

for all $0 \leq \mathbf{c}(t) \leq \mathbf{c}(0)$. The boundary conditions are $V_t(\mathbf{0}) = 0$ for all t and $V_T(\mathbf{c}) = 0$ for all $\mathbf{c} \in \mathbb{Z}_+^m$.

4. The Spiked-MNL Choice Model

In this section, we define the *spiked-MNL* choice model and discuss its properties. To simplify notation, in this section we omit the market index h, since each customer is associated with one market h.

In the spiked-MNL choice model, every product $j \in \mathcal{J}$ is associated with two parameters $w_j > 0$, $v_j > 0$. The parameter w_j represents the *special attractiveness* of product j when it is the cheapest available product for its itinerary; otherwise, product j has a *regular attractiveness* of v_j. The attractiveness of the null alternative is denoted by v_0. Throughout the paper (except in Online Appendix D), we assume that each product's special attractiveness is greater than its regular attractiveness.

Assumption 1. *The cheapest fare spikes are nonnegative, i.e., $w_j \geq v_j > 0$ for all products $j \in \mathcal{J}$.*

For any product $j \in \mathcal{J}$, let $g(j)$ denote the itinerary that product j is associated with, and let $J(j) := \{j' \in \mathcal{J}^{g(j)} : r_{j'} > r_j\}$ denote the set of products associated with the same itinerary as product j and that have higher fares than product j. Let $\bar{J}(j) := J(j) \cup \{j\}$, and let $\underline{J}(j) := \{j' \in \mathcal{J}^{g(j)} : r_{j'} < r_j\}$ denote the set of products associated with the same itinerary as product j and that have lower fares than product j. Suppose the customer's choice set is $S \cup \{0\}$. Let $\mathbb{1}(j,S)$ be an indicator function, such that $\mathbb{1}(j,S) = 1$ if j is the cheapest available product in S for its itinerary, and $\mathbb{1}(j,S) = 0$ otherwise. That is, $\mathbb{1}(j,S) = 1$ if $j \in S$ and $\big(\mathcal{J}^{g(j)} \cap S\big) \subset \bar{J}(j)$, and $\mathbb{1}(j,S) = 0$ otherwise. The spiked-MNL model specifies that product $j \in S$ is chosen with probability

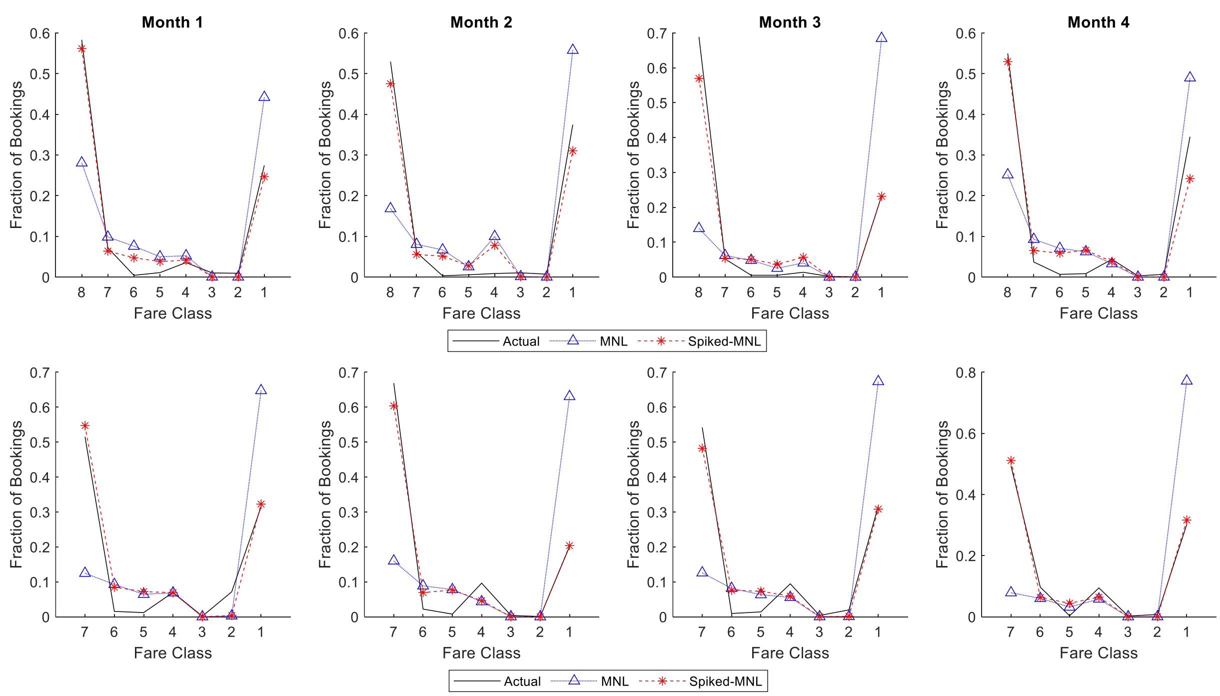
$$P_{j:S} \;=\; \frac{v_j(1 - \mathbb{1}(j,S)) + w_j \mathbb{1}(j,S)}{v_0 + \sum_{j'\in S} \left[v_{j'}(1 - \mathbb{1}(j',S)) + w_{j'}\mathbb{1}(j',S)\right]}.$$

The probability that the customer chooses the null alternative is given by

$$P_{0:S} \;=\; \frac{v_0}{v_0 + \sum_{j' \in S} \left[v_{j'} (1 - \mathbb{1}(j', S)) + w_{j'} \mathbb{1}(j', S) \right]}.$$

Choice Model Properties. In Online Appendix A, we study some fundamental properties of the spiked-MNL model. We show that the spiked-MNL model can be explained by a *context-dependent* utility model, which was introduced to explain empirical observations of context effects, that is, observations that the relative attractiveness of two alternatives depends on the presence of other alternatives (Tversky and Simonson 1993, Pompilio and Kacelnik 2010, Rooderkerk et al. 2011, Wang 2018). We also show that the spiked-MNL model does not always possess the regularity property and the submodularity property, and therefore it is not always explained by a random utility model with context-independent utilities.

Figure 2 Fractions of bookings among open fare classes predicted by the MNL and the spiked-MNL models.



Note. Each column corresponds to aggregated bookings in one month. The top row shows the fractions of bookings when fare classes 1–8 are offered; the bottom row shows the fractions of bookings when fare classes 1–7 are offered. “Actual” represents the fractions of bookings in the airline data, “MNL” and “Spiked-MNL” represent out-of-sample predictions by MNL and spiked-MNL models, respectively, calibrated with data from the previous month.

Empirical Observations. We compared the prediction accuracy of MNL and spiked-MNL models using airline data. Figure 2 shows out-of-sample actual bookings and predictions of bookings for all direct flights for one origin-destination market that departed during a 4-month period. (Empirical results for additional markets can be found in Online Appendix B.) The model parameters were calibrated on a rolling month basis; that is, we estimated the choice models using data of bookings for flights that departed during month $k \in \{0,1,2,3\}$, and then compared the booking predictions of the models with the data of actual bookings for flights that departed during month $k+1$. The "fraction of bookings" (vertical axis) represents the fraction of bookings for each open fare class; the open fare classes are shown along the horizontal axis. It is clear from Figure 2 that the predictions by the spiked-MNL model are much closer to the actual booking data than the predictions by the MNL model. We present additional numerical results in Section 8.

5. Efficient Assortments under the Spiked-MNL Model

In this section, we consider *efficient assortments* under the spiked-MNL model. In Online Appendix C, we consider an assortment optimization problem (without resource and time constraints) under the spiked-MNL model and derive the structure of the optimal assortment. This provides insight into the structure of efficient assortments in network revenue management (with resource and time constraints). Below, we present our results for network revenue management problems.

Definition 1 (Efficient Assortments). An assortment $S \subset \mathcal{J}$ is *inefficient* if a mixture of other assortments has strictly greater expected revenue with the same or less expected resource consumption. That is, there exists a set of weights $\{\mu(A) : A \subset \mathcal{J}\}$ satisfying $\sum_{A \subset \mathcal{J}} \mu(A) = 1$ and $\mu(A) \geq 0$ for all $A \subset \mathcal{J}$, such that

$$R(S) \quad < \quad \sum_{A \subset \mathcal{J}} \mu(A) R(A) \qquad \text{and} \quad Q_f(S) \quad \geq \quad \sum_{A \subset \mathcal{J}} \mu(A) Q_f(A) \quad \text{for all } f \in \mathcal{F}.$$

An assortment that is not inefficient is *efficient*.

An efficient assortment offers a Pareto-optimal tradeoff between expected revenue and expected resource consumption. Efficient assortments play an important role in RM. For the single-flight RM problem, Talluri and Van Ryzin (2004a) showed that an optimal policy always offers efficient assortments; furthermore, under the MNL choice model, the efficient assortments are nested-by-revenue. Liu and Van Ryzin (2008) show that the same result holds for the parallel-flight RM problem. The nested-by-revenue property is important because it motivates the use of a nested booking limit policy, a type of policy that is widely used in airline RM.

Next we give the definition of the nested-by-revenue property for the network RM problem, and we establish that the efficient assortments under the spiked-MNL model are nested-by-revenue. The nested-by-revenue property will be important for showing that the large-scale CDLP under the spiked-MNL model can be solved by solving a concise SBLP.

Definition 2 (Nested-by-revenue Assortments for Network RM). An assortment S is *nested-by-revenue* if for any product $j \in S$, all products associated with the same itinerary as j and with higher revenues than j are also in the assortment. That is, S is nested-by-revenue if for any $j \in S$, it holds that $J(j) \subset S$.

Theorem 1. *Under Assumption 1, every efficient assortment under the spiked-MNL model is nested-by-revenue.*

Remark 1. The nested-by-revenue result in Theorem 1 does not follow from the existing result for the MNL model. Specifically, the proof in Talluri and Van Ryzin (2004a) for the nested-by-revenue property under the MNL model requires that the expected resource consumption satisfies $Q_f(S) \leq Q_f(T)$ for any $S \subset T$. This condition does not hold in general under the spiked-MNL model.

In Online Appendix D, we give an example in which Assumption 1 does not hold, i.e., if $w_j < v_j$ for some product j, and for which there is an efficient assortment that is not nested-by-revenue. Therefore, the nonnegative spike condition in Assumption 1 is needed to establish the nested-by-revenue property.

6. Deterministic Approximation and Static Booking Limit Control

The DP (1) is intractable for large networks due to the curse of dimensionality. This motivates us to consider an approximation of the DP. A deterministic fluid approximation used in the RM literature is the choice-based deterministic linear program (CDLP), that we present in Section 6.1. A serious shortcoming of the CDLP in general is that the number of decision variables is exponential in the number of products. In Section 6.2 we present a concise problem with only n decision variables that can be used to solve the CDLP in time that is polynomial in the number of products. Solutions of the concise problem can also be used to construct various revenue management policies, including static booking limit policies that are studied in Section 7.

6.1. Choice-based Deterministic Linear Program

The choice-based deterministic linear program (CDLP) is an approximation of DP (1) in which customer arrivals and choices are replaced by their means, and capacity and demand are modeled as real-valued rather than integer valued (Gallego et al. 2004). Let decision variable $\alpha(A)$ denote the fraction of time that assortment $A \subset \mathcal{J}$ is offered, and let $\alpha := (\alpha(A), A \subset \mathcal{J})$. The CDLP is given by

$$z^{\text{CDLP}} \quad := \quad \max_{\alpha \geq 0} \sum_{A \subset \mathcal{J}} \alpha(A) T \sum_{h \in \mathcal{H}} \lambda_h \sum_{j \in A} r_j P^h_{j:A \cap \mathcal{J}(h)} \tag{2a}$$

$$\text{s.t.} \sum_{A\subset\mathcal{J}} \alpha(A) \;\; \leq \;\; 1, \tag{2b}$$

$$\sum_{A\subset\mathcal{J}} \alpha(A) T \sum_{h\in\mathcal{H}} \lambda_h \sum_{j\in A} a_f^j P^h_{j:A\cap\mathcal{J}(h)} \;\; \leq \;\; c_f \qquad \forall\, f\in\mathcal{F}. \tag{2c}$$

The objective (2a) of the CDLP is the expected total revenue over the time horizon. Constraint (2b) specifies that the sum of the fractions of time that different assortments are offered is less than 1. In the remaining $1-\sum_{A\subset\mathcal{J}}\alpha(A)$ fraction of time, an empty set is offered. Constraint (2c) enforce resource capacity constraints.

Liu and Van Ryzin (2008) showed that optimal solutions of problem (2) use efficient assortments only, that is, problem (2) has an optimal solution α^* such that $\alpha^*(A)=0$ for all inefficient assortments A. Theorem 1 established that every efficient assortment under the spiked-MNL model is nested-by-revenue. Thus, under the spiked-MNL model, if assortment A is not nested-by-revenue, then decision variable $\alpha(A)$ can be omitted. Therefore the number of decision variables is reduced from $2^n = 2^{\sum_{g\in\mathcal{G}} n(g)} = \prod_{g\in\mathcal{G}} 2^{n(g)}$ to $\prod_{g\in\mathcal{G}} (n(g)+1)$. However, the reduced number of decision variables still is exponential in the number of itineraries. This motivated us to develop a more concise LP formulation in the next section. This formulation also uses the result in Theorem 1.

6.2. Sales-Based Linear Program

Under the MNL model, there is an LP formulation called the sales-based linear program (SBLP), which has a polynomial number of decision variables and constraints, and which can be used to solve the CDLP (Gallego et al. 2015). Dai et al. (2014) developed an SBLP formulation for the spiked-MNL model with a number of decision variables that is quadratic in the number of products, and presented a polynomial-time algorithm to convert optimal solutions for the SBLP into optimal solutions for the CDLP and vice versa. Next we use the result of Theorem 1 to develop an SBLP formulation for the spiked-MNL choice model with a number of decision variables that is linear in the number of products.

Our SBLP formulation takes into account that all assortments offered by an optimal solution are nested-by-revenue. Let x_j denote the sales of product j when j is the cheapest available product for its itinerary, and let $\mathbf{x} := (x_j, j\in\mathcal{J})$. Consider any assortment A that is nested-by-revenue, and such that j is the cheapest available product in A for its itinerary. Note that $\bar{J}(j)$, the set of products for the same itinerary as product j and that have equal or higher fares than product j, satisfies $\bar{J}(j)\subset A$. Then, for the market h such that $j\in\mathcal{J}(h)$ and for any $j'\in J(j)$, it holds that $P^h_{j':A\cap\mathcal{J}(h)}/P^h_{j:A\cap\mathcal{J}(h)} = v_{j'}/w_j$, which is the same for all A satisfying the conditions above. Therefore, at the same time that x_j units of product j is sold when j is the cheapest available product for its itinerary, $(v_{j'}/w_j)\,x_j$ units of each product $j'\in J(j)$ is sold. Let x_0^h denote the number of customers in market h who choose the no-purchase alternative, and let $\mathbf{x}_0 := (x_0^h, h \in \mathcal{H})$. The SBLP under the spiked-MNL model is given by

$$
\begin{aligned}
z^{\text{SBLP}} \quad = \quad & \max_{\mathbf{x},\mathbf{x}_0} \sum_{h\in\mathcal{H}} \sum_{j\in\mathcal{J}(h)} \left(r_j + \sum_{j'\in J(j)} r_{j'} \frac{v_{j'}}{w_j} \right) x_j && && \text{(3a)} \\
& \text{s.t. } x_0^h + \sum_{j\in\mathcal{J}(h)} \left(1 + \sum_{j'\in J(j)} \frac{v_{j'}}{w_j} \right) x_j \quad = \quad \lambda_h T && \forall\ h \in \mathcal{H} && \text{(3b)} \\
& \sum_{h\in\mathcal{H}} \sum_{j\in\mathcal{J}(h)} \left(1 + \sum_{j'\in J(j)} \frac{v_{j'}}{w_j} \right) a_f^j x_j \quad \le \quad c_f && \forall\ f \in \mathcal{F} && \text{(3c)} \\
& \sum_{j\in\mathcal{J}^g} \frac{x_j}{w_j} \quad \le \quad \frac{x_0^h}{v_0} && \forall\ h \in \mathcal{H},\ g \in \mathcal{J}(h) && \text{(3d)} \\
& \mathbf{x} \quad \ge \quad \mathbf{0}, \quad \mathbf{x}_0 \quad \ge \quad \mathbf{0}.
\end{aligned}
$$

The objective (3a) is the total revenue. Constraint (3b) represents the fact that, for each market, the number of bookings plus the number of no-purchase customers equals the number of arrivals. Constraint (3c) is the capacity constraint for each resource. Constraint (3d) is a generalization of the scale constraint in Gallego et al. (2015) to include the spike effect. The quantity x_j/w_j is proportional to the amount of time that product j is the cheapest available product for its itinerary. Since the null alternative is always available, the constraint states that the total amount of time that different products are the cheapest available products for an itinerary cannot exceed the total amount of time that the null alternative is available. The SBLP formulation above applies to a time-homogeneous demand model and a single booking channel. It is easy to extend the SBLP formulation to a piecewise constant time-varying demand model and multiple booking channels. This extension is used in our numerical experiments based on real-world airline data (Section 8).

The following result establishes that the SBLP (3) can be used to solve the CDLP (2) in polynomial time under the spiked-MNL model.

Theorem 2. *Under the spiked-MNL model, given an optimal solution of the CDLP* (2)*, an optimal solution of the SBLP* (3) *can be constructed in polynomial time, and vice versa. Moreover, the CDLP* (2) *has an optimal solution that consists of a nested sequence of assortments, each of which is nested-by-revenue.*

Specifically, in Online Appendix E.3, we give algorithms that convert optimal solutions between the two problems in polynomial time. We show how an optimal solution of the SBLP (3) can be converted to an optimal solution of the CDLP (2) that consists of a *nested sequence of assortments*, each of which is nested-by-revenue. That is, the algorithm constructs a sequence of assortments $S_1 \supset S_2 \supset \cdots \supset S_k$, each of which is nested-by-revenue, and an optimal CDLP solution $(\alpha^*(A), A \subset \mathcal{J})$, such that $\alpha^*(A) > 0$ only if $A \in \{S_1, \ldots, S_k\}$. Note that, unlike the case for a single flight, for a network of flights, a set of assortments that are nested-by-revenue might *not* form a nested sequence — a simple counterexample is the following two nested-by-revenue assortments for two parallel flights: one assortment contains the highest fare class for the first flight only, and the other assortment contains the highest fare class for the second flight only. The observation that there exists an optimal CDLP solution that consists of a nested sequence of assortments plays an important role in the construction of static booking limit controls that we discuss next.

7. Static Booking Limit Controls

Booking limits are widely used by airline reservation systems for controlling availability of fare classes. With a *partitioned* booking limit policy, a number of seats, called the partitioned booking limit (or just booking limit), is allocated to each product, and a product is closed for bookings once the number of units of that product sold reaches its booking limit. With a *nested* booking limit policy, a number of seats, called the nested booking limit, is allocated to each subset $\underline{J}(j) \cup \{j\}$, $j \in \mathcal{J}$, of products that is nested-by-revenue. A nested booking limit policy can be implemented using either standard nesting or theft nesting (see, e.g., Talluri and Van Ryzin 2004b, Bertsimas and De Boer 2005). Under standard nesting, product j is closed for booking if, for any product $j' \in \bar{J}(j)$, it holds that the number of units sold in subset $\underline{J}(j') \cup \{j'\}$ has reached the nested booking limit of subset $\underline{J}(j') \cup \{j'\}$. Under theft nesting, product j is closed for booking if the total number of units of all products in $\mathcal{J}^{g(j)}$ sold has reached the booking limit of subset $\underline{J}(j) \cup \{j\}$. Under both nested booking limit policies, a higher revenue product is available whenever a lower revenue product is available. If there were no cancellations, then under any of the three booking limit policies above, once a product is closed for booking, it would remain closed until the end of the time horizon.

7.1. Booking Limits from the SBLP Solution

By Theorem 2, an optimal solution for the SBLP (3) can be used to obtain booking limits, where the booking limit for each product in the case of a partitioned booking limit policy (or each nested subset of products in the case of a nested booking limit policy) is given by the optimal sales of that product (or that nested subset of products) for the SBLP (3). In particular, let $\mathbf{x}^* = (x_j^* : j \in \mathcal{J})$ be an optimal solution for the SBLP. The resulting amount of product j sold, denoted by b_j^*, is given by

$$b_j^* \quad = \quad x_j^* + \sum_{j' \in \underline{J}(j)} \frac{v_j}{w_{j'}} x_{j'}^*. \tag{4}$$

By (3c), $\sum_{j\in\mathcal{J}} a_f^j b_j^* \le c_f$. We thus define a (static) partitioned booking limit policy by setting the booking limit of product j to b_j^*. We also define a (static) nested booking limit policy, where the booking limit b_j^{nested} for subset $\underline{J}(j)\cup\{j\}$ is given by

$$b_j^{\text{nested}} \;=\; \sum_{j'\in\underline{J}(j)\cup\{j\}} b_{j'}^*. \tag{5}$$

In this way, an optimal solution for SBLP (3) provides three static booking limit policies:

- a partitioned booking limit policy, using the sales given by (4) as booking limits;
- a standard nested booking limit policy, using booking limits given by (5);
- a theft nested booking limit policy, also using booking limits given by (5).

Under any of the static booking limit policies, and under any sample path of customer arrivals and choices, a sequence of assortments $S_1, S_2, \ldots, S_K$ are offered such that $S_1 \supset S_2 \supset \cdots \supset S_K$. If all the random variables in the system associated with customer arrivals and choices were replaced by their means, then the resulting sequence of assortments would correspond to an optimal CDLP solution, arranged to form a nested sequence of assortments (Theorem 2).

In contrast with the booking limit policies described above, other researchers (e.g., Liu and Van Ryzin 2008) have proposed using an optimal CDLP solution $(\alpha^*(A), A\subset\mathcal{J})$ to construct a static time-based policy, as follows: The time horizon is partitioned into intervals with lengths proportional to $\alpha^*(A)$, $A\subset\mathcal{J}$, and during each such time interval the assortment A is offered (as long as sufficient resources are available).

7.2. Asymptotic Optimality of the Static Partitioned Booking Limit Policy

In this section we study the asymptotic properties of the partitioned booking limit policy defined in Section 7. In the asymptotic setting, it is convenient to consider the continuous time version of the problem. Thus, in this section, we assume that customers arrive according to a Poisson process instead of a Bernoulli process; the Bernoulli process considered in Section 3 can be viewed as an approximation of the Poisson process if the probability that more than one customer arrives in a period is negligible. We study the partitioned booking limit policy under the following asymptotic regime often considered in the RM literature. Let $\mathbf{c}$ denote the baseline capacity, and let λ denote the baseline customer arrival rate. Consider a sequence of RM problems indexed by $\theta = 1, 2, \ldots$, with capacity $\theta\mathbf{c}$ and arrival rate $\theta\lambda$, respectively. Other model parameters remain constant when θ grows. We refer to an RM problem scaled by θ as the *θ-scaled problem*. Let z_{OPT}^{θ} denote the optimal expected revenue for the θ-scaled problem. Note that for the θ-scaled problem, the optimal objective value of the corresponding CDLP is θz^{CDLP}, where z^{CDLP} denotes the optimal objective value of the baseline CDLP (2). Then $z_{OPT}^{\theta} \le \theta z^{\text{CDLP}}$ (Gallego et al. 2015, Liu and Van Ryzin 2008). Let

Z^θ denote the objective value, that is the (random) revenue, for the θ-scaled problem under the partitioned booking limit policy. Then the following result establishes that the partitioned booking limit policy is asymptotically optimal under fluid scaling.

THEOREM 3. *The expected revenue* $\mathsf{E}[Z^\theta]$ *of the partitioned booking limit policy defined by* (4) *satisfies*

$$\lim_{\theta\to\infty} \frac{\mathsf{E}[Z^\theta]}{\theta} \quad = \quad z^{\text{CDLP}}.$$

Theorem 3 implies that

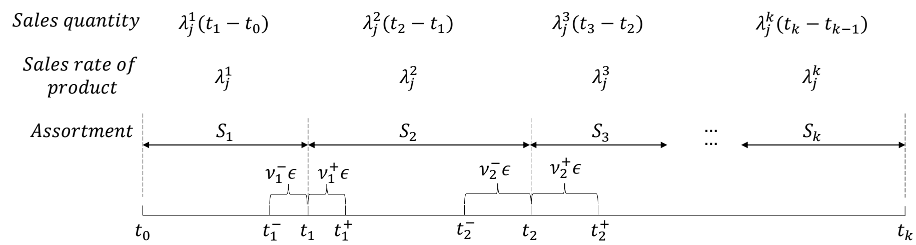
$$\lim_{\theta\to\infty} \frac{\mathsf{E}[Z^\theta]}{z^\theta_{OPT}} \quad = \quad \lim_{\theta\to\infty} \frac{\mathsf{E}[Z^\theta]}{\theta z^{\text{CDLP}}} \quad = \quad 1.$$

Therefore, when customer demand and seat capacities are large, the partitioned booking limit policy is near optimal.

The proof of Theorem 3 can be found in Online Appendix E.4. The asymptotic optimality result for booking limit policies under a choice model is the first of its kind. It required a proof technique that is different from the standard technique for time-based policies (described at the end of Section 7.1). Theorem 3 required a different technique because time-based policies result in a deterministic sequence of assortments, and therefore a deterministic sequence of choice probabilities. In contrast, under booking limit policies, random customer choices cause products to close for booking in random order, thereby resulting in a random sequence of assortments, and therefore a random sequence of choice probabilities. This difference introduced complications. These complications do not exist for the independent demand model, for example, in the analysis of a static booking limit policy under the independent demand model in Cooper (2002). Therefore, a new approach was used to prove Theorem 3.

Our proof is based on the following approach. First, based on an optimal CDLP solution $(\alpha^*(A), A \in \{S_1, \ldots, S_k\})$, the time horizon is partitioned into intervals given by time points $t_0 < t_1 < \cdots < t_k$. Then, an amount of "padding" is added around each time t_i, given by $t_i^- = t_i - \nu_i^- \varepsilon$ and $t_i^+ = t_i + \nu_i^+ \varepsilon$ for some small $\varepsilon > 0$ and appropriately chosen values of ν_i^- and ν_i^+, $i = 1, \ldots, k$ (see Figure 3). We show that with high probability, the assortments offered by the booking limit policy outside the intervals (t_i^-, t_i^+) are the same as the assortments offered by the CDLP solution. The booking process within intervals (t_i^-, t_i^+) can be complex, but we derive upper and lower bounds on the deviation of the booking process from the CDLP prediction. We show by induction that the booking quantities at the end of the time horizon is $O(\varepsilon)$ away from the static booking limits given by (4), and since we can choose ε to be arbitrarily small as $\theta \to \infty$, this establishes the asymptotic optimality result.

Figure 3



8. Numerical Experiments

In this section, we examine the spiked-MNL model in terms of how well it predicts customer choices on out-of-sample data and in terms of revenue performance. First we conduct an experiment with synthetic data in which we calibrate three choice models and then solve assortment optimization problems under the calibrated models. The results show that the spiked-MNL model outperforms the other two choice models. We also consider a dynamic learning setting in which a vendor iteratively collects sales data, calibrates choice models, optimizes controls, and implements the controls. We demonstrate that the revenue gap between the MNL model and the spiked-MNL model may increase as learning proceeds.

We then examine the prediction performance of the spiked-MNL model using real airline data. We compare the out-of-sample prediction performance of three models: the MNL model, the spiked-MNL model, and a recently proposed random forest (RF) model. We also compare the revenue performance of different RM policies and examine their robustness with respect to parameter perturbations. (The data and source code for this section can be found at `https://github.com/cyf-sjtu/spikedMNL`.)

8.1. Experiments with Synthetic Data

In this section, we use synthetic data to compare the performance of three choice models: the MNL model, the spiked-MNL model, and the general attraction model (GAM) proposed in Gallego et al. (2015). For the MNL model, we estimate an alternative-specific constant for each product, i.e., its attractiveness. For the spiked-MNL model, we estimate two parameters for each product: the regular attractiveness and the special attractiveness when the product is the cheapest available product in the assortment. For the GAM model, we also estimate two parameters: the regular attractiveness and the shadow attractiveness when the product is absent from the assortment. Given that there are n possible products, the numbers of parameters of the MNL, spiked-MNL, and GAM models are n, $2n-1$, and $2n$ respectively. We test the prediction and revenue performance of the above three models. Finally, we consider a simple one-resource two-product example to demonstrate the evolution of revenue when the model parameters are updated iteratively.

8.1.1. Experiment Setup. We randomly generate the choice data using a complex ground choice model that is very different from either the MNL, the spiked-MNL, or the GAM model, but that demonstrates a spike phenomenon. In the ground choice model, there are p customer types. Each customer type $\ell \in \{1,\ldots,p\}$ is characterized by a consideration set of products $\mathcal{J}_\ell \subset \mathcal{J}$. All types of customers have equal arrival rates. An arriving customer chooses the cheapest product in her consideration set. If no product in her consideration set is available, then the customer chooses the null alternative.

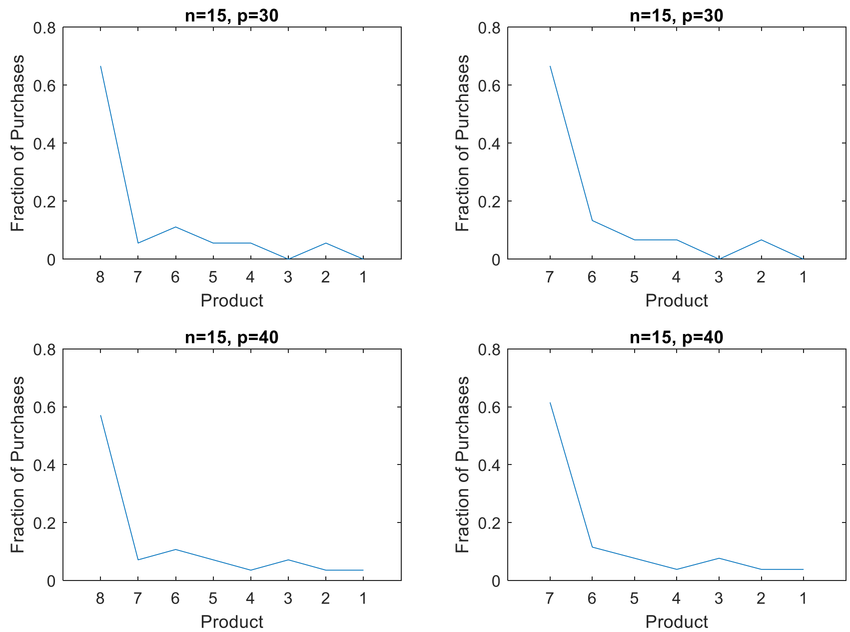
To generate the consideration sets, we consider the case in which the products have an inherent ordering, such that each product $j \in \{1,\ldots,n-1\}$ has a higher quality and a higher price than product $j+1$. Each type of customers has a maximum price she can afford and a minimum quality she accepts, that is, each type ℓ of customers has a range $\{i_\ell,\ldots,j_\ell\}$ of products as consideration set $\mathcal{J}_\ell$, so that a customer of type ℓ does not consider products with price higher than that of product i_ℓ and does not consider products with quality lower than that of product j_ℓ. In this case, each type of customers chooses the cheapest available product that is within this range.

In each experiment, we randomly generated p consideration sets $\mathcal{J}_\ell = \{i_\ell,\ldots,j_\ell\}$, $\ell = 1,\ldots,p$, as follows: For each ℓ, first i_ℓ is generated according to the discrete uniform distribution on $\{1,\ldots,n\}$, and then j_ℓ is generated according to the discrete uniform distribution on $\{i_\ell,\ldots,n\}$. If $\{i_\ell,\ldots,j_\ell\} = \{i_{\ell'},\ldots,j_{\ell'}\}$ for some $\ell' < \ell$, then a new realization of $\{i_\ell,\ldots,j_\ell\}$ is generated, until there is no duplication. Thereafter we generated the i.i.d. choice data according to the ground choice model. Specifically, for each $t = 1,\ldots,\tau$, we generated assortment $A(t) = \{1,\ldots,m(t)\}$, where τ denotes the number of customers in the data set, and each $m(t)$ is generated according to the discrete uniform distribution on $\{1,\ldots,n\}$. The type $\ell(t)$ of customer t is generated according to the discrete uniform distribution on $\{1,\ldots,p\}$. Then the alternative $j(t) = \max\{j \in A(t) \cap \mathcal{J}_{\ell(t)}\}$ chosen by customer t is determined by the ground choice model described above. If $A(t) \cap \mathcal{J}_{\ell(t)} = \varnothing$, then $j(t) = 0$. Thus the choice data consists of the pairs $\{(A(t), j(t)) : t = 1,\ldots,\tau\}$.

Figure 4 shows the choice probabilities $P_{j:A}/(1 - P_{0:A})$, conditional on a product being chosen, under a realization of the ground choice model with $n = 15$ products, for different assortments A and different products j. Specifically, the top two figures show the choice probabilities for one random realization of the ground choice model with $p = 30$ customer types, and the bottom two figures show the choice probabilities for one random realization of the ground choice model with $p = 40$ customer types. In the left two figures the assortment $A = \{1,\ldots,8\}$, and in the right two figures the assortment $A = \{1,\ldots,7\}$. Note that the ground choice model exhibits a cheapest fare spike effect.

8.1.2. Prediction Performance. We conducted 20 independent tests of the prediction performance of the MNL, the spiked-MNL, and the GAM models. In each of these tests we fixed $n = 15$,

Figure 4 Choice probabilities $P_{j:A}/(1-P_{0:A})$, conditional on a product being chosen, for different assortments A and products j, for one realization of the ground choice model.



and we randomly generated a ground choice model for each $p \in \{30, 40\}$. For each of these ground choice models we generated the choices of τ customers, for each $\tau \in \{1500, 3000, 4500, 6000, 7500\}$. Each of these data sets was used as training data, and for each data set we used maximum likelihood estimation to fit the MNL, the GAM, and the spiked-MNL models. For each data set we generated the choices of another 3000 customers as testing data. Then we used the testing data to compare the fitted choice models in terms of their *out-of-sample log-likelihood* values.

Table 1 Out-of-sample log-likelihoods of the fitted choice models ($n = 15$, $p = 30$)

τ	MNL LogLik.	GAM LogLik.	Spiked-MNL LogLik.	Impr. over MNL	Impr. over GAM
1500	-4293.6 (70.9)	-4279.7 (70.3)	**-3851.8** (75.4)	10.29%	10.00%
3000	-4297.5 (70.6)	-4281.9 (70.2)	**-3848.5** (72.5)	10.47%	10.14%
4500	-4300.6 (76.2)	-4283.7 (76.0)	**-3855.6** (74.7)	10.35%	10.00%
6000	-4293.4 (73.2)	-4277.7 (73.0)	**-3836.6** (72.6)	10.65%	10.32%
7500	-4296.9 (65.5)	-4282.0 (65.6)	**-3836.5** (64.7)	10.71%	10.40%

The results are summarized in Table 1 and Table 2. The first column gives the size τ of the training data. The second to the fourth columns report the average out-of-sample log-likelihood values of

Table 2 Out-of-sample log-likelihoods of the fitted choice models ($n = 15$, $p = 40$)

τ	MNL LogLik.	GAM LogLik.	Spiked-MNL LogLik.	Impr. over MNL	Impr. over GAM
1500	-4414.5 (49.0)	-4404.4 (48.1)	**-4018.3** (54.6)	8.97%	8.77%
3000	-4411.4 (51.1)	-4397.9 (51.2)	**-4011.0** (55.5)	9.10%	8.82%
4500	-4417.1 (51.9)	-4403.4 (51.6)	**-4013.3** (56.8)	9.18%	8.89%
6000	-4407.1 (49.9)	-4393.6 (49.4)	**-4007.5** (54.7)	9.09%	8.81%
7500	-4409.0 (43.0)	-4394.8 (42.9)	**-4001.7** (50.1)	9.26%	8.96%

the MNL, the GAM, and the spiked-MNL models, respectively, over the 20 tests; the numbers in parentheses are the corresponding standard errors of the 20 out-of-sample log-likelihood values. The fifth and sixth columns are the percentage improvement of the average out-of-sample log-likelihood values of the spiked-MNL model over the MNL and the GAM models respectively. The spiked-MNL model outperformed both the MNL and the GAM models significantly in terms of prediction power. Taking into account the fact that the GAM model has double the number of parameters of the MNL model, and one more parameter than the spiked-MNL model, we conclude that the superior prediction power of the spiked-MNL model is due to the structure of the spiked-MNL model and not just its number of parameters being greater than that of the MNL model.

8.1.3. Assortment Revenue Performance. We compared the MNL, GAM, and spiked-MNL models in terms of the revenue performance of their chosen assortments. For each test $q = 1, \ldots, 20$, one ground choice model P^q was generated. For each test q, each assortment $A \subset \mathcal{J}$, and each product $j \in \mathcal{J}$, let $P^q_j(A)$ denote the choice probability under the ground choice model of test q of product j given that assortment A is offered. Also, let $\phi^q_{j,z}(A)$ denote the choice probability under the fitted model $z \in \{$MNL, GAM, spiked-MNL$\}$ of test q of product j given that assortment A is offered. For each test we also generate 20 samples of the product revenues. Specifically, for each test q, and for each revenue sample $k = 1, \ldots, 20$, we generate $n = 15$ i.i.d. uniform $(0, 100)$ random variables, and then we sort them to give the product revenues $r^{qk}_1 > \cdots > r^{qk}_n$. For each test q, fitted model $z \in \{$MNL, GAM, spiked-MNL$\}$, and revenue sample k, let $\hat{A}^{qk}_z$ denote an optimal solution of the assortment optimization problem $\max_{A \subset \mathcal{J}} \sum_{j \in A} r^{qk}_j \, \phi^q_{j,z}(A)$. Then the expected revenue of assortment $\hat{A}^{qk}_z$ under the ground choice model is given by $R^{qk}_z := \sum_{j \in \hat{A}^{qk}_z} r^{qk}_j \, P^q_j(\hat{A}^{qk}_z)$. Similarly, the optimal expected revenue under the ground choice model is given by $R^{qk}_{\text{opt}} := \max_{A \subset \mathcal{J}} \sum_{j \in A} r^{qk}_j \, P^q_j(A)$. R^{qk}_{opt} can be computed using the method described in Aouad et al. (2015). We compute the ratios $R^{qk}_z / R^{qk}_{\text{opt}}$, and average these ratios over the revenue samples $k = 1, \ldots, 20$ and tests $q = 1, \ldots, 20$.

The results are summarized in Table 3 and Table 4. The layout of these tables is the same as that of Table 1. The three choice models performed similarly in these tests, capturing between 90% and 93% of the optimal expected revenues. The spiked-MNL model performed slightly better than the other two models especially when the types of customers are not very diverse and when more training data are available.

Table 3 Revenue performance of the fitted choice models ($n = 15$, $p = 30$)

τ	MNL Rev. Ratio	GAM Rev. Ratio	Spiked-MNL Rev. Ratio	Impr. over MNL	Impr. over GAM
1500	.9118 (.0037)	**.9146** (.0037)	.9081 (.0040)	-0.41%	-0.71%
3000	.9127 (.0038)	**.9186** (.0036)	.9182 (.0035)	0.64%	-0.01%
4500	.9066 (.0034)	.9118 (.0032)	**.9225** (.0033)	1.77%	1.20%
6000	.9114 (.0037)	.9135 (.0036)	**.9239** (.0032)	1.40%	1.16%
7500	.9093 (.0037)	.9142 (.0035)	**.9241** (.0031)	1.67%	1.14%

Table 4 Revenue performance of the fitted choice models ($n = 15$, $p = 40$)

τ	MNL Rev. Ratio	GAM Rev. Ratio	Spiked-MNL Rev. Ratio	Impr. over MNL	Impr. over GAM
1500	.9227 (.0032)	**.9243** (.0031)	0.9154 (.0036)	-0.79%	-0.96%
3000	.9179 (.0031)	**.9210** (.0031)	.9180 (.0031)	0.05%	-0.28%
4500	.9230 (.0031)	**.9289** (.0030)	.9239 (.0030)	0.13%	-0.50%
6000	.9164 (0.0034)	.9204 (0.0034)	**.9237** (0.0031)	0.80%	0.37%
7500	.9193 (0.0032)	.9232 (0.0031)	**.9254** (0.0029)	0.70%	0.28%

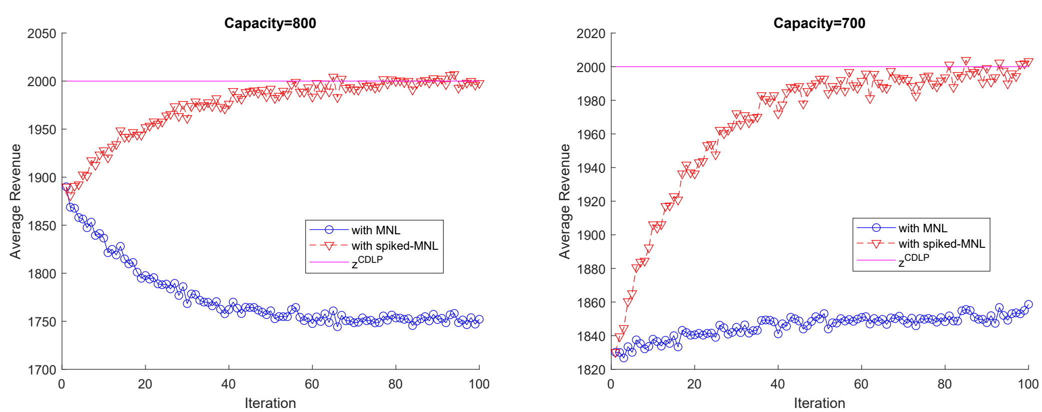
In summary, the spiked-MNL model outperforms the MNL and the GAM models in a synthetic setting in which customers prefer the cheapest product in their consideration sets.

8.1.4. Evolution of Average Revenue. We present a simple one-resource two-product example to demonstrate the evolution of revenue when choice models are calibrated with observed choice data, and revenue management decisions are based on the calibrated models. A seller sells two types of products H and L, each of which uses one unit of a resource. Products H and L have revenues

$r_H = 3$ and $r_L = 2$ respectively. During each selling season customers arrive according to a homogeneous Poisson process with rate $\lambda T = 1000$. Customers make their choices according to probabilities $P_{H:\{H,L\}} = 1/4$, $P_{L:\{H,L\}} = 1/2$, and $P_{H:\{H\}} = 2/3$. Here we consider only nested-by-revenue assortments as they are the efficient assortments under the MNL and the spiked-MNL models. The seller controls the fractions of time that assortments $\{H, L\}$ and $\{H\}$ are offered. More specifically, assortment $\{H, L\}$ is offered for a fraction α_{HL} of the selling season, and then assortment $\{H\}$ is offered for the remaining time until the seller runs out of stock or the selling season ends. Initially, the fraction of time that $\{H, L\}$ is offered is set at $\alpha_{HL} = 0.5$. After each selling season the seller calibrates the choice model being used (either MNL or spiked-MNL) using maximum likelihood estimation with the data generated so far, and then the seller solves the CDLP under the choice model being used to get the control α^*_{HL}. The control for the next selling season is given by $\alpha^{new}_{HL} = (1-\kappa)\alpha^{old}_{HL} + \kappa\alpha^*_{HL}$, where parameter $\kappa = 0.05$ is a smoothing parameter; and α^{old}_{HL} denotes the control used in the most recent selling season. The process is simulated for 100 successive selling seasons, and the revenue collected in each selling season is recorded. This gives a trajectory of evolving revenue for one sample path. We simulated 200 i.i.d. sample paths, and calculated the average revenue trajectory.

Figure 5 shows the average revenue trajectories under the MNL and the spiked-MNL models, for two values of capacity. The figure on the left shows a case in which the average revenue under the MNL model decreases over time while the average revenue under the spiked-MNL model increases. The figure on the right shows a case in which the average revenue under the MNL model slightly increases over time. In both cases a non-negligible gap opens up between the average revenue under the MNL model and that under the spiked-MNL model. In Online Appendix A.4, we provide an analysis of such a dynamical system and the difference between the revenue performance of the MNL model and the spiked-MNL model.

Figure 5 Trajectories of average revenue, with the average revenue under the MNL model deteriorating.



8.2. Experiments with Airline Data

In this section, we evaluate the performance of the spiked-MNL model with airline data.

8.2.1. Data and Models. We considered a busy origin-destination market with more than 30 flights per day. The data set includes booking data of all airlines that operate in this market. There are 5 booking channels. For demand modeling purposes, the selling horizon is divided into 200 booking time intervals. Each channel-interval combination is associated with a different customer segment. That is, the calibrated models allow the choice parameters to be different for different combinations of booking channel and booking time interval. More details about the airline data and how we processed it can be found in Online Appendix F.

We model and estimate customer demand as follows. Let $\mathcal{N}$ denote the index set of the booking requests in the data set for the specific origin-destination market for a specific departure date. For each request $i \in \mathcal{N}$, let c_i denote the booking channel used, let ℓ_i denote the index of the booking time interval, and let A_i denote the assortment offered to customer i. Let $\mathbf{x}^{i,j}$ denote an attribute vector containing information about request i and product j. For example, to allow different price sensitivity estimates for customers who use different channels and who arrive during different time periods, there are attributes of the form $\text{price}_j \times \mathbb{1}\{c_i = c, \ell_i = \ell\}$ for each booking channel c and each booking time interval ℓ. Other attributes include departure time period, change fees, and frequent flyer mileage gain. We compared the prediction performance of the MNL model, the spiked-MNL model, as well as several variants of the random forest (RF) choice model, which is proposed in Lhéritier et al. (2019) specifically for the airline context. For the spiked-MNL model and some versions of the RF model, $\mathbf{x}^{i,j}$ also contains a binary variable indicating whether product j is the cheapest available product for its itinerary.

For the MNL model and the spiked-MNL model, let $v(\mathbf{x}^{i,j})$ denote the attractiveness of product j for customer i given attribute vector $\mathbf{x}^{i,j}$. Quantity $v^0(c_i, t_i)$ denotes the attractiveness of the null alternative. Then request i chooses alternative $j \in A_i$ with probability

$$P_{j:A_i} \quad = \quad \frac{v(\mathbf{x}^{i,j})}{v^0(c_i, \ell_i) + \sum_{j' \in A_i} v(\mathbf{x}^{i,j'})}. \tag{6}$$

The parameters in (6) are estimated with the airline data using maximum likelihood estimation.

The RF model proposed in Lhéritier et al. (2019) predicts for each alternative, given its attribute values, whether it will be chosen or not. We implemented the original RF model proposed in Lhéritier et al. (2019) as well as two variants that include additional information about the assortment. The details of the RF models are given below:

- RF-1: This is the original RF model proposed in Lhéritier et al. (2019). It takes the raw booking data as an attribute.

- RF-2: This is a variant of RF-1, which takes as input the raw booking data as well as an indicator whether a product is the cheapest available product for its itinerary.
- RF-3: It is the same as RF-2, but further includes the number of products in the assortment as input.

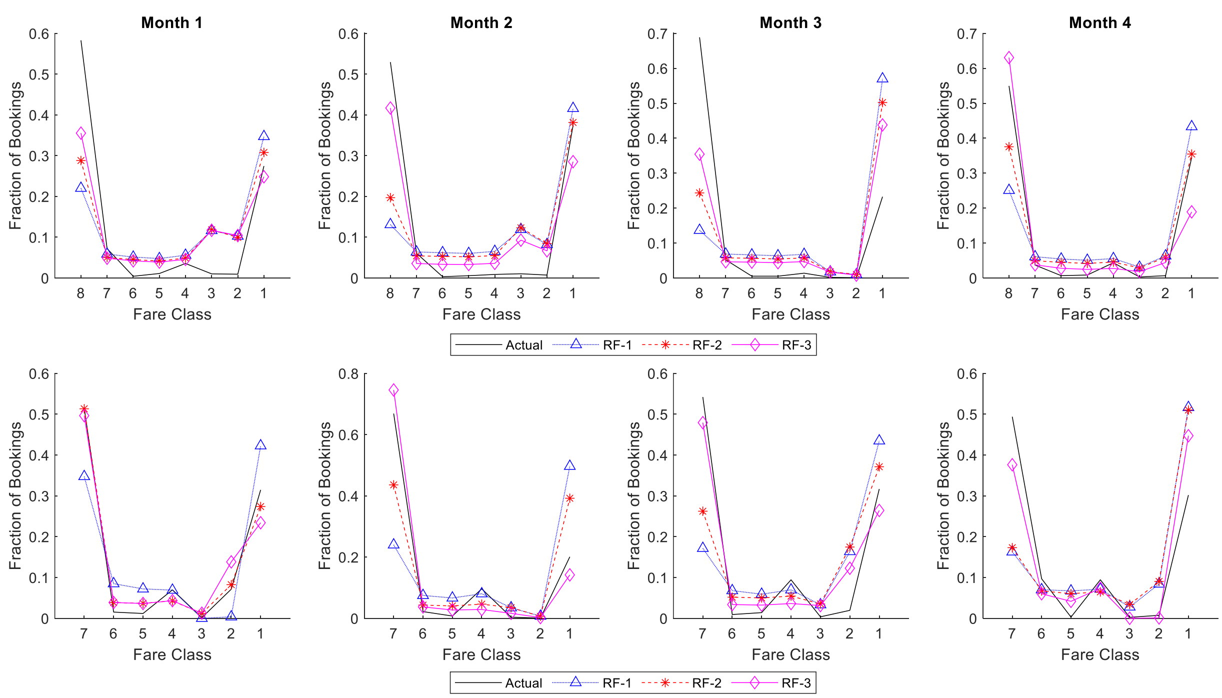
The RF models were calibrated using `R Studio` with `R 3.5.1` and the `wsrf` package.

8.2.2. Prediction Performance. We calibrated and tested all the models on a rolling month basis. That is, we calibrated the models using the data of one month and then tested their prediction performance for the next month. Table 5 reports the (scaled) *out-of-sample* log-likelihoods of these models over 10 months. The spiked-MNL model achieved on average the best out-of-sample prediction and it often gave the best or second best performance.

Table 5 Out-of-sample log-likelihoods of different models calibrated on a rolling month basis.

	1	2	3	4	5	6	7	8	9	10	Avg.
Spiked-MNL	**-1.470**	-1.476	**-1.477**	-1.494	-1.477	**-1.476**	**-1.478**	**-1.417**	**-1.459**	**-1.475**	**-1.470**
MNL	-1.533	-1.553	-1.546	-1.551	-1.540	-1.528	-1.539	-1.484	-1.533	-1.527	-1.533
RF-1	-1.522	-1.522	-1.513	-1.521	-1.526	-1.511	-1.520	-1.564	-1.516	-1.516	-1.523
RF-2	-1.487	-1.475	-1.478	-1.468	-1.485	-1.478	-1.480	-1.525	-1.474	-1.489	-1.484
RF-3	-1.488	**-1.474**	-1.486	**-1.466**	**-1.472**	-1.507	-1.482	-1.511	-1.518	-1.492	-1.490

Figure 6 Actual fractions of bookings and predicted fractions of bookings given by the RF models.



We also checked whether these RF models captured the spike effect we observed in the booking data. Figure 6 reports the actual fractions of bookings and the predictions by the RF models for four consecutive months, given that the assortment offered fare classes $1, \ldots, 8$ for a particular itinerary (top row), and given that the assortment offered fare classes $1, \ldots, 7$ for the same itinerary (bottom row). The original RF model (RF-1) in Lhéritier et al. (2019) does not capture the spike effect well. With the inclusion of additional information about the assortment, as in RF-2 and RF-3, the RF models capture the spike effect better, but still the performance of the RF models is not as good as the spiked-MNL model in terms of capturing the spike effect and fitting the data. See Figure 2 for comparison. Also, Table 6 and Table 7 report the Wasserstein distance and Kullback-Leibler divergence between the actual fractions of bookings and the predicted fractions of bookings given by different choice models for different months. In most cases, the spiked-MNL model gives predicted fractions that are the closest to the empirical distribution, despite having a much smaller number of parameters than the random forest models. In addition, the RF models take much more computational effort to calibrate than the other two models. The MNL and spiked-MNL models were calibrated within a few minutes with the given data set, but the RF models took several hours to calibrate.

Table 6 Wasserstein distance between the actual fractions of bookings and the predicted fractions of bookings given by different choice models for different months.

	1	2	3	4
Spiked-MNL	**0.018**	**0.031**	**0.024**	**0.033**
MNL	0.039	0.053	0.025	0.041
RF-1	0.073	0.089	0.054	0.053
RF-2	0.069	0.081	0.047	0.044
RF-3	0.064	0.051	0.064	0.039

(a) when fare classes 1–8 are open

	1	2	3	4
Spiked-MNL	**0.017**	**0.025**	**0.028**	**0.020**
MNL	0.058	0.026	0.062	0.086
RF-1	0.029	0.054	0.072	0.050
RF-2	0.020	0.081	0.064	0.047
RF-3	0.036	0.034	0.033	0.031

(b) when fare classes 1–7 are open

Table 7 Kullback-Leibler divergence between the actual fractions of bookings and the predicted fractions of bookings given by different choice models for different months.

	1	2	3	4
Spiked-MNL	**0.098**	**0.142**	**0.075**	0.127
MNL	0.299	0.448	0.816	0.291
RF-1	0.428	0.614	0.838	0.272
RF-2	0.322	0.443	0.481	0.139
RF-3	0.265	0.192	0.271	**0.114**

(a) when fare classes 1–8 are open

	1	2	3	4
Spiked-MNL	0.129	0.107	**0.137**	**0.068**
MNL	0.702	0.772	0.639	0.743
RF-1	0.279	0.468	0.464	0.410
RF-2	**0.038**	0.184	0.310	0.397
RF-3	0.067	**0.085**	0.147	0.098

(b) when fare classes 1–7 are open

8.2.3. Comparison of Different Revenue Management Policies. We compare the performance of various RM policies using the airline data. We calibrated demand models with the booking data for the origin-destination market mentioned above, for itineraries with Monday departures in year 1. Then we used the calibrated demand models to derive various RM policies. Next we calibrated demand models with the booking data for itineraries with Monday departures in year 2, and we used these demand models in a simulation to evaluate the performance of these RM policies. We used the following performance metric. For a given RM policy ψ, let $E[Z^\psi]$ denote the expected revenue achieved using policy ψ. Since the CDLP optimal value z^{CDLP} is an upper bound for the optimal expected revenue of optimization problem (1), we use the *revenue ratio* $\rho^\psi := \mathsf{E}[Z^\psi]/z^{\mathrm{CDLP}}$ as a metric for evaluating the performance of policy ψ.

We evaluate the following RM policies.

- **EMSR-b**: The nested booking limit heuristic proposed by Belobaba (1989), which is a popular heuristic used in airline reservation systems.
- **SBLP**: The nested booking limit heuristic proposed in Section 7, where the booking limits are obtained from the optimal solution of the SBLP.
- **CDLP**: This policy offers specified assortments for fractions of time specified by an optimal solution of the CDLP.

Note that EMSR-b and SBLP are both nested booking limit policies. There are two variants of nested booking limit policies, i.e., standard nesting and theft nesting. A detailed discussion on standard versus theft nesting can be found in Talluri and Van Ryzin (2004b) and Haerian et al. (2006). We implemented both variants for both EMSR-b and SBLP, and we use "-s" and "-t" to distinguish them.

Revenue performance of different policies. Figure 7 shows the revenue ratios ρ^ψ of the policies, with their 95% confidence intervals, obtained with 100 simulation runs. The CDLP policy has the best average performance among the policies tested. The SBLP policies perform slightly worse, but they outperform the EMSR-b policies widely used in the airline industry, capturing an additional 2–4% revenue.

Robustness of different policies. Next, we perturb the model parameter values to evaluate the robustness of different policies. We follow the approach proposed in Liu and Van Ryzin (2008) and evaluate the policies under different load factors. Specifically, we scale the capacity by a factor $k_1 \in \{0.8, 1.0, 1.2\}$ and the no-purchase attractiveness by a factor $k_2 \in \{0.8, 1.0, 1.2\}$. Table 8 reports the revenue ratios ρ^ψ under these perturbations, averaged over 100 simulation runs. In most cases the SBLP policies perform better than the EMSR-b policies, and the performance of the SBLP policies is less variable than the performance of the CDLP policy.

Figure 7 95% confidence intervals of ρ^ψ over a real-world dataset under different policies.

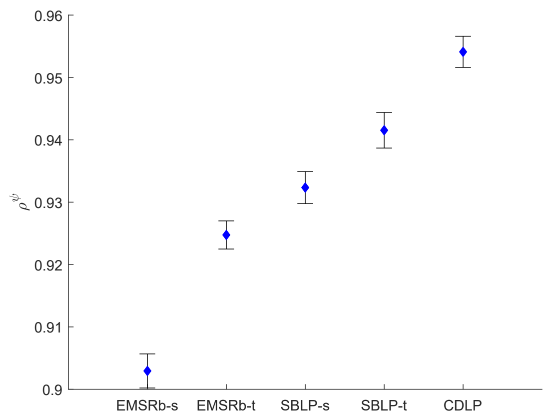


Table 8 Revenue ratios ρ^ψ under different capacity and null attractiveness scaling.

k_1	k_2	EMSRb-s	EMSRb-t	SBLP-s	SBLP-t	CDLP
0.8	0.8	0.929	0.941	0.928	0.939	0.888
	1.0	0.911	0.925	0.920	0.929	0.916
	1.2	0.894	0.915	0.916	0.924	0.930
1.0	0.8	0.909	0.932	0.929	0.943	0.955
	1.0	0.903	0.925	0.932	0.942	0.956
	1.2	0.905	0.921	0.934	0.938	0.958
1.2	0.8	0.882	0.911	0.909	0.929	0.961
	1.0	0.886	0.907	0.911	0.924	0.965
	1.2	0.895	0.907	0.917	0.921	0.967

9. Conclusion

The spiked-MNL model is primarily motivated by the empirical observation that a very large fraction of customers who book an airline ticket for a particular itinerary, choose the cheapest available fare class for that itinerary, and that this behavior is not captured well by familiar choice models such as the MNL model. The spiked-MNL model was proposed as a choice model that captures this cheapest spike phenomenon and that has many of the desirable properties of the MNL model, such as tractability of estimation and assortment optimization.

In this paper we considered a network revenue management problem under the spiked-MNL choice model. We showed that efficient assortments under the spiked-MNL model are nested-by-revenue assortments. This property was used to obtain a more concise sales-based linear programming (SBLP) formulation under the spiked-MNL model. We proposed a booking limit policy that is based on an optimal solution of the SBLP. We showed that this booking limit policy is asymptotically optimal under fluid scaling. Unlike familiar time-based policies obtained from the CDLP, the booking limit policy obtained from the SBLP results in a random sequence of assortments, which required us to develop a new technique to prove asymptotic optimality.

Acknowledgments

The authors thank the Area Editor, the Associate Editor and two anonymous referees for their constructive comments that improved our paper substantially.

References

Aouad A, Farias VF, Levi R (2015) Assortment optimization under consider-then-choose choice models. *available at SSRN 2618823* .

Belobaba PP (1989) OR practice—Application of a probabilistic decision model to airline seat inventory control. *Operations Research* 37(2):183–197.

Berbeglia G, Joret G (2020) Assortment optimisation under a general discrete choice model: A tight analysis of revenue-ordered assortments. *Algorithmica* 82(4):681–720.

Bertsimas D, De Boer S (2005) Simulation-based booking limits for airline revenue management. *Operations Research* 53(1):90–106.

Bertsimas D, Mišic VV (2015) Data-driven assortment optimization. Working paper, Operations Research Center, MIT.

Bertsimas D, Popescu I (2003) Revenue management in a dynamic network environment. *Transportation Science* 37(3):257–277.

Boyd EA, Kallesen R (2004) Practice papers: The science of revenue management when passengers purchase the lowest available fare. *Journal of Revenue and Pricing Management* 3(2):171–177.

Bront JJM, Méndez-Díaz I, Vulcano G (2009) A column generation algorithm for choice-based network revenue management. *Operations Research* 57(3):769–784.

Cooper WL (2002) Asymptotic behavior of an allocation policy for revenue management. *Operations Research* 50(4):720–727.

Cooper WL, Homem-de Mello T, Kleywegt AJ (2006) Models of the spiral-down effect in revenue management. *Operations Research* 54(5):968–987.

Cooper WL, Li L (2012) On the use of buy up as a model of customer choice in revenue management. *Production and Operations Management* 21(5):833–850.

Dai J, Ding W, Kleywegt A, Wang X, Zhang Y (2014) Choice based revenue management for parallel flights. Working paper, available at SSRN 2404193.

Davis JM, Gallego G, Topaloglu H (2014) Assortment optimization under variants of the nested logit model. *Operations Research* 62(2):250–273.

Ding W (2017) *Estimation and Optimization Problems in Revenue Management with Customer Choice Behavior*. Ph.D. thesis, Georgia Institute of Technology.

Draief M, Massouli L (2010) *Epidemics and Rumours in Complex Networks* (Cambridge University Press, Cambridge, United Kindom).

Farias VF, Jagabathula S, Shah D (2013) A nonparametric approach to modeling choice with limited data. *Management Science* 59(2):305–322.

Feldman JB, Topaloglu H (2015) Capacity constraints across nests in assortment optimization under the nested logit model. *Operations Research* 63(4):812–822.

Feldman JB, Topaloglu H (2017) Revenue management under the Markov chain choice model. *Operations Research* 65(5):1322–1342.

Gallego G, Iyengar G, Phillips R, Dubey A (2004) Managing flexible products on a network. Working paper, available at SSRN 3567371.

Gallego G, Li L, Ratliff R (2009) Choice-based emsr methods for single-leg revenue management with demand dependencies. *Journal of Revenue and Pricing Management* 8(2-3):207–240.

Gallego G, Ratliff R, Shebalov S (2015) A general attraction model and sales-based linear program for network revenue management under customer choice. *Operations Research* 63(1):212–232.

Golrezaei N, Nazerzadeh H, Rusmevichientong P (2014) Real-time optimization of personalized assortments. *Management Science* 60(6):1532–1551.

Haerian L, Homem-de Mello T, Mount-Campbell CA (2006) Modeling revenue yield of reservation systems that use nested capacity protection strategies. *International Journal of Production Economics* 104(2):340–353.

Hübner AH, Kuhn H (2012) Retail category management: State-of-the-art review of quantitative research and software applications in assortment and shelf space management. *Omega* 40(2):199–209.

Keller PW, Levi R, Perakis G (2014) Efficient formulations for pricing under attraction demand models. *Mathematical Programming* 145:223–261.

Kök AG, Fisher ML, Vaidyanathan R (2015) Assortment planning: Review of literature and industry practice. *Retail Supply Chain Management*, 175–236 (Springer).

Lhéritier A, Bocamazo M, Delahaye T, Acuna-Agost R (2019) Airline itinerary choice modeling using machine learning. *Journal of Choice Modelling* 31:198–209.

Littlewood K (1972) Forecasting and control of passenger bookings. *Airline Group International Federation of Operational Research Societies Proceedings, 1972* 12:95–117.

Liu Q, Van Ryzin G (2008) On the choice-based linear programming model for network revenue management. *Manufacturing & Service Operations Management* 10(2):288–310.

Luce RD (1959) *Individual Choice Behavior* (John Wiley, New York, NY).

Manski CF, McFadden D, et al. (1981) *Structural Analysis of Discrete Data with Econometric Applications* (MIT Press, Cambridge, MA).

McFadden D (2001) Economic choices. *American Economic Review* 91(3):351–378.

McGill JI, Van Ryzin GJ (1999) Revenue management: Research overview and prospects. *Transportation Science* 33(2):233–256.

Pompilio L, Kacelnik A (2010) Context-dependent utility overrides absolute memory as a determinant of choice. *Proceedings of the National Academy of Sciences* 107(1):508–512.

Rooderkerk RP, Van Heerde HJ, Bijmolt TH (2011) Incorporating context effects into a choice model. *Journal of Marketing Research* 48(4):767–780.

Rusmevichientong P, Shen ZJM, Shmoys DB (2010) Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Operations Research* 58(6):1666–1680.

Rusmevichientong P, Shmoys D, Tong C, Topaloglu H (2014) Assortment optimization under the multinomial logit model with random choice parameters. *Production and Operations Management* 23(11):2023–2039.

Rusmevichientong P, Topaloglu H (2012) Robust assortment optimization in revenue management under the multinomial logit choice model. *Operations Research* 60(4):865–882.

Strauss AK, Klein R, Steinhardt C (2018) A review of choice-based revenue management: Theory and methods. *European Journal of Operational Research* 271(2):375–387.

Talluri K (2014) New formulations for choice network revenue management. *INFORMS Journal on Computing* 26(2):401–413.

Talluri K, Van Ryzin G (1998) An analysis of bid-price controls for network revenue management. *Management Science* 44(11-part-1):1577–1593.

Talluri K, Van Ryzin G (2004a) Revenue management under a general discrete choice model of consumer behavior. *Management Science* 50(1):15–33.

Talluri KT, Van Ryzin GJ (2004b) *The Theory and Practice of Revenue Management*. International Series in Operations Research & Management Science (Kluwer Academic Publishers).

Train KE (2009) *Discrete Choice Methods with Simulation* (Cambridge University Press, Cambridge, United Kingdom).

Tversky A, Simonson I (1993) Context-dependent preferences. *Management Science* 39(10):1179–1189.

van Ryzin G, Vulcano G (2008) Computing virtual nesting controls for network revenue management under customer choice behavior. *Manufacturing & Service Operations Management* 10(3):448–467.

Walczak D, Mardan S, Kallesen R (2010) Customer choice, fare adjustments and the marginal expected revenue data transformation: A note on using old yield management techniques in the brave new world of pricing. *Journal of Revenue and Pricing Management* 9(1-2):94–109.

Wang R (2018) When prospect theory meets consumer choice models: Assortment and pricing management with reference prices. *Manufacturing & Service Operations Management* 20(3):583–600.

Zhang D, Adelman D (2009) An approximate dynamic programming approach to network revenue management with customer choice. *Transportation Science* 43(3):381–394.

Zhang D, Cooper WL (2005) Revenue management for parallel flights with customer-choice behavior. *Operations Research* 53(3):415–431.

Yufeng Cao is an assistant professor in the Antai College of Economics and Management at Shanghai Jiao Tong University. His research interests include revenue management and marketplace analytics.

Anton J. Kleywegt is an associate professor in the School of Industrial and Systems Engineering at Georgia Tech. His research focuses on optimization and stochastic modeling with applications in transportation, distribution, and logistics.

He Wang is Colonel John B. Day Early Career Professor and an assistant professor in the School of Industrial and Systems Engineering at Georgia Tech. His research focuses on the interface between machine learning and operations management. He is interested in data-driven methods for applications in supply chain, pricing, and transportation.

Online Appendix

Appendix A: Properties of the Spiked-MNL Model

There are several differences between the fundamental properties of the MNL model and the spiked-MNL model. Here we explore some of these properties of the spiked-MNL model.

A.1. Regularity.

The regularity property means that the probability of choosing any alternative, including the null alternative, from an assortment does not increase if the assortment is enlarged (Manski et al. 1981). More formally, the definition of a regular choice model is as follows.

Definition EC.1. A choice model P is *regular* if for any two assortments S_1 and S_2 satisfying $S_1 \subset S_2 \subset \mathcal{J}$ and any alternative $j \in S_1 \cup \{0\}$, it holds that $P_{j:S_1} \geq P_{j:S_2}$.

Regularity is a property commonly held by choice models used in the assortment optimization literature (see, e.g., Golrezaei et al. 2014, Berbeglia and Joret 2020). It is easily verified that the MNL choice model is regular. However, the spiked-MNL choice model is not necessarily regular, as shown by the following example:

Example EC.1. A seller sells three products H, M, and L with revenues $r_H > r_M > r_L$. The attractiveness parameters of these products are $v_H = v_M = w_L = 1$ and $w_M = 8$ (we don't need to specify v_L or w_H in this example), and the null attractiveness is $v_0 = 1$. Then $P_{H:\{H,M\}} = v_H/(v_H + w_M + v_0) = 1/10$ and $P_{H:\{H,M,L\}} = v_H/(v_H + v_M + w_L + v_0) = 1/4$, which violates the regularity property.

The following necessary and sufficient condition can be used to check whether a spiked-MNL model is regular, or to enforce regularity when estimating a spiked-MNL model.

Proposition EC.1. *The spiked-MNL model is regular if and only if for any two products j and j' for the same itinerary, with j' more expensive than j, i.e., for any $j' \in J(j)$, it holds that*

$$w_j + v_{j'} \quad \geq \quad w_{j'}.$$

Remark EC.1. The complexity of checking the regularity of a spiked-MNL model is $O(\sum_{g \in \mathcal{G}} n(g)^2)$.

A.2. Submodularity.

Given a choice model P, let the *demand function* $d : 2^{\mathcal{J}} \mapsto \mathbb{R}$ of the choice model be given by $d(S) := \sum_{j \in S} P_{j:S}$ for any assortment $S \subset \mathcal{J}$. Another property of many choice models is the submodularity of their demand functions, which means that the marginal increment in total choice probability decreases as the assortment enlarges (Berbeglia and Joret 2020). More formally, the definition of a submodular demand function is as follows.

Definition EC.2. The demand function d of a choice model is *submodular* if

$$d(S_2 \cup \{k\}) - d(S_2) \quad \leq \quad d(S_1 \cup \{k\}) - d(S_1), \quad \forall \ S_1 \subset S_2 \subset \mathcal{J}, \ k \in \mathcal{J} \setminus S_2. \tag{EC.1}$$

The demand function of the MNL choice model is submodular, but the demand function of the spiked-MNL choice model is not necessarily submodular, as shown by the following example:

Example EC.2. A seller sells three products H, M, and L with revenues $r_H > r_M > r_L$. Let the attractiveness parameters of the products be $v_H = 1$, $w_H = 3$, and $v_M = w_M = w_L = 2$ (we don't need to specify v_L); and let the null attractiveness be $v_0 = 1$. Consider set $S_1 = \{H\}$, set $S_2 = \{H, L\}$, and product $k = M$. Then $d(S_2 \cup \{k\}) - d(S_2) = d(\{H, M, L\}) - d(\{H, L\}) = 5/6 - 3/4 = 1/12$, and $d(S_1 \cup \{k\}) - d(S_1) = d(\{H, M\}) - d(\{H\}) = 3/4 - 3/4 = 0$. Therefore, the demand function is not submodular.

Berbeglia and Joret (2020) showed that any random utility model with context-independent utilities is equivalent to a stochastic preference model and has a submodular demand function. The fact that a spiked-MNL model may not have a submodular demand function implies that it is not always representable by a random utility model with context-independent utilities or a stochastic preference model.

A.3. Representation by Context-Dependent Utility Models

The results in Section A.2 imply that the spiked-MNL model cannot always be represented by a random utility model with context-independent utilities. In this section we consider a more general family of random utility models known as *context-dependent* utility models. This family of utility models was introduced to describe empirical observations of "context effects," that is, the phenomenon that the relative attractiveness of alternatives depends on the presence of other alternatives (Tversky and Simonson 1993, Pompilio and Kacelnik 2010, Rooderkerk et al. 2011, Wang 2018). Specifically, in a context-dependent utility model the utility of an alternative j given choice set S is represented by (see e.g. Rooderkerk et al. 2011, Equation (1))

$$U_{j:S} \quad = \quad V_j + V_{j:S} + \epsilon_{j:S},$$

where V_j and $V_{j:S}$ are the context-independent and the context-dependent parts of the deterministic component of the utility, respectively, and $\epsilon_{j:S}$ is the random error term. In particular, $V_{j:S}$ can depend on the attributes of the other alternatives in S.

Next we show that the spiked-MNL model can be represented by a context-dependent utility model. Consider any spiked-MNL model with parameters $v_0 > 0$, and $v_j, w_j > 0$ for $j \in \mathcal{J}$, and with i.i.d. Gumbel distributed error terms ε_j for $j \in \mathcal{J}$. Let $V_0 := \log(v_0)$, $V_j := \log(v_j)$, $V_{j:S} := \big[\log(w_j) -$

$\log(v_j)]\mathbb{1}(j,S)$, and $\epsilon_{j:S} := \varepsilon_j$ for $j \in \mathcal{J}$. Then, the probability that a customer chooses alternative j from a choice set S is given by

$$\begin{aligned} P_{j:S} &= \frac{v_j(1-\mathbb{1}(j,S)) + w_j\mathbb{1}(j,S)}{\sum_{j'\in S}[v_{j'}(1-\mathbb{1}(j',S)) + w_{j'}\mathbb{1}(j',S)] + v_0} \\ &= \frac{v_j + (w_j - v_j)\mathbb{1}(j,S)}{\sum_{j'\in S}[v_{j'} + (w_{j'} - v_{j'})\mathbb{1}(j',S)] + v_0} \\ &= \frac{\exp(V_j + V_{j:S})}{\sum_{j'\in S}\exp(V_{j'} + V_{j':S}) + \exp(V_0)} \end{aligned}$$

that is, the choice probability $P_{j:S}$ is the same as the choice probability given by a random utility model with context-dependent utilities $U_{j:S} = V_j + V_{j:S} + \epsilon_{j:S}$, where $\epsilon_{j:S} = \varepsilon_j$, $j \in \mathcal{J}$, are i.i.d. Gumbel distributed error terms.

A.4. Revenue Gap between the MNL and the Spiked-MNL Choice Models

Next we show that using an MNL model instead of a spiked-MNL model in the presence of the cheapest-fare spike effect can lead to arbitrarily bad relative revenue performance. More specifically, we consider a sequence of selling seasons, indexed by $k = 1, 2, 3, \ldots$. At the end of each selling season k, the available data are used to calibrate a choice model, and then an assortment that is optimal for the calibrated model is offered during the next selling season. Let R_{MNL} denote the long-run expected revenue per selling season when calibrating an MNL demand model with data and choosing an optimal assortment for the calibrated model, and let R_{SMNL} denote the long-run expected revenue per selling season when calibrating a spiked-MNL demand model with data and choosing an optimal assortment for the calibrated model. We provide an example such that for any $\varepsilon \in (0,1)$, the long-run loss ratio is

$$\text{Loss} \;=\; \frac{R_{\text{SMNL}} - R_{\text{MNL}}}{R_{\text{SMNL}}} \;\geq\; 1 - \varepsilon.$$

Setting. A seller can offer two products H and L with no capacity limits. The products H and L sell at prices r_H and $r_L > 0$ respectively, with $0 < r_L/r_H \leq \varepsilon$. In each season k, the seller offers one assortment $A^{(k)} \subset \{H, L\}$. As shown in Section 5, every efficient assortment under the spiked-MNL model (and the MNL model) is nested-by-revenue, and therefore it suffices to consider either $\{H\}$ or $\{H, L\}$ for $A^{(k)}$. Customers make choices according to a spiked-MNL model with parameters $v_0 = 1$, $v_H = 0$, and

$$0 \;<\; w_H \;=\; w_L \;<\; \eta \;:=\; \frac{r_L}{r_H - r_L}.$$

Note that $v_H = 0$ corresponds to the so-called 100% buydown effect, where customers buy only product L when both H and L are offered. It follows that $r_H w_H/(v_0 + w_H) > (r_H v_H + r_L w_L)/(v_0 + v_H + w_L)$, and thus it is optimal to offer assortment $\{H\}$. As process primitives, consider the following 4 independent, i.i.d. sequences of random variables: $N_1^{(k)}$ with mean $\lambda w_H/(v_0 + w_H)$, representing the number of customers who would choose H in season k if assortment $\{H\}$ is offered; $N_2^{(k)}$ with mean $\lambda v_0/(v_0+w_H)$, representing the number of customers who would choose 0 in season k if assortment $\{H\}$ is offered; $N_3^{(k)}$ with mean $\lambda w_L/(v_0+v_H+w_L)$, representing the number of customers who would choose L in season k if assortment $\{H,L\}$ is offered; $N_4^{(k)}$ with mean $\lambda v_0/(v_0+v_H+w_L)$, representing the number of customers who would choose 0 in season k if assortment $\{H,L\}$ is offered.

Dynamics. After each season, the revenue manager calibrates an MNL model using maximum likelihood estimation (MLE) with all historical sales data (including no-purchase customers), and decides which assortment to offer in the next season based on the estimated MNL model. The MNL model is specified with attractiveness parameters $\tilde{v}_H$, $\tilde{v}_L$, and $v_0=1$. For each season k, let $n_H^{(k)}$ denote the sales of product H and let $n_0^{(k)}$ denote the number of customers who choose not to purchase. One can show the following:

- Regardless of the assortments offered, the MLE estimated attractiveness parameter of product H is given by $\tilde{v}_H^{(k)}=\sum_{k'=1}^k n_H^{(k')}/\sum_{k'=1}^k n_0^{(k')}$. (To deal with the possibility that the denominator may be zero for the first few seasons, assume that the denominator is set to 1 if $\sum_{k'=1}^k n_0^{(k')}=0$.)
- According to the estimated MNL model, it is optimal to offer assortment $\{H\}$ if $\tilde{v}_H^{(k)}\geq\eta$; otherwise, it is optimal to offer assortment $\{H,L\}$.

Next, we show that when using the MNL model, the offered assortment converges to $\{H,L\}$ w.p.1. The long-run revenue loss ratio under this assortment is greater than $1-\varepsilon$. We consider two cases.

Case 1: The revenue manager offers assortment $A^{(1)}=\{H,L\}$ in the first season. Then, due to 100% buydown, it follows that $n_H^{(1)}=0$. Thus, the estimated parameter $\tilde{v}_H^{(1)}=n_H^{(1)}/n_0^{(1)}=0<\eta$. It follows by induction that the revenue manager will offer assortment $\{H,L\}$ and $\tilde{v}_H^{(k)}=0$ for all k.

Case 2: The revenue manager offers assortment $A^{(1)}=\{H\}$ in the first season. Note that, if in any season k it holds that $\tilde{v}_H^{(k)}<\eta$, then the revenue manager will offer assortment $A^{(k+1)}=\{H,L\}$ in season $k+1$. Subsequently, it follows that $n_H^{(k+1)}=0$ and thus $\tilde{v}_H^{(k+1)}=\sum_{k'=1}^{k+1}n_H^{(k')}/\sum_{k'=1}^{k+1}n_0^{(k')}=\left(\sum_{k'=1}^k n_H^{(k')}+0\right)/\left(\sum_{k'=1}^k n_0^{(k')}+N_4^{(k+1)}\right)\leq\sum_{k'=1}^k n_H^{(k')}/\sum_{k'=1}^k n_0^{(k')}=\tilde{v}_H^{(k)}<\eta$, and it follows by induction that $A^{(k')}=\{H,L\}$ for all $k'>k$. Thus, either there is a K such that $A^{(k)}=\{H,L\}$ for all $k>K$, or $\tilde{v}_H^{(k)}\geq\eta$ and $A^{(k)}=\{H\}$ for all k. Next we show that the event that $\tilde{v}_H^{(k)}\geq\eta$ and $A^{(k)}=\{H\}$ for all k has probability 0. Note that if $A^{(k)}=\{H\}$ for all k, then $n_H^{(k)}=N_1^{(k)}$ and $n_0^{(k)}=N_2^{(k)}$ for all k. By the Strong Law of Large Numbers, w.p.1, $\sum_{k'=1}^k N_1^{(k')}/k\to\lambda w_H/(v_0+w_H)$ and $\sum_{k'=1}^k N_2^{(k')}/k\to\lambda v_0/(v_0+w_H)$ as $k\to\infty$. Thus, if $A^{(k)}=\{H\}$ for all k, then, except for a subset B with probability 0, it holds that $\sum_{k'=1}^k n_H^{(k')}/k=\sum_{k'=1}^k N_1^{(k')}/k\to\lambda w_H/(v_0+w_H)$ and $\sum_{k'=1}^k n_0^{(k')}/k=\sum_{k'=1}^k N_2^{(k')}/k\to\lambda v_0/(v_0+w_H)$ as $k\to\infty$, and hence

$$\tilde{v}_H^{(k)} \;=\; \frac{\sum_{k'=1}^k n_H^{(k')}}{\sum_{k'=1}^k n_0^{(k')}} \;\to\; \frac{w_H}{v_0} \;<\; \eta.$$

Therefore, the event that $\tilde{v}_H^{(k)}\geq\eta$ and $A^{(k)}=\{H\}$ for all k is contained in the subset B and thus has probability 0.

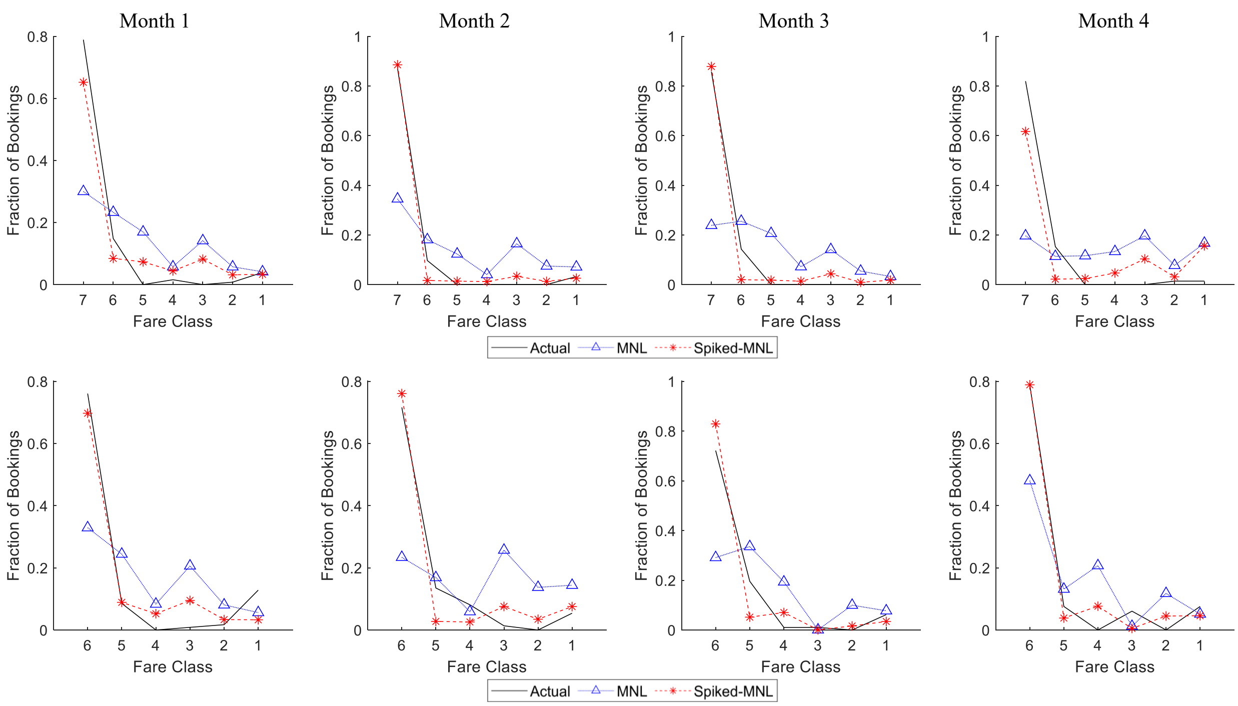
Loss Ratio. In both cases above, the long-run loss ratio is

$$\text{Loss} \;=\; \frac{R_{\text{SMNL}} - R_{\text{MNL}}}{R_{\text{SMNL}}} \;=\; \frac{R(\{H\}) - R(\{H,L\})}{R(\{H\})} \;=\; 1 - \frac{r_L}{r_H} \;\geq\; 1 - \varepsilon.$$

Appendix B: Additional Examples of the Spike Effect

Figure EC.1 shows more examples of the spike effect and how the MNL and the spiked-MNL models predict. The figures are similar to those in Figure 2 but are based on airline booking data for another origin-destination market.

Figure EC.1 Actual fraction of bookings and predicted fraction of bookings under the MNL model and the spiked-MNL model.



Note. The fare classes shown on the horizontal axis are the available fare classes for the itinerary when the booking data were recorded. “Actual” shows the actual fraction of bookings in each available fare class, “MNL” shows the fraction of bookings in each available fare class predicted by a MNL choice model calibrated with booking data, and “Spiked-MNL” shows the fraction of bookings in each available fare class predicted by a spiked-MNL model calibrated with the same data.

Appendix C: Assortment Optimization

In this section we consider assortment optimization under the spiked-MNL model. Let $\mathcal{N} := \{1,2,\ldots,n\}$ denote the seller's set of products. The products are indexed such that $r_1 > r_2 > \cdots > r_n$. A seller selects an assortment $S \subset \mathcal{N}$ to offer to customers to maximize the seller's expected revenue per customer. Each customer chooses a product or the no-purchase alternative according to the spiked-MNL model. Recall that each product $i \in \mathcal{N}$ is associated with two parameters, a standard attractiveness v_i and a spiked attractiveness w_i. In this setting, if the seller offers anything then there is exactly one cheapest available product in the assortment, and the attractiveness of the cheapest available product is spiked.

The seller's assortment optimization problem is formulated as follows:

$$R^* \;:=\; \max_{S \subset \mathcal{N}} \left\{ R(S) \;:=\; \frac{\sum_{i\in S} r_i \left[v_i(1-\mathbb{1}(i,S)) + w_i \mathbb{1}(i,S)\right]}{v_0 + \sum_{i\in S}\left[v_i(1-\mathbb{1}(i,S)) + w_i\mathbb{1}(i,S)\right]} \right\} \qquad \text{(Assortment)}$$

Characterization and Efficient Computation of Optimal Assortments

Note that problem Assortment has 2^n decision variables. First consider the general case in which it is not required that $w_i \geq v_i$ for all $i \in \mathcal{N}$. We show that Assortment can be solved by solving a linear program with $O(n^2)$ variables and $O(n^2)$ constraints. Then we consider the structure of the optimal assortment when $w_i \geq v_i$ for all $i \in \mathcal{N}$, and it follows that Assortment can be solved by solving a linear program with $O(n)$ variables and 2 constraints.

Using the decision variables $z_0 \in \mathbb{R}$ and $\mathbf{z} := (z_i^k \in \mathbb{R}_+,\, i,k \in \mathcal{N}$ and $i \leq k)$, consider the linear program

$$R^*_{LP} \;:=\; \max_{(z_0,\mathbf{z})\in\mathbb{R}_+\times\mathbb{R}_+^{n(n+1)/2}} \left\{ R_{LP}(z_0,\mathbf{z}) \;:=\; \sum_{k\in\mathcal{N}} \left(\sum_{i=1}^{k-1} r_i v_i z_i^k + r_k w_k z_k^k \right) \right\} \qquad \text{(Assortment LP)}$$

$$\text{s.t.}\;\; v_0 z_0 + \sum_{k\in\mathcal{N}} \left(\sum_{i=1}^{k-1} v_i z_i^k + w_k z_k^k \right) \;=\; 1 \qquad \text{(EC.2a)}$$

$$\sum_{k\in\mathcal{N}} z_k^k \;\leq\; z_0 \qquad \text{(EC.2b)}$$

$$z_i^k \;\leq\; z_k^k \qquad \forall\, i,k \in \mathcal{N},\, i<k \qquad \text{(EC.2c)}$$

We will show that an optimal solution of Assortment LP can be converted to an optimal solution of Assortment in $O(n)$ steps. First we show that any feasible solution of Assortment can be converted to a feasible solution of Assortment LP with the same objective value. Let $W(S) := \sum_{i\in S}\left[v_i(1-\mathbb{1}(i,S)) + w_i\mathbb{1}(i,S)\right]$. Given a feasible solution $\hat{S} \subset \mathcal{N}$ of Assortment, let $\hat{z}_0 = \frac{1}{v_0 + W(\hat{S})}$, and $\hat{z}_i^k = \mathbb{1}(k,\hat{S})\mathbb{1}\{i \in \hat{S}\}\hat{z}_0$ for all $i,k \in \mathcal{N},\ i \leq k$. Then $\sum_{k\in\mathcal{N}}\left(\sum_{i=1}^{k-1} v_i \hat{z}_i^k + w_k \hat{z}_k^k\right) = \sum_{k\in\mathcal{N}} \mathbb{1}(k,\hat{S})\left(\sum_{i=1}^{k-1} v_i \mathbb{1}\{i\in\hat{S}\} + w_k\mathbb{1}\{k\in\hat{S}\}\right)\hat{z}_0 = \sum_{i\in\hat{S}}\left[v_i(1-\mathbb{1}(i,\hat{S})) + w_i\mathbb{1}(i,\hat{S})\right]\hat{z}_0 = W(\hat{S})\hat{z}_0$. Thus $v_0\hat{z}_0 + \sum_{k\in\mathcal{N}}\left(\sum_{i=1}^{k-1} v_i\hat{z}_i^k + w_k\hat{z}_k^k\right) = (v_0 + W(\hat{S}))\hat{z}_0 = 1$, and thus $(\hat{z}_0,\hat{\mathbf{z}})$ satisfies (EC.2a). Also, $\sum_{k\in\mathcal{N}} \hat{z}_k^k = 0$ if $\hat{S} = \varnothing$ and $\sum_{k\in\mathcal{N}}\hat{z}_k^k = \hat{z}_0$ if $\hat{S} \neq \varnothing$, and thus $(\hat{z}_0,\hat{\mathbf{z}})$ satisfies (EC.2b). Also, $\hat{z}_i^k = \mathbb{1}(k,\hat{S})\mathbb{1}\{i\in\hat{S}\}\hat{z}_0 \leq \mathbb{1}(k,\hat{S})\hat{z}_0 = \hat{z}_k^k$, and thus $(\hat{z}_0,\hat{\mathbf{z}})$ satisfies (EC.2c). In addition, the objective value of $(\hat{z}_0,\hat{\mathbf{z}})$ in Assortment LP is

$$R_{LP}(\hat{z}_0,\hat{\mathbf{z}}) \;=\; \sum_{k\in\mathcal{N}} \mathbb{1}(k,\hat{S}) \left(\sum_{i=1}^{k-1} r_i v_i \mathbb{1}\{i\in\hat{S}\} + r_k w_k \mathbb{1}\{k\in\hat{S}\} \right) \hat{z}_0$$

$$= \sum_{i\in\hat{S}} r_i\left[v_i(1-\mathbb{1}(i,\hat{S}))+w_i\mathbb{1}(i,\hat{S})\right]\hat{z}_0$$

$$= \frac{\sum_{i\in\hat{S}} r_i\left[v_i(1-\mathbb{1}(i,\hat{S}))+w_i\mathbb{1}(i,\hat{S})\right]}{v_0+W(\hat{S})} = R(\hat{S}). \tag{EC.3}$$

Thus, any feasible solution $\hat{S}$ of ASSORTMENT can be converted to a feasible solution $(\hat{z}_0,\hat{\mathbf{z}})$ of ASSORTMENT LP such that $R_{LP}(\hat{z}_0,\hat{\mathbf{z}})=R(\hat{S})$, and hence $R^*_{LP}\ge R^*$.

To show that any basic feasible solution of ASSORTMENT LP can be converted to a feasible solution of ASSORTMENT with the same objective value, we first establish the following lemma.

LEMMA EC.1 **(Extreme Point Solutions)**. *Let $(\hat{z}_0,\hat{\mathbf{z}})$ be any basic feasible solution of* ASSORTMENT LP*. Then,*

(1) $\hat{z}_i^k\in\{0,\hat{z}_k^k\}$ for all $i,k\in\mathcal{N}$ with $i<k$;
(2) there is at most one $k\in\mathcal{N}$ such that $\hat{z}_k^k>0$; and
(3) if $\hat{z}_k^k>0$, then $\hat{z}_k^k=\hat{z}_0$.

Proof. We show claim (1) by contradiction. Let $H_k:=\{i\in\mathcal{N}: i<k, \hat{z}_i^k=\hat{z}_k^k\}$ and $M_k:=\{i\in\mathcal{N}: i<k, 0<\hat{z}_i^k<\hat{z}_k^k\}$. Suppose that there exists some k' such that $M_{k'}\neq\varnothing$. Then we construct two distinct feasible solutions $(\tilde{z}_0,\tilde{\mathbf{z}})$ and $(\bar{z}_0,\bar{\mathbf{z}})$ such that $(\hat{z}_0,\hat{\mathbf{z}})=\frac{1}{2}(\tilde{z}_0,\tilde{\mathbf{z}})+\frac{1}{2}(\bar{z}_0,\bar{\mathbf{z}})$. This contradicts the assumption that $(\hat{z}_0,\hat{\mathbf{z}})$ is basic. Let $V(A):=\sum_{i\in A}v_i$. Let $\varepsilon:=\min\left\{\left(\hat{z}_{k'}^{k'}-\max\{\hat{z}_i^{k'}: i\in M_{k'}\}\right)/(v_0+V(H_{k'})+V(M_{k'})+w_{k'}), \min\{\hat{z}_i^{k'}: i\in M_{k'}\}/(v_0+V(H_{k'})+w_{k'})\right\}>0$. Let

$$\tilde{z}_0 = \hat{z}_0-V(M_{k'})\varepsilon$$

$$\tilde{z}_i^k = \begin{cases}\hat{z}_k^k-V(M_{k'})\varepsilon & \text{if } k=k' \text{ and } i\in H_{k'}\cup\{k'\}\\ \hat{z}_i^k+[v_0+V(H_{k'})+w_{k'}]\varepsilon & \text{if } k=k' \text{ and } i\in M_{k'}\\ \hat{z}_i^k & \text{otherwise.}\end{cases}$$

Then $(\tilde{z}_0,\tilde{\mathbf{z}})$ satisfies (EC.2b) and (EC.2c). To see that it also satisfies (EC.2a), let $g(z_0,\mathbf{z}):=v_0z_0+\sum_{k\in\mathcal{N}}\left(\sum_{i=1}^{k-1}v_iz_i^k+w_kz_k^k\right)$. Then $g(\hat{z}_0,\hat{\mathbf{z}})-g(\tilde{z}_0,\tilde{\mathbf{z}})=v_0V(M_{k'})\varepsilon+\sum_{i\in H_{k'}}v_iV(M_{k'})\varepsilon+w_{k'}V(M_{k'})\varepsilon-\sum_{i\in M_{k'}}v_i(v_0+V(H_{k'})+w_{k'})\varepsilon=0$. So, $(\tilde{z}_0,\tilde{\mathbf{z}})$ also satisfies (EC.2a). Similarly, let

$$\bar{z}_0 = \hat{z}_0+V(M_{k'})\varepsilon$$

$$\bar{z}_i^k = \begin{cases}\hat{z}_k^k+V(M_{k'})\varepsilon & \text{if } k=k' \text{ and } i\in H_{k'}\cup\{k'\}\\ \hat{z}_i^k-[v_0+V(H_{k'})+w_{k'}]\varepsilon & \text{if } k=k' \text{ and } i\in M_{k'}\\ \hat{z}_i^k & \text{otherwise.}\end{cases}$$

Then $(\bar{z}_0,\bar{\mathbf{z}})$ satisfies (EC.2a)–(EC.2c). Moreover, $(\hat{z}_0,\hat{\mathbf{z}})=\frac{1}{2}(\tilde{z}_0,\tilde{\mathbf{z}})+\frac{1}{2}(\bar{z}_0,\bar{\mathbf{z}})$, which contradicts the assumption that $(\hat{z}_0,\hat{\mathbf{z}})$ is basic. Thus $M_{k'}=\varnothing$. Therefore, $\hat{z}_i^k\in\{0,\hat{z}_k^k\}$ for all $i,k\in\mathcal{N}$ with $i<k$.

Next we show claim (2) by contradiction. Suppose that there are distinct $k_1,k_2\in\mathcal{N}$ such that $\hat{z}_{k_1}^{k_1}>0$ and $\hat{z}_{k_2}^{k_2}>0$. It follows from (EC.2b) that $\hat{z}_0>0$. It follows from claim (1) that

$M_{k_1} = M_{k_2} = \varnothing$. Without loss of generality, assume that $V(H_{k_1}) + w_{k_1} \le V(H_{k_2}) + w_{k_2}$. Let $\varepsilon :=$ $\min\left\{\hat{z}_0/[V(H_{k_2})+w_{k_2}-V(H_{k_1})-w_{k_1}],\, \hat{z}_{k_1}^{k_1}/[V(H_{k_2})+w_{k_2}+v_0],\, \hat{z}_{k_2}^{k_2}/[V(H_{k_1})+w_{k_1}+v_0]\right\} > 0$. Let

$$
\begin{aligned}
\tilde{z}_0 &= \hat{z}_0 + [V(H_{k_1})+w_{k_1}-V(H_{k_2})-w_{k_2}]\,\varepsilon \\
\tilde{z}_i^k &= \begin{cases} \hat{z}_k^k - [V(H_{k_2})+w_{k_2}+v_0]\varepsilon & \text{if } k=k_1 \text{ and } i\in H_{k_1}\cup\{k_1\} \\ \hat{z}_k^k + [V(H_{k_1})+w_{k_1}+v_0]\varepsilon & \text{if } k=k_2 \text{ and } i\in H_{k_2}\cup\{k_2\} \\ \hat{z}_i^k & \text{otherwise} \end{cases}
\end{aligned}
$$

and

$$
\begin{aligned}
\bar{z}_0 &= \hat{z}_0 - [V(H_{k_1})+w_{k_1}-V(H_{k_2})-w_{k_2}]\,\varepsilon \\
\bar{z}_i^k &= \begin{cases} \hat{z}_k^k + [V(H_{k_2})+w_{k_2}+v_0]\varepsilon & \text{if } k=k_1 \text{ and } i\in H_{k_1}\cup\{k_1\} \\ \hat{z}_k^k - [V(H_{k_1})+w_{k_1}+v_0]\varepsilon & \text{if } k=k_2 \text{ and } i\in H_{k_2}\cup\{k_2\} \\ \hat{z}_i^k & \text{otherwise.} \end{cases}
\end{aligned}
$$

It follows that $(\tilde{z}_0,\tilde{\mathbf{z}})$ and $(\bar{z}_0,\bar{\mathbf{z}})$ satisfy (EC.2a)–(EC.2c). Moreover, $(\hat{z}_0,\hat{\mathbf{z}}) = \frac{1}{2}(\tilde{z}_0,\tilde{\mathbf{z}}) + \frac{1}{2}(\bar{z}_0,\bar{\mathbf{z}})$, which contradicts the assumption that $(\hat{z}_0,\hat{\mathbf{z}})$ is basic. Therefore, there is at most one $k\in\mathcal{N}$ such that $\hat{z}_k^k > 0$.

Next we show claim (3) by contradiction. Suppose that $0 < \hat{z}_{k'}^{k'} < \hat{z}_0$. It follows from claim (1) that $\hat{z}_i^{k'} = \hat{z}_{k'}^{k'}$ for $i\in H_{k'}$, $\hat{z}_i^{k'} = 0$ for $i\notin H_{k'}\cup\{k'\}$, and $\hat{z}_i^k = 0$ for all $k\in\mathcal{N}\setminus\{k'\}$ and $i\le k$. Let $\varepsilon := \min\left\{(\hat{z}_0-\hat{z}_{k'}^{k'})/[v_0+V(H_{k'})+w_{k'}],\, \hat{z}_{k'}^{k'}/v_0\right\} > 0$. Let

$$
\begin{aligned}
\tilde{z}_0 &= \hat{z}_0 - [V(H_{k'})+w_{k'}]\varepsilon \\
\tilde{z}_i^k &= \begin{cases} \hat{z}_i^k + v_0\varepsilon & \text{if } k=k' \text{ and } i\in H_{k'}\cup\{k'\} \\ \hat{z}_i^k & \text{otherwise} \end{cases}
\end{aligned}
$$

and

$$
\begin{aligned}
\bar{z}_0 &= \hat{z}_0 + [V(H_{k'})+w_{k'}]\varepsilon \\
\bar{z}_i^k &= \begin{cases} \hat{z}_i^k - v_0\varepsilon & \text{if } k=k' \text{ and } i\in H_{k'}\cup\{k'\} \\ \hat{z}_i^k & \text{otherwise} \end{cases}
\end{aligned}
$$

It follows that $(\tilde{z}_0,\tilde{\mathbf{z}})$ and $(\bar{z}_0,\bar{\mathbf{z}})$ satisfy (EC.2a)–(EC.2c). Moreover, $(\hat{z}_0,\hat{\mathbf{z}}) = \frac{1}{2}(\tilde{z}_0,\tilde{\mathbf{z}}) + \frac{1}{2}(\bar{z}_0,\bar{\mathbf{z}})$, which contradicts the assumption that $(\hat{z}_0,\hat{\mathbf{z}})$ is basic. Therefore, if $\hat{z}_k^k > 0$, then $\hat{z}_k^k = \hat{z}_0$. □

Theorem EC.1 **(LP Formulation)**. *Consider any basic optimal solution* $(z_0^*,\mathbf{z}^*)$ *of* Assortment LP. *Then* $S^* := \{i\in\mathcal{N} : \hat{z}_i^k > 0$ *for some* $k\in\mathcal{N}\}$ *is an optimal solution of* Assortment.

Proof. It follows from Lemma EC.1 that either $\mathbf{z}^* = 0$ or there is one $k'\in\mathcal{N}$ such that $z_{k'}^{*k'} = z_0^*$, $z_i^{*k'} \in \{0, z_0^*\}$ for all $i<k'$, and $z_i^{*k} = 0$ for all $k\neq k'$ and $i\le k$.

If $\mathbf{z}^* = 0$, then $S^* = \varnothing$, and thus $R(S^*) = 0 = R_{LP}(z_0^*,\mathbf{z}^*) = R_{LP}^*$. Since $R_{LP}^* \ge R^*$, it follows that S^* is an optimal solution of Assortment.

Otherwise, $S^* = H_{k'} \cup \{k'\}$, where $H_{k'}$ is as defined in the proof of Lemma EC.1. It follows from constraint (EC.2a) that $1 = v_0 z_0^* + \sum_{i\in H_{k'}} v_i z_i^{*k'} + w_{k'} z_{k'}^{*k'} = v_0 z_0^* + \sum_{i\in H_{k'}} v_i z_0^* + w_{k'} z_0^* = v_0 z_0^* + \sum_{i\in S^*} [v_i(1-\mathbb{1}(i,S^*)) + w_i \mathbb{1}(i,S^*)]\, z_0^* = [v_0 + W(S^*)] z_0^*$, and thus $z_0^* = \frac{1}{v_0+W(S^*)} = z_i^{*k'}$ for all $i \in S^*$, and $\mathbb{1}(k',S^*) = 1$. Hence

$$\begin{aligned} R_{LP}^* &= R_{LP}(z_0^*, \mathbf{z}^*) = \sum_{i\in H_{k'}} r_i v_i z_i^{*k'} + r_{k'} w_{k'} z_{k'}^{*k'} \\ &= \frac{\sum_{i\in H_{k'}} r_i v_i + r_{k'} w_{k'}}{v_0 + W(S^*)} = \frac{\sum_{i\in S^*} r_i\, [v_i(1-\mathbb{1}(i,S^*)) + w_i \mathbb{1}(i,S^*)]}{v_0 + W(S^*)} = R(S^*) \end{aligned}$$

As before, since $R_{LP}^* \geq R^*$, it follows that S^* is an optimal solution of Assortment. □

Characterization of an Optimal Assortment

First we show that in general an optimal assortment is almost nested by revenue, but that an optimal assortment is not nested by revenue in general. Then we show that in typical practical settings with positive spikes, i.e., with $w_i \geq v_i$ for all $i \in \mathcal{N}$, optimal assortments are nested by revenue. The latter observation enables us to derive a simplified LP formulation for the assortment optimization problem.

Theorem EC.2 **(General Optimal Assortments)**. *An optimal assortment S^* satisfies $S^* = \varnothing$ or $S^* = \{1\}$ or $S^* = \{1,\ldots,i^*\} \cup \{k^*\}$ for some $i^*, k^* \in \mathcal{N}$ with $i^* < k^*$.*

Proof. We show that $S = \{k\}$ with $k > 1$ cannot be optimal, and $S \subset \mathcal{N}$ with $i_2, k \in S$ and $i_1 \in \mathcal{N} \setminus S$ such that $i_1 < i_2 < k$ cannot be optimal.

Consider $\hat{S} = \{\hat{k}\}$ with $\hat{k} > 1$. Then $R(\hat{S}) = r_{\hat{k}} w_{\hat{k}}/(v_0 + w_{\hat{k}})$. If $r_{\hat{k}} < 0$, then $R(\varnothing) = 0 > R(\hat{S})$, and thus $\hat{S}$ cannot be optimal. If $r_{\hat{k}} \geq 0$, then consider $\tilde{S} = \{1,\ldots,\hat{k}\}$. Then

$$R(\tilde{S}) = \frac{\sum_{i=1}^{\hat{k}-1} r_i v_i + r_{\hat{k}} w_{\hat{k}}}{v_0 + \sum_{i=1}^{\hat{k}-1} v_i + w_{\hat{k}}} > \frac{\sum_{i=1}^{\hat{k}-1} r_{\hat{k}} v_i + r_{\hat{k}} w_{\hat{k}}}{v_0 + \sum_{i=1}^{\hat{k}-1} v_i + w_{\hat{k}}} \geq \frac{r_{\hat{k}} w_{\hat{k}}}{v_0 + w_{\hat{k}}} = R(\hat{S})$$

and thus $\hat{S}$ cannot be optimal.

Consider $\hat{S} \subset \mathcal{N}$ with $i_2, \hat{k} \in \hat{S}$ and $i_1 \in \mathcal{N} \setminus \hat{S}$ such that $i_1 < i_2 < \hat{k}$ and $\mathbb{1}(\hat{k}, \hat{S}) = 1$. Let $\hat{z}_0 = \frac{1}{v_0+W(\hat{S})}$, $\hat{z}_i^{\hat{k}} = \mathbb{1}\{i \in \hat{S}\}\hat{z}_0$ for all $i \in \mathcal{N}$, $i \leq \hat{k}$, and $\hat{z}_i^k = 0$ for all $k \in \mathcal{N} \setminus \{\hat{k}\}$ and $i \leq k$. As shown before, $R_{LP}(\hat{z}_0, \hat{\mathbf{z}}) = R(\hat{S})$. If $v_{i_1} \leq v_{i_2}$, then let $\tilde{z}_0 = \tilde{z}_{i_1}^{\hat{k}} = \hat{z}_0$, $\tilde{z}_{i_2}^{\hat{k}} = \hat{z}_0 - v_{i_1}\hat{z}_0/v_{i_2}$, and $\tilde{z}_i^k = \hat{z}_i^k$ for all $i, k \in \mathcal{N}$ such that $(i,k) \neq (i_1, \hat{k})$ and $(i,k) \neq (i_2, \hat{k})$. Then $(\tilde{z}_0, \tilde{\mathbf{z}})$ is feasible for Assortment LP, and $R_{LP}(\tilde{z}_0, \tilde{\mathbf{z}}) - R_{LP}(\hat{z}_0, \hat{\mathbf{z}}) = (r_{i_1} - r_{i_2}) v_{i_1} \hat{z}_0 > 0$, and thus $\hat{S}$ cannot be optimal. If $v_{i_1} > v_{i_2}$, then let $\tilde{z}_0 = \hat{z}_0$, $\tilde{z}_{i_1}^{\hat{k}} = v_{i_2}\hat{z}_0/v_{i_1}$, $\tilde{z}_{i_2}^{\hat{k}} = 0$, and $\tilde{z}_i^k = \hat{z}_i^k$ for all $i, k \in \mathcal{N}$ such that $(i,k) \neq (i_1, \hat{k})$ and $(i,k) \neq (i_2, \hat{k})$. Then $(\tilde{z}_0, \tilde{\mathbf{z}})$ is feasible for Assortment LP, and $R_{LP}(\tilde{z}_0, \tilde{\mathbf{z}}) - R_{LP}(\hat{z}_0, \hat{\mathbf{z}}) = (r_{i_1} - r_{i_2}) v_{i_2} \hat{z}_0 > 0$, and thus $\hat{S}$ cannot be optimal. □

Theorem EC.2 suggests that, instead of considering all possible subsets of the products, it suffices to consider only $O(n^2)$ assortments characterized by a lowest-priced product k^* and a nested-by-revenue set $\{1,\ldots,i^*\}$ with $i^* < k^*$. To see that it is possible to have a nontrivial gap between i^* and k^*, i.e., $k^* - i^* \geq 2$, consider the following example.

EXAMPLE EC.3. A seller sells three products with revenues $r_1 = 5$, $r_2 = 3$, and $r_3 = 2$. The attractiveness parameters are $v_1 = 5$, $v_2 = 10$, $w_1 = 2$, $w_2 = 4$, and $w_3 = 1$ (we don't need to specify v_3); the null attractiveness is $v_0 = 1$. Note that the spike effects are negative in this example. It is not difficult to verify that the optimal assortment for this instance of ASSORTMENT is $S^* = \{1,3\}$.

Next we consider the setting with positive spike effects, that is, $w_i \geq v_i$ for all $i \in \mathcal{N}$. We show that the optimal assortment is nested by revenue, which implies that we only need to examine n candidate assortments in this setting.

THEOREM EC.3 **(Nested-by-Revenue Optimal Assortments)**. *Under Assumption 1, the optimal solution S^* of* ASSORTMENT *is nested-by-revenue; that is, if $i_2 \in S^*$ and $r_{i_1} > r_{i_2}$, then $i_1 \in S^*$.*

Proof. Consider any optimal solution S^* for ASSORTMENT. First we show by contradiction that if $i \in S^*$ then $r_i \geq R(S^*)$. Suppose that $B := \{i \in S^* : r_i < R(S^*)\} \neq \varnothing$. If $S^* = B$, then

$$\begin{aligned}
& R(S^*) \;=\; \frac{\sum_{i \in S^*} r_i \left[v_i(1 - \mathbb{1}(i,S^*)) + w_i \mathbb{1}(i,S^*)\right]}{v_0 + W(S^*)} \;<\; \frac{\sum_{i \in S^*} R(S^*) \left[v_i(1 - \mathbb{1}(i,S^*)) + w_i \mathbb{1}(i,S^*)\right]}{v_0 + W(S^*)} \\
\Rightarrow \quad & R(S^*) \frac{v_0}{v_0 + W(S^*)} \;<\; 0 \quad \Rightarrow \quad R(S^*) \;<\; 0 \;=\; R(\varnothing)
\end{aligned}$$

which contradicts S^* being optimal. Otherwise, $S^* \setminus B \neq \varnothing$. Let $i^* := \max\{i \in S^* \setminus B\}$. Then

$$\begin{aligned}
& R(S^*)\left[v_0 + W(S^*)\right] \;=\; \sum_{i \in S^*} r_i \left[v_i(1 - \mathbb{1}(i,S^*)) + w_i \mathbb{1}(i,S^*)\right] \\
\Rightarrow \quad & R(S^*)\left[v_0 + W(S^*)\right] + R(S^*)\left(w_{i^*} - v_{i^*}\right) \;\leq\; \sum_{i \in S^*} r_i \left[v_i(1 - \mathbb{1}(i,S^*)) + w_i \mathbb{1}(i,S^*)\right] + r_{i^*}\left(w_{i^*} - v_{i^*}\right) \\
\Rightarrow \quad & R(S^*) \;\leq\; \frac{\sum_{i \in S^*} r_i \left[v_i(1 - \mathbb{1}(i,S^*)) + w_i \mathbb{1}(i,S^*)\right] + r_{i^*}\left(w_{i^*} - v_{i^*}\right)}{v_0 + W(S^*) + w_{i^*} - v_{i^*}} \\
& \qquad < \; \frac{\sum_{i \in S^* \setminus (B \cup \{i^*\})} r_i v_i + r_{i^*} w_{i^*}}{v_0 + W(S^* \setminus B) + W(B)} + \frac{\sum_{i \in B} R(S^*) \left[v_i(1 - \mathbb{1}(i,S^*)) + w_i \mathbb{1}(i,S^*)\right]}{v_0 + W(S^* \setminus B) + W(B)} \\
\Rightarrow \quad & R(S^*) \frac{v_0 + W(S^* \setminus B)}{v_0 + W(S^* \setminus B) + W(B)} \;<\; \frac{\sum_{i \in S^* \setminus (B \cup \{i^*\})} r_i v_i + r_{i^*} w_{i^*}}{v_0 + W(S^* \setminus B) + W(B)} \\
\Rightarrow \quad & R(S^*) \;<\; \frac{\sum_{i \in S^* \setminus B} r_i \left[v_i(1 - \mathbb{1}(i,S^* \setminus B)) + w_i \mathbb{1}(i,S^* \setminus B)\right]}{v_0 + W(S^* \setminus B)} \;=\; R(S^* \setminus B)
\end{aligned}$$

which contradicts S^* being optimal. Thus, $B = \varnothing$, and hence if $i \in S^*$ then $r_i \geq R(S^*)$.

Next we show by contradiction that if S^* is optimal and $i_2 \in S^*$ and $r_{i_1} > r_{i_2}$, then $i_1 \in S^*$. Suppose that S^* is optimal and $i_2 \in S^*$ and $r_{i_1} > r_{i_2}$, but $i_1 \notin S^*$. As shown above, $r_{i_2} \geq R(S^*)$, and thus $r_{i_1} > R(S^*)$. Hence

$$R(S^*)\left[v_0 + W(S^*)\right] \;=\; \sum_{i \in S^*} r_i \left[v_i(1 - \mathbb{1}(i,S^*)) + w_i \mathbb{1}(i,S^*)\right]$$

$$\Rightarrow \quad R(S^*)\,[v_0 + W(S^*)] + R(S^*)v_{i_1} \quad < \quad \sum_{i\in S^*} r_i\,[v_i(1-\mathbb{1}(i,S^*)) + w_i\mathbb{1}(i,S^*)] + r_{i_1}v_{i_1}$$

$$\Rightarrow \quad R(S^*) \quad < \quad \frac{\sum_{i\in S^*\cup\{i_1\}} r_i\,[v_i(1-\mathbb{1}(i,S^*)) + w_i\mathbb{1}(i,S^*)]}{v_0 + W(S^*\cup\{i_1\})} \quad = \quad R(S^*\cup\{i_1\})$$

which contradicts S^* being optimal. Therefore, an optimal assortment S^* is nested by revenue. □

Remark EC.2. In the first part of the proof it was shown that an optimal assortment S^* satisfies the inequality $S^* \subset \{i \in \mathcal{N} : r_i \geq R(S^*)\}$. One may conjecture that S^* satisfies the equality $S^* = \{i \in \mathcal{N} : r_i \geq R(S^*)\}$. However, this is not the case under the spiked-MNL model. Consider the following example with two products $\{H, L\}$ such that $r_H = 2$, $r_L = 1.2$ and $v_0 = w_H = 1$, $v_H = w_L = 0.5$. Then the optimal assortment is $S^* = \{H\}$ with $R(S^*) = 1 < r_L$ and $L \notin S^*$. This observation differs from a related result under the classic MNL model (see Rusmevichientong and Topaloglu 2012, Theorem 3.2).

Remark EC.3. Note that the proof of Theorem EC.3 does not require an assumption that $r_i \geq 0$ for all $i \in \mathcal{N}$. Thus, the nested-by-revenue result also holds if a common amount (such as a resource cost) is subtracted from all r_i. The observation suggests that for a revenue management problem that includes resource consumption, the efficient sets, that offer the most favorable trade-off between revenue earned and resource consumption, are nested by revenue. This nested-by-revenue result is generalized to revenue management problems in Section 5.

Compact Assortment LP with Positive Spike Effects

Theorem EC.3 shows that if products exhibit positive spike effects, i.e., $w_i \geq v_i$ for all $i \in \mathcal{N}$, then it suffices to consider nested-by-revenue assortments. Such assortments are specified by their cheapest products. This property can be used to further simplify Assortment LP. Let $x_0 = v_0 z_0$ and $x_k = w_k z_k^k$ for all $k \in \mathcal{N}$. Lemma EC.1 and Theorem EC.3 imply that an optimal solution satisfies $v_i z_i^k = v_i z_k^k = v_i x_k / w_k$ for all $i < k$. That gives the following linear program for assortment optimization:

$$\max_{(x_0,\mathbf{x})\in\mathbb{R}\times\mathbb{R}^n_+} \sum_{k\in\mathcal{N}} \left(r_k + \sum_{i=1}^{k-1} r_i \frac{v_i}{w_k} \right) x_k \qquad \text{(Compact Assortment LP)}$$

$$\text{s.t.} \quad x_0 + \sum_{k\in\mathcal{N}} \left(1 + \sum_{i=1}^{k-1} \frac{v_i}{w_k} \right) x_k \quad = \quad 1 \qquad \text{(EC.4a)}$$

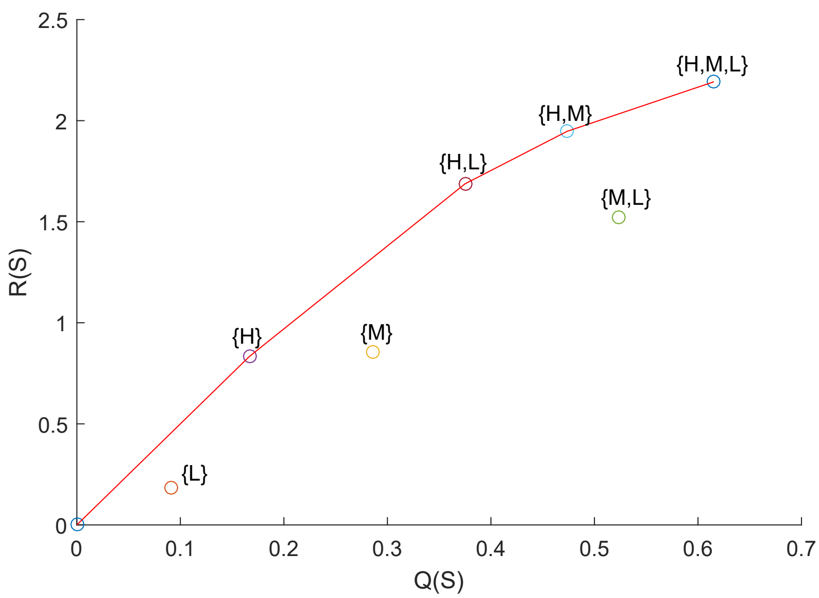
$$\sum_{k\in\mathcal{N}} \frac{x_k}{w_k} \quad \leq \quad \frac{x_0}{v_0} \qquad \text{(EC.4b)}$$

Problem Compact Assortment LP has only $n+1$ decision variables and 2 constraints.

Appendix D: Example of Efficient Assortment that is Not Nested-by-Revenue

Example EC.4. A seller sells three products: H, M, and L, with revenues $r_H = 5$, $r_M = 3$, and $r_L = 2$, using the same resource. The attractiveness parameters are $v_H = 5$, $v_M = 10$, $w_H = 2$, $w_M = 4$, and $w_L = 1$ (we don't need to specify v_L); the null attractiveness is $v_0 = 10$. Note that the spike effect is negative in this example. Figure EC.2 shows the plot of $(Q(S), R(S))$ for different assortments S, and the convex envelope of Pareto-optimal assortments. By Definition 1, the efficient assortments are on the convex envelope. Note that assortment $\{H, L\}$ is on the convex envelope and hence it is an efficient assortment, but it is not nested-by-revenue.

Figure EC.2 Example with $w_j < v_j$ in which efficient assortment $\{H, L\}$ under the spiked-MNL model is not nested-by-revenue.



Appendix E: Proofs

For any alternative $j \in \mathcal{J} \cup \{0\}$ and assortment $A \subset \mathcal{J}$, let

$$\tilde{v}(j, A) \quad := \quad \begin{cases} w_j \mathbb{1}(j, A) + v_j (1 - \mathbb{1}(j, A)) & \text{if } j \in A \\ v_0 & \text{if } j = 0 \\ 0 & \text{otherwise} \end{cases}$$

denote the attractiveness of j when A is offered, and let $W(A) := \sum_{j \in A} \tilde{v}(j, A)$ denote the total attractiveness of A.

E.1. Proof of Proposition EC.1

Proof of Proposition EC.1. (*Sufficiency.*) Suppose that, for any two products j and j' for the same itinerary, with $j' \in J(j)$, it holds that

$$w_j + v_{j'} \quad \geq \quad w_{j'}.$$

We show that the spiked-MNL model is regular, that is, we show that for any two sets S' and S satisfying $S' \subset S \subset \mathcal{J}$, it holds that $P_{j:S'} \geq P_{j:S}$ for all $j \in S' \cup \{0\}$. Note that there is a nested sequence of sets $S' = S_0 \subset S_1 \subset S_2 \subset \cdots \subset S_{|S\setminus S'|-1} \subset S_{|S\setminus S'|} = S$ such that $|S_i \setminus S_{i-1}| = 1$ for all $i = 1, \ldots, |S \setminus S'|$. Therefore, we consider the case in which $|S \setminus S'| = 1$, and the regularity follows for general S' and S by induction. Let $S \setminus S' = \{k\}$.

<u>Case 1:</u> $k \in J(j)$ for some $j \in S'$. Then the cheapest available fare class on each itinerary remains the same when k is added to assortment S'. Then regularity holds as for the MNL model.

<u>Case 2.1:</u> $k \notin J(j)$ for any $j \in S'$, and $J(k) \cap S' = \varnothing$. This is similar to Case 1. Alternative k is the cheapest (and only) fare class for its itinerary in S. The cheapest available fare class on each other itinerary remains the same when k is added to assortment S'. Then regularity holds as for the MNL model.

<u>Case 2.2:</u> $k \notin J(j)$ for any $j \in S'$, and $J(k) \cap S' \neq \varnothing$. Alternative k is the cheapest fare class for its itinerary in S. Let $l \in J(k) \cap S'$ denote the cheapest fare class for itinerary $g(k)$ in assortment S'. Recall that $w_k + v_l \geq w_l$ by assumption, and thus $W(S) - W(S') = w_k + v_l - w_l \geq 0$, i.e., $W(S) \geq W(S') \geq 0$. Next we consider three cases:

If $j = l$, then $\tilde{v}(j, S') = w_l \geq v_l = \tilde{v}(j, S)$.

If $j \in S' \setminus \{l\}$, then $\tilde{v}(j, S') = v_j(1 - \mathbb{1}(j, S')) + w_j \mathbb{1}(j, S') = v_j(1 - \mathbb{1}(j, S)) + w_j \mathbb{1}(j, S) = \tilde{v}(j, S)$.

If $j = 0$, then $\tilde{v}(j, S') = \tilde{v}(j, S) = v_0$.

Therefore, for any $j \in S' \cup \{0\}$, it holds that $\tilde{v}(j, S') \geq \tilde{v}(j, S)$. It follows that

$$P_{j:S'} \quad = \quad \frac{\tilde{v}(j, S')}{W(S') + v_0} \quad \geq \quad \frac{\tilde{v}(j, S)}{W(S) + v_0} \quad = \quad P_{j:S} \quad \forall\, j \in S' \cup \{0\}.$$

(*Necessity.*) Suppose that there are two products j and j' for the same itinerary, with $j' \in J(j)$, such that

$$w_j + v_{j'} \quad < \quad w_{j'}.$$

We show that in such a case the spiked-MNL model is not regular. Let $S' = \{j'\}$ and $S = \{j, j'\}$. Then

$$P_{0:S'} \quad = \quad \frac{v_0}{v_0 + w_{j'}} \quad < \quad \frac{v_0}{v_0 + w_j + v_{j'}} \quad = \quad P_{0:S},$$

and thus the spiked-MNL model is not regular. $\square$

E.2. Proof of Theorem 1

Liu and Van Ryzin (2008) provided the following necessary and sufficient condition for an assortment to be efficient.

Proposition EC.2 **(Liu and Van Ryzin (2008))**. *An assortment $S \subset \mathcal{J}$ is efficient if and only if for some $\pi \in \mathbb{R}_+^m$, set S is an optimal solution of the problem* $\max_{A \subset \mathcal{J}} \{R(A) - \pi^\top \mathbf{Q}(A)\}$.

The following lemma gives a necessary condition for an assortment to be efficient, and will be used to prove Theorem 1.

LEMMA EC.2. *If an assortment S is efficient, then there exists a $\gamma \in \mathbb{R}^n$, satisfying $\gamma_j > \gamma_{j'}$ for all $j \in \mathcal{J}$ and $j' \in \underline{J}(j)$, such that S is an optimal solution of the problem*

$$\max_{A \subset \mathcal{J}} \left\{ \sum_{j \in A} \gamma_j P_{j:A} \right\}. \tag{EC.5}$$

Proof. If a set S is efficient, then by Proposition EC.2, there exists $\pi \in \mathbb{R}^m_+$ such that S is an optimal solution of $\max_{A \subset \mathcal{J}} \{R(A) - \pi^\top \mathbf{Q}(A)\}$. Note that

$$R(A) - \pi^\top \mathbf{Q}(A) \quad = \quad \sum_{j \in A} \left(r_j P_{j:A} - \pi^\top \mathbf{a}^j P_{j:A} \right) \quad = \quad \sum_{j \in A} \gamma_j P_{j:A},$$

where $\gamma_j := r_j - \pi^\top \mathbf{a}^j$. Note that if $j \in \mathcal{J}$ and $j' \in \underline{J}(j)$, then j and j' are associated with the same itinerary, and thus $\mathbf{a}^j = \mathbf{a}^{j'}$, hence $\pi^\top \mathbf{a}^j = \pi^\top \mathbf{a}^{j'}$, and $\gamma_j > \gamma_{j'}$ since $r_j > r_{j'}$. Therefore, optimization problem $\max_{A \subset \mathcal{J}} \{R(A) - \pi^\top \mathbf{Q}(A)\}$ is reduced to the optimization problem in (EC.5). $\square$

Next we state the definition of efficient sets with specific reference to the set of products considered.

DEFINITION EC.3 (RELATIVELY EFFICIENT SETS). An assortment $S \subset \mathcal{J}$ is said to be *inefficient relative to* $\mathcal{J}$ if a mixture of other assortments in $\mathcal{J}$ has strictly higher expected revenue with the same or lower expected resource consumption. That is, there exists a set of weights $\{\mu(A) : A \subset \mathcal{J}\}$ satisfying $\sum_{A \subset \mathcal{J}} \mu(A) = 1$ and $\mu(A) \geq 0$ for all $A \subset \mathcal{J}$ such that

$$\begin{aligned}
R(S) &:= \sum_{h \in \mathcal{H}} \frac{\lambda_h}{\sum_{h' \in \mathcal{H}} \lambda_{h'}} \sum_{j \in S \cap \mathcal{J}(h)} r_j P^h_{j:S \cap \mathcal{J}(h)} \\
&< \sum_{A \subset \mathcal{J}} \mu(A) R(A) := \sum_{A \subset \mathcal{J}} \mu(A) \sum_{h \in \mathcal{H}} \frac{\lambda_h}{\sum_{h' \in \mathcal{H}} \lambda_{h'}} \sum_{j \in A \cap \mathcal{J}(h)} r_j P^h_{j:A \cap \mathcal{J}(h)} \qquad \text{and} \\
Q_f(S) &:= \sum_{h \in \mathcal{H}} \frac{\lambda_h}{\sum_{h' \in \mathcal{H}} \lambda_{h'}} \sum_{j \in S \cap \mathcal{J}(h)} a^j_f P^h_{j:S \cap \mathcal{J}(h)} \\
&\geq \sum_{A \subset \mathcal{J}} \mu(A) Q_f(A) := \sum_{A \subset \mathcal{J}} \mu(A) \sum_{h \in \mathcal{H}} \frac{\lambda_h}{\sum_{h' \in \mathcal{H}} \lambda_{h'}} \sum_{j \in A \cap \mathcal{J}(h)} a^j_f P^h_{j:A \cap \mathcal{J}(h)} \quad \text{for all } f \in \mathcal{F}.
\end{aligned}$$

An assortment in $\mathcal{J}$ that is not inefficient relative to $\mathcal{J}$ is said to be *efficient relative to* $\mathcal{J}$.

Next we show that a necessary (but not sufficient) condition for an assortment S to be efficient relative to $\mathcal{J}$ is that $S \cap \mathcal{J}(h)$ is efficient relative to $\mathcal{J}(h)$ for each h.

LEMMA EC.3. *If an assortment S is efficient relative to $\mathcal{J}$, then $S \cap \mathcal{J}(h)$ is efficient relative to $\mathcal{J}(h)$ for each $h \in \mathcal{H}$.*

Proof. Suppose that there is an $h' \in \mathcal{H}$ such that $S \cap \mathcal{J}(h')$ is inefficient relative to $\mathcal{J}(h')$, that is, there is a set of weights $\{\mu'(A') : A' \subset \mathcal{J}(h')\}$ satisfying $\sum_{A' \subset \mathcal{J}(h')} \mu'(A') = 1$ and $\mu'(A') \geq 0$ for all $A' \subset \mathcal{J}(h')$ such that

$$\begin{aligned}
\sum_{j \in S \cap \mathcal{J}(h')} r_j P^{h'}_{j:S \cap \mathcal{J}(h')} &< \sum_{A' \subset \mathcal{J}(h')} \mu'(A') \sum_{j \in A'} r_j P^{h'}_{j:A'} \qquad \text{and} \\
\sum_{j \in S \cap \mathcal{J}(h')} a^j_f P^{h'}_{j:S \cap \mathcal{J}(h')} &\geq \sum_{A' \subset \mathcal{J}(h')} \mu'(A') \sum_{j \in A'} a^j_f P^{h'}_{j:A'} \quad \text{for all } f \in \mathcal{F}.
\end{aligned}$$

Then we show that S is inefficient relative to $\mathcal{J}$. Consider the set of weights $\{\mu(A) : A \subset \mathcal{J}\}$ constructed as follows. For each $A' \subset \mathcal{J}(h')$, let $A := A' \cup (S \setminus \mathcal{J}(h'))$ and let $\mu(A) = \mu'(A')$. For all other $B \subset \mathcal{J}$, let $\mu(B) = 0$. Then $\sum_{A \subset \mathcal{J}} \mu(A) = \sum_{A' \subset \mathcal{J}(h')} \mu'(A') = 1$ and $\mu(A) \geq 0$ for all $A \subset \mathcal{J}$. Also,

$$\begin{aligned}
R(S) &:= \frac{\lambda_{h'}}{\sum_{h'' \in \mathcal{H}} \lambda_{h''}} \sum_{j \in S \cap \mathcal{J}(h')} r_j P^{h'}_{j:S \cap \mathcal{J}(h')} + \sum_{\{h \in \mathcal{H} : h \neq h'\}} \frac{\lambda_h}{\sum_{h'' \in \mathcal{H}} \lambda_{h''}} \sum_{j \in S \cap \mathcal{J}(h)} r_j P^h_{j:S \cap \mathcal{J}(h)} \\
&< \frac{\lambda_{h'}}{\sum_{h'' \in \mathcal{H}} \lambda_{h''}} \sum_{A' \subset \mathcal{J}(h')} \mu'(A') \sum_{j \in A'} r_j P^{h'}_{j:A'} + \sum_{\{h \in \mathcal{H} : h \neq h'\}} \frac{\lambda_h}{\sum_{h'' \in \mathcal{H}} \lambda_{h''}} \sum_{j \in S \cap \mathcal{J}(h)} r_j P^h_{j:S \cap \mathcal{J}(h)} \\
&= \frac{\lambda_{h'}}{\sum_{h'' \in \mathcal{H}} \lambda_{h''}} \sum_{A \subset \mathcal{J}} \mu(A) \sum_{j \in A \cap \mathcal{J}(h')} r_j P^{h'}_{j:A \cap \mathcal{J}(h')} \\
&\qquad + \sum_{\{h \in \mathcal{H} : h \neq h'\}} \frac{\lambda_h}{\sum_{h'' \in \mathcal{H}} \lambda_{h''}} \sum_{A \subset \mathcal{J}} \mu(A) \sum_{j \in A \cap \mathcal{J}(h)} r_j P^h_{j:A \cap \mathcal{J}(h)} \\
&= \sum_{A \subset \mathcal{J}} \mu(A) \sum_{h \in \mathcal{H}} \frac{\lambda_h}{\sum_{h'' \in \mathcal{H}} \lambda_{h''}} \sum_{j \in A \cap \mathcal{J}(h)} r_j P^h_{j:A \cap \mathcal{J}(h)} \qquad \text{and} \\
Q_f(S) &:= \frac{\lambda_{h'}}{\sum_{h'' \in \mathcal{H}} \lambda_{h''}} \sum_{j \in S \cap \mathcal{J}(h')} a^j_f P^{h'}_{j:S \cap \mathcal{J}(h')} + \sum_{\{h \in \mathcal{H} : h \neq h'\}} \frac{\lambda_h}{\sum_{h'' \in \mathcal{H}} \lambda_{h''}} \sum_{j \in S \cap \mathcal{J}(h)} a^j_f P^h_{j:S \cap \mathcal{J}(h)} \\
&\geq \frac{\lambda_{h'}}{\sum_{h'' \in \mathcal{H}} \lambda_{h''}} \sum_{A' \subset \mathcal{J}(h')} \mu'(A') \sum_{j \in A'} a^j_f P^{h'}_{j:A'} + \sum_{\{h \in \mathcal{H} : h \neq h'\}} \frac{\lambda_h}{\sum_{h'' \in \mathcal{H}} \lambda_{h''}} \sum_{j \in S \cap \mathcal{J}(h)} a^j_f P^h_{j:S \cap \mathcal{J}(h)} \\
&= \frac{\lambda_{h'}}{\sum_{h'' \in \mathcal{H}} \lambda_{h''}} \sum_{A \subset \mathcal{J}} \mu(A) \sum_{j \in A \cap \mathcal{J}(h')} a^j_f P^{h'}_{j:A \cap \mathcal{J}(h')} \\
&\qquad + \sum_{\{h \in \mathcal{H} : h \neq h'\}} \frac{\lambda_h}{\sum_{h'' \in \mathcal{H}} \lambda_{h''}} \sum_{A \subset \mathcal{J}} \mu(A) \sum_{j \in A \cap \mathcal{J}(h)} a^j_f P^h_{j:A \cap \mathcal{J}(h)} \\
&= \sum_{A \subset \mathcal{J}} \mu(A) \sum_{h \in \mathcal{H}} \frac{\lambda_h}{\sum_{h'' \in \mathcal{H}} \lambda_{h''}} \sum_{j \in A \cap \mathcal{J}(h)} a^j_f P^h_{j:A \cap \mathcal{J}(h)} \quad \text{for all } f \in \mathcal{F}
\end{aligned}$$

and thus S is inefficient relative to $\mathcal{J}$. □

Lemma EC.3 shows that a necessary condition for an assortment S to be efficient relative to $\mathcal{J}$ is that $S \cap \mathcal{J}(h)$ is efficient relative to $\mathcal{J}(h)$ for each h. Next we show that a necessary condition for $S \cap \mathcal{J}(h)$ to be efficient relative to $\mathcal{J}(h)$ is that $S \cap \mathcal{J}(h)$ is nested-by-revenue.

Lemma EC.4 **(Nested-by-Revenue Efficient Assortments)**. *For any assortment $S \subset \mathcal{J}$ and any market $h \in \mathcal{H}$, if $S \cap \mathcal{J}(h)$ is efficient relative to $\mathcal{J}(h)$, then $S \cap \mathcal{J}(h)$ is nested-by-revenue.*

Proof. Suppose that $S \cap \mathcal{J}(h)$ is efficient relative to $\mathcal{J}(h)$. Then it follows from Lemma EC.2 that there is a $\gamma \in \mathbb{R}^n$ such that $\gamma_j > \gamma_{j'}$ for all $j \in \mathcal{J}(h)$ and all $j' \in \underline{J}(j)$, and

$$A \;:=\; S \cap \mathcal{J}(h) \;\in\; \operatorname*{arg\,max}_{S' \subset \mathcal{J}(h)} \left\{ \Gamma(S') \;:=\; \sum_{j \in S'} \gamma_j P_{j:S'} \right\}. \tag{EC.6}$$

First we show by contradiction that if $i \in A$, then $\gamma_i \geq \Gamma(A)$. Suppose that $B := \{j \in A : \gamma_j < \Gamma(A)\} \neq \varnothing$. If $A = B$, then

$$\begin{aligned}
\Gamma(A) \;&=\; \sum_{j \in A} \gamma_j \frac{v_j(1 - \mathbb{1}(j,A)) + w_j \mathbb{1}(j,A)}{v_0 + W(A)} \;<\; \sum_{j \in A} \Gamma(A) \frac{v_j(1 - \mathbb{1}(j,A)) + w_j \mathbb{1}(j,A)}{v_0 + W(A)} \\
\Rightarrow \quad \Gamma(A) \frac{v_0}{v_0 + W(A)} \;&<\; 0 \quad \Rightarrow \quad \Gamma(A) \;<\; 0 \;=\; \Gamma(\varnothing)
\end{aligned}$$

which contradicts (EC.6). Otherwise, $A \setminus B \neq \varnothing$. Let $i^* := \arg\min\{\gamma_i : i \in A \setminus B\}$. Then

$$\begin{aligned}
& \Gamma(A)\,[v_0 + W(A)] \;=\; \sum_{j \in A} \gamma_j \,[v_j(1 - \mathbb{1}(j,A)) + w_j \mathbb{1}(j,A)] \\
\Rightarrow \quad & \Gamma(A)\,[v_0 + W(A)] + \Gamma(A)\,(w_{i^*} - v_{i^*}) \;\leq\; \sum_{j \in A} \gamma_j \,[v_j(1 - \mathbb{1}(j,A)) + w_j \mathbb{1}(j,A)] + \gamma_{i^*}\,(w_{i^*} - v_{i^*}) \\
\Rightarrow \quad & \Gamma(A) \;\leq\; \sum_{j \in A \setminus B} \gamma_j \frac{v_j(1 - \mathbb{1}(j,A \setminus B)) + w_j \mathbb{1}(j,A \setminus B)}{v_0 + W(A \setminus B) + W(B)} + \sum_{j \in B} \gamma_j \frac{v_j(1 - \mathbb{1}(j,A)) + w_j \mathbb{1}(j,A)}{v_0 + W(A \setminus B) + W(B)} \\
& \qquad < \; \sum_{j \in A \setminus B} \gamma_j \frac{v_j(1 - \mathbb{1}(j,A \setminus B)) + w_j \mathbb{1}(j,A \setminus B)}{v_0 + W(A \setminus B) + W(B)} + \sum_{j \in B} \Gamma(A) \frac{v_j(1 - \mathbb{1}(j,A)) + w_j \mathbb{1}(j,A)}{v_0 + W(A \setminus B) + W(B)} \\
\Rightarrow \quad & \Gamma(A) \frac{v_0 + W(A \setminus B)}{v_0 + W(A \setminus B) + W(B)} \;<\; \sum_{j \in A \setminus B} \gamma_j \frac{v_j(1 - \mathbb{1}(j,A \setminus B)) + w_j \mathbb{1}(j,A \setminus B)}{v_0 + W(A \setminus B) + W(B)} \\
\Rightarrow \quad & \Gamma(A) \;<\; \sum_{j \in A \setminus B} \gamma_j \frac{v_j(1 - \mathbb{1}(j,A \setminus B)) + w_j \mathbb{1}(j,A \setminus B)}{v_0 + W(A \setminus B)} \;=\; \Gamma(A \setminus B)
\end{aligned}$$

which contradicts (EC.6). Thus, $B = \varnothing$, and hence if $i \in A$, then $\gamma_i \geq \Gamma(A)$.

Next we show by contradiction that if $S \cap \mathcal{J}(h)$ is efficient relative to $\mathcal{J}(h)$, then $S \cap \mathcal{J}(h)$ is nested-by-revenue. Suppose that $A := S \cap \mathcal{J}(h)$ is efficient relative to $\mathcal{J}(h)$ and $i_2 \in A$ and $i_1 \in J(i_2)$, but $i_1 \notin A$. Since $i_1 \in J(i_2)$, it holds that $r_{i_1} > r_{i_2}$ and $i_2 \in \underline{J}(i_1)$, and thus $\gamma_{i_1} > \gamma_{i_2} \geq \Gamma(A)$. Hence

$$\begin{aligned}
& \Gamma(A)\,[v_0 + W(A)] \;=\; \sum_{j \in A} \gamma_j \,[v_j(1 - \mathbb{1}(j,A)) + w_j \mathbb{1}(j,A)] \\
\Rightarrow \quad & \Gamma(A)\,[v_0 + W(A)] + \Gamma(A) v_{i_1} \;<\; \sum_{j \in A} \gamma_j \,[v_j(1 - \mathbb{1}(j,A)) + w_j \mathbb{1}(j,A)] + \gamma_{i_1} v_{i_1} \\
\Rightarrow \quad & \Gamma(A) \;<\; \sum_{j \in A \cup \{i_1\}} \gamma_j \frac{v_j(1 - \mathbb{1}(j, A \cup \{i_1\})) + w_j \mathbb{1}(j, A \cup \{i_1\})}{v_0 + W(A \cup \{i_1\})} \;=\; \Gamma(A \cup \{i_1\})
\end{aligned}$$

which contradicts (EC.6). Thus $S \cap \mathcal{J}(h)$ is nested-by-revenue. □

Proof of Theorem 1. Lemma EC.3 shows that if assortment $S \subset \mathcal{J}$ is efficient (relative to $\mathcal{J}$), then $S \cap \mathcal{J}(h)$ is efficient relative to $\mathcal{J}(h)$ for each h. Lemma EC.4 shows that if $S \cap \mathcal{J}(h)$ is efficient relative to $\mathcal{J}(h)$, then $S \cap \mathcal{J}(h)$ is nested-by-revenue. Therefore, all efficient sets under the spiked-MNL choice model are nested-by-revenue. □

E.3. Proof of Theorem 2

E.3.1. From SBLP to CDLP The proof of Theorem 2 is constructive. First we propose a polynomial time algorithm (Algorithm 1) to convert a feasible solution $(\mathbf{x}, \mathbf{x}_0)$ of the SBLP (3) into a feasible solution α of the CDLP (2), with the same objective value. The vector $\mathbf{x}$ can be regarded as the planned sales of each product while the product is the cheapest available product for its itinerary. Algorithm 1 constructs a sequence of assortments by translating these planned sales quantities into the amounts of time that each successive assortment is available. The following example illustrates the algorithm.

Example EC.5. Itinerary 1 and Itinerary 2 serve the same origin-destination pair. Itinerary 1 has three products $\{1,2,3\}$ and Itinerary 2 has two products $\{4,5\}$. The given SBLP solution is denoted by $(\mathbf{x}, x_0)$. The length of the time horizon is proportional to x_0/v_0, represented by the horizontal distance in Figure EC.3. For each itinerary g, the quantity x_j/w_j for $j \in \mathcal{J}^g$ is proportional to the amount of time that product j is the cheapest product for itinerary g. By constraint (3d), for each itinerary g, the time intervals of length x_j/w_j for $j \in \mathcal{J}^g$ can be placed in the overall time interval of length x_0/v_0 in such a way that they do not overlap; any such placement of the time intervals for $j \in \mathcal{J}^g$ will do. Thereby the time intervals are determined during which the cheapest products on all the itineraries remain unchanged. This, together with the nested-by-revenue property, determine the assortments and the amounts of time that each assortment is offered for the CDLP. In addition, if, for each itinerary g, the time intervals for $j \in \mathcal{J}^g$ are placed next to each other in sequence from lowest fare product in $\mathcal{J}^g$ to highest fare product in $\mathcal{J}^g$, then the resulting sequence of assortments will be a nested sequence. In the example portrayed in Figure EC.3, this results in assortments $S_1 = \{1,2,3,4,5\} \supset S_2 = \{1,2,3,4\} \supset S_3 = \{1,2,4\} \supset S_4 = \{1,4\} \supset S_5 = \{1\} \supset S_6 = \varnothing$.

Note that the SBLP (3) is based on the *nested-by-revenue* structure. Therefore the optimal solution of the CDLP produced by Algorithm 1 will also have the nested-by-revenue structure. Moreover, by first offering the cheapest products j for which $x_j > 0$, the algorithm constructs a *nested sequence* of assortments.

Lemma EC.5. *Given a feasible solution $(\mathbf{x}, \mathbf{x}_0)$ for SBLP (3), Algorithm 1 terminates in $O(|\mathcal{G}|n)$ steps.*

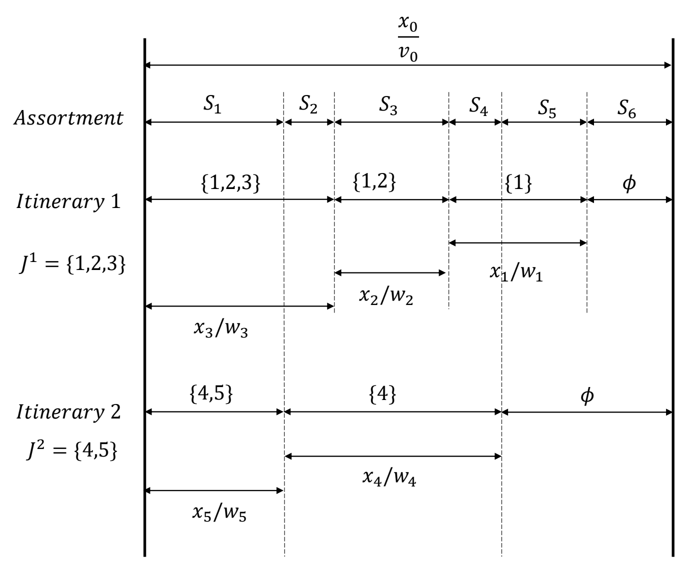
Proof. In each iteration of the while-loop (line 2–line 34), according to the definition of Y_k^g, $Y_k(h)$, $\alpha_k(h)$, and $\alpha(A_k)$, at least one of the positive components of $\mathbf{x}$ is reduced to 0. Therefore, after at most n iterations of the while-loop, it holds that $\mathbf{x} = \mathbf{0}$ and the algorithm terminates. The for-loops in each iteration (line 4–line 20, line 24–line 32) require at most $O(|\mathcal{G}|)$ steps. (The fare classes of each itinerary g can be sorted by revenue in advance so that line 12 can be executed in constant

Algorithm 1 Converting a SBLP solution to a CDLP solution

Require: $\mathcal{J} \leftarrow$ set of products, $\mathcal{H} \leftarrow$ set of markets, $\mathcal{G} \leftarrow$ set of itineraries, $\mathcal{J}(h) \leftarrow$ set of itineraries for each market $h \in \mathcal{H}$, $\mathcal{J}^g \leftarrow$ set of products for each itinerary $g \in \mathcal{G}$, $\bar{J}(j) \leftarrow$ set of products j' for the same itinerary as j with revenue $r_{j'} \geq r_j$, SBLP solution $(\mathbf{x}, \mathbf{x}_0)$

set vector of fractions of time $\alpha \leftarrow \mathbf{0}$ and iteration count $k \leftarrow 1$
while there exists a product $j \in \mathcal{J}$ with $x_j > 0$ **do**
  # form an assortment indexed by k and determine the fraction of time that it is offered #
  for all $h \in \mathcal{H}$ # all markets # **do**
    if $x_j = 0$ for all $j \in \mathcal{J}(h)$ # no products with remaining sales in market h # **then**
      $A_k(h) \leftarrow \varnothing$ and $\alpha_k(h) \leftarrow 0$ # offer nothing in market h #
    else
      for all $g \in \mathcal{J}(h)$ # all itineraries g in market h # **do**
        if $x_j = 0$ for all $j \in \mathcal{J}^g$ # no products for itinerary g with remaining sales # **then**
          set $Y_k^g \leftarrow 0$ # ignore itinerary g #
        else
          $j_k^g \leftarrow \arg\min\{r_j : j \in \mathcal{J}^g, x_j > 0\}$ # pick the cheapest product with remaining sales in g #
          $Y_k^g \leftarrow \frac{x_{j_k^g}}{w_{j_k^g}}$ # fraction of time until product j_k^g will run out of sales #
        end if
      end for
      $A_k(h) \leftarrow \cup_{\{g \in \mathcal{J}(h) : Y_k^g > 0\}} \bar{J}(j_k^g)$ # determine the next assortment in market h #
      $Y_k(h) \leftarrow \min\{Y_k^g : g \in \mathcal{J}(h), Y_k^g > 0\}$ # determine the smallest fraction of time until product j_k^g for any itinerary g in market h will run out of sales #
      $\alpha_k(h) \leftarrow \frac{W(A_k(h)) + v_0}{\lambda_h T} Y_k(h)$ # set the fraction of time to offer assortment $A_k(h)$ in market h #
    end if
  end for
  $A_k \leftarrow \cup_{h \in \mathcal{H}} A_k(h)$ # set the overall assortment for all markets #
  $\alpha(A_k) \leftarrow \min\{\alpha_k(h) : h \in \mathcal{H}, \alpha_k(h) > 0\}$ # set the fraction of time to offer assortment A_k #
  # reduce the remaining sales by the expected sales when offering assortment A_k #
  for all $h \in \mathcal{H}$ # all markets # **do**
    if $A_k(h) \neq \varnothing$ **then**
      for all $g \in \mathcal{J}(h)$ # all itineraries g in market h # **do**
        if $Y_k^g > 0$ # any products for itinerary g in assortment A_k # **then**
          $x_{j_k^g} \leftarrow x_{j_k^g} - \lambda_h \alpha(A_k) T \frac{w_{j_k^g}}{W(A_k(h)) + v_0}$ # reduce the remaining sales of product j_k^g by the expected sales of product j_k^g when offering assortment A_k #
        end if
      end for
    end if
  end for
  $k \leftarrow k + 1$ # increment iteration count #
end while
output vector α

Figure EC.3 An example of converting a SBLP solution to a CDLP solution.



time for each g. Also, for each $j \in \mathcal{J}$, one can compute the values of $W(\bar{J}(j)) = w_j + \sum_{j' \in J(j)} v_{j'}$ in advance, which can be done inductively from the highest fare class for each itinerary to the lowest fare class for the itinerary, in a total of $O(n)$ steps. Then, in line 18, $W(A_k(h))$ can be computed for all h in $O(|\mathcal{G}|)$ steps as follows: $W(A_k(h)) = \sum_{\{g \in \mathcal{J}(h) \,:\, Y_k^g > 0\}} W(\bar{J}(j_k^g))$.) So Algorithm 1 terminates in $O(|\mathcal{G}|n)$ steps. $\square$

LEMMA EC.6. *The output of Algorithm 1 satisfies the following properties:*

(1) In each iteration k, the assortment A_k defined in line 21 is nested-by-revenue.

(2) In each iteration k, for each market h and each itinerary $g \in \mathcal{J}(h)$, the amount $\lambda_h \alpha(A_k) T \frac{w_{j_k^g}}{W(A_k(h)) + v_0}$ subtracted from $x_{j_k^g}$ is equal to the expected sales quantity of product j_k^g while assortment A_k is offered for $\alpha(A_k)T$ units of time.

(3) The CDLP solution α produced by Algorithm 1 satisfies

$$x_j \quad = \quad \lambda_h T \sum_{\{k \,:\, I(j, A_k) = 1\}} \alpha(A_k) P^h_{j: A_k \cap \mathcal{J}(h)}$$

for all $h \in \mathcal{H}$ and all $j \in \mathcal{J}(h)$. That is, for each product j, the sales quantity x_j while j is the cheapest available fare class for its itinerary specified by the SBLP solution $(\mathbf{x}, \mathbf{x}_0)$ is equal to the sales quantity of j while j is the cheapest available fare class for its itinerary resulting from CDLP solution α.

Proof. (1) For each k and each h, $A_k(h)$ is either $\varnothing$ or a union over itineraries g of nested-by-revenue assortments $\bar{J}(j_k^g)$, so each $A_k(h)$ is a nested-by-revenue assortment. Each A_k is a union over markets h of nested-by-revenue assortments $A_k(h)$, so each A_k is a nested-by-revenue assortment.

(2) The expected sales quantity of j_k^g while assortment A_k is offered for $\alpha(A_k)T$ units of time is equal to $\lambda_h \alpha(A_k) T P^h_{j_k^g : A_k \cap \mathcal{J}(h)} = \lambda_h \alpha(A_k) T \frac{w_{j_k^g}}{W(A_k \cap \mathcal{J}(h)) + v_0} = \lambda_h \alpha(A_k) T \frac{w_{j_k^g}}{W(A_k(h)) + v_0}$.

(3) It follows from above that

$$x_j \;=\; \sum_{\{k\,:\,I(j,A_k)=1\}} \lambda_h \alpha(A_k) T \frac{w_j}{W(A_k(h)) + v_0} \;=\; \lambda_h T \sum_{\{k\,:\,I(j,A_k)=1\}} \alpha(A_k) P^h_{j:A_k \cap \mathcal{J}(h)}. \qquad \square$$

PROPOSITION EC.3. *Given a feasible solution* $(\mathbf{x}, \mathbf{x}_0)$ *of SBLP* (3)*, Algorithm 1 computes a feasible solution* α *of CDLP* (2)*, such that the sales quantity of each product is the same in both solutions, and the two solutions have the same objective value.*

Proof. First we show that $(\mathbf{x}, \mathbf{x}_0)$ satisfies SBLP constraint (3c) if and only if α produced by Algorithm 1 satisfies CDLP constraint (2c). The left side of SBLP constraint (3c) is

$$\begin{aligned}
& \sum_{h \in \mathcal{H}} \sum_{j \in \mathcal{J}(h)} \left(1 + \sum_{j' \in J(j)} \frac{v_{j'}}{w_j}\right) a_f^j x_j \\
=\; & \sum_{h \in \mathcal{H}} \sum_{j \in \mathcal{J}(h)} \left(1 + \sum_{j' \in J(j)} \frac{v_{j'}}{w_j}\right) a_f^j \sum_{\{k\,:\,I(j,A_k)=1\}} \lambda_h \alpha(A_k) T \frac{w_j}{W(A_k(h)) + v_0} \\
=\; & \sum_k \alpha(A_k) T \sum_{h \in \mathcal{H}} \lambda_h \sum_{\{j \in \mathcal{J}(h)\,:\,I(j,A_k)=1\}} a_f^j \frac{w_j + \sum_{j' \in J(j)} v_{j'}}{W(A_k(h)) + v_0} \\
=\; & \sum_k \alpha(A_k) T \sum_{h \in \mathcal{H}} \lambda_h \sum_{\{j \in \mathcal{J}(h)\,:\,I(j,A_k)=1\}} \left(a_f^j \frac{w_j}{W(A_k(h)) + v_0} + \sum_{j' \in J(j)} a_f^{j'} \frac{v_{j'}}{W(A_k(h)) + v_0} \right) \\
=\; & \sum_k \alpha(A_k) T \sum_{h \in \mathcal{H}} \lambda_h \sum_{\{j \in \mathcal{J}(h)\,:\,I(j,A_k)=1\}} \left(a_f^j P^h_{j:A_k \cap \mathcal{J}(h)} + \sum_{j' \in J(j)} a_f^{j'} P^h_{j':A_k \cap \mathcal{J}(h)} \right) \\
=\; & \sum_{A \subset \mathcal{J}} \alpha(A) T \sum_{h \in \mathcal{H}} \lambda_h \sum_{j \in \mathcal{J}(h)} a_f^j P^h_{j:A \cap \mathcal{J}(h)} \\
=\; & \sum_{A \subset \mathcal{J}} \alpha(A) T \sum_{h \in \mathcal{H}} \lambda_h \sum_{j \in A} a_f^j P^h_{j:A \cap \mathcal{J}(h)}.
\end{aligned}$$

which is the left side of CDLP constraint (2c). The third equality holds because for any $j \in \mathcal{J}$ and any $j' \in J(j)$ it holds that $a_f^j = a_f^{j'}$ for all $f \in \mathcal{F}$.

Next we show that if $(\mathbf{x}, \mathbf{x}_0)$ satisfies SBLP constraints (3b) and (3d), then α satisfies CDLP constraint (2b). Consider any h and any $g \in \mathcal{J}(h)$. In each iteration k, if there is a product $j \in \mathcal{J}^g$ such that $x_j > 0$, then one such product j_k^g is chosen. Then the quantity $x_{j_k^g}$ is reduced by $\lambda_h \alpha(A_k) T \frac{w_{j_k^g}}{W(A_k(h)) + v_0}$, and for all the other products $j' \in \mathcal{J}^g \setminus \{j_k^g\}$, $x_{j'}$ remains unchanged. Otherwise, if there is no product $j \in \mathcal{J}^g$ such that $x_j > 0$, then x_j remains 0 for all $j \in \mathcal{J}^g$. Thus, for any h, any $g \in \mathcal{J}(h)$, and any $j \in \mathcal{J}^g$, it holds that

$$\sum_{\{k\,:\,j_k^g = j\}} \lambda_h \alpha(A_k) T \frac{1}{W(A_k(h)) + v_0} \;=\; \frac{x_j}{w_j},$$

and hence

$$\sum_{\{k\,:\,Y_k^g>0\}} \lambda_h \alpha(A_k) T \frac{1}{W(A_k(h))+v_0} \;=\; \sum_{j\in\mathcal{J}^g} \sum_{\{k\,:\,j_k^g=j\}} \lambda_h \alpha(A_k) T \frac{1}{W(A_k(h))+v_0} \;=\; \sum_{j\in\mathcal{J}^g} \frac{x_j}{w_j} \;\le\; \frac{x_0^h}{v_0},$$

where the inequality follows from SBLP constraint (3d), and

$$\begin{aligned}
&\sum_{\{k\,:\,Y_k^g>0\}} W(\bar{J}(j_k^g)) \lambda_h \alpha(A_k) T \frac{1}{W(A_k(h))+v_0} \\
&\qquad = \sum_{j\in\mathcal{J}^g} \sum_{\{k\,:\,j_k^g=j\}} W(\bar{J}(j)) \lambda_h \alpha(A_k) T \frac{1}{W(A_k(h))+v_0} \;=\; \sum_{j\in\mathcal{J}^g} W(\bar{J}(j)) \frac{x_j}{w_j}.
\end{aligned}$$

Note that there is at least one $g\in\mathcal{G}$ such that $Y_k^g>0$ for all k. Let $g^*\in\mathcal{G}$ be such that $Y_k^{g^*}>0$ for all k, and let $h^*\in\mathcal{H}$ be such that $g^*\in\mathcal{J}(h^*)$. It follows from line 18 and line 21 of Algorithm 1 that

$$\begin{aligned}
&\sum_{A\subset\mathcal{J}} \alpha(A) \;=\; \sum_k \alpha(A_k) \\
=\;& \sum_k \alpha(A_k) \frac{\lambda_{h^*} T}{\lambda_{h^*} T} \frac{W(A_k(h^*))+v_0}{W(A_k(h^*))+v_0} \\
=\;& \frac{1}{\lambda_{h^*} T} \sum_k \lambda_{h^*} \alpha(A_k) T \frac{\sum_{\{g\in\mathcal{J}(h^*)\,:\,Y_k^g>0\}} W(\bar{J}(j_k^g))+v_0}{W(A_k(h^*))+v_0} \\
=\;& \frac{1}{\lambda_{h^*} T} \left[\sum_{g\in\mathcal{J}(h^*)} \sum_{\{k\,:\,Y_k^g>0\}} W(\bar{J}(j_k^g)) \lambda_{h^*} \alpha(A_k) T \frac{1}{W(A_k(h^*))+v_0} + \sum_k \lambda_{h^*} \alpha(A_k) T \frac{v_0}{W(A_k(h^*))+v_0} \right] \\
=\;& \frac{1}{\lambda_{h^*} T} \left[\sum_{g\in\mathcal{J}(h^*)} \sum_{j\in\mathcal{J}^g} W(\bar{J}(j)) \frac{x_j}{w_j} + \sum_{\{k\,:\,Y_k^{g^*}>0\}} \lambda_{h^*} \alpha(A_k) T \frac{v_0}{W(A_k(h^*))+v_0} \right] \\
\le\;& \frac{1}{\lambda_{h^*} T} \left[\sum_{g\in\mathcal{J}(h^*)} \sum_{j\in\mathcal{J}^g} \left(w_j + \sum_{j'\in J(j)} v_{j'} \right) \frac{x_j}{w_j} + v_0 \frac{x_0^{h^*}}{v_0} \right] \\
=\;& \frac{1}{\lambda_{h^*} T} \left[\sum_{j\in\mathcal{J}(h^*)} \left(1 + \sum_{j'\in J(j)} \frac{v_{j'}}{w_j} \right) x_j + x_0^{h^*} \right] \;=\; 1,
\end{aligned}$$

where the last equality follows from SBLP constraint (3b).

Next we show that the objective values of the SBLP solution $(\mathbf{x},\mathbf{x}_0)$ and the CDLP solution α are the same.

$$\begin{aligned}
&\sum_{h\in\mathcal{H}} \sum_{j\in\mathcal{J}(h)} \left(r_j + \sum_{j'\in J(j)} r_{j'} \frac{v_{j'}}{w_j} \right) x_j \\
=\;& \sum_{h\in\mathcal{H}} \sum_{g\in\mathcal{J}(h)} \sum_{j\in\mathcal{J}^g} \left(r_j + \sum_{j'\in J(j)} r_{j'} \frac{v_{j'}}{w_j} \right) \sum_{\{k\,:\,j_k^g=j\}} \lambda_h \alpha(A_k) T \frac{w_j}{W(A_k(h))+v_0} \\
=\;& \sum_k \alpha(A_k) T \sum_{h\in\mathcal{H}} \lambda_h \sum_{g\in\mathcal{J}(h)} \sum_{\{j\in\mathcal{J}^g\,:\,j_k^g=j\}} \left(r_j \frac{w_j}{W(A_k(h))+v_0} + \sum_{j'\in J(j)} r_{j'} \frac{v_{j'}}{W(A_k(h))+v_0} \right)
\end{aligned}$$

$$= \sum_{k} \alpha(A_k) T \sum_{h \in \mathcal{H}} \lambda_h \sum_{j \in A_k \cap \mathcal{J}(h)} r_j P^h_{j:A_k \cap \mathcal{J}(h)}$$
$$= \sum_{A \subset \mathcal{J}} \alpha(A) T \sum_{h \in \mathcal{H}} \lambda_h \sum_{j \in A \cap \mathcal{J}(h)} r_j P^h_{j:A \cap \mathcal{J}(h)}.$$

Thus we have established that Algorithm 1 converts any feasible solution of SBLP (3) into a feasible solution of CDLP (2), such that the sales quantity of each product is the same in both solutions, and the two solutions have the same objective value. □

E.3.2. From CDLP to SBLP. In this section we address the opposite direction: converting a CDLP solution into a SBLP solution.

PROPOSITION EC.4. *Consider any feasible solution α of CDLP* (2) *with support on assortments that are nested-by-revenue. Then there is a feasible solution $(\mathbf{x}, \mathbf{x}_0)$ of SBLP* (3) *such that the sales quantity of each product is the same in both solutions, and the two solutions have the same objective value.*

Proof. Since α has support on assortments that are nested-by-revenue, it follows that for any $A \subset \mathcal{J}$ such that $\alpha(A) > 0$ and for any $j \in A \cap \mathcal{J}(h)$, it holds that $J(j) \subset A \cap \mathcal{J}(h)$. For every $h \in \mathcal{H}$ and $j \in \mathcal{J}(h)$, let

$$x_j := \lambda_h T \sum_{\{A \subset \mathcal{J} \,:\, \mathbb{1}(j,A)=1\}} P^h_{j:A \cap \mathcal{J}(h)} \alpha(A)$$
$$\text{and} \quad x^h_0 := \lambda_h T \left\{ \sum_{A \subset \mathcal{J}} P^h_{0:A \cap \mathcal{J}(h)} \alpha(A) + \left[1 - \sum_{A \subset \mathcal{J}} \alpha(A) \right] \right\}.$$

Next we show that $(\mathbf{x}, \mathbf{x}_0)$ is feasible for SBLP (3). The balance constraint (3b) in the SBLP holds, since

$$\begin{aligned}
& x^h_0 + \sum_{j \in \mathcal{J}(h)} \left(1 + \sum_{j' \in J(j)} \frac{v_{j'}}{w_j} \right) x_j \\
= \;& \lambda_h T \left\{ \sum_{A \subset \mathcal{J}} P^h_{0:A \cap \mathcal{J}(h)} \alpha(A) + \left[1 - \sum_{A \subset \mathcal{J}} \alpha(A) \right] \right\} \\
& \quad + \sum_{j \in \mathcal{J}(h)} \left(1 + \sum_{j' \in J(j)} \frac{v_{j'}}{w_j} \right) \lambda_h T \sum_{\{A \subset \mathcal{J} \,:\, \mathbb{1}(j,A)=1\}} P^h_{j:A \cap \mathcal{J}(h)} \alpha(A) \\
= \;& \lambda_h T \Bigg\{ \sum_{A \subset \mathcal{J}} \left[P^h_{0:A \cap \mathcal{J}(h)} + \sum_{\{j \in \mathcal{J}(h) \,:\, \mathbb{1}(j,A)=1\}} \left(1 + \sum_{j' \in J(j)} \frac{v_{j'}}{w_j} \right) P^h_{j:A \cap \mathcal{J}(h)} \right] \alpha(A) \\
& \quad + \left[1 - \sum_{A \subset \mathcal{J}} \alpha(A) \right] \Bigg\} \\
= \;& \lambda_h T \Bigg\{ \sum_{A \subset \mathcal{J}} \left[P^h_{0:A \cap \mathcal{J}(h)} + \sum_{\{j \in \mathcal{J}(h) \,:\, \mathbb{1}(j,A)=1\}} \left(P^h_{j:A \cap \mathcal{J}(h)} + \sum_{j' \in J(j)} \frac{v_{j'}}{w_j} \frac{w_j}{W(A \cap \mathcal{J}(h)) + v_0} \right) \right] \alpha(A)
\end{aligned}$$

$$
\begin{aligned}
&\quad + \left[1 - \sum_{A \subset \mathcal{J}} \alpha(A)\right]\Bigg\} \\
&= \lambda_h T \Bigg\{ \sum_{A \subset \mathcal{J}} \left[P^h_{0:A\cap\mathcal{J}(h)} + \sum_{\{j \in \mathcal{J}(h) \,:\, \mathbb{1}(j,A)=1\}} \left(P^h_{j:A\cap\mathcal{J}(h)} + \sum_{j' \in J(j)} P^h_{j':A\cap\mathcal{J}(h)} \right) \right] \alpha(A) \\
&\quad + \left[1 - \sum_{A \subset \mathcal{J}} \alpha(A)\right]\Bigg\} \\
&= \lambda_h T \left\{ \sum_{A \subset \mathcal{J}} \alpha(A) + \left[1 - \sum_{A \subset \mathcal{J}} \alpha(A)\right] \right\} \;=\; \lambda T.
\end{aligned}
$$

For each $f \in \mathcal{F}$, the SBLP capacity constraint (3c) also holds, since

$$
\begin{aligned}
&\sum_{h\in\mathcal{H}} \sum_{j \in \mathcal{J}(h)} \left(1 + \sum_{j' \in J(j)} \frac{v_{j'}}{w_j}\right) a^j_f x_j \\
&= \sum_{h\in\mathcal{H}} \sum_{j \in \mathcal{J}(h)} \left(1 + \sum_{j' \in J(j)} \frac{v_{j'}}{w_j}\right) a^j_f \lambda_h T \sum_{\{A \subset \mathcal{J} \,:\, \mathbb{1}(j,A)=1\}} P^h_{j:A\cap\mathcal{J}(h)} \alpha(A) \\
&= \sum_{A \subset \mathcal{J}} \alpha(A) T \sum_{h\in\mathcal{H}} \lambda_h \sum_{\{j \in \mathcal{J}(h) \,:\, \mathbb{1}(j,A)=1\}} \left(1 + \sum_{j' \in J(j)} \frac{v_{j'}}{w_j}\right) a^j_f P^h_{j:A\cap\mathcal{J}(h)} \\
&= \sum_{A \subset \mathcal{J}} \alpha(A) T \sum_{h\in\mathcal{H}} \lambda_h \sum_{\{j \in \mathcal{J}(h) \,:\, \mathbb{1}(j,A)=1\}} a^j_f \left(P^h_{j:A\cap\mathcal{J}(h)} + \sum_{j' \in J(j)} \frac{v_{j'}}{w_j} \frac{w_j}{W(A \cap \mathcal{J}(h)) + v_0} \right) \\
&= \sum_{A \subset \mathcal{J}} \alpha(A) T \sum_{h\in\mathcal{H}} \lambda_h \sum_{\{j \in \mathcal{J}(h) \,:\, \mathbb{1}(j,A)=1\}} a^j_f \left(P^h_{j:A\cap\mathcal{J}(h)} + \sum_{j' \in J(j)} P^h_{j':A\cap\mathcal{J}(h)} \right) \\
&= \sum_{A \subset \mathcal{J}} \alpha(A) T \sum_{h\in\mathcal{H}} \lambda_h \sum_{j \in \mathcal{J}(h)} a^j_f P^h_{j:A\cap\mathcal{J}(h)} \;\le\; c_f,
\end{aligned}
$$

where the last equality holds because for any $j \in \mathcal{J}$ and any $j' \in J(j)$ it holds that $a^j_f = a^{j'}_f$ for all $f \in \mathcal{F}$. Also, for any $h \in \mathcal{H}$ and $g \in \mathcal{J}(h)$,

$$
\begin{aligned}
\sum_{j \in \mathcal{J}^g} \frac{x_j}{w_j} &= \sum_{j \in \mathcal{J}^g} \frac{1}{w_j} \lambda_h T \sum_{\{A \subset \mathcal{J} \,:\, \mathbb{1}(j,A)=1\}} P^h_{j:A\cap\mathcal{J}(h)} \alpha(A) \\
&= \lambda_h T \sum_{A \subset \mathcal{J}} \sum_{\{j \in \mathcal{J}^g \,:\, \mathbb{1}(j,A)=1\}} \frac{1}{w_j} \frac{w_j}{W(A \cap \mathcal{J}(h)) + v_0} \alpha(A) \\
&= \lambda_h T \sum_{A \subset \mathcal{J}} \sum_{\{j \in \mathcal{J}^g \,:\, \mathbb{1}(j,A)=1\}} \frac{1}{v_0} \frac{v_0}{W(A \cap \mathcal{J}(h)) + v_0} \alpha(A) \\
&= \lambda_h T \sum_{\{A \subset \mathcal{J} \,:\, A \cap \mathcal{J}^g \neq \varnothing\}} P^h_{0:A\cap\mathcal{J}(h)} \alpha(A) \frac{1}{v_0} \;\le\; \frac{x^h_0}{v_0},
\end{aligned}
$$

which is SBLP constraint (3d). Finally, the objective value of SBLP is

$$
\sum_{h\in\mathcal{H}} \sum_{j \in \mathcal{J}(h)} \left(r_j + \sum_{j' \in J(j)} r_{j'} \frac{v_{j'}}{w_j} \right) x_j
$$

$$
\begin{aligned}
&= \sum_{h\in\mathcal{H}}\sum_{j\in\mathcal{J}(h)}\left(r_j+\sum_{j'\in J(j)} r_{j'}\frac{v_{j'}}{w_j}\right)\lambda_h T \sum_{\{A\subset\mathcal{J}\,:\,\mathbb{1}(j,A)=1\}} P^h_{j:A\cap\mathcal{J}(h)}\alpha(A) \\
&= \sum_{A\subset\mathcal{J}}\alpha(A)T\sum_{h\in\mathcal{H}}\lambda_h \sum_{\{j\in\mathcal{J}(h)\,:\,\mathbb{1}(j,A)=1\}}\left(r_j+\sum_{j'\in J(j)} r_{j'}\frac{v_{j'}}{w_j}\right)P^h_{j:A\cap\mathcal{J}(h)} \\
&= \sum_{A\subset\mathcal{J}}\alpha(A)T\sum_{h\in\mathcal{H}}\lambda_h \sum_{\{j\in\mathcal{J}(h)\,:\,\mathbb{1}(j,A)=1\}}\left(r_jP^h_{j:A\cap\mathcal{J}(h)}+\sum_{j'\in J(j)} r_{j'}\frac{v_{j'}}{w_j}\frac{w_j}{W(A\cap\mathcal{J}(h))+v_0}\right) \\
&= \sum_{A\subset\mathcal{J}}\alpha(A)T\sum_{h\in\mathcal{H}}\lambda_h \sum_{j\in\mathcal{J}(h)} r_jP^h_{j:A\cap\mathcal{J}(h)} \\
&= \sum_{A\subset\mathcal{J}}\alpha(A)T\sum_{h\in\mathcal{H}}\lambda_h \sum_{j\in A} r_jP^h_{j:A\cap\mathcal{J}(h)}
\end{aligned}
$$

which is equal to the objective value of CDLP. □

Theorem 2 follows from Lemma EC.5, Lemma EC.6, Proposition EC.3, and Proposition EC.4.

E.4. Proof of Theorem 3

Proof of Theorem 3. First, we review some properties of the CDLP, and describe the associated partitioned booking limit policy. By Theorem 2, there is an optimal solution α^* for the CDLP that is supported on a nested sequence of assortments $S_1 \supset S_2 \supset \cdots \supset S_k$, with $\alpha^*(S_i) > 0$ for $i = 1, \ldots, k$, and with each S_i nested-by-revenue. Let $t_0 := 0$ and $t_i := \sum_{i'=1}^{i} \alpha^*(S_{i'})T$. Thus, an optimal solution for the CDLP is to offer each assortment S_i during $(t_{i-1}, t_i]$. For the CDLP, the sales rate of product j during $(t_{i-1}, t_i]$ is given by $\lambda_j^i := \lambda P_{j:S_i}$, and the corresponding sales quantity is $\lambda_j^i(t_i - t_{i-1})$. (Note that $\lambda_j^i = 0$ if $j \notin S_i$.) It follows from (4) that the booking limit b_j^* of the partitioned booking limit policy satisfies

$$
b_j^* \;=\; \sum_{i=1}^{k} \lambda_j^i (t_i - t_{i-1}),
$$

and thus the optimal objective value z^{CDLP} of the CDLP satisfies

$$
z^{\text{CDLP}} \;=\; \sum_{j\in\mathcal{J}} r_j \sum_{i=1}^{k} \lambda_j^i (t_i - t_{i-1}) \;=\; \sum_{j\in\mathcal{J}} r_j b_j^*.
$$

For each $i = 1, \ldots, k$ and each $j \in \mathcal{J}$, let

$$
\begin{aligned}
\underline{\lambda}_j^i \;&:=\; \min\{\lambda P_{j:S} \,:\, S_{i+1} \subseteq S \subseteq S_i, j \in S\} \\
\text{and} \qquad \overline{\lambda}_j^i \;&:=\; \max\{\lambda P_{j:S} \,:\, S_{i+1} \subseteq S \subseteq S_i, j \in S\},
\end{aligned}
$$

where $S_{k+1} := \varnothing$. Note that under the spiked-MNL model, $\underline{\lambda}_j^i > 0$ if $j \in S_i$. Also note that $\overline{\lambda}_j^i \geq \lambda P_{j:S_{i+1}} = \lambda_j^{i+1}$ for all i and all $j \in S_i$.

Given the times $0 = t_0 < t_1 < \cdots < t_k = T$ resulting from the CDLP solution, and $\varepsilon > 0$, consider the times $t_i^- := t_i - \nu_i^- \varepsilon$ and $t_i^+ := t_i + \nu_i^+ \varepsilon$, where $\nu_i^-, \nu_i^+ > 0$ are specified inductively as follows: Let $\nu_0^- = \nu_0^+ = 0$, and let $\overline{\lambda}_j^0 = 0$ for all $j \in \mathcal{J}$. Then, for each $i = 1, \ldots, k$, let

$$\nu_i^- \quad := \quad \max_{j \in S_i} \left\{ \frac{\overline{\lambda}_j^{i-1} \nu_{i-1}^- + \left(\overline{\lambda}_j^{i-1} - \lambda_j^i \right) \nu_{i-1}^+ + \overline{\lambda}_j^{i-1} + \lambda_j^i}{\lambda_j^i} \right\}, \tag{EC.7a}$$

$$\nu_i^+ \quad := \quad \max_{j \in S_i} \left\{ \frac{\lambda_j^i \nu_{i-1}^+ + \left(\lambda_j^i - \underline{\lambda}_j^i \right) \nu_i^- + \lambda_j^i + \underline{\lambda}_j^i}{\underline{\lambda}_j^i} \right\}. \tag{EC.7b}$$

Also, let $\zeta_j^0 = 0$ for all j, and for each $i = 1, \ldots, k$, and $j \in S_i$, let

$$\eta_j^i \quad := \quad \lambda_j^i \left(\nu_{i-1}^+ + \nu_i^- + 1 \right), \qquad \zeta_j^i \quad := \quad \overline{\lambda}_j^i \left(\nu_i^- + \nu_i^+ + 1 \right). \tag{EC.8}$$

Note that (EC.7) and (EC.8) imply that

$$\lambda_j^i \left(\nu_{i-1}^+ + \nu_i^- + 1 \right) \quad = \quad \eta_j^i \quad \leq \quad \underline{\lambda}_j^i \left(\nu_i^- + \nu_i^+ - 1 \right), \tag{EC.9a}$$

$$\overline{\lambda}_j^i \left(\nu_i^- + \nu_i^+ + 1 \right) \quad = \quad \zeta_j^i \quad \leq \quad \lambda_j^{i+1} \left(\nu_i^+ + \nu_{i+1}^- - 1 \right). \tag{EC.9b}$$

We consider $\varepsilon > 0$ sufficiently small such that $t_{i-1}^+ < t_i^-$ for all $i = 1, \ldots, k$, and $\zeta_j^i \varepsilon < \lambda_j^{i+1} (t_{i+1} - t_i)$ for all $i = 1, \ldots, k-1$ and all $j \in S_{i+1}$. Since $t_k^+ > t_k = T$, for convenience, the analysis below considers a continuation of the booking process after time T, but we will only count the total bookings up to time T towards the total revenue. Figure 3 in the paper illustrates the quantities defined above.

Next we define the stochastic sales process for Poisson demand. For each $i \in \{0, \ldots, k\}$ and $j \in \mathcal{J}$, let $\hat{N}_j^{i-} := \{\hat{N}_j^{i-}(t) : t \geq 0\}$ and $\hat{N}_j^{i+} := \{\hat{N}_j^{i+}(t) : t \geq 0\}$ denote Poisson processes with rate 1, with all the Poisson processes $\{\hat{N}_j^{i\pm} : i \in \{0, \ldots, k\}, j \in \mathcal{J}\}$ independent. For any scaling factor θ, let $S^\theta(t)$ denote the assortment offered at time t under the considered policy. Then, for each product $j \in \mathcal{J}$, each $i \in \{0, \ldots, k-1\}$, and each time $t \in (t_i^+, t_{i+1}^-]$, the total sales of product j over $(0, t]$ is

$$\sum_{i'=0}^{i-1} \hat{N}_j^{i'+} \left(\theta\lambda \int_{\tau \in (t_{i'}^+, t_{i'+1}^-]} P_{j:S^\theta(\tau)} d\tau \right) + \sum_{i'=1}^{i} \hat{N}_j^{i'-} \left(\theta\lambda \int_{\tau \in (t_{i'}^-, t_{i'}^+]} P_{j:S^\theta(\tau)} d\tau \right) + \hat{N}_j^{i+} \left(\theta\lambda \int_{\tau \in (t_i^+, t]} P_{j:S^\theta(\tau)} d\tau \right)$$

and similarly, for each $i \in \{1, \ldots, k\}$, and each time $t \in (t_i^-, t_i^+]$, the total sales of product j over $(0, t]$ is

$$\sum_{i'=0}^{i-1} \hat{N}_j^{i'+} \left(\theta\lambda \int_{\tau \in (t_{i'}^+, t_{i'+1}^-]} P_{j:S^\theta(\tau)} d\tau \right) + \sum_{i'=1}^{i-1} \hat{N}_j^{i'-} \left(\theta\lambda \int_{\tau \in (t_{i'}^-, t_{i'}^+]} P_{j:S^\theta(\tau)} d\tau \right) + \hat{N}_j^{i-} \left(\theta\lambda \int_{\tau \in (t_i^-, t]} P_{j:S^\theta(\tau)} d\tau \right).$$

Note that while the assortment is S, product j is sold according to a Poisson process with rate $\theta\lambda P_{j:S}$. Let π denote a policy that, for each scaling factor θ and at each time t, prescribes the assortment $S^\theta(t)$ to be offered at time t. Then the objective value under policy π for the θ-scaled problem is given by

$$\mathsf{E}^\pi \Bigg[\sum_{j \in \mathcal{J}} r_j \Bigg\{ \sum_{i=0}^{k-1} \hat{N}_j^{i+} \left(\theta\lambda \int_{\tau \in (t_i^+, t_{i+1}^-]} P_{j:S^\theta(\tau)} d\tau \right) + \sum_{i=1}^{k-1} \hat{N}_j^{i-} \left(\theta\lambda \int_{\tau \in (t_i^-, t_i^+]} P_{j:S^\theta(\tau)} d\tau \right) + \hat{N}_j^{k-} \left(\theta\lambda \int_{\tau \in (t_k^-, T]} P_{j:S^\theta(\tau)} d\tau \right) \Bigg\} \Bigg].$$

Now we describe the stochastic sales process and the offered assortments under the partitioned booking limit policy. We ignore the probability 0 event that more than 1 arrival of the Poisson processes $\{\hat{N}_j^{i\pm} : i \in \{0,\ldots,k\}, j \in \mathcal{J}\}$ take place at the same time. Let S_ℓ^θ denote the ℓth assortment offered under the partitioned booking limit policy (note that, except for $\ell = 1$, it does not hold in general that $S_\ell^\theta = S_\ell$), and let assortment S_ℓ^θ be offered over time period $(\tau_{\ell-1}^\theta, \tau_\ell^\theta]$. That is, $S^\theta(t) = S_\ell^\theta$ for $t \in (\tau_{\ell-1}^\theta, \tau_\ell^\theta]$. For any $i = 0,\ldots,k-1$ and $\tau \in (\tau_{\ell-1}^\theta, T]$, let

$$T_{i,\ell}^{\theta+} := \begin{cases} \min\{t_{i+1}^-, \tau_\ell^\theta\} - \max\{t_i^+, \tau_{\ell-1}^\theta\} & \text{if } \tau_\ell^\theta > t_i^+ \text{ and } \tau_{\ell-1}^\theta < t_{i+1}^- \\ 0 & \text{otherwise} \end{cases}$$

$$T_{i,\ell}^{\theta+}(\tau) := \begin{cases} \min\{t_{i+1}^-, \tau\} - \max\{t_i^+, \tau_{\ell-1}^\theta\} & \text{if } \tau > t_i^+ \text{ and } \tau_{\ell-1}^\theta < t_{i+1}^- \\ 0 & \text{otherwise} \end{cases}$$

denote the duration of overlap between $(t_i^+, t_{i+1}^-]$ and $(\tau_{\ell-1}^\theta, \tau_\ell^\theta]$, and between $(t_i^+, t_{i+1}^-]$ and $(\tau_{\ell-1}^\theta, \tau]$. Similarly, for any $i = 1,\ldots,k-1$ and $\tau \in (\tau_{\ell-1}^\theta, T]$, let

$$T_{i,\ell}^{\theta-} := \begin{cases} \min\{t_i^+, \tau_\ell^\theta\} - \max\{t_i^-, \tau_{\ell-1}^\theta\} & \text{if } \tau_\ell^\theta > t_i^- \text{ and } \tau_{\ell-1}^\theta < t_i^+ \\ 0 & \text{otherwise} \end{cases}$$

$$T_{i,\ell}^{\theta-}(\tau) := \begin{cases} \min\{t_i^+, \tau\} - \max\{t_i^-, \tau_{\ell-1}^\theta\} & \text{if } \tau > t_i^- \text{ and } \tau_{\ell-1}^\theta < t_i^+ \\ 0 & \text{otherwise} \end{cases}$$

denote the duration of overlap between $(t_i^-, t_i^+]$ and $(\tau_{\ell-1}^\theta, \tau_\ell^\theta]$, and between $(t_i^-, t_i^+]$ and $(\tau_{\ell-1}^\theta, \tau]$, and let

$$T_{k,\ell}^{\theta-} := \begin{cases} \tau_\ell^\theta - \max\{t_k^-, \tau_{\ell-1}^\theta\} & \text{if } \tau_\ell^\theta > t_k^- \\ 0 & \text{otherwise} \end{cases}$$

$$T_{k,\ell}^{\theta-}(\tau) := \begin{cases} \tau - \max\{t_k^-, \tau_{\ell-1}^\theta\} & \text{if } \tau > t_k^- \\ 0 & \text{otherwise} \end{cases}$$

denote the duration of overlap between $(t_k^-, T]$ and $(\tau_{\ell-1}^\theta, \tau_\ell^\theta]$, and between $(t_k^-, T]$ and $(\tau_{\ell-1}^\theta, \tau]$. Specifically, let $\tau_0^\theta := 0$, and let $S_1^\theta := S_1$ denote the first assortment under the partitioned booking limit policy. For each $\ell \in \{1, 2, \ldots\}$ such that $\tau_{\ell-1}^\theta < T$ and $S_\ell^\theta \neq \varnothing$, let

$$\begin{aligned} \tau_\ell^\theta \;:=\; \min\Bigg\{T,\ \min\Bigg\{\inf\Bigg\{\tau \in (\tau_{\ell-1}^\theta, T] \,:\, &\sum_{\ell'=1}^{\ell-1}\left[\sum_{i=0}^{k-1} \hat{N}_j^{i+}\left(\theta\lambda P_{j:S_{\ell'}^\theta} T_{i,\ell'}^{\theta+}\right) + \sum_{i=1}^{k} \hat{N}_j^{i-}\left(\theta\lambda P_{j:S_{\ell'}^\theta} T_{i,\ell'}^{\theta-}\right)\right] \\ &+ \left[\sum_{i=0}^{k-1} \hat{N}_j^{i+}\left(\theta\lambda P_{j:S_\ell^\theta} T_{i,\ell}^{\theta+}(\tau)\right) + \sum_{i=1}^{k} \hat{N}_j^{i-}\left(\theta\lambda P_{j:S_\ell^\theta} T_{i,\ell}^{\theta-}(\tau)\right)\right] = \theta b_j^*\Bigg\} \,:\, j \in S_\ell^\theta\Bigg\}\Bigg\} \end{aligned}$$

denote the last time that assortment S_ℓ^θ is offered (with the convention that $\inf\varnothing = \infty$). If $\tau_\ell^\theta < T$, then let j_ℓ^θ be the (unique) $j \in S_\ell^\theta$ such that

$$\begin{aligned} &\sum_{\ell'=1}^{\ell-1}\left[\sum_{i=0}^{k-1} \hat{N}_j^{i+}\left(\theta\lambda P_{j:S_{\ell'}^\theta} T_{i,\ell'}^{\theta+}\right) \;+\; \sum_{i=1}^{k} \hat{N}_j^{i-}\left(\theta\lambda P_{j:S_{\ell'}^\theta} T_{i,\ell'}^{\theta-}\right)\right] \\ &\qquad + \left[\sum_{i=0}^{k-1} \hat{N}_j^{i+}\left(\theta\lambda P_{j:S_\ell^\theta} T_{i,\ell}^{\theta+}(\tau_\ell^\theta)\right) \;+\; \sum_{i=1}^{k} \hat{N}_j^{i-}\left(\theta\lambda P_{j:S_\ell^\theta} T_{i,\ell}^{\theta-}(\tau_\ell^\theta)\right)\right] \;=\; \theta b_j^*, \end{aligned}$$

that is, j_ℓ^θ denotes the first product in S_ℓ^θ that sells out. Then $S_{\ell+1}^\theta := S_\ell^\theta \setminus \{j_\ell^\theta\}$.

Let $N_j^\theta(t)$ denote the quantity of product j sold up to time t under the partitioned booking limit policy, that is, for $t \in (\tau_{\ell-1}^\theta, \tau_\ell^\theta]$,

$$N_j^\theta(t) \;:=\; \sum_{\ell'=1}^{\ell-1}\left[\sum_{i=0}^{k-1}\hat{N}_j^{i+}\left(\theta\lambda P_{j:S_{\ell'}^\theta}T_{i,\ell'}^{\theta+}\right) \;+\; \sum_{i=1}^{k}\hat{N}_j^{i-}\left(\theta\lambda P_{j:S_{\ell'}^\theta}T_{i,\ell'}^{\theta-}\right)\right] + \left[\sum_{i=0}^{k-1}\hat{N}_j^{i+}\left(\theta\lambda P_{j:S_\ell^\theta}T_{i,\ell}^{\theta+}(t)\right) \;+\; \sum_{i=1}^{k}\hat{N}_j^{i-}\left(\theta\lambda P_{j:S_\ell^\theta}T_{i,\ell}^{\theta-}(t)\right)\right]$$

Thus, if $\tau_\ell^\theta < T$, then τ_ℓ^θ satisfies

$$\tau_\ell^\theta \;=\; \min\left\{\inf\left\{\tau > \tau_{\ell-1}^\theta \,:\, N_j^\theta(\tau) = \theta b_j^*\right\} \,:\, j \in S_\ell^\theta\right\} \;=\; \inf\left\{\tau > \tau_{\ell-1}^\theta \,:\, N_{j_\ell^\theta}^\theta(\tau) = \theta b_{j_\ell^\theta}^*\right\}.$$

Next we define the following events for each θ and each $i = 1, \ldots, k$:

$$\begin{aligned}
E_i^{\theta-} &:= \bigcap_{j\in S_i}\left\{\sum_{i'=1}^{i}\theta\lambda_j^{i'}(t_{i'}-t_{i'-1}) - \theta\eta_j^i\varepsilon \;<\; N_j^\theta(t_i^-) \;<\; \sum_{i'=1}^{i}\theta\lambda_j^{i'}(t_{i'}-t_{i'-1})\right\}\\
E_i^{\theta+} &:= \bigcap_{j\in S_i}\left\{\sum_{i'=1}^{i}\theta\lambda_j^{i'}(t_{i'}-t_{i'-1}) \;\le\; N_j^\theta(t_i^+) \;\le\; \sum_{i'=1}^{i}\theta\lambda_j^{i'}(t_{i'}-t_{i'-1}) + \theta\zeta_j^i\varepsilon\right\}\\
F_i^{\theta+} &:= \cap_{i'=1}^{i}\left(E_{i'}^{\theta-}\cap E_{i'}^{\theta+}\right)\\
F_i^{\theta-} &:= F_{i-1}^{\theta+}\cap E_i^{\theta-}
\end{aligned}$$

Similarly, let $F_0^{\theta+} = E_0^{\theta+} := \cap_{j\in\mathcal{J}}\left\{0 \le N_j^\theta(t_0^+) \le \theta\zeta_j^0\varepsilon\right\}$. That is, $E_i^{\theta-}$ is the event that for all products in S_i, the booking quantity at time t_i^- is slightly below the CDLP quantity at time t_i, and $E_i^{\theta+}$ is the event that for all products in S_i, the booking quantity at time t_i^+ is equal to or slightly above the CDLP quantity at time t_i. Note that, because the booking limit of product $j \in S_i \setminus S_{i+1}$ is $\theta b_j^* = \theta\sum_{i'=1}^{i}\lambda_j^{i'}(t_{i'}-t_{i'-1})$, and the booking limit of product $j \in S_{i+1}$ is at least $\theta\sum_{i'=1}^{i+1}\lambda_j^{i'}(t_{i'}-t_{i'-1})$, the event $F_{i+1}^{\theta-} := F_i^{\theta+}\cap E_{i+1}^{\theta-}$ implies that $S^\theta(t) = S_{i+1}$ for all $t \in (t_i^+, t_{i+1}^-]$; i.e., the assortment S_{i+1} is offered during $(t_i^+, t_{i+1}^-]$.

Lemma EC.7. *For all $i = 0, \ldots, k-1$, the following holds:*

$$\begin{aligned}
&F_i^{\theta+}\bigcap_{j\in S_{i+1}}\left\{\theta\lambda_j^{i+1}(t_{i+1}-t_i) - \theta\eta_j^{i+1}\varepsilon \,<\, \hat{N}_j^{i+}(\theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+)) \,<\, \theta\lambda_j^{i+1}(t_{i+1}-t_i) - \theta\zeta_j^i\varepsilon\right\}\\
\subset\;\; &F_i^{\theta+}\bigcap_{j\in S_{i+1}}\left\{\sum_{i'=1}^{i+1}\theta\lambda_j^{i'}(t_{i'}-t_{i'-1}) - \theta\eta_j^{i+1}\varepsilon \;<\; N_j^\theta(t_{i+1}^-) \;<\; \sum_{i'=1}^{i+1}\theta\lambda_j^{i'}(t_{i'}-t_{i'-1})\right\}.
\end{aligned}$$

Proof. Consider any sample path in the event on the left. We show that the sample path is in the event on the right. Recall that since the sample path is in $F_i^{\theta+}$, it holds that $S^\theta(t_i^+) = S_{i+1}$. Let ℓ be such that $S^\theta(t_i^+) = S_{i+1} = S_\ell^\theta$.

We show by contradiction that, for the considered sample path, no product can reach its booking limit during $[t_i^+, t_{i+1}^-)$. Suppose that $\tau_\ell^\theta \in [t_i^+, t_{i+1}^-)$. Thus $j_\ell^\theta \in S_{i+1}$ satisfies $N^\theta_{j_\ell^\theta}(\tau_\ell^\theta) = \theta b^*_{j_\ell^\theta}$. Then

$$\begin{aligned}
N^\theta_{j_\ell^\theta}(\tau_\ell^\theta) &= N^\theta_{j_\ell^\theta}(t_i^+) + \hat{N}^{i+}_{j_\ell^\theta}(\theta\lambda^{i+1}_{j_\ell^\theta}(\tau_\ell^\theta - t_i^+)) \\
&\le \left(\sum_{i'=1}^{i} \theta\lambda^{i'}_{j_\ell^\theta}(t_{i'} - t_{i'-1}) + \theta\zeta^i_{j_\ell^\theta}\varepsilon\right) + \hat{N}^{i+}_{j_\ell^\theta}(\theta\lambda^{i+1}_{j_\ell^\theta}(\tau_\ell^\theta - t_i^+)) \\
&\le \left(\sum_{i'=1}^{i} \theta\lambda^{i'}_{j_\ell^\theta}(t_{i'} - t_{i'-1}) + \theta\zeta^i_{j_\ell^\theta}\varepsilon\right) + \hat{N}^{i+}_{j_\ell^\theta}(\theta\lambda^{i+1}_{j_\ell^\theta}(t_{i+1}^- - t_i^+)) \\
&< \left(\sum_{i'=1}^{i} \theta\lambda^{i'}_{j_\ell^\theta}(t_{i'} - t_{i'-1}) + \theta\zeta^i_{j_\ell^\theta}\varepsilon\right) + \left(\theta\lambda^{i+1}_{j_\ell^\theta}(t_{i+1} - t_i) - \theta\zeta^i_{j_\ell^\theta}\varepsilon\right) \le \theta b^*_{j_\ell^\theta}
\end{aligned}$$

The first inequality holds because the sample path is in $F_i^{\theta+}$, the second inequality holds because $\tau_\ell^\theta < t_{i+1}^-$, the third inequality holds because the sample path is in the event on the left, and the fourth inequality holds because $j_\ell^\theta \in S_{i+1}$. This contradicts $N^\theta_{j_\ell^\theta}(\tau_\ell^\theta) = \theta b^*_{j_\ell^\theta}$. Thus $S^\theta(t) = S_{i+1}$ for all $t \in [t_i^+, t_{i+1}^-]$.

Thus, for each $j \in S_{i+1}$ it holds that

$$\begin{aligned}
N_j^\theta(t_{i+1}^-) &= N_j^\theta(t_i^+) + \hat{N}_j^{i+}(\theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+)) \\
&\ge \sum_{i'=1}^{i} \theta\lambda_j^{i'}(t_{i'} - t_{i'-1}) + \hat{N}_j^{i+}(\theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+)) > \sum_{i'=1}^{i+1} \theta\lambda_j^{i'}(t_{i'} - t_{i'-1}) - \theta\eta_j^{i+1}\varepsilon.
\end{aligned}$$

The inequalities hold because the sample path is in the event on the left. Also,

$$\begin{aligned}
N_j^\theta(t_{i+1}^-) &= N_j^\theta(t_i^+) + \hat{N}_j^{i+}(\theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+)) \\
&\le \left(\sum_{i'=1}^{i} \theta\lambda_j^{i'}(t_{i'} - t_{i'-1}) + \theta\zeta_j^i\varepsilon\right) + \hat{N}_j^{i+}(\theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+)) \\
&< \left(\sum_{i'=1}^{i} \theta\lambda_j^{i'}(t_{i'} - t_{i'-1}) + \theta\zeta_j^i\varepsilon\right) + \left(\theta\lambda_j^{i+1}(t_{i+1} - t_i) - \theta\zeta_j^i\varepsilon\right) \\
&= \sum_{i'=1}^{i+1} \theta\lambda_j^{i'}(t_{i'} - t_{i'-1}).
\end{aligned}$$

Therefore the sample path is in the event on the right. $\square$

LEMMA EC.8. *For all $i = 1, \ldots, k$, the following holds:*

$$\begin{aligned}
& F_i^{\theta-} \bigcap_{j \in S_i} \left\{\theta\eta_j^i\varepsilon \le \hat{N}_j^{i-}\left(\theta\underline{\lambda}_j^i(t_i^+ - t_i^-)\right)\right\} \bigcap_{j \in S_{i+1}} \left\{\hat{N}_j^{i-}\left(\theta\overline{\lambda}_j^i(t_i^+ - t_i^-)\right) \le \theta\zeta_j^i\varepsilon\right\} \\
\subset\ & F_i^{\theta-} \bigcap_{j \in S_i \setminus S_{i+1}} \left\{\sum_{i'=1}^{i} \theta\lambda_j^{i'}(t_{i'} - t_{i'-1}) - N_j^\theta(t_i^-) = \hat{N}_j^{i-}\left(\theta\lambda \int_{\tau \in (t_i^-, t_i^+]} P_{j:S^\theta(\tau)} d\tau\right)\right\} \\
& \bigcap_{j \in S_{i+1}} \left\{\theta\eta_j^i\varepsilon \le \hat{N}_j^{i-}\left(\theta\lambda \int_{\tau \in (t_i^-, t_i^+]} P_{j:S^\theta(\tau)} d\tau\right) \le \theta\zeta_j^i\varepsilon\right\}.
\end{aligned}$$

Proof. Consider any sample path in the event on the left. We show that the sample path is in the event on the right. Recall that since the sample path is in $F_i^{\theta-}$, it holds that $S^\theta(t_i^-) = S_i$. Let ℓ be such that $S^\theta(t_i^-) = S_i = S_\ell^\theta$.

First we show by contradiction that it cannot hold that $S^\theta(t) = S_i$ for all $t \in (t_i^-, t_i^+]$, that is, at least one product $j \in S_i$ must reach its booking limit during $(t_i^-, t_i^+]$. Suppose that $S^\theta(t) = S_i$ for all $t \in (t_i^-, t_i^+]$. Consider any $j \in S_i \setminus S_{i+1}$. Then

$$
\begin{aligned}
N_j^\theta(t_i^+) &= N_j^\theta(t_i^-) + \hat{N}_j^{i-}\left(\theta\lambda P_{j:S_i}(t_i^+ - t_i^-)\right) \\
&\geq N_j^\theta(t_i^-) + \hat{N}_j^{i-}\left(\theta\underline{\lambda}_j^i(t_i^+ - t_i^-)\right) \;>\; \left(\sum_{i'=1}^{i}\theta\lambda_j^{i'}(t_{i'} - t_{i'-1}) - \theta\eta_j^i\varepsilon\right) + \theta\eta_j^i\varepsilon \;=\; \theta b_j^*.
\end{aligned}
$$

The first inequality follows because the definition of $\underline{\lambda}_j^i$ implies that $\lambda P_{j:S_i} \geq \underline{\lambda}_j^i$, and the second inequality follows from the definition of the event on the left. Under the partitioned booking limit policy it cannot hold that $N_j^\theta(t_i^+) > \theta b_j^*$, and therefore for any sample path in the event on the left it cannot hold that $S^\theta(t) = S_i$ for all $t \in (t_i^-, t_i^+]$. Thus, $\tau_\ell^\theta \in [t_i^-, t_i^+)$.

Next we show that $j_\ell^\theta \notin S_{i+1}$, that is, a product $j \in S_{i+1}$ cannot be the first to reach its booking limit during $(t_i^-, t_i^+]$. Consider any $j \in S_{i+1}$. Then

$$
\begin{aligned}
N_j^\theta(\tau_\ell^\theta) &= N_j^\theta(t_i^-) + \hat{N}_j^{i-}\left(\theta\lambda P_{j:S_i}(\tau_\ell^\theta - t_i^-)\right) \\
&\leq N_j^\theta(t_i^-) + \hat{N}_j^{i-}\left(\theta\overline{\lambda}_j^i(\tau_\ell^\theta - t_i^-)\right) \\
&\leq N_j^\theta(t_i^-) + \hat{N}_j^{i-}\left(\theta\lambda P_{j:S_i}(t_i^+ - t_i^-)\right) \\
&< \sum_{i'=1}^{i}\theta\lambda_j^{i'}(t_{i'} - t_{i'-1}) + \theta\zeta_j^i\varepsilon \\
&< \sum_{i'=1}^{i}\theta\lambda_j^{i'}(t_{i'} - t_{i'-1}) + \theta\lambda_j^{i+1}(t_{i+1} - t_i) \;\leq\; \theta b_j^*.
\end{aligned}
$$

The first inequality follows from the definition of $\overline{\lambda}_j^i$, the second inequality follows since $\tau_\ell^\theta < t_i^+$, the third inequality follows from the definition of the event on the left, the fourth inequality follows from the assumption that $\varepsilon > 0$ is sufficiently small such that $\zeta_j^i\varepsilon < \lambda_j^{i+1}(t_{i+1} - t_i)$ for all $i = 1, \ldots, k-1$ and all $j \in S_{i+1}$, and the fifth inequality follows from $j \in S_{i+1}$. Thus $j_\ell^\theta \in S_i \setminus S_{i+1}$, and hence $S_{\ell+1}^\theta = S_\ell^\theta \setminus \{j_\ell^\theta\}$ satisfies $S_{i+1} \subset S_{\ell+1}^\theta \subset S_i$.

Next, we continue by induction on ℓ. Suppose that for some $\tilde{\ell} \geq \ell$ it holds that $\tau_{\tilde{\ell}}^\theta < t_i^+$ and $S_{i+1} \subset S_{\tilde{\ell}+1}^\theta \subset S_i$. We consider two cases: either $S_{\tilde{\ell}+1}^\theta \setminus S_{i+1} \neq \varnothing$ or $S_{\tilde{\ell}+1}^\theta \setminus S_{i+1} = \varnothing$.
Case $S_{\tilde{\ell}+1}^\theta \setminus S_{i+1} \neq \varnothing$: In this case we repeat the argument above. First we show by contradiction that it cannot hold that $S^\theta(t) = S_{\tilde{\ell}+1}^\theta$ for all $t \in (\tau_{\tilde{\ell}}^\theta, t_i^+]$, that is, at least one product $j \in S_{\tilde{\ell}+1}^\theta$ must

reach its booking limit during $(\tau^\theta_{\tilde{\ell}}, t_i^+]$. Suppose that $S^\theta(t) = S^\theta_{\tilde{\ell}+1}$ for all $t \in (\tau^\theta_{\tilde{\ell}}, t_i^+]$. Consider any $j \in S^\theta_{\tilde{\ell}+1} \setminus S_{i+1}$. Note that $j \in S_i \setminus S_{i+1}$. Then

$$\begin{aligned}
N_j^\theta(t_i^+) &= N_j^\theta(t_i^-) + \hat{N}_j^{i-}\left(\theta\lambda P_{j:S_i}(\tau^\theta_\ell - t_i^-)\right) \\
&\qquad + \sum_{\ell'=\ell+1}^{\tilde{\ell}} \hat{N}_j^{i-}\left(\theta\lambda P_{j:S^\theta_{\ell'}}(\tau^\theta_{\ell'} - \tau^\theta_{\ell'-1})\right) + \hat{N}_j^{i-}\left(\theta\lambda P_{j:S^\theta_{\tilde{\ell}+1}}(t_i^+ - \tau^\theta_{\tilde{\ell}})\right) \\
&\geq N_j^\theta(t_i^-) + \hat{N}_j^{i-}\left(\theta\underline{\lambda}_j^i(t_i^+ - t_i^-)\right) \\
&> \left(\sum_{i'=1}^{i} \theta\lambda_j^{i'}(t_{i'} - t_{i'-1}) - \theta\eta_j^i\varepsilon\right) + \theta\eta_j^i\varepsilon \quad = \quad \theta b_j^*.
\end{aligned}$$

The first inequality holds because $S_{i+1} \subset S^\theta_{\ell'} \subset S_i$ and thus $\lambda P_{j:S^\theta_{\ell'}} \geq \underline{\lambda}_j^i$ for all $\ell' = \ell, \ldots, \tilde{\ell}+1$, and the second inequality follows from the definition of the event on the left. Under the partitioned booking limit policy it cannot hold that $N_j^\theta(t_i^+) > \theta b_j^*$, and therefore it cannot hold that $S^\theta(t) = S^\theta_{\tilde{\ell}+1}$ for all $t \in (\tau^\theta_{\tilde{\ell}}, t_i^+]$. Thus, $\tau^\theta_{\tilde{\ell}+1} \in (\tau^\theta_{\tilde{\ell}}, t_i^+)$.

Next we show that $j^\theta_{\tilde{\ell}+1} \notin S_{i+1}$, that is, a product $j \in S_{i+1}$ cannot be the first to reach its booking limit during $(\tau^\theta_{\tilde{\ell}}, t_i^+]$. Consider any $j \in S_{i+1}$. Then

$$\begin{aligned}
N_j^\theta(\tau^\theta_{\tilde{\ell}+1}) &= N_j^\theta(t_i^-) + \hat{N}_j^{i-}\left(\theta\lambda P_{j:S_i}(\tau^\theta_\ell - t_i^-)\right) + \sum_{\ell'=\ell+1}^{\tilde{\ell}+1} \hat{N}_j^{i-}\left(\theta\lambda P_{j:S^\theta_{\ell'}}(\tau^\theta_{\ell'} - \tau^\theta_{\ell'-1})\right) \\
&\leq N_j^\theta(t_i^-) + \hat{N}_j^{i-}\left(\theta\overline{\lambda}_j^i(\tau^\theta_{\tilde{\ell}+1} - t_i^-)\right) \\
&\leq N_j^\theta(t_i^-) + \hat{N}_j^{i-}\left(\theta\overline{\lambda}_j^i(t_i^+ - t_i^-)\right) \\
&< \sum_{i'=1}^{i} \theta\lambda_j^{i'}(t_{i'} - t_{i'-1}) + \theta\zeta_j^i\varepsilon \\
&< \sum_{i'=1}^{i} \theta\lambda_j^{i'}(t_{i'} - t_{i'-1}) + \theta\lambda_j^{i+1}(t_{i+1} - t_i) \quad \leq \quad \theta b_j^*.
\end{aligned}$$

The first inequality holds because $S_{i+1} \subset S^\theta_{\ell'} \subset S_i$ and thus $\lambda P_{j:S^\theta_{\ell'}} \leq \overline{\lambda}_j^i$ for all $\ell' = \ell, \ldots, \tilde{\ell}+1$, the second inequality holds since $\tau^\theta_{\tilde{\ell}+1} < t_i^+$, the third inequality follows from the definition of the event on the left, the fourth inequality follows from the assumption that $\varepsilon > 0$ is sufficiently small such that $\zeta_j^i\varepsilon < \lambda_j^{i+1}(t_{i+1} - t_i)$ for all $i = 1, \ldots, k-1$ and all $j \in S_{i+1}$, and the fifth inequality follows from $j \in S_{i+1}$. Thus $j^\theta_{\tilde{\ell}+1} \in S^\theta_{\tilde{\ell}+1} \setminus S_{i+1}$, and hence $S^\theta_{\tilde{\ell}+2} = S^\theta_{\tilde{\ell}+1} \setminus \{j^\theta_{\tilde{\ell}+1}\}$ satisfies $S_{i+1} \subset S^\theta_{\tilde{\ell}+2} \subset S_i$. Hence, in the case $S^\theta_{\tilde{\ell}+1} \setminus S_{i+1} \neq \varnothing$, the induction continues.

Case $S^\theta_{\tilde{\ell}+1} \setminus S_{i+1} = \varnothing$: Then $S^\theta_{\tilde{\ell}+1} = S_{i+1}$. That is, for each $j \in S_i \setminus S_{i+1}$, it holds that $N_j^\theta(t_i^+) = N_j^\theta(\tau^\theta_{\tilde{\ell}}) = \theta b_j^* = \sum_{i'=1}^{i} \theta\lambda_j^{i'}(t_{i'} - t_{i'-1})$, and thus $N_j^\theta(t_i^-) + \hat{N}_j^{i-}\left(\theta\lambda\int_{\tau\in(t_i^-, t_i^+]} P_{j:S^\theta(\tau)} d\tau\right) = \sum_{i'=1}^{i} \theta\lambda_j^{i'}(t_{i'} - t_{i'-1})$.

Next we show by contradiction that $S^\theta(t) = S_{i+1}$ for all $t \in (\tau^\theta_{\tilde{\ell}}, t_i^+]$. Suppose that $\tau^\theta_{\tilde{\ell}+1} < t_i^+$. Thus $j^\theta_{\tilde{\ell}+1} \in S_{i+1}$ satisfies $N^\theta_{j^\theta_{\tilde{\ell}+1}}(\tau^\theta_{\tilde{\ell}+1}) = \theta b^*_{j^\theta_{\tilde{\ell}+1}}$. Then

$$N^\theta_{j^\theta_{\tilde{\ell}+1}}(\tau^\theta_{\tilde{\ell}+1}) = N^\theta_{j^\theta_{\tilde{\ell}+1}}(t_i^-) + \hat{N}^{i-}_{j^\theta_{\tilde{\ell}+1}}\left(\theta\lambda P_{j^\theta_{\tilde{\ell}+1}:S_i}(\tau^\theta_\ell - t_i^-)\right) + \sum_{\ell'=\ell+1}^{\tilde{\ell}+1} \hat{N}^{i-}_{j^\theta_{\tilde{\ell}+1}}\left(\theta\lambda P_{j^\theta_{\tilde{\ell}+1}:S^\theta_{\ell'}}(\tau^\theta_{\ell'} - \tau^\theta_{\ell'-1})\right)$$

$$
\begin{aligned}
&\le\ N^{\theta}_{j^{\theta}_{\tilde\ell+1}}(t_i^-) + \hat N^{i-}_{j^{\theta}_{\tilde\ell+1}}\left(\theta\overline{\lambda}^i_{j^{\theta}_{\tilde\ell+1}}(\tau^{\theta}_{\tilde\ell+1} - t_i^-)\right) \\
&\le\ N^{\theta}_{j^{\theta}_{\tilde\ell+1}}(t_i^-) + \hat N^{i-}_{j^{\theta}_{\tilde\ell+1}}\left(\theta\overline{\lambda}^i_{j^{\theta}_{\tilde\ell+1}}(t_i^+ - t_i^-)\right) \\
&<\ \sum_{i'=1}^{i}\theta\lambda^{i'}_{j^{\theta}_{\tilde\ell+1}}(t_{i'} - t_{i'-1}) + \theta\zeta^i_{j^{\theta}_{\tilde\ell+1}}\varepsilon \\
&<\ \sum_{i'=1}^{i}\theta\lambda^{i'}_{j^{\theta}_{\tilde\ell+1}}(t_{i'} - t_{i'-1}) + \theta\lambda^{i+1}_{j^{\theta}_{\tilde\ell+1}}(t_{i+1} - t_i)\ \ \le\ \ \theta b^*_{j^{\theta}_{\tilde\ell+1}}.
\end{aligned}
$$

The first inequality holds because $S_{i+1} \subset S^{\theta}_{\ell'} \subset S_i$ and thus $\lambda P_{j^{\theta}_{\tilde\ell+1}:S^{\theta}_{\ell'}} \le \overline{\lambda}^i_{j^{\theta}_{\tilde\ell+1}}$ for all $\ell' = \ell, \ldots, \tilde\ell + 1$, the second inequality holds since $\tau^{\theta}_{\tilde\ell+1} < t_i^+$, the third inequality follows from the definition of the event on the left, the fourth inequality follows from the assumption that $\varepsilon > 0$ is sufficiently small such that $\zeta^i_j\varepsilon < \lambda^{i+1}_j(t_{i+1} - t_i)$ for all $i = 1, \ldots, k-1$ and all $j \in S_{i+1}$, and the fifth inequality follows from $j^{\theta}_{\tilde\ell+1} \in S_{i+1}$. This contradicts $N^{\theta}_{j^{\theta}_{\tilde\ell+1}}(\tau^{\theta}_{\tilde\ell+1}) = \theta b^*_{j^{\theta}_{\tilde\ell+1}}$.

Thus, for each $\tau \in [t_i^-, t_i^+]$ it holds that $S_{i+1} \subset S^{\theta}(\tau) \subset S_i$. Hence, for each $j \in S_{i+1}$, it holds that $\hat N^{i-}_j\left(\theta\underline{\lambda}^i_j(t_i^+ - t_i^-)\right) \le \hat N^{i-}_j\left(\theta\lambda\int_{\tau\in(t_i^-,t_i^+]}P_{j:S^{\theta}(\tau)}d\tau\right) \le \hat N^{i-}_j\left(\theta\overline{\lambda}^i_j(t_i^+ - t_i^-)\right)$. Thus it follows from the event on the left that $\theta\eta^i_j\varepsilon\ \le\ \hat N^{i-}_j\left(\theta\lambda\int_{\tau\in(t_i^-,t_i^+]}P_{j:S^{\theta}(\tau)}d\tau\right)\ \le\ \theta\zeta^i_j\varepsilon$. Thereby it has been established that the sample path is in the event on the right. ☐

We will use the following concentration inequality for Poisson process (Draief and Massouli 2010, Proposition 5.3):

Lemma EC.9 **(Poisson tail bound)**. *Let $\{N(t), t \ge 0\}$ be a Poisson process with unit rate. For any $x > 0$, $t > 0$, $\varepsilon > 0$, it holds that*

$$
\mathsf{P}[N(xt) - xt \ge x\varepsilon]\ \le\ \exp\big(-xth(\varepsilon/t)\big) \quad \text{and} \quad \mathsf{P}[N(xt) - xt \le -x\varepsilon]\ \le\ \exp\big(-xth(\varepsilon/t)\big),
$$

where $h(y) := (1+y)\log(1+y) - y$. (Note that $h(y) > 0$ for all $y > 0$.)

For each $i = 1, \ldots, k$, let

$$
\begin{aligned}
\delta^{\theta-}_i\ &:=\ 2\sum_{j\in S_i}\exp\left(-\theta\lambda^i_j(t_i^- - t^+_{i-1})\ h\left(\frac{\varepsilon}{t_i^- - t^+_{i-1}}\right)\right), \\
\delta^{\theta+}_i\ &:=\ \sum_{j\in S_i}\exp\left(-\theta\underline{\lambda}^i_j(t_i^+ - t_i^-)\ h\left(\frac{\varepsilon}{t_i^+ - t_i^-}\right)\right) + \sum_{j\in S_i}\exp\left(-\theta\overline{\lambda}^i_j(t_i^+ - t_i^-)\ h\left(\frac{\varepsilon}{t_i^+ - t_i^-}\right)\right).
\end{aligned}
$$

Next, we prove by induction on i that

$$
\mathsf{P}\left[F^{\theta+}_i\right]\ \ge\ 1 - \sum_{i'=1}^{i}\left(\delta^{\theta-}_{i'} + \delta^{\theta+}_{i'}\right). \tag{EC.10}
$$

Base Case: $i = 0$**.** Since $t_0^+ = 0$ and $\zeta^0_j = 0$ for all $j \in \mathcal{J}$, it holds that $0 \le N^{\theta}_j(t_0^+) \le \theta\zeta^0_j$, so $\mathsf{P}[F^{\theta+}_0] = 1$.

Induction Step: from i to $i+1$. Suppose the result holds for $i \in \{0,\dots,k-1\}$. Then

$$
\begin{aligned}
&\mathsf{P}\left[F_{i+1}^{\theta-} \mid F_i^{\theta+}\right] \;=\; \mathsf{P}\left[E_{i+1}^{\theta-} \mid F_i^{\theta+}\right] \\
=\;& \mathsf{P}\left[\bigcap_{j\in S_{i+1}} \left\{\sum_{i'=1}^{i+1} \theta\lambda_j^{i'}(t_{i'}-t_{i'-1}) - \theta\eta_j^{i+1}\varepsilon \;<\; N_j^{\theta}(t_{i+1}^-) \;<\; \sum_{i'=1}^{i+1}\theta\lambda_j^{i'}(t_{i'}-t_{i'-1})\right\} \mid F_i^{\theta+}\right] \\
\geq\;& \mathsf{P}\left[\cap_{j\in S_{i+1}} \left\{\theta\lambda_j^{i+1}(t_{i+1}-t_i) - \theta\eta_j^{i+1}\varepsilon < \hat{N}_j^{i+}(\theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+)) < \theta\lambda_j^{i+1}(t_{i+1}-t_i) - \theta\zeta_j^i\varepsilon\right\} \mid F_i^{\theta+}\right] \\
=\;& \mathsf{P}\left[\cap_{j\in S_{i+1}} \left\{\theta\lambda_j^{i+1}(t_{i+1}-t_i) - \theta\eta_j^{i+1}\varepsilon < \hat{N}_j^{i+}(\theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+)) < \theta\lambda_j^{i+1}(t_{i+1}-t_i) - \theta\zeta_j^i\varepsilon\right\}\right] \\
=\;& 1 - \mathsf{P}\Big[\cup_{j\in S_{i+1}} \Big(\left\{\theta\lambda_j^{i+1}(t_{i+1}-t_i) - \theta\eta_j^{i+1}\varepsilon \geq \hat{N}_j^{i+}(\theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+))\right\} \\
&\qquad\qquad \cup \left\{\hat{N}_j^{i+}(\theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+)) \geq \theta\lambda_j^{i+1}(t_{i+1}-t_i) - \theta\zeta_j^i\varepsilon\right\}\Big)\Big] \\
\geq\;& 1 - \sum_{j\in S_{i+1}} \mathsf{P}\left[\hat{N}_j^{i+}(\theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+)) \leq \theta\lambda_j^{i+1}(t_{i+1}-t_i) - \theta\eta_j^{i+1}\varepsilon\right] \\
&\qquad - \sum_{j\in S_{i+1}} \mathsf{P}\left[\hat{N}_j^{i+}(\theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+)) \geq \theta\lambda_j^{i+1}(t_{i+1}-t_i) - \theta\zeta_j^i\varepsilon\right] \\
=\;& 1 - \sum_{j\in S_{i+1}} \mathsf{P}\left[\hat{N}_j^{i+}(\theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+)) \leq \theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+) + \theta\lambda_j^{i+1}(\nu_{i+1}^- + \nu_i^+)\varepsilon - \theta\eta_j^{i+1}\varepsilon\right] \\
&\qquad - \sum_{j\in S_{i+1}} \mathsf{P}\left[\hat{N}_j^{i+}(\theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+)) \geq \theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+) + \theta\lambda_j^{i+1}(\nu_{i+1}^- + \nu_i^+)\varepsilon - \theta\zeta_j^i\varepsilon\right] \\
\geq\;& 1 - \sum_{j\in S_{i+1}} \Big\{\mathsf{P}\left[\hat{N}_j^{i+}(\theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+)) - \theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+) \leq -\theta\lambda_j^{i+1}\varepsilon\right] \\
&\qquad + \mathsf{P}\left[\hat{N}_j^{i+}(\theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+)) - \theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+) \geq \theta\lambda_j^{i+1}\varepsilon\right]\Big\} \\
\geq\;& 1 - 2\sum_{j\in S_{i+1}} \exp\left(-\theta\lambda_j^{i+1}(t_{i+1}^- - t_i^+)h\left(\frac{\varepsilon}{t_{i+1}^- - t_i^+}\right)\right) \;=\; 1 - \delta_{i+1}^{\theta-}.
\end{aligned}
$$

The first inequality follows from Lemma EC.7, the third equality holds because $\hat{N}_j^{i+}$ and $F_i^{\theta+}$ are independent, the second inequality applies the union bound, the third inequality follows from (EC.9), and the fourth inequality follows from Lemma EC.9.

Next we consider the conditional probability of $F_i^{\theta+}$ given $F_i^{\theta-}$.

$$
\begin{aligned}
&\mathsf{P}\left[F_i^{\theta+} \mid F_i^{\theta-}\right] \;=\; \mathsf{P}\left[E_i^{\theta+} \mid F_i^{\theta-}\right] \\
=\;& \mathsf{P}\left[\bigcap_{j\in S_i} \left\{\sum_{i'=1}^{i} \theta\lambda_j^{i'}(t_{i'}-t_{i'-1}) \;\leq\; N_j^{\theta}(t_i^+) \;\leq\; \sum_{i'=1}^{i}\theta\lambda_j^{i'}(t_{i'}-t_{i'-1}) + \theta\zeta_j^i\varepsilon\right\} \mid F_i^{\theta-}\right] \\
=\;& \mathsf{P}\Big[\bigcap_{j\in S_i\setminus S_{i+1}} \left\{\sum_{i'=1}^{i} \theta\lambda_j^{i'}(t_{i'}-t_{i'-1}) - N_j^{\theta}(t_i^-) \;=\; N_j^{\theta}(t_i^+) - N_j^{\theta}(t_i^-)\right\} \\
&\qquad \bigcap_{j\in S_{i+1}} \Big\{\sum_{i'=1}^{i} \theta\lambda_j^{i'}(t_{i'}-t_{i'-1}) - N_j^{\theta}(t_i^-) \;\leq\; N_j^{\theta}(t_i^+) - N_j^{\theta}(t_i^-) \\
&\qquad\qquad \leq\; \sum_{i'=1}^{i} \theta\lambda_j^{i'}(t_{i'}-t_{i'-1}) - N_j^{\theta}(t_i^-) + \theta\zeta_j^i\varepsilon\Big\} \mid F_i^{\theta-}\Big]
\end{aligned}
$$

$$
\begin{aligned}
&\geq \; \mathsf{P}\Bigg[\bigcap_{j\in S_i\setminus S_{i+1}}\left\{\sum_{i'=1}^{i}\theta\lambda_j^{i'}(t_{i'}-t_{i'-1})-N_j^\theta(t_i^-) \;=\; N_j^\theta(t_i^+)-N_j^\theta(t_i^-)\right\} \\
&\qquad\qquad \bigcap_{j\in S_{i+1}}\left\{\theta\eta_j^i\varepsilon \;\leq\; N_j^\theta(t_i^+)-N_j^\theta(t_i^-) \;\leq\; \theta\zeta_j^i\varepsilon\right\} \mid F_i^{\theta-}\Bigg] \\
&= \; \mathsf{P}\Bigg[\bigcap_{j\in S_i\setminus S_{i+1}}\left\{\sum_{i'=1}^{i}\theta\lambda_j^{i'}(t_{i'}-t_{i'-1})-N_j^\theta(t_i^-) \;=\; \hat{N}_j^{i-}\left(\theta\lambda\int_{\tau\in(t_i^-,t_i^+]}P_{j:S^\theta(\tau)}d\tau\right)\right\} \\
&\qquad\qquad \bigcap_{j\in S_{i+1}}\left\{\theta\eta_j^i\varepsilon \;\leq\; \hat{N}_j^{i-}\left(\theta\lambda\int_{\tau\in(t_i^-,t_i^+]}P_{j:S^\theta(\tau)}d\tau\right) \;\leq\; \theta\zeta_j^i\varepsilon\right\} \mid F_i^{\theta-}\Bigg] \\
&\geq \; \mathsf{P}\left[\bigcap_{j\in S_i}\left\{\theta\eta_j^i\varepsilon \;\leq\; \hat{N}_j^{i-}\left(\theta\underline{\lambda}_j^i(t_i^+-t_i^-)\right)\right\}\bigcap_{j\in S_{i+1}}\left\{\hat{N}_j^{i-}\left(\theta\overline{\lambda}_j^i(t_i^+-t_i^-)\right) \;\leq\; \theta\zeta_j^i\varepsilon\right\} \mid F_i^{\theta-}\right] \\
&= \; \mathsf{P}\left[\bigcap_{j\in S_i}\left\{\theta\eta_j^i\varepsilon \;\leq\; \hat{N}_j^{i-}\left(\theta\underline{\lambda}_j^i(t_i^+-t_i^-)\right)\right\}\bigcap_{j\in S_{i+1}}\left\{\hat{N}_j^{i-}\left(\theta\overline{\lambda}_j^i(t_i^+-t_i^-)\right) \;\leq\; \theta\zeta_j^i\varepsilon\right\}\right] \\
&= \; 1-\mathsf{P}\left[\bigcup_{j\in S_i}\left\{\theta\eta_j^i\varepsilon \;>\; \hat{N}_j^{i-}\left(\theta\underline{\lambda}_j^i(t_i^+-t_i^-)\right)\right\}\bigcup_{j\in S_{i+1}}\left\{\hat{N}_j^{i-}\left(\theta\overline{\lambda}_j^i(t_i^+-t_i^-)\right) \;>\; \theta\zeta_j^i\varepsilon\right\}\right] \\
&\geq \; 1-\sum_{j\in S_i}\mathsf{P}\left[\hat{N}_j^{i-}\left(\theta\underline{\lambda}_j^i(t_i^+-t_i^-)\right) \;<\; \theta\eta_j^i\varepsilon\right]-\sum_{j\in S_{i+1}}\mathsf{P}\left[\hat{N}_j^{i-}\left(\theta\overline{\lambda}_j^i(t_i^+-t_i^-)\right) \;>\; \theta\zeta_j^i\varepsilon\right] \\
&\geq \; 1-\sum_{j\in S_i}\mathsf{P}\left[\hat{N}_j^{i-}\left(\theta\underline{\lambda}_j^i(t_i^+-t_i^-)\right) \;<\; \theta\eta_j^i\varepsilon\right]-\sum_{j\in S_i}\mathsf{P}\left[\hat{N}_j^{i-}\left(\theta\overline{\lambda}_j^i(t_i^+-t_i^-)\right) \;>\; \theta\zeta_j^i\varepsilon\right] \\
&\geq \; 1-\sum_{j\in S_i}\mathsf{P}\left[\hat{N}_j^{i-}\left(\theta\underline{\lambda}_j^i(t_i^+-t_i^-)\right)-\theta\underline{\lambda}_j^i(\nu_i^-+\nu_i^+)\varepsilon \;<\; -\theta\underline{\lambda}_j^i\varepsilon\right] \\
&\qquad -\sum_{j\in S_i}\mathsf{P}\left[\hat{N}_j^{i-}\left(\theta\overline{\lambda}_j^i(t_i^+-t_i^-)\right)-\theta\overline{\lambda}_j^i(\nu_i^-+\nu_i^+)\varepsilon \;>\; \theta\overline{\lambda}_j^i\varepsilon\right] \\
&= \; 1-\sum_{j\in S_i}\mathsf{P}\left[\hat{N}_j^{i-}\left(\theta\underline{\lambda}_j^i(t_i^+-t_i^-)\right)-\theta\underline{\lambda}_j^i(t_i^+-t_i^-) \;<\; -\theta\underline{\lambda}_j^i\varepsilon\right] \\
&\qquad -\sum_{j\in S_i}\mathsf{P}\left[\hat{N}_j^{i-}\left(\theta\overline{\lambda}_j^i(t_i^+-t_i^-)\right)-\theta\overline{\lambda}_j^i(t_i^+-t_i^-) \;>\; \theta\overline{\lambda}_j^i\varepsilon\right] \\
&\geq \; 1-\sum_{j\in S_i}\exp\left(-\theta\underline{\lambda}_j^i(t_i^+-t_i^-)h\left(\frac{\varepsilon}{t_i^+-t_i^-}\right)\right)-\sum_{j\in S_i}\exp\left(-\theta\overline{\lambda}_j^i(t_i^+-t_i^-)h\left(\frac{\varepsilon}{t_i^+-t_i^-}\right)\right) \\
&= \; 1-\delta_i^{\theta+}.
\end{aligned}
$$

The first inequality holds because $F_i^{\theta-}\subset E_i^{\theta-}$, and for all sample paths in $E_i^{\theta-}$ it holds that $\sum_{i'=1}^{i}\theta\lambda_j^{i'}(t_{i'}-t_{i'-1})-\theta\eta_j^i\varepsilon < N_j^\theta(t_i^-) < \sum_{i'=1}^{i}\theta\lambda_j^{i'}(t_{i'}-t_{i'-1})$ for all $j\in S_i$, the second inequality follows from Lemma EC.8, the fifth equality holds because $\hat{N}_j^{i-}$ and $F_i^{\theta-}$ are independent, the third inequality applies the union bound, the fourth inequality holds because $S_{i+1}\subset S_i$, the fifth inequality follows from (EC.9), and the sixth inequality follows from Lemma EC.9.

Therefore, using the induction hypothesis, it follows that

$$
\mathsf{P}\left[F_{i+1}^{\theta+}\right] \;=\; \mathsf{P}\left[F_{i+1}^{\theta+}\mid F_{i+1}^{\theta-}\right]\mathsf{P}\left[F_{i+1}^{\theta-}\mid F_i^{\theta+}\right]\mathsf{P}\left[F_i^{\theta+}\right]
$$

$$\geq \quad (1-\delta_{i+1}^{\theta+})\,(1-\delta_{i+1}^{\theta-})\left(1-\sum_{i'=1}^{i}\left(\delta_{i'}^{\theta-}+\delta_{i'}^{\theta+}\right)\right) \quad\geq\quad 1-\sum_{i'=1}^{i+1}\left(\delta_{i'}^{\theta-}+\delta_{i'}^{\theta+}\right).$$

Thus we have established (EC.10). Since the expected sales quantity of any product during $[t_k, t_k^+]$ is bounded by $\theta\lambda(t_k^+ - t_k) = \theta O(\varepsilon)$, it follows that

$$\begin{aligned}
\mathsf{E}\left[N_j^\theta(T)\right] \quad &\geq \quad \mathsf{E}\left[N_j^\theta(t_k^+)\right] - \theta O(\varepsilon) \\
&\geq \quad \mathsf{E}\left[N_j^\theta(t_k^+) \mid F_k^{\theta+}\right] \mathsf{P}\left[F_k^{\theta+}\right] - \theta O(\varepsilon) \quad\geq\quad \theta b_j^*\left[1-\sum_{i=1}^{k}(\delta_i^{\theta-}+\delta_i^{\theta+})\right] - \theta O(\varepsilon).
\end{aligned}$$

For any fixed $\varepsilon > 0$, it holds that $\delta_i^{\theta-} \to 0$ and $\delta_i^{\theta+} \to 0$ for all i as $\theta \to \infty$. Thus

$$\begin{aligned}
\liminf_{\theta\to\infty} \frac{1}{\theta}\mathsf{E}\left[Z^\theta\right] \quad &= \quad \liminf_{\theta\to\infty} \frac{1}{\theta}\sum_{j\in\mathcal{J}} r_j\,\mathsf{E}\left[N_j^\theta(T)\right] \\
\geq \quad \liminf_{\theta\to\infty} \sum_{j\in\mathcal{J}} r_j &\left\{ b_j^*\left[1-\sum_{i=1}^{k}(\delta_i^{\theta-}+\delta_i^{\theta+})\right] - O(\varepsilon)\right\} \quad=\quad \sum_{j\in\mathcal{J}} r_j b_j^* - O(\varepsilon) \quad=\quad z^{\mathrm{CDLP}} - O(\varepsilon).
\end{aligned}$$

Since ε can be arbitrarily small, it follows that $\liminf_{\theta\to\infty}\frac{1}{\theta}\mathsf{E}\left[Z^\theta\right] \geq z^{\mathrm{CDLP}}$. Also,

$$\limsup_{\theta\to\infty}\frac{1}{\theta}\mathsf{E}\left[Z^\theta\right] \leq \limsup_{\theta\to\infty}\frac{1}{\theta} z_{OPT}^\theta \leq z^{\mathrm{CDLP}}.$$

Therefore $\lim_{\theta\to\infty}\frac{1}{\theta}\mathsf{E}\left[Z^\theta\right] = z^{\mathrm{CDLP}}$. $\square$

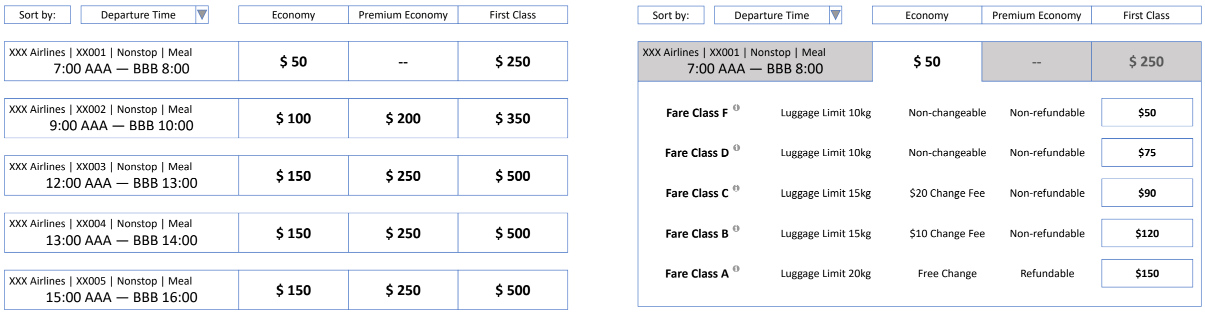
Appendix F: Airline Data, Preprocessing and Demand Calibration

The data and source code used in all the numerical studies in the paper can be downloaded at `https://github.com/cyf-sjtu/spikedMNL`. Here we give a concise description of the airline data set provided to us by an airline, and how we processed the data and calibrated the demand models.

F.1. Products Presentation

We describe how products are presented on the airline's website to show that potential customers can see and can choose among all the available fare classes. Figure EC.4 shows an example of relevant data displayed on the airline's own ticket-booking website. The left panel in Figure EC.4 shows what a customer sees after the customer has entered search data, such as origin, destination, and date, in the website's search engine. Different itineraries with their departure times and arrival times are presented to the customer, with the lowest available price for each combination of itinerary and cabin (Economy/Premium Economy/First Class). If the customer clicks on a combination of itinerary and cabin, then the website displays all the available fare classes for that combination of itinerary and cabin, with their associated booking rules and prices, as shown in the right panel. So, indeed, a customer can see and choose among all the available fare products offered by the airline on the airline's website.

Figure EC.4 Airline Ticket-Booking Website Mock-ups



Note. After a customer has entered search data, such as origin AAA, destination BBB, and date, in the website's search engine, the website displays the available combinations of itinerary and cabin (left panel). If the customer clicks on a combination of itinerary and cabin, e.g. the first itinerary in the Economy cabin with a price of $50, the website displays all the available fare classes for this combination of itinerary and cabin, with their corresponding booking rules and prices (right panel).

F.2. Booking and Availability Data

We received individual passenger-level ticket booking data, including booking time and channel, as well as periodic (typically daily) snapshots of the assortments of available products, from the airline. For most bookings, the assortment of available products at the time of the booking can be inferred. For each itinerary, the booking and availability data covered a three-month period before the departure time. There was sufficient diversity in the offered assortments over this booking horizon to estimate choice models including the no-purchase alternative. To represent correlation between booking time and customer preferences, we partitioned the booking horizon into 200 intervals based on combinations of factors such as weekday/weekend, work/off-work hours, and the number of weeks before the departure time; we then calibrated different price sensitivity and booking rule sensitivity parameters for different intervals to account for time-varying preference behavior.

F.3. Competitors' Booking Data

There were three major competitors in the market that we considered. We obtained booking data of all three competitors, some of which was collected by one of the airlines and some of which was purchased by the airline from a Global Distribution System (GDS). We also obtained availability data of the products offered by all three competitors, which was collected by the same airline.

F.4. Demand Model Calibration

Using the booking data and availability data of the airlines competing in the market, we calibrated the demand models for this project. The demand models consisted of two parts: the arrival process of potential customers, and the customer choice model. The demand models were calibrated using maximum likelihood estimation.