

The Value of Robust Assortment Optimization Under Ranking-based Choice Models

Bradley Sturt

Department of Information and Decision Sciences
University of Illinois at Chicago, bsturt@uic.edu

We study a class of robust assortment optimization problems that was proposed by Farias, Jagabathula, and Shah (2013). The goal in these problems is to find an assortment that maximizes a firm’s worst-case expected revenue under all ranking-based choice models that are consistent with the historical sales data generated by the firm’s past assortments. We establish for various settings that these robust optimization problems can either be solved in polynomial-time or can be reformulated as compact mixed-integer optimization problems. To establish our results, we prove that optimal assortments for these robust optimization problems have a simple structure that is closely related to the structure of revenue-ordered assortments. We use our results to show how robust optimization can be used to overcome the risks of estimate-then-optimize and the need for experimentation with ranking-based choice models in the overparameterized regime.

Key words: assortment planning; robust optimization; nonparametric choice modeling.

History: First version: Dec. 9, 2021. Revisions submitted on Feb. 16, 2023 and Nov. 6, 2023. Accepted for publication on Feb. 21, 2024.

“First, do no harm” - Hippocratic Oath

1. Introduction

The ranking-based choice model is one of the most fundamental and influential discrete choice models in revenue management. It is used by firms in industries such as e-commerce and brick-and-mortar retail to predict the demand for the firm’s products as a function of the subset of products that the firm offers to their customers. The popularity of the ranking-based choice model can be attributed to its generality: it can represent any random utility maximization model, and thus encompasses many other popular discrete choice models such as the multinomial logit model (Block and Marschak 1959, Farias et al. 2013). The ranking-based choice model posits that customers who visit the firm have preferences that are represented by randomly-chosen rankings, and, based on the subset of products offered by the firm, each customer will purchase the product that is most preferred according to their personal ranking.

Unfortunately, accurately estimating a ranking-based choice model from a firm’s historical sales data is notoriously challenging. The key issue is that the ranking-based choice model is comprised of around $n!$ parameters, where n is the number of product alternatives that a firm can elect to offer to their customers. Because the number of past subsets of products M that the firm has previously offered to their customers usually satisfies $M \ll n!$, many selections of these parameters can yield a ranking-based choice model that is consistent (has low or zero prediction error) with the firm’s historical sales data. This leaves firms with the challenge of selecting *which* ranking-based choice model, out of all of those that are consistent with their historical sales data, to use when making operational planning decisions.

The challenge of estimation with ranking-based choice models is particularly acute in the context of *assortment planning*. Here, the typical goal of a firm is to identify a new subset of products

(referred to as an assortment) to offer to their customers in order to increase the firm’s expected revenue. Because the true relationship between assortments and expected revenue is unknown, firms will typically interpret one selection of a ranking-based choice model as the “ground truth” and subsequently solve an optimization problem to find an assortment which maximizes the “predicted” expected revenue (Aouad et al. 2018, Bertsimas and Mišić 2019, Honhon et al. 2012, van Ryzin and Vulcano 2015, 2017, Feldman et al. 2019, Aouad et al. 2021, Désir et al. 2021). This widely-used technique for identifying a new assortment is referred to in the revenue management literature as *estimate-then-optimize*. The estimate-then-optimize technique can be attractive from a computational standpoint, due to its decoupling of the combinatorial problems related to estimation and optimization. But the assortment that is optimal under one selection of a ranking-based choice model that is consistent with the historical sales data may be highly *suboptimal* under another ranking-based choice model that is also consistent with the historical sales data. This raises concerns about whether estimate-then-optimize can be trusted to identify a new assortment for the firm with an expected revenue that, at the very least, is no less than the firm’s highest expected revenue from their past assortments.

To overcome the risks of estimate-then-optimize in the overparameterized regime¹, one potential approach available to firms is to perform experimentation. That is, a firm can offer various new assortments to their customers for short durations, and then use the sales data obtained by this experimentation to further constrain the set of ranking-based choice models that are consistent with the firm’s historical sales data. But engaging in experimentation may be undesirable for various practical reasons, ranging from a firm’s aversion to short-term losses incurred by experimentation to the implementation challenges for performing experimentation in settings such as brick-and-mortar retail (Ariely 2010, Ignat 2022). Moreover, the large number of parameters in ranking-based choice models implies that experimentation with an extensive array of new assortments may not guarantee the identification of a unique ranking-based choice model (Sher et al. 2011, Susan et al. 2022). Given that experimentation may be costly and may not lead to the identification of a unique ranking-based choice model, our paper investigates an alternative approach to experimentation centered around *robust optimization*.

1.1. A Robust Optimization Approach

Our paper investigates a class of robust optimization problems proposed by Farias, Jagabathula, and Shah (2013, henceforth abbreviated as FJS13) that aims to circumvent the risks of estimate-then-optimize with ranking-based choice models in the overparameterized regime. The goal of the class of robust optimization problems, stated succinctly, is to find an assortment that has high expected revenue under *all* of the ranking-based choice models that are consistent with the historical sales data generated by the firm’s past assortments. The robust optimization problems achieve this goal by yielding an assortment that maximizes the predicted expected revenue under the worst-case ranking-based choice model that is consistent with the historical sales data generated by the firm’s past assortments. Formally, the class of robust optimization problems is given by

$$\max_{S \in \mathcal{S}} \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S), \quad (\text{RO})$$

where \mathcal{S} is the set of all assortments that can be offered by the firm to its customers, \mathcal{U} is the set of all ranking-based choice models that are consistent with the historical sales data generated by the firm’s past assortments, and $\mathcal{R}^\lambda(S)$ is the expected revenue for a new assortment S under a ranking-based choice model $\lambda \in \mathcal{U}$.

¹We use the term ‘overparameterized regime’ to refer to assortment planning problems in which many selections of ranking-based choice models are consistent with the historical sales data generated by the firm’s past assortments.

The class of robust optimization problems (RO) can be attractive from a managerial perspective for a variety of reasons. First, in contrast to estimate-then-optimize, the robust optimization problem (RO) aims to find an assortment that can be trusted to improve the firm’s expected revenue in a way that is not exclusive to just one of the many ranking-based choice models that are consistent with the firm’s historical sales data. Second, the robust optimization problem (RO) can potentially obviate the need for experimentation—which can be costly or otherwise undesirable to firms—by identifying assortments with desirable expected revenue guarantees even though a unique ranking-based choice model cannot be identified using the firm’s historical sales data. Third, FJS13 showed using real-world data that the worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ can be a significantly more accurate prediction of expected revenue compared to the predictions made by parametric methods. Hence, there is empirical evidence that the assortments obtained from solving (RO) may be desirable to firms from a revenue-improving perspective.

Despite the appeal of the class of robust optimization problems (RO), very limited progress has been made on solving these robust optimization problems until now. As far as we can tell, the only prior work that considers the computational tractability of (RO) is that of FJS13, which presented algorithms for computing the worst-case expected revenue of a fixed assortment. However, no algorithms to date have been developed for solving this class of robust optimization problems (Rusmevichientong and Topaloglu (2012, p.867), Jagabathula (2014, p.8)), and “it is not clear how one may formulate the problem of optimizing the worst-case revenue as an efficiently solvable mathematical optimization reformulation.” (Mišić 2016, p.118). This lack of progress is especially notable, in view of the fact that algorithms and structural results have been developed for other classes of robust assortment optimization problems (Rusmevichientong and Topaloglu 2012, Bertsimas and Mišić 2017, Désir et al. 2023, Wang et al. 2020).

The lack of practical algorithms and structural results for (RO) has ultimately prevented a formal understanding of the settings in which the class of robust optimization problems can provide managerial value. On one hand, the robust optimization problem (RO) promises to help firms circumvent the risks of estimate-then-optimize by finding an assortment that delivers high expected revenue across all of the ranking-based choice models that are consistent with the firm’s historical sales data. But on the other hand, no previous work has established whether assortments $S \in \mathcal{S}$ that have high expected revenue $\mathcal{R}^\lambda(S)$ across all ranking-based choice models $\lambda \in \mathcal{U}$ can ever *exist*. Because the question of whether (RO) is too conservative to be practically valuable has remained open, the extent to which the robust optimization approach can actually overcome the risks of estimate-then-optimize or serve as a viable alternative to experimentation has been unknown.

1.2. Main Questions

In this paper, we pursue an agenda of understanding *if* and *when* the class of robust optimization problems (RO) can be computationally tractable and provide managerial value as an alternative to estimate-then-optimize and experimentation in the overparameterized regime. In the pursuit of this agenda, we focus in this paper on four open questions.

Our first question concerns the conservatism of the class of robust optimization problem (RO):

QUESTION 1. *Can the optimal objective value of (RO) ever be strictly greater than the expected revenue of the firm’s best past assortment?*

We show in §2 that Question 1 is equivalent to asking whether there can ever exist an assortment with an expected revenue that is strictly greater than the expected revenue of the firm’s best past

assortment across all of the ranking-based choice models that are consistent with the historical sales data generated by the firm’s past assortments. An affirmative answer to Question 1 is thus necessary for it to be possible for the robust optimization problem (RO) to yield new assortments that can be trusted to improve the firm’s expected revenue in a way that is not exclusive to just a subset of the many ranking-based choice models that are consistent with the firm’s historical sales data.

Our second question asks about the relationship between the conservatism of (RO) and the past assortments that the firm has offered to its customers:

QUESTION 2. If the answer to Question 1 is yes, then what structural properties of a firm’s past assortments are necessary for the optimal objective value of (RO) to be strictly greater than the expected revenue of the firm’s best past assortment?

Recall that Question 1 is an existential question, in the sense that Question 1 asks whether there exists a problem instance (meaning a collection of past assortments, prices, and historical sales data) for which the optimal objective value of (RO) is strictly greater than the expected revenue of the firm’s best past assortment. But even if the answer to Question 1 is yes, there exist collections of past assortments for which the optimal objective value of (RO) will never be strictly greater than the expected revenue of the firm’s best past assortment.² In response, Question 2 asks whether there are structural properties of the firm’s collection of past assortments that are necessary in order for (RO) to have a chance of not being overly conservative.

Our third question relates the computational tractability of the robust optimization problem (RO) to the collection of past assortments that the firm has offered to its customers:

QUESTION 3. What structural properties of a firm’s past assortments are sufficient for designing efficient algorithms for solving the robust optimization problem (RO)?

Even if the optimal objective value of (RO) is strictly greater than the expected revenue of the firm’s best past assortment for a given problem instance, the practical value of the robust optimization problem (RO) for that problem instance is limited if an optimal assortment for (RO) cannot be found in practical computation times. Moreover, as discussed in §1.1, the class of robust optimization problems (RO) is widely believed in the revenue management community to be computationally intractable, and no practical algorithms for solving these problems for any collection of past assortments have been developed thus far. Against this backdrop, Question 3 asks whether there are structural properties of a firm’s past assortments that make it possible to design theoretically- and practically-efficient algorithms for solving (RO).

Our fourth question asks how the robust optimization problem (RO) can be deployed in settings in which the robust optimization approach is overly conservative:

QUESTION 4. If the optimal objective value of (RO) is not strictly greater than the expected revenue of the firm’s best past assortment for a particular problem instance, then what assortment should the firm offer to its customers?

If the optimal objective value of (RO) is strictly greater than the expected revenue of the firm’s best past assortment for a given problem instance, then we recommend that the firm offer the

² For example, we readily observe that the optimal objective value of (RO) will never be strictly greater than the expected revenue of the firm’s best past assortment if every assortment $S \in \mathcal{S}$ is a past assortment (i.e., has already been offered by the firm to its customers).

optimal assortment for (RO) to its customers, as such an assortment can be trusted to improve the firm’s expected revenue in a way that is not exclusive to just one of the many ranking-based choice models that are consistent with the firm’s historical sales data. But if the optimal objective value of the robust optimization problem (RO) is not strictly greater than the expected revenue of the firm’s best past assortment for the given problem instance, then can the robust optimization problem (RO) provide any guidance as to what should a firm do?

1.3. Contributions

Our main contributions of this paper consist of providing the first answers to Questions 1-4.

First, we establish an affirmative answer to Question 1. Specifically, we present an example in §3 of an assortment planning problem with two past assortments in which the optimal objective value of the robust optimization problem (RO) is strictly greater than the expected revenue of the firm’s best past assortment. The example thus shows that the robust optimization problem (RO) can yield practical value to firms, even when a small number of assortments have been offered in the past.

Second, we address Question 2 by characterizing the structure of optimal assortments for the class of robust optimization problems (RO) (Theorem 1 in §4). Our characterization reveals that optimal assortments for these robust optimization problems have a simple and interpretable structure that is similar to the structure of the widely-studied class of revenue-ordered assortments. We use Theorem 1 to show that it is necessary for the firm’s past assortments to deviate from revenue-ordered assortments in order for it be possible for the optimal objective value of (RO) to be strictly greater than the expected revenue of the firm’s best past assortment (§4.1). We further show that Theorem 1 can be used to drastically reduce the number of candidate assortments that need to be checked in order to solve the robust optimization problem (RO) when the number of past assortments is small (§4.2-§4.3). Our proof of Theorem 1 is based on a simple yet intricate analysis of reachability conditions for vertices in a data-driven class of directed acyclic graphs (§4.4).

Third, we provide an affirmative answer to Question 3 by developing the first polynomial-time algorithms and compact mixed-integer optimization reformulations for the class of robust optimization problems (RO). The existence of polynomial-time algorithms resolves the open question in the literature regarding the theoretical tractability of (RO) for settings with small numbers of past assortments, and we provide numerical evidence that our algorithms and mixed-integer optimization reformulations are practical. In greater detail, we make the following algorithmic developments:

- (i) We use our structure of optimal assortments to develop a strongly polynomial-time algorithm for solving the robust optimization problem when the firm has offered two past assortments (Theorem 3 in §5.1). Our algorithm reduces the robust optimization problem to a sequence of minimum-cost network flow problems, and our algorithm has a total running time of $\mathcal{O}(n^5 \log(nr_n))$, where n is the number of available products and r_n is the integral price of the most expensive product. We provide numerical experiments showing that our algorithm can require less than 30 seconds for problem instances with $n = 100$ products.
- (ii) We generalize our algorithm for two past assortments to show that the robust optimization problem can be solved in weakly polynomial-time for any fixed number of past assortments (Theorem 4 in §5.2). The algorithm provides evidence that it can be possible to develop practical algorithms for solving the robust optimization problem (RO) for real-world problem instances, where the composition of products in the past assortments do not exhibit any convenient structure and where there may be no ranking-based choice models that have zero prediction error on the historical sales data.

- (iii) To address applications in which the number of past assortments is large, we consider a common structural assumption in the revenue management literature in which the past assortments are nested (see §5.3). When the past assortments are nested, we show that our technical developments from §4 can be used to develop a mixed-integer optimization reformulation of the robust optimization problem (RO) in which the number of decision variables and constraints scales polynomially in both the number of products as well as the number of past assortments (Theorem 5 in §5.3). We show using numerical experiments that our mixed-integer linear optimization problem can require less than 30 seconds to solve (RO) to optimality on problem instances with up to $M = 20$ past assortments.

Finally, we address Question 4 by showing that our algorithms can yield managerial value even when the optimal objective value of (RO) is not strictly greater than the expected revenue of the firm’s best past assortment. Specifically, we show that firms can compute an upper bound on the benefits of experimentation by extending our algorithms from §5 to solve an optimistic version of the robust optimization problem (RO) (Theorems 6 and 7 in §6.1). If the gap between the upper bound and the expected revenue of the firm’s best past assortment is small, then we recommend that the firm should not perform experimentation and continue offering their best past assortment. If the gap is large, then the firm can potentially benefit from experimentation, and we show in this case that the robust and optimistic optimization problems can be combined into a Pareto approach for finding new assortments for experimentation that have potential to increase the firm’s best-case expected revenue while maintaining bounds on the worst-case decrease in expected revenue. We illustrate the practical value of the aforementioned techniques using synthetic and real data.

In summary, the techniques proposed in this paper can provide value to firms such as brick-and-mortar and small e-commerce retailers who change their assortments infrequently, wish to find new assortments that generate higher expected revenue than their past assortments, and lack the infrastructure or willingness to experiment with new assortments that lead to short-run declines in expected revenue. Indeed, we show using synthetic and real data in Appendix A that our proposed algorithms require practical computation times and yield new assortments with desirable worst-case revenue guarantees when the number of past assortments is small or when the past assortments are nested. The numerical experiments thus show that the methodology developed in this paper is well positioned to provide value for the sizes and types of historical sales data faced by real-world firms.

The rest of this paper is organized as follows. In §2, we present our problem setting. In §3, we give an affirmative answer to Question 1 by presenting an example of an assortment planning problem with two past assortments in which the optimal objective value of the robust optimization problem (RO) is strictly greater than the expected revenue of the firm’s best past assortment. In §4, we develop our key structural result which characterizes the optimal solutions for robust optimization problems under ranking-based choice models. In §5, we develop polynomial-time algorithms and mixed-integer optimization reformulations for the robust optimization problem. In §6, we show how our algorithms can provide value in problem instances in which the optimal objective value of the robust optimization problem (RO) is not strictly greater than the expected revenue of the firm’s best past assortment. In §7, we offer concluding thoughts and directions for future research. The numerical experiments from this paper are found in Appendix A, and the proofs of the technical results from the paper are found in Appendices B–J.

Notation and Terminology. We use \mathbb{R} to denote the real numbers, \mathbb{R}_+ to denote the nonnegative real numbers, and $y^\top x$ to denote the inner product of two vectors. We use the phrase ‘collection’ to refer to a set of sets. We let the set of all probability distributions which are supported on a finite set \mathcal{A} be denoted by $\Delta_{\mathcal{A}} \triangleq \{\lambda : \sum_{a \in \mathcal{A}} \lambda_a = 1, \lambda_a \geq 0 \forall a \in \mathcal{A}\}$. We assume throughout that a norm $\|\cdot\|$ is either the ℓ_1 -norm or ℓ_∞ -norm, and so it follows that optimization problems of the form $\min_{x,y} \{c^\top x + d^\top y \mid Ax + By \leq b, \|y\| \leq \eta\}$ can be referred to as linear optimization problems. We let $\mathbb{I}\{\cdot\}$ denote the indicator function, which equals one if \cdot is true and equals zero otherwise.

Code Availability. The code for conducting the numerical experiments in this paper is freely available and can be accessed at <https://github.com/brad-sturt/IdentificationQuestion>.

2. Problem Setting

We adopt the perspective of a firm that must select a subset of products to offer to their customers. Let the universe of products available to the firm be denoted by $\mathcal{N} \triangleq \{1, \dots, n\}$, where the no-purchase option is denoted by index 0 and $\mathcal{N}_0 \triangleq \mathcal{N} \cup \{0\}$. The revenue generated by selling one unit of product $i \in \mathcal{N}$ is represented by $r_i > 0$, and the revenue associated with the no-purchase option is $r_0 = 0$. We assume throughout the paper that the revenues of the products are unique and that products have been sorted in ascending order by revenue, $0 < r_1 < \dots < r_n$. An assortment is defined as any subset of products $S \subseteq \mathcal{N}_0$, and we let $\mathcal{S} \triangleq \{S \subseteq \mathcal{N}_0 : 0 \in S\}$ denote the collection of all assortments that include the no-purchase option.

We study a problem setting in which the underlying relationship between assortment and customer demand is unknown, and our only information on this relationship comes from historical sales data generated by the firm's past assortments. Let the past assortments be denoted by $\mathcal{M} \triangleq \{S_1, \dots, S_M\} \subseteq \mathcal{N}_0$, and let the indices of these past assortments be denoted by $\mathcal{M} \triangleq \{1, \dots, M\}$. Unless stated otherwise, we will make no assumptions on the mechanism by which the firm selected the assortments to offer in the past. That is, the firm could have chosen the past assortments by drawing products randomly; alternatively, the past assortments could have been chosen using managerial intuition or some other systematic approach. We assume that the firm offered each past assortment $S_m \in \mathcal{M}$ to their customers for a sufficient duration to obtain an accurate estimate of the purchase frequencies, *i.e.*, the fraction of customers $v_{m,i} \in [0, 1]$ that purchase product $i \in S_m$ when offered assortment S_m . This historical sales data is assumed to be normalized such that $\sum_{i \in \mathcal{N}_0} v_{m,i} = 1$, and the purchase frequencies for products that are not in an assortment are defined equal to zero, that is, $v_{m,i} = 0$ for all $i \notin S_m$ and $m \in \mathcal{M}$. The expected revenue generated by the past assortment S_m is written compactly as $\sum_{i \in \mathcal{N}_0} r_i v_{m,i} = r^\top v_m$.

A discrete choice model is a function that predicts purchase frequencies for the firm based on the assortment that the firm offers to their customers. A *ranking-based choice model* is a type of choice model which is parameterized by a probability distribution λ over the set of all distinct rankings of the products, where a ranking refers to a one-to-one mapping of the form $\sigma : \{0, \dots, n\} \rightarrow \{0, \dots, n\}$. Specifically, a ranking σ encodes a preference for product i over product j if and only if $\sigma(i) < \sigma(j)$. Let the set of all distinct rankings over the products be denoted by Σ , and we readily observe that the number of distinct rankings in this set satisfies $|\Sigma| = (n+1)!$. Given a probability distribution over rankings $\lambda \in \Delta_\Sigma$ and an assortment $S \in \mathcal{S}$, the prediction made by the ranking-based choice model for the purchase frequency of each product $i \in \mathcal{N}_0$ is given by

$$\mathcal{D}_i^\lambda(S) \triangleq \sum_{\sigma \in \Sigma} \mathbb{I} \left\{ i = \arg \min_{j \in S} \sigma(j) \right\} \lambda_\sigma.$$

It is straightforward to see from the above definition that a ranking-based choice model always satisfies the equality $\mathcal{D}_i^\lambda(S) = 0$ for all products i that are not in the assortment S . The predicted expected revenue for a firm that offers assortment S under the ranking-based choice model with parameter λ is given by

$$\mathcal{R}^\lambda(S) \triangleq \sum_{i \in \mathcal{N}_0} r_i \mathcal{D}_i^\lambda(S) = r^\top \mathcal{D}^\lambda(S).$$

We say that a ranking-based choice model is *consistent* with the historical sales data generated by the firm's past assortments if the difference between the predicted purchase frequency $\mathcal{D}_i^\lambda(S_m)$ and the historical sales data $v_{m,i}$ is small for each of the products $i \in S_m$ that were offered in each of the past assortments $m \in \mathcal{M}$. We define the set of all ranking-based choice models that are consistent with the historical sales data as

$$\mathcal{U} \triangleq \left\{ \lambda \in \Delta_\Sigma : \begin{array}{l} \text{there exists a vector } \epsilon \text{ such that } \|\epsilon\| \leq \eta \text{ and} \\ \mathcal{D}_i^\lambda(S_m) - v_{m,i} = \epsilon_{m,i} \text{ for all } i \in S_m \text{ and } m \in \mathcal{M} \end{array} \right\},$$

where the radius $\eta \geq 0$ of the set \mathcal{U} is a parameter that is selected by the firm, and where $\|\cdot\|$ is defined at the end of §1 as either the ℓ_1 or ℓ_∞ norm.

To develop an understanding for the above set of ranking-based choice models \mathcal{U} , let us consider the case in which the radius η of the above set is equal to zero. In that case, we observe that the above set contains exactly the probability distributions for which the corresponding ranking-based choice models have perfect accuracy on the historical sales data generated by the past assortments. In other words, if $\eta = 0$, then the set \mathcal{U} is comprised of all of the probability distributions $\lambda \in \Delta_\Sigma$ that satisfy $\mathcal{D}_i^\lambda(S_m) = v_{m,i}$ for each of the products $i \in S_m$ that were offered in each of the past assortments $m \in \mathcal{M}$. From a theoretical perspective, it is known that the set \mathcal{U} with $\eta = 0$ is guaranteed to be nonempty if the firm's customers' behavior is captured by a random utility maximization model and if the historical sales data has been observed without noise; see Block and Marschak (1959). From a practical perspective, it can be reasonable to expect that the set \mathcal{U} will be nonempty with $\eta = 0$ in problem instances in which the number of past assortments M is much smaller than the number of parameters in the ranking-based choice model, $|\Sigma| = (n+1)!$. Nonetheless, if there are no ranking-based choice models that have perfect accuracy on the historical sales data, then the radius η can always be made sufficiently large to ensure that the set of ranking-based choice models \mathcal{U} is nonempty.³ For simplicity, we make the standing assumption throughout our paper that the set of ranking-based choice models \mathcal{U} is nonempty for the firm's selection of the radius η .

We conclude the present section with a theoretical derivation of the class of robust optimization problems (RO). Indeed, consider a firm that wishes to identify an assortment with an expected revenue that is strictly greater than expected revenues of the firm's past assortments under all of the ranking-based choice models that are consistent with firm's historical sales data. In this case, the firm seeks an assortment $S \in \mathcal{S}$ that satisfies

$$\mathcal{R}^\lambda(S) > \max \{ \mathcal{R}^\lambda(S_1), \dots, \mathcal{R}^\lambda(S_M) \} \quad \forall \lambda \in \mathcal{U}. \quad (\text{IMP}')$$

If $\eta = 0$, then it follows from algebra⁴ that an assortment $S \in \mathcal{S}$ satisfies (IMP) if and only if

$$\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) > \max_{m \in \mathcal{M}} r^\top v_m. \quad (\text{IMP})$$

The above analysis shows that there exists an assortment that satisfies (IMP') if and only if the optimal objective value of the robust optimization problem (RO) is strictly greater than $\max_{m \in \mathcal{M}} r^\top v_m$, i.e., the expected revenue generated by the firm's best past assortment. Moreover, if there exists an assortment that satisfies (IMP), then every optimal solution of the robust optimization problem (RO) satisfies (IMP). Hence, we conclude that the robust optimization problem (RO) enjoys theoretical justification as a tool for identifying a new assortment with an expected revenue that is strictly greater than expected revenues of the firm's past assortments across all of the ranking-based choice models that are consistent with firm's historical sales data.

³ A sufficiently large choice of η which ensures that \mathcal{U} is nonempty can easily be found by applying binary search over the choice of η , provided that one has an algorithm that solves (RO).

⁴ If $\eta = 0$, then we observe for each past assortment $m \in \mathcal{M}$ that $\mathcal{R}^\lambda(S_m) = r^\top \mathcal{D}^\lambda(S_m) = r^\top v_m$ for all $\lambda \in \mathcal{U}$.

3. Answer to Question 1

In this section, we present an example of an assortment planning problem with two past assortments in which the optimal objective value of the robust optimization problem (RO) is strictly greater than the expected revenue of the firm's best past assortment. The example is insightful for three reasons. First, the example shows that there exist problem instances in which the assortment that satisfies (IMP) is not a revenue-ordered assortment (the formal definition of revenue-ordered assortments is found in §4). Second, the example shows that estimate-then-optimize can perform poorly—that is, yield an assortment with an expected revenue that is strictly less than the expected revenue of the firm's best past assortment—even in problem instances for which there exist assortments that satisfy (IMP). Third, the example illustrates that assortments that satisfy (IMP) can have a non-trivial structure, and discovering such assortments appears to be challenging without the aid of structural results and algorithms like those developed in §4 and §5.

We begin by defining the assortment planning problem that will be our focus throughout this section. The problem instance is comprised by a universe of four products $\mathcal{N} \equiv \{1, 2, 3, 4\}$, and the two past assortments that the firm has previously offered to its customers are denoted by $S_1 = \{0, 2, 3, 4\}$ and $S_2 = \{0, 1, 2, 4\}$. The revenues corresponding to the four products are $r_1 = \$10$, $r_2 = \$20$, $r_3 = \$30$, and $r_4 = \$100$. The historical sales data generated by the past assortments are the purchase frequencies $(v_{1,0}, v_{1,2}, v_{1,3}, v_{1,4}) = (0.3, 0.3, 0.3, 0.1)$ and $(v_{2,0}, v_{2,1}, v_{2,2}, v_{2,4}) = (0.3, 0.3, 0.1, 0.3)$. We analyze the robust optimization problem (RO) in which the uncertainty set is constructed as all ranking-based choice models that perfectly fit this historical sales data, i.e., the radius of \mathcal{U} is $\eta = 0$.

Let us offer some useful facts about the problem instance. We first observe that the expected revenues generated by the two past assortments $S_1 = \{0, 2, 3, 4\}$ and $S_2 = \{0, 1, 2, 4\}$ are

$$\begin{aligned} r^\top v_1 &= \$0 \times 0.3 + \$20 \times 0.3 + \$30 \times 0.3 + \$100 \times 0.1 = \$25, \\ r^\top v_2 &= \$0 \times 0.3 + \$10 \times 0.3 + \$20 \times 0.1 + \$100 \times 0.3 = \$35. \end{aligned}$$

Hence, the expected revenue generated by the firm's best past assortment is $\max\{\$25, \$35\} = \$35$. We next show that the set of ranking-based choice models \mathcal{U} is nonempty when the radius is $\eta = 0$. To show this, we will construct a ranking-based choice model λ that perfectly fits the historical sales data. Specifically, consider the following five rankings:

$$\begin{array}{ccccc} \sigma^1(0) = 0, & \sigma^2(1) = 0, & \sigma^3(1) = 0, & \sigma^4(2) = 0, & \sigma^5(3) = 0, \\ \sigma^1() = 1, & \sigma^2(2) = 1, & \sigma^3(4) = 1, & \sigma^4(4) = 1, & \sigma^5(4) = 1, \\ \sigma^1() = 2, & \sigma^2(4) = 2, & \sigma^3(0) = 2, & \sigma^4(0) = 2, & \sigma^5(0) = 2, \\ \sigma^1() = 3, & \sigma^2(0) = 3, & \sigma^3() = 3, & \sigma^4() = 3, & \sigma^5() = 3, \\ \sigma^1() = 4, & \sigma^2() = 4, & \sigma^3() = 4, & \sigma^4() = 4, & \sigma^5() = 4. \end{array}$$

We recall from §2 that a ranking $\sigma \in \Sigma$ prefers product $i \in \mathcal{N}_0$ over product $j \in \mathcal{N}_0$ if $\sigma(i) < \sigma(j)$. For the sake of simplicity, we have chosen without loss of generality to not specify the preference ordering among products that are less preferred than the no-purchase option in each ranking. We observe that the most preferred products from the past assortments $S_1 = \{0, 2, 3, 4\}$ and $S_2 = \{0, 1, 2, 4\}$ under each of the five rankings are

$$\begin{array}{ccccc} \arg \min_{j \in S_1} \sigma^1(j) = 0, & \arg \min_{j \in S_1} \sigma^2(j) = 2, & \arg \min_{j \in S_1} \sigma^3(j) = 4, & \arg \min_{j \in S_1} \sigma^4(j) = 2, & \arg \min_{j \in S_1} \sigma^5(j) = 3, \\ \arg \min_{j \in S_2} \sigma^1(j) = 0, & \arg \min_{j \in S_2} \sigma^2(j) = 1, & \arg \min_{j \in S_2} \sigma^3(j) = 1, & \arg \min_{j \in S_2} \sigma^4(j) = 2, & \arg \min_{j \in S_2} \sigma^5(j) = 4. \end{array}$$

It follows from the above observations that the historical sales data $(v_{1,0}, v_{1,2}, v_{1,3}, v_{1,4}) = (0.3, 0.3, 0.3, 0.1)$ and $(v_{2,0}, v_{2,1}, v_{2,2}, v_{2,4}) = (0.3, 0.3, 0.1, 0.3)$ is perfectly fit by the ranking-based choice model $\hat{\lambda} \in \Delta_{\Sigma}$ that satisfies the following equalities:

$$\hat{\lambda}_{\sigma^1} = 0.3, \quad \hat{\lambda}_{\sigma^2} = 0.2, \quad \hat{\lambda}_{\sigma^3} = 0.1, \quad \hat{\lambda}_{\sigma^4} = 0.1, \quad \hat{\lambda}_{\sigma^5} = 0.3.$$

Equipped with the above useful facts, we now consider the performance of estimate-then-optimize for this specific problem instance. Specifically, suppose the firm uses estimate-then-optimize in which the selected ranking-based choice model is the probability distribution $\hat{\lambda}$. In that case, estimate-then-optimize will output a new assortment $S^{\hat{\lambda}}$ that is an optimal solution to the combinatorial optimization problem $\max_{S \in \mathcal{S}} \mathcal{R}^{\hat{\lambda}}(S)$. To determine the optimal solution to this combinatorial optimization problem, let us inspect the five rankings $\sigma^1, \dots, \sigma^5$ that have nonzero probability under the estimated probability distribution. We observe that the first of these five rankings satisfies $\arg \min_{j \in \mathcal{S}} \sigma^1(j) = 0$ for all assortments $S \in \mathcal{S}$. Moreover, we observe that the assortment $\{0, 4\} \in \mathcal{S}$ satisfies $\arg \min_{j \in \{0,4\}} \sigma^2(j) = \dots = \arg \min_{j \in \{0,4\}} \sigma^5(j) = 4$. We conclude that $\{0, 4\}$ is an optimal solution for the above combinatorial optimization problem, and the predicted expected revenue for this assortment is $\mathcal{R}^{\hat{\lambda}}(\{0, 4\}) = 0.7 \times \$100 = \$70$.

We will now show that the assortment $S^{\hat{\lambda}} \triangleq \{0, 4\}$ obtained from estimate-then-optimize can have an expected revenue that is strictly less than the expected revenue of the firm's best past assortment under a ranking-based choice model that is consistent with the historical sales data. To show this, we consider the following seven rankings:

$$\begin{aligned} \sigma^6(0) &= 0, & \sigma^7(1) &= 0, & \sigma^8(2) &= 0, & \sigma^9(3) &= 0, & \sigma^{10}(4) &= 0, & \sigma^{11}(1) &= 0, & \sigma^{12}(3) &= 0, \\ \sigma^6(\cdot) &= 1, & \sigma^7(0) &= 1, & \sigma^8(0) &= 1, & \sigma^9(0) &= 1, & \sigma^{10}(0) &= 1, & \sigma^{11}(2) &= 1, & \sigma^{12}(4) &= 1, \\ \sigma^6(\cdot) &= 2, & \sigma^7(\cdot) &= 2, & \sigma^8(\cdot) &= 2, & \sigma^9(\cdot) &= 2, & \sigma^{10}(\cdot) &= 2, & \sigma^{11}(0) &= 2, & \sigma^{12}(0) &= 2, \\ \sigma^6(\cdot) &= 3, & \sigma^7(\cdot) &= 3, & \sigma^8(\cdot) &= 3, & \sigma^9(\cdot) &= 3, & \sigma^{10}(\cdot) &= 3, & \sigma^{11}(\cdot) &= 3, & \sigma^{12}(\cdot) &= 3, \\ \sigma^6(\cdot) &= 4, & \sigma^7(\cdot) &= 4, & \sigma^8(\cdot) &= 4, & \sigma^9(\cdot) &= 4, & \sigma^{10}(\cdot) &= 4, & \sigma^{11}(\cdot) &= 4, & \sigma^{12}(\cdot) &= 4. \end{aligned}$$

We observe that the most preferred products from the first past assortment $S_1 = \{0, 2, 3, 4\}$ are

$$\begin{aligned} \arg \min_{j \in S_1} \sigma^6(j) &= 0, & \arg \min_{j \in S_1} \sigma^7(j) &= 0, & \arg \min_{j \in S_1} \sigma^8(j) &= 2, & \arg \min_{j \in S_1} \sigma^9(j) &= 3, & \arg \min_{j \in S_1} \sigma^{10}(j) &= 4, \\ & & \arg \min_{j \in S_1} \sigma^{11}(j) &= 2, & \arg \min_{j \in S_1} \sigma^{12}(j) &= 3, \end{aligned}$$

the most preferred products from the second past assortment $S_2 = \{0, 1, 2, 4\}$ are

$$\begin{aligned} \arg \min_{j \in S_2} \sigma^6(j) &= 0, & \arg \min_{j \in S_2} \sigma^7(j) &= 1, & \arg \min_{j \in S_2} \sigma^8(j) &= 2, & \arg \min_{j \in S_2} \sigma^9(j) &= 0, & \arg \min_{j \in S_2} \sigma^{10}(j) &= 4, \\ & & \arg \min_{j \in S_2} \sigma^{11}(j) &= 1, & \arg \min_{j \in S_2} \sigma^{12}(j) &= 4, \end{aligned}$$

and the most preferred products from the assortment $S^{\hat{\lambda}} = \{0, 4\}$ are

$$\begin{aligned} \arg \min_{j \in S^{\hat{\lambda}}} \sigma^6(j) &= 0, & \arg \min_{j \in S^{\hat{\lambda}}} \sigma^7(j) &= 0, & \arg \min_{j \in S^{\hat{\lambda}}} \sigma^8(j) &= 0, & \arg \min_{j \in S^{\hat{\lambda}}} \sigma^9(j) &= 0, & \arg \min_{j \in S^{\hat{\lambda}}} \sigma^{10}(j) &= 4, \\ & & \arg \min_{j \in S^{\hat{\lambda}}} \sigma^{11}(j) &= 0, & \arg \min_{j \in S^{\hat{\lambda}}} \sigma^{12}(j) &= 4. \end{aligned}$$

It follows from the above observations that the historical sales data $(v_{1,0}, v_{1,2}, v_{1,3}, v_{1,4}) = (0.3, 0.3, 0.3, 0.1)$ and $(v_{2,0}, v_{2,1}, v_{2,2}, v_{2,4}) = (0.3, 0.3, 0.1, 0.3)$ are perfectly fit by the ranking-based choice model $\bar{\lambda} \in \Delta_{\Sigma}$ that satisfies the following equalities:

$$\bar{\lambda}_{\sigma^6} = 0.2, \quad \bar{\lambda}_{\sigma^7} = 0.1, \quad \bar{\lambda}_{\sigma^8} = 0.1, \quad \bar{\lambda}_{\sigma^9} = 0.1, \quad \bar{\lambda}_{\sigma^{10}} = 0.1, \quad \bar{\lambda}_{\sigma^{11}} = 0.2, \quad \bar{\lambda}_{\sigma^{12}} = 0.2.$$

However, we observe that the expected revenue for the assortment $S^{\hat{\lambda}} = \{0, 4\}$ under the ranking-based choice model $\bar{\lambda}$ is $\mathcal{R}^{\bar{\lambda}}(S^{\hat{\lambda}}) = 0.3 \times \$100 = \$30$. Since this is strictly less than the expected revenue generated by the firm's best past assortment, \$35, we have thus shown that estimate-then-optimize can yield an assortment for the problem instance with an expected revenue that is strictly worse than the expected revenue generated by the firm's best past assortment under a ranking-based choice model that is consistent with the historical sales data.

We conclude by showing that $\{0, 2, 4\}$ is the unique optimal solution for the robust optimization problem (RO) and is the unique assortment that satisfies (IMP). To show this, we present below an exhaustive list of the worst-case expected revenues $\min_{\lambda \in \mathcal{U}} \mathcal{R}^{\lambda}(S)$ for all assortments $S \in \mathcal{S}$ that satisfy $4 \in S$.⁵ The following quantities are computed using the algorithm from Lemma 7 in §5.1.

$$\begin{aligned} \min_{\lambda \in \mathcal{U}} \mathcal{R}^{\lambda}(\{0, 4\}) &= \$30, & \min_{\lambda \in \mathcal{U}} \mathcal{R}^{\lambda}(\{0, 1, 4\}) &= \$33, \\ \min_{\lambda \in \mathcal{U}} \mathcal{R}^{\lambda}(\{0, 2, 4\}) &= \$36, & \min_{\lambda \in \mathcal{U}} \mathcal{R}^{\lambda}(\{0, 3, 4\}) &= \$19, \\ \min_{\lambda \in \mathcal{U}} \mathcal{R}^{\lambda}(\{0, 1, 2, 4\}) &= \$35, & \min_{\lambda \in \mathcal{U}} \mathcal{R}^{\lambda}(\{0, 1, 3, 4\}) &= \$12, \\ \min_{\lambda \in \mathcal{U}} \mathcal{R}^{\lambda}(\{0, 2, 3, 4\}) &= \$25, & \min_{\lambda \in \mathcal{U}} \mathcal{R}^{\lambda}(\{0, 1, 2, 3, 4\}) &= \$14. \end{aligned}$$

We observe that the assortment $\{0, 2, 4\}$ satisfies (IMP), since this assortment satisfies $\mathcal{R}^{\lambda}(\{0, 2, 4\}) \geq \$36 > \$35$ for all ranking-based choice models $\lambda \in \mathcal{U}$. In particular, we note that $\{0, 2, 4\}$ is not a revenue-ordered assortment. We remark that the correctness of the above equalities $\min_{\lambda \in \mathcal{U}} \mathcal{R}^{\lambda}(\{0, 1, 2, 4\}) = \35 and $\min_{\lambda \in \mathcal{U}} \mathcal{R}^{\lambda}(\{0, 2, 3, 4\}) = \25 for the two past assortments follows immediately from the fact that $\eta = 0$, which implies that $\mathcal{D}^{\lambda}(S_m) = v_m$ for all $m \in \{1, 2\}$ and all $\lambda \in \mathcal{U}$. We thus conclude that there exist problem instances for which there exists an assortment that satisfies (IMP), and have therefore provided an affirmative answer to Question 1.

4. Characterization of Optimal Assortments for (RO)

In this section, we establish the first characterization of optimal assortments for the robust optimization problem (RO). In particular, we show that there are optimal assortments for (RO) with a simple structure that is closely related to the structure of revenue-ordered assortments. Recall the following definition of the collection of revenue-ordered assortments:

$$\bar{\mathcal{S}} \triangleq \{S \in \mathcal{S} : \text{if } i^* \in S \text{ and } r_{i^*} < r_i, \text{ then } i \in S\}.$$

A fundamental result in the theory of assortment optimization is that revenue-ordered assortments are optimal under the multinomial logit choice model (Talluri and Van Ryzin 2004, Gallego et al. 2004, Rusmevichientong et al. 2014). Revenue-ordered assortments also have attractive approximation guarantees for assortment optimization problems under mixture-of-logits and ranking-based choice models with known parameters (Rusmevichientong et al. 2014, Aouad et al. 2018, Berbeglia and Joret 2020). Due to their simplicity and strong theoretical and empirical performance, revenue-ordered assortments are widely recommended in the revenue management literature.

In view of the above background, we proceed to develop our main result regarding the structure of optimal assortments for the robust optimization problem (RO). To this end, we first define the following set of past assortments for each product $i \in \mathcal{N}_0$:

$$\mathcal{M}_i \triangleq \{m \in \mathcal{M} : i \in S_m\}.$$

⁵ For each assortment $S \in \mathcal{S}$ that satisfies $4 \notin S$ and for every ranking-based choice model $\lambda \in \Delta_{\Sigma}$, we observe that $\mathcal{R}^{\lambda}(S) \leq 1 \times r_3 = \30 . Because \$30 is strictly less than the expected revenue generated by the firm's best past assortment, \$35, we can without loss of generality restrict our analysis to assortments that satisfy $4 \in S$.

The above set contains all of the past assortments in which the firm offered product i to their customers. In particular, we observe from this definition that the statement $\mathcal{M}_i \subseteq \mathcal{M}_j$ holds if, for all of the past assortments in which the firm offered product i , the firm also offered product j . We now introduce the following new collection of assortments:

$$\widehat{\mathcal{S}} \triangleq \{S \in \mathcal{S} : \text{if } i^* \in S, r_{i^*} < r_i, \text{ and } \mathcal{M}_{i^*} \subseteq \mathcal{M}_i, \text{ then } i \in S\}.$$

The above collection has a natural interpretation as the collection of all assortments which, speaking informally, can be viewed as revenue-ordered *relative to* the firm's past assortments. Indeed, consider two products $i^*, i \in \mathcal{N}_0$ for which the revenue r_{i^*} from the first product i^* is strictly less than the revenue r_i from the second product i . Then every assortment $S \in \widehat{\mathcal{S}}$ which offers the first product must also offer the second product *unless* there is historical sales data from a past assortment in which the first product i^* was offered and the second product i was not offered. Our main result is the following:

THEOREM 1. *There exists an assortment $S \in \widehat{\mathcal{S}}$ that is optimal for (RO).*

The remainder of §4 focuses on the implications and derivation of Theorem 1. In §4.1, we use Theorem 1 to show that the past assortments must deviate from revenue-ordered assortments for the optimal objective value of (RO) to be strictly greater than the expected revenue of the firm's best past assortment. In §4.2, we develop upper bounds on the cardinality of $\widehat{\mathcal{S}}$ based on the number of past assortments and number of products. In §4.3, we prove that $\widehat{\mathcal{S}}$ can be an exact description of the assortments that are optimal for (RO). In §4.4, we provide an overview of the proof of Theorem 1.

4.1. Answer to Question 2

Due to revenue-ordered assortment's simplicity and desirable theoretical guarantees, a large body of literature has advocated to firms for offering revenue-ordered assortments across numerous application domains. Equipped with Theorem 1, we show in the following Lemma 1 and Corollary 1 that a firm which has offered the revenue-ordered assortments has operated under a worst-possible behavior from the perspective of finding an assortment that satisfies (IMP).

LEMMA 1. *If $\mathcal{M} = \bar{\mathcal{S}}$, then $\widehat{\mathcal{S}} = \mathcal{M}$.*

COROLLARY 1. *If $\mathcal{M} = \bar{\mathcal{S}}$ and $\eta = 0$, then $\max_{S \in \mathcal{S}} \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) = \max_{m \in \mathcal{M}} r^\top v_m$.*

The above lemma and corollary establish that it is impossible to find an assortment that satisfies (IMP) when the firm's past assortments are the revenue-ordered assortments. Specifically, it follows from Lemma 1 and Theorem 1 that if the firm's past assortments are the revenue-ordered assortments, then there must exist a revenue-ordered assortment that is an optimal solution for the robust optimization problem (RO), regardless of the prices of the products r and the observations of the historical sales data v_1, \dots, v_M . It thus follows as an immediate corollary of Lemma 1 and Theorem 1 in the case of $\eta = 0$ that the past assortments must deviate from revenue-ordered assortments in order for it to be possible for the optimal objective value of (RO) to be strictly greater than the expected revenue of the firm's best past assortment.

Motivated by Lemma 1 and Corollary 1, we perform numerical experiments in Appendix A.1 to assess the performance of assortments obtained using estimate-then-optimize when the historical sales data is randomly generated from revenue-ordered assortments. The results of the experiments are striking: in more than 98% of the problem instances in which estimate-then-optimize

recommended a new assortment, the worst-case decline in expected revenue from implementing the new assortment (relative to the expected revenue from the best past assortment) exceeded the best-case increase in expected revenue. The difference in magnitude between the worst-case and best-case change in expected revenue from implementing the new assortment found from estimate-then-optimize is also significant; the average best-case improvement of the new assortment over the best past assortment is 6.56%, while the average worst-case improvement of the new assortment over the best past assortment is -21.71%. Our results in Appendix A.1 thus demonstrate that estimate-then-optimize with ranking-based choice models can risk leading to significant declines in a firm's expected revenue, even in settings in which the firm has implemented a celebrated and widely-recommended class of assortments.

4.2. Upper Bounds

In §5, we will design algorithms for solving the class of robust optimization problems (RO) that consist of exhaustive search over the assortments in the collection $\widehat{\mathcal{S}}$. To motivate these algorithms, we show in the present subsection that the cardinality of $\widehat{\mathcal{S}}$ scales as a polynomial in n for any fixed number of past assortments.

We begin by considering the case of two past assortments. In the following Lemma 2, we develop a closed-form representation of the collection of assortments $\widehat{\mathcal{S}}$. This representation will be useful in §5.1 because it provides an efficient procedure for iterating over the assortments in $\widehat{\mathcal{S}}$. Moreover, the following Lemma 2 is useful because it immediately implies that the number of assortments in the collection $\widehat{\mathcal{S}}$ scales quadratically in the number of products n when there are two past assortments.

LEMMA 2. *If $M = 2$, then*

$$\widehat{\mathcal{S}} = \left\{ S \in \mathcal{S} : \begin{array}{l} \text{there exists } i_1 \in S_1 \setminus S_2 \text{ and } i_2 \in S_2 \setminus S_1 \text{ such that} \\ S = (S_1 \cap S_2) \cup \{j \in S_1 \setminus S_2 : j \geq i_1\} \cup \{j \in S_2 \setminus S_1 : j \geq i_2\} \end{array} \right\}. \quad (1)$$

Lemma 2 shows that the assortments in the collection $\widehat{\mathcal{S}}$ can be parameterized by the pairs of products from the sets $S_1 \setminus S_2$ and $S_2 \setminus S_1$. In particular, Lemma 2 shows for the case of two past assortments that each assortment $S \in \widehat{\mathcal{S}}$ contains all of the products in $S_1 \cap S_2$, contains all of the products in $S_1 \setminus S_2$ with revenue greater than or equal to r_{i_1} for some $i_1 \in S_1 \setminus S_2$, and contains all of the products in $S_2 \setminus S_1$ with revenue greater than or equal to r_{i_2} for some $i_2 \in S_1 \setminus S_2$. Hence, the number of assortments in the collection $\widehat{\mathcal{S}}$ is at most $|S_1 \setminus S_2| \times |S_2 \setminus S_1| = \mathcal{O}(n^2)$.

We now develop an upper bound on the cardinality of $\widehat{\mathcal{S}}$ for general numbers of past assortments. To motivate our upper bound, we recall that the number of assortments in the collection $\widehat{\mathcal{S}}$ can indeed be small in special cases of problem instances: namely, we recall from Lemma 1 in §4.1 that $|\widehat{\mathcal{S}}| = \mathcal{O}(n)$ when the past assortments are the revenue-ordered assortments, and we showed in Lemma 2 that $|\widehat{\mathcal{S}}| = \mathcal{O}(n^2)$ when there are two past assortments. In the next intermediary result, denoted by Lemma 3, we develop a more general result along these lines. Specifically, the following lemma establishes that the number of assortments in the collection $\widehat{\mathcal{S}}$ can be upper bounded by a polynomial of the number of products n for *any* fixed number of past assortments M . While the bound in the following lemma is not the tightest possible, the bound will be sufficient for its theoretical purpose in designing polynomial-time algorithms in §5.2 for solving the robust optimization problem (RO) when the number of past assortments M is fixed.

LEMMA 3. $|\widehat{\mathcal{S}}| \leq (n + 2)^{2^M}$.

We conclude the present §4.2 by developing an algorithm for computing the collection $\widehat{\mathcal{S}}$. The algorithm which achieves the specified running time of the following lemma is based on dynamic programming over a compact graphical representation of the collection of assortments $\widehat{\mathcal{S}}$. In particular, the running time of the algorithm from Lemma 4 can be viewed as attractive due to its mild dependence on the number of assortments in the collection $\widehat{\mathcal{S}}$. For example, it follows from Lemma 3 that the algorithm from the following lemma has a running time that is polynomial in the number of products n for any fixed number of past assortments M .

LEMMA 4. *The collection of assortments $\widehat{\mathcal{S}}$ can be constructed in $\mathcal{O}(n^2(M + |\widehat{\mathcal{S}}|))$ time.*

4.3. Discussion of Upper Bound

In Lemma 3 of the previous subsection, we developed an upper bound on $|\widehat{\mathcal{S}}|$ that grows exponentially with respect to the number of past assortments. In this subsection, we provide an example that shows that $|\widehat{\mathcal{S}}|$ can indeed grow exponentially with respect to the number of past assortments, and we show for that example that $\widehat{\mathcal{S}}$ is an exact description of the assortments that can be optimal for the robust optimization problem (RO). Our discussion will focus on the following specific collection of past assortments:

DEFINITION 1. We say that the firm's past assortments are *reverse revenue-ordered* if $S_1 \triangleq \{0, n\}$, $S_m \triangleq \{0, 1, \dots, m-1, n\}$ for each $m \in \{2, \dots, n\}$, and $\mathcal{M} = \widetilde{\mathcal{S}} \triangleq \{S_1, \dots, S_n\}$.

We readily observe from the above definition that the reverse revenue-ordered assortments are nested, in the sense that $S_1 \subset \dots \subset S_n$. Moreover, we observe that the following lemma follows immediately from the definition of the collection of assortments $\widehat{\mathcal{S}}$ from §4.

LEMMA 5. *If $\mathcal{M} = \widetilde{\mathcal{S}}$, then $\widehat{\mathcal{S}} = \{S \subseteq \mathcal{N}_0 : \{0, n\} \subseteq S\}$.*

The above lemma shows that if the past assortments \mathcal{M} are the reverse revenue-ordered assortments, then the corresponding collection of assortments $\widehat{\mathcal{S}}$ is the collection of all subsets of products that include the most expensive product n and the no-purchase option 0. It follows immediately from Lemma 5 and from the fact that $M = n$ that $|\widehat{\mathcal{S}}| = 2^{M-2}$ when the past assortments are the reversed revenue-ordered assortments. We have thus shown that the cardinality of the collection of assortments $\widehat{\mathcal{S}}$ can grow exponentially in the number of past assortments.

Our main contribution of §4.3, which is presented below as Theorem 2, consists of establishing that the characterization of optimal assortments from Theorem 1 cannot be refined for reverse revenue-ordered assortments. More precisely, Theorem 2 shows that $\widehat{\mathcal{S}}$ is exactly the collection of assortments that can be optimal for the robust optimization problem (RO) when the past assortments are reverse revenue-ordered. This theorem is significant because it proves that the collection of assortments $\widehat{\mathcal{S}}$ is an accurate description of the range of possible optimal solutions for the robust optimization problem (RO), at least for the case when the past assortments are reverse revenue-ordered. It also is significant because it implies that the exponential growth of the cardinality of the collection of assortments $\widehat{\mathcal{S}}$ that is established in Lemma 5 is not the result of constructing an overly-conservative collection of assortments $\widehat{\mathcal{S}}$.

THEOREM 2. *Suppose that $\mathcal{M} = \widetilde{\mathcal{S}}$, that $\eta = 0$, and that the revenues $r_1 < \dots < r_n$ are fixed. Then for every assortment $\bar{S} \in \widehat{\mathcal{S}}$, there exists a realization of the historical data $v \equiv (v_{m,i} : m \in \mathcal{M}, i \in S_m)$ such that \bar{S} is the unique optimal solution of the robust optimization problem (RO).*

In summary, we have shown in the present §4.3 that there exist examples in which $|\widehat{\mathcal{S}}|$ grows exponentially in the number of past assortments and in which $\widehat{\mathcal{S}}$ is an accurate description of the

range of possible optimal solutions for the robust optimization problem (RO). Therefore, unless the firm's past assortments have a special structure such as revenue-ordered assortments (see Lemma 1), we conclude that solving (RO) via an exhaustive search over the assortments in $\widehat{\mathcal{S}}$ will be computationally prohibitive when the number of past assortments is large. For such settings, we show in §5.3 that our proof techniques from the following §4.4 can nonetheless be extended to develop compact mixed-integer optimization reformulations of (RO) when the collection of past assortments enjoys a common structure.

4.4. Overview of Proof of Theorem 1

Our proof of Theorem 1 is split into two steps. In the first step of the proof of Theorem 1, which is formalized below as Proposition 1, we show that computing the worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ corresponding to any fixed assortment $S \in \mathcal{S}$ can be reformulated as a linear optimization problem with $\mathcal{O}(n^M)$ decision variables and $\mathcal{O}(nM)$ constraints. Indeed, for each assortment $S \in \mathcal{S}$ and each product in the assortment $i \in S$, let the set of rankings that prefer product i to all other products in the assortment S be defined as follows:

DEFINITION 2. $\mathcal{D}_i(S) \triangleq \{\sigma \in \Sigma : i = \arg \min_{j \in S} \sigma(j)\}$.

Given the past assortments S_1, \dots, S_M , we also define the following set of tuples of products:

DEFINITION 3. $\mathcal{L} \triangleq \{(i_1, \dots, i_M) \in S_1 \times \dots \times S_M : \bigcap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m) \neq \emptyset\}$.

We readily observe from the above definitions that a tuple of products satisfies $(i_1, \dots, i_M) \in \mathcal{L}$ if and only if there exists a ranking $\sigma \in \Sigma$ for which i_1 is the most preferred product under that ranking in assortment S_1 , i_2 is the most preferred product under that ranking in assortment S_2 , and so on and so forth. For convenience, we henceforth say that a ranking $\sigma \in \Sigma$ corresponds to a tuple $(i_1, \dots, i_M) \in \mathcal{L}$ if and only if $\sigma \in \bigcap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m)$. Finally, for each assortment $S \in \mathcal{S}$ and each tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$, we let $\rho_{i_1 \dots i_M}(S)$ be defined as the minimum revenue among the products in the assortment S that can be the most preferred product in S under a ranking that corresponds to the tuple (i_1, \dots, i_M) .

DEFINITION 4. $\rho_{i_1 \dots i_M}(S) \triangleq \min_{i \in S : \bigcap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m) \cap \mathcal{D}_i(S) \neq \emptyset} r_i$.

Equipped with the above notation, the following result shows that the worst-case expected revenue of any fixed assortment can be reformulated as a linear optimization problem:

PROPOSITION 1. *For each $S \in \mathcal{S}$, $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ is equal to the optimal objective value of the following linear optimization problem:*

$$\begin{aligned}
& \underset{\lambda, \epsilon}{\text{minimize}} && \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \rho_{i_1 \dots i_M}(S) \lambda_{i_1 \dots i_M} \\
& \text{subject to} && \sum_{(i_1, \dots, i_M) \in \mathcal{L} : i_m = i} \lambda_{i_1 \dots i_M} - \epsilon_{m,i} = v_{m,i} \quad \forall m \in \mathcal{M}, i \in S_m \\
& && \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \lambda_{i_1 \dots i_M} = 1 \\
& && \|\epsilon\| \leq \eta \\
& && \lambda_{i_1 \dots i_M} \geq 0 \quad \forall (i_1, \dots, i_M) \in \mathcal{L}.
\end{aligned} \tag{2}$$

The decision variables $\lambda_{i_1 \dots i_M} \geq 0$ in the linear optimization problem (2) can be interpreted as the likelihood that a randomly-arriving customer will have preferences that follow a ranking that

corresponds to the tuple $(i_1, \dots, i_M) \in \mathcal{L}$. The following two remarks articulate the key properties of (2) that will be important to our developments in the subsequent discussions.

REMARK 1. The assortment S does not appear in any of the constraints of (2).

REMARK 2. (2) has $|\mathcal{L}| = \mathcal{O}(n^M)$ decision variables and $\mathcal{O}(nM)$ constraints.

We note that the asymptotic upper bound in Remark 2 on the number of decision variables in the linear optimization problem (2) follows immediately from the fact that the set of tuples of products \mathcal{L} is a subset of $S_1 \times \dots \times S_M$.

The second step of the proof of Theorem 1, which is formalized below as Proposition 2, consists of analyzing the behavior of the functions $S \mapsto \rho_{i_1 \dots i_M}(S)$. To motivate this second step, let us reflect on the structure of the linear optimization problem from Proposition 1. We recall from Remark 1 that the set of feasible solutions of the linear optimization problem (2) is independent of the assortment $S \in \mathcal{S}$. We thus observe that if two assortments $S, S' \in \mathcal{S}$ satisfy $\rho_{i_1 \dots i_M}(S) \leq \rho_{i_1 \dots i_M}(S')$ for all tuples of products $(i_1, \dots, i_M) \in \mathcal{L}$, then Proposition 1 implies that the worst-case expected revenue of the first assortment, $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$, will be less than or equal to the worst-case expected revenue of the second assortment, $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S')$. Against this backdrop, the following Proposition 2 readily implies that every assortment $S \in \mathcal{S}$ can be exchanged for an assortment $S' \in \widehat{\mathcal{S}}$ with the same or better worst-case expected revenue.

PROPOSITION 2. *Let $S \in \mathcal{S}$ and $i \notin S$. If there exists $i^* \in S$ that satisfies $r_{i^*} < r_i$ and $\mathcal{M}_{i^*} \subseteq \mathcal{M}_i$, then $\rho_{i_1 \dots i_M}(S) \leq \rho_{i_1 \dots i_M}(S \cup \{i\})$ for each $(i_1, \dots, i_M) \in \mathcal{L}$.*

Our proof of Proposition 2 consists of developing a graphical interpretation of the quantities $\rho_{i_1 \dots i_M}(S)$. Specifically, we show in Appendix B.2 that every tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$ can be represented as a directed acyclic graph in which there is a vertex for each product $i \in \mathcal{N}_0$, and we show that $\rho_{i_1 \dots i_M}(S)$ can be computed by searching over the products in the directed acyclic graph corresponding to the tuple $(i_1, \dots, i_M) \in \mathcal{L}$ that do not have a directed path to any vertices corresponding to the products in the assortment S . We then exploit this graphical interpretation of $\rho_{i_1 \dots i_M}(S)$ in Appendix B.3 to relate the quantities $\rho_{i_1 \dots i_M}(S)$ and $\rho_{i_1 \dots i_M}(S \cup \{i\})$ and complete the proof of Proposition 2. The proof of Theorem 1, which follows readily from Propositions 1 and 2, is found in Appendix B.4.

5. Algorithms

In this section, we use our developments from §4 to develop the first polynomial-time algorithms and the first compact mixed-integer optimization reformulations for the robust optimization problem (RO).

REMARK 3. We assume throughout §5 that every product $i \in \mathcal{N}_0$ has been offered by the firm to its customers in at least one of the past assortments, that is, $\cup_{m \in \mathcal{M}} S_m = \mathcal{N}_0$. This assumption simplifies our notation and can always be satisfied by redefining \mathcal{N}_0 to be equal to $\cup_{m \in \mathcal{M}} S_m$. Such a redefining of \mathcal{N}_0 does not impact the optimal objective value of the robust optimization problem (RO); see Appendix F for further details.

5.1. Polynomial-Time Algorithm for Two Past Assortments

We begin by leveraging Theorem 1 to develop the first algorithm for solving the robust optimization problem (RO) with running time that is strongly polynomial in the number of products n for two

past assortments. Our development of an algorithm for the case of $M = 2$ is important because it marks the first step towards the development of efficient *general* algorithms for solving the robust optimization problem (RO) in settings with large numbers of products. We will show that the algorithms from this subsection extend to general settings with any fixed number of past assortments in the following §5.2. In Appendix A.2, we provide numerical experiments showing that our algorithm for two past assortments can be practically efficient, running in less than 30 seconds for problem instances with $n = 100$ products.

At a high level, our algorithm for solving the robust optimization problem (RO) in the case of $M = 2$ consists of reducing (RO) to solving a sequence of minimum-cost network flow problems. Stated formally, the main contribution of §5.1 is the following:

THEOREM 3. *If $M = 2$ and $\eta = 0$, then (RO) can be solved in $\mathcal{O}(n^5 \log(nr_n))$ time.*

In the above theorem and throughout the rest of §5.1, we assume that the revenues r_1, \dots, r_n are represented as positive integers. We also assume without any loss of generality that $n \in S_1 \cap S_2$.⁶

To prove Theorem 3, and to develop an algorithm with the desired computation time, we utilize three intermediary results. The first intermediary result is Lemma 2 from §4.2. In that lemma, we developed a closed-form representation of the collection of assortments $\widehat{\mathcal{S}}$. That representation is useful in the proof of Theorem 3 because it provides an efficient procedure for iterating over the assortments in $\widehat{\mathcal{S}}$ and established that $|\widehat{\mathcal{S}}| = \mathcal{O}(n^2)$. In our second intermediary result, denoted by Lemma 6, we develop a closed-form representation of the set of tuples of products \mathcal{L} . The following representation of \mathcal{L} will be useful in the proof of Theorem 3 because it will allow us to show that the worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ for each assortment $S \in \mathcal{S}$ can be computed by solving a minimum-cost network flow problem.

LEMMA 6. *If $M = 2$, then $\mathcal{L} = ((S_1 \setminus S_2) \times S_2) \cup (S_1 \times (S_2 \setminus S_1)) \cup \{(i, i) : i \in S_1 \cap S_2\}$.*

In our third and final intermediary result, denoted by Lemma 7, we show that the worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ for each assortment $S \in \mathcal{S}$ can be computed by solving a minimum-cost network flow problem.

LEMMA 7. *If $M = 2$ and $\eta = 0$, then $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ can be evaluated for each assortment $S \in \mathcal{S}$ by solving a minimum-cost network flow problem in a directed acyclic graph with $\mathcal{O}(n)$ vertices and $\mathcal{O}(n^2)$ directed edges.*

Using the above three intermediary results, we now describe our algorithm for solving the robust optimization problem (RO), which follows a brute-force strategy. Namely, our algorithm iterates over each of the assortments $S \in \widehat{\mathcal{S}}$, and, for each such assortment, the algorithm computes the corresponding worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ by solving a minimum-cost network flow problem. The algorithm concludes by returning the maximum value of $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ across the assortments $S \in \widehat{\mathcal{S}}$. The correctness of this algorithm for solving the robust optimization problem (RO) follows immediately from Theorem 1. The proof that this algorithm achieves the desired running time of Theorem 3 can be found in Appendix G.3.

⁶ To see why the assumption that $n \in S_1 \cap S_2$ can be made without loss of generality, suppose for the sake of argument that the product n is not contained in $S_1 \cap S_2$. In that case, we can create a fictitious product with index $n + 1$ that is associated with any arbitrary revenue in the range $r_{n+1} = (r_n, \infty)$, and we can augment the historical sales data to be $v'_1 \triangleq (v_1, 0) \in \mathbb{R}_+^{n+2}$ and $v'_2 \triangleq (v_2, 0) \in \mathbb{R}_+^{n+2}$. It is straightforward to see that any feasible solution to (2) with the augmented historical sales data will satisfy $\lambda_{n+1, i_2} = \lambda_{i_1, n+1} = 0$ for all $i_1 \in S_1$ and $i_2 \in S_2$ that satisfy $(n + 1, i_2) \in \mathcal{L}$ and $(i_1, n + 1) \in \mathcal{L}$. This implies that the optimal objective value of (2) will be unchanged using the augmented historical sales data for all assortments $S \in \mathcal{S}$. Hence, we have shown that we can assume without loss of generality that the past assortments satisfy $n \in S_1 \cap S_2$.

5.2. Polynomial-Time Algorithm for Fixed Number of Past Assortments

We next develop a polynomial-time algorithm for solving the robust optimization problem (RO) for any fixed number of past assortments M . The algorithm developed in this subsection is particularly attractive due to its generality: it does not require any assumptions on the composition of products in the past assortments. Moreover, the algorithm allows for the radius η in the set of ranking-based choice models to be strictly positive (see §2); hence, the algorithm can be applied in settings in which there are no ranking-based choice models that perfectly fit the firm’s historical sales data.

Our polynomial-time algorithm for solving the robust optimization problem (RO) for any fixed number of past assortments follows the same strategy as developed in §5.1. Specifically, our algorithm reduces the robust optimization problem (RO) to computing the worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ for each candidate assortment $S \in \widehat{\mathcal{S}}$ and outputting the assortment that has the maximum worst-case expected revenue. The correctness of such a brute-force algorithm for solving the robust optimization problem (RO) follows immediately from Theorem 1. In the remainder of §5.2, we prove that there exists an implementation of the aforementioned brute-force algorithm which yields the following theoretical guarantee on its computational tractability:

THEOREM 4. (RO) can be solved in weakly $\mathcal{O}(\text{poly}(n))$ time for every fixed M .

The above theorem holds for any choice of the radius $\eta \geq 0$ in the set of ranking-based choice models and any composition of products in the past assortments. We note that our algorithm is guaranteed to run in weakly, as opposed to strongly, polynomial-time due to its reduction to solving linear optimization problems.

Our proof of Theorem 4 makes use of several intermediary results. First, we recall from Lemmas 3 and 4 in §4.2 that if the number of past assortments is fixed, then the collection of assortments $\widehat{\mathcal{S}}$ contains $\mathcal{O}(\text{poly}(n))$ assortments and can be constructed in $\mathcal{O}(\text{poly}(n))$ time. Therefore, all that remains is showing that the worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ can be computed for each assortment $S \in \widehat{\mathcal{S}}$ in weakly polynomial-time for every fixed number of past assortments M . The following Lemmas 8 and 9, together with Proposition 1 from §4.4, show for every fixed $M \geq 2$ and assortment $S \in \widehat{\mathcal{S}}$ that the worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ can be computed by solving a linear optimization problem that can be constructed in polynomial-time and has a polynomial number of decision variables and constraints.

LEMMA 8. $|\mathcal{L}| = \mathcal{O}(n^M)$, and \mathcal{L} can be constructed in $\mathcal{O}(Mn^{M+1})$ time.

LEMMA 9. For each assortment $S \in \widehat{\mathcal{S}}$ and each tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$, the quantity $\rho_{i_1 \dots i_M}(S)$ can be computed in $\mathcal{O}(M^2n)$ time.

Combining the above intermediary lemmas with Lemmas 3 and 4 from §4.2 and Proposition 1 from §4.4, we obtain an algorithm that satisfies the running time requirements of Theorem 4. The detailed description of the algorithm and the proof of its running time can be found in Appendix H.3.

5.3. Mixed-Integer Optimization Reformulation

In §5.1 and §5.2, we developed algorithms for solving the robust optimization problem (RO) that are theoretically efficient in applications in which the number of products n is large and the number of past assortments M is small. The crux of those algorithms was the observation that solving the robust optimization problem (RO) can be reduced to an exhaustive search over the collection of assortments $\widehat{\mathcal{S}}$. However, we showed in Theorem 2 from §4.3 that the cardinality of $\widehat{\mathcal{S}}$ can grow exponentially in the number of past assortments, thereby preventing the algorithm from §5.2 from

being computationally efficient in general in applications with large numbers of past assortments. In this subsection, we show nonetheless that our technical developments from §4.4 can also be used to develop practical algorithms for solving the robust optimization problem in applications in which the number of past assortments and number of products are both reasonably large.

Our discussion in the present subsection focuses on applications in which the collection of past assortments has a widely-studied structure called *nested*. Formally, the past assortments are nested if and only if there is an ordering of the past assortments such as that each assortment is a subset of the subsequent assortment:

ASSUMPTION 1 (Nested). *The past assortments $\mathcal{M} \equiv \{S_1, \dots, S_M\} \subseteq \mathcal{N}_0$ satisfy $S_1 \subset \dots \subset S_M$.*

Applications in which historical sales data is obtained from nested past assortments are “common in retail and revenue management settings” (Jagabathula and Rusmevichientong 2019, p.2202). For example, the past assortments will be nested if a retailer initially offers a large assortment of products to its customers and does not replenish inventory, in which the case the initial assortment will dwindle to smaller and smaller assortments as products run out of inventory. Nested past assortments also arise as new products become available in the marketplace, as well as in online retail as customers scroll through products across multiple pages of a website (Davis et al. 2015). Finally, nested assortments are a generalization of revenue-ordered assortments, and thus arise in any situations in which a firm has only offered revenue-ordered assortments.

Our main contribution of the present subsection is showing that our technical developments from §4.4 can be used to obtain a compact mixed-integer optimization reformulation of the robust optimization problem (RO) when the past assortments are nested. This reformulation can offer significant value compared to the algorithm from §5.2 when the number of past assortments is large; indeed, the number of decision variables and constraints of the mixed-integer optimization reformulation scale as a polynomial in the number of past assortments as well as the number of products. As a result, we show through numerical experiments in Appendix A.3 that our compact mixed-integer optimization reformulation can solve the robust optimization problem (RO) in less than 30 seconds in applications with as many as $M = 20$ past assortments. More generally, this subsection lays a roadmap for combining the technical developments behind Theorem 1 in §4.4 with structural properties of a collection of past assortments to develop practical algorithms for solving the robust optimization problem (RO) when $|\hat{S}|$ is large.

To motivate the construction of our mixed-integer optimization reformulation of the robust optimization problem (RO), we begin by showing for the case of nested past assortments that the linear optimization problem (2) has a number of decision variables that grows exponentially in the number of past assortments. Indeed, for each $m \in \mathcal{M}$, let the set of products that appear in the past assortment S_m and do not appear in any preceding past assortments be denoted by

$$\mathcal{B}_m \triangleq S_m \setminus \left(\bigcup_{m'=1}^{m-1} S_{m'} \right) = \begin{cases} S_1, & \text{if } m = 1, \\ S_m \setminus S_{m-1}, & \text{otherwise,} \end{cases} \quad (3)$$

where the right-most equality follows from algebra and from the assumption in §5.3 that the past assortments are nested. We observe from the above definition that the sets $\mathcal{B}_1, \dots, \mathcal{B}_M$ are disjoint. In view of this notation, the following lemma and its corollary show that the number of decision variables in the linear optimization problem (2) grows exponentially with respect to the number of past assortments M when the past assortments are nested.

LEMMA 10. *If Assumption 1 holds, then*

$$\mathcal{L} = \{(i_1, \dots, i_M) \in S_1 \times \dots \times S_M : \text{for each } m \in \{1, \dots, M-1\}, i_{m+1} \in \{i_m\} \cup \mathcal{B}_{m+1}\}.$$

COROLLARY 2. *If Assumption 1 holds, then $|\mathcal{L}| \geq 2^{M-1}$.*

With the above motivation, we now demonstrate that our technical developments behind Theorem 1 in §4.4 can be used to reformulate (2) as a linear optimization problem with numbers of decision variables and constraints that scale polynomially in both n and M . To explain our reformulation strategy, we observe in the case of $\eta = 0$ that the linear optimization problem (2) can be rewritten equivalently as

$$\underset{y \in \mathcal{Y}(S)}{\text{minimize}} \quad \sum_{i \in \mathcal{N}_0} r_i y_i, \quad (4)$$

where

$$\mathcal{Y}(S) \triangleq \left\{ y \in \mathbb{R}^{\mathcal{N}_0} : \left. \begin{array}{l} \text{there exists a vector } \lambda \equiv (\lambda_{i_1 \dots i_M} : (i_1, \dots, i_M) \in \mathcal{L}) \geq 0 \text{ such that:} \\ \sum_{(i_1, \dots, i_M) \in \mathcal{L} : \rho_{i_1 \dots i_M}(S) = r_i} \lambda_{i_1 \dots i_M} = y_i \quad \text{for all } i \in S \\ \sum_{(i_1, \dots, i_M) \in \mathcal{L} : i_m = i} \lambda_{i_1 \dots i_M} = v_{m,i} \quad \text{for all } m \in \mathcal{M}, i \in S_m \\ \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \lambda_{i_1 \dots i_M} = 1 \end{array} \right\}.$$

The number of decision variables in the linear optimization problem (4) is equal to $|\mathcal{N}_0| = n + 1$. Hence, we observe that the computational tractability of the linear optimization problem (4) depends on our ability to develop a compact representation of the set $\mathcal{Y}(S)$.⁷ In view of this observation, we proceed to show that the constraints on the decision variables $y \in \mathbb{R}^{\mathcal{N}_0}$ can be represented compactly when the past assortments are nested. The key insight that will enable the compact reformulation of the constraints on the decision variables $y \in \mathbb{R}^{\mathcal{N}_0}$ is the following Lemma 11, which shows that the quantities $\rho_{i_1 \dots i_M}(S \cap S_m)$ can be written *recursively*. Similarly to the proof of Theorem 1, the proof of the following Lemma 11 is based on the graphical interpretation of $\rho_{i_1 \dots i_M}(S)$ that is developed in Appendix B.2.

LEMMA 11. *If Assumption 1 holds and the assortment $S \subseteq \mathcal{N}_0$ satisfies $S \cap S_1 \neq \emptyset$,⁸ then the following equality holds for each tuple $(i_1, \dots, i_M) \in \mathcal{L}$ and $m \in \mathcal{M}$:*

$$\rho_{i_1 \dots i_M}(S \cap S_m) = \begin{cases} r_{i_m}, & \text{if } i_m \in S, \\ \min_{i \in \mathcal{B}_m \cap S} r_i, & \text{if } i_m \notin S \text{ and } m = 1, \\ \min \left\{ \min_{i \in \mathcal{B}_m \cap S} r_i, \rho_{i_1 \dots i_M}(S \cap S_{m-1}) \right\}, & \text{if } i_m \notin S \text{ and } m \geq 2. \end{cases} \quad (5)$$

Equipped with the observation that the quantities $\rho_{i_1 \dots i_M}(S \cap S_m)$ can be written recursively, we now introduce a directed acyclic graph that will serve as a low-dimensional network representation

⁷ A linear optimization problem over the feasible set $\mathcal{Y}(S)$ bears resemblance to the RANKAGGREGATIONLP optimization problem that is studied by Jagabathula and Rusmevichientong (2019). In contrast, our set $\mathcal{Y}(S)$ requires an aggregation of the decision variables $\lambda_{i_1 \dots i_M}$ in such a way that preserves the quantity $\rho_{i_1 \dots i_M}(S)$, where we recall from §4.4 that $\rho_{i_1 \dots i_M}(S)$ can be interpreted as the worst-case revenue that is earned by the firm when assortment S is offered to a customer whose preferences are represented by a ranking that corresponds to the tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$. Because of this difference, the reformulations developed by Jagabathula and Rusmevichientong (2019, §4 and §5) are not readily applicable to the set $\mathcal{Y}(S)$.

⁸ The condition that $S \cap S_1 \neq \emptyset$ in Lemma 11 is satisfied, for example, when $S, S_1 \in \mathcal{S}$.

of the decision variables of the linear optimization problem (2). Indeed, we define the vertices and directed edges of the graph as follows:

$$\begin{aligned} \mathfrak{V} &\triangleq \{(m, i, \kappa) : m \in \mathcal{M}, i \in S_m, \kappa \in \{r_j : j \in S_m\}\}; \\ \mathfrak{E} &\triangleq \{((m, i, \kappa), (m', i', \kappa')) \in \mathfrak{V} \times \mathfrak{V} : m' = m + 1, i' \in \{i\} \cup \mathcal{B}_{m+1}, \kappa' \in \{\kappa\} \cup \{r_j : j \in \mathcal{B}_{m+1}\}\}. \end{aligned}$$

We observe that each of the directed edges in the graph $\mathfrak{G} \equiv (\mathfrak{V}, \mathfrak{E})$ is of the form $((m, i, \kappa), (m + 1, i', \kappa'))$, which readily implies that the directed graph $\mathfrak{G} \equiv (\mathfrak{V}, \mathfrak{E})$ is acyclic. Moreover, the number of vertices $|\mathfrak{V}|$ and the number of directed edges $|\mathfrak{E}|$ clearly scale as a polynomial with respect to both the number of products n as well as the number of past assortments M . Using this directed acyclic graph, we now state our compact reformulation of the linear optimization problem (2).

PROPOSITION 3. *Let Assumption 1 hold. Then for each $S \subseteq \mathcal{N}_0$ that satisfies $S \cap S_1 \neq \emptyset$, $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ is equal to the optimal objective value of the following linear optimization problem:*

$$\begin{aligned} &\underset{f, g, \epsilon}{\text{minimize}} && \sum_{i, \kappa : (M, i, \kappa) \in \mathfrak{V}} \kappa g_{M, i, \kappa} \\ &\text{subject to} && \sum_{\kappa : (m, i, \kappa) \in \mathfrak{V}} g_{m, i, \kappa} - \epsilon_{m, i} = v_{m, i} && \forall m \in \mathcal{M}, i \in S_m \\ &&& \sum_{i', \kappa' : ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E}} f_{m, i, \kappa, i', \kappa'} = g_{m, i, \kappa} && \forall (m, i, \kappa) \in \mathfrak{V} : m \in \{1, \dots, M-1\} \\ &&& \sum_{i', \kappa' : ((m-1, i', \kappa'), (m, i, \kappa)) \in \mathfrak{E}} f_{m-1, i', \kappa', i, \kappa} = g_{m, i, \kappa} && \forall (m, i, \kappa) \in \mathfrak{V} : m \in \{2, \dots, M\} \\ &&& g_{m, i, \kappa} = 0 && \forall (m, i, \kappa) \in \mathfrak{V} : i \in S, i \in \mathcal{B}_m, \kappa \neq r_i \\ &&& g_{m, i, \kappa} = 0 && \forall (m, i, \kappa) \in \mathfrak{V} : i \in S, i \notin \mathcal{B}_m, \kappa \in \{r_j : j \in \mathcal{B}_m\} \\ &&& g_{m, i, \kappa} = 0 && \forall (m, i, \kappa) \in \mathfrak{V} : \kappa \in \{r_j : j \in \mathcal{B}_m \setminus S\} \\ &&& \sum_{i, \kappa : (M, i, \kappa) \in \mathfrak{V}} g_{M, i, \kappa} = 1 \\ &&& \|\epsilon\| \leq \eta \\ &&& f_{m, i, \kappa, i', \kappa'} \geq 0 && \forall ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E} \\ &&& g_{m, i, \kappa} \geq 0 && \forall (m, i, \kappa) \in \mathfrak{V}. \end{aligned} \tag{6}$$

Let us make several remarks about the linear optimization problem (6). First, we remark that the decision variables f, g can be interpreted as a network flow through the directed acyclic graph $\mathfrak{G} \equiv (\mathfrak{V}, \mathfrak{E})$. Indeed, the decision variable $g_{m, i, \kappa}$ represents the amount of flow that passes through each vertex $(m, i, \kappa) \in \mathfrak{V}$, the decision variable $f_{m, i, \kappa, i', \kappa'}$ represents the amount of flow through the directed edge from vertex $(m, i, \kappa) \in \mathfrak{V}$ to vertex $(m+1, i', \kappa') \in \mathfrak{V}$, and the constraints of the form $\sum_{i', \kappa' : ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E}} f_{m, i, \kappa, i', \kappa'} = g_{m, i, \kappa}$ and $\sum_{i', \kappa' : ((m-1, i', \kappa'), (m, i, \kappa)) \in \mathfrak{E}} f_{m-1, i', \kappa', i, \kappa} = g_{m, i, \kappa}$ ensure that flow is conserved through each vertex $(m, i, \kappa) \in \mathfrak{V}$ with $m \in \{2, \dots, M-1\}$. Second, we show in the proof of Proposition 3 that the decision variables in the linear optimization problem (6) can be interpreted as an aggregation of the decision variables in the linear optimization problem (2). In particular, we show in Appendices I.4 and I.5 that each decision variable $\lambda_{i_1 \dots i_M}$ can be interpreted as flow through the path $(1, i_1, \rho_{i_1 \dots i_M}(S \cap S_1)), \dots, (M, i_M, \rho_{i_1 \dots i_M}(S \cap S_M))$ in the directed acyclic graph $\mathfrak{G} \equiv (\mathfrak{V}, \mathfrak{E})$. Hence, any optimal solution for the linear optimization problem (6) can be transformed back into the decision variables of the linear optimization problem (2) by applying the flow decomposition theorem (Ahuja et al. 1988, Theorem 3.5).

In summary, we have shown in Proposition 3 that the worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ in the case of nested past assortments can be evaluated by solving a linear optimization problem

with size that is polynomial in the number of products n and the number of past assortments M . The key step behind developing the compact reformulation is exploiting the graphical interpretation of the quantities $\rho_{i_1 \dots i_M}(S)$ that is developed in Appendix B.2 to design a low-dimensional network representation of the decision variables in the linear optimization problem (2). By applying strong duality to the linear optimization problem from Proposition 3 and introducing binary decision variables to represent the assortment S , we readily obtain a mixed-integer optimization reformulation of the robust optimization problem (RO):

THEOREM 5. *Let Assumption 1 hold and $S_1, \dots, S_M \in \mathcal{S}$. Then the robust optimization problem (RO) can be reformulated as a mixed-integer linear optimization problem with $\mathcal{O}(n)$ binary decision variables, $\mathcal{O}(\text{poly}(n, M))$ continuous decision variables, and $\mathcal{O}(\text{poly}(n, M))$ constraints.*

The description of the mixed-integer optimization reformulation from Theorem 5 is found at the end of Appendix I.6.

REMARK 4. The reformulation in Appendix I.6 is presented for cases in which the firm does not have any business constraints on the assortments $S \in \mathcal{S}$. In many applications of assortment optimization, firms have additional business constraints (e.g., due to limited shelf space or bundling requirements) that limit the space of possible new assortments; see Mišić (2016, p.113). Because our algorithm from Theorem 5 consists of solving a mixed-integer optimization problem with a binary decision variable $x_i \in \{0, 1\}$ for each product $i \in \mathcal{N}_0$ that satisfies $x_i = 1$ if and only if product i is included in the assortment S , such business constraints can be easily integrated directly into our mixed-integer optimization problem. In particular, as long as the number of such business constraints is not exponential in n or M , the size of our mixed-integer optimization reformulation of (RO) will remain polynomial in the number of products and the number of past assortments.

6. Answer to Question 4

In the previous sections, we showed that the robust optimization problem (RO) can find assortments that outperform the past assortments (see §3), and we developed algorithms for solving the robust optimization problem (RO) that are theoretically and practically efficient (see §5). Therefore, when confronted with an assortment planning problem, we recommend that firms begin by solving the robust optimization problem (RO). If the optimal objective value of the robust optimization problem (RO) is strictly greater than the expected revenue of the firm's best past assortment, then the optimal assortment for (RO) should be offered by the firm to its customers, as such an assortment can be trusted to improve the firm's expected revenue in a way that is not exclusive to just one of the many ranking-based choice models that are consistent with the firm's historical sales data.

In this section, we provide guidance to firms for what assortment to offer to customers in cases in which the optimal objective value of the robust optimization problem (RO) is *not* strictly greater than the expected revenue of the firm's best past assortment. We begin in §6.1 by recommending that the firm first solve an optimistic version of the robust optimization problem (RO) to obtain a data-driven upper bound on the potential increase in expected revenue that a firm can hope to gain by performing experimentation. If the gap between the optimal objective value of the optimistic optimization problem and the expected revenue of the firm's best past assortment is small, then we can conclude that the potential increase in expected revenue that a firm can hope to gain by exploring with new assortments is small, thereby guiding the firm to avoid performing experimentation. If the gap is large, then we conclude that the firm can potentially benefit from experimentation. For this case, we show in §6.2 that the robust and optimistic optimization problems can be combined into a Pareto optimization problem for finding new assortments for experimentation that have potential to increase the firm's best-case expected revenue while maintaining bounds on the worst-case decrease in expected revenue.

6.1. Upper Bound on the Value of Experimentation

If the optimal objective value of the robust optimization problem (RO) is not strictly greater than the expected revenue of the firm’s best past assortment, then there are several options available to the firm. One option is for the firm to continue offering their best past assortment, i.e., the past assortment with the highest expected revenue. This option can be interpreted as ‘maintaining the status-quo’. A second option is for the firm to conduct experimentation by offering new assortments to their customers, with the goal of further constraining the set of ranking-based choice models that are consistent with the firm’s historical sales data. However, experimentation can be costly or otherwise undesirable to firms as discussed in §1. As such, a natural question faced by firms is whether the potential benefits of conducting experimentation will outweigh the costs.

In this subsection, we show that our algorithms from §5 for solving the robust optimization problem (RO) can be repurposed to quantify the potential benefits of experimentation. Specifically, we observe that an upper bound on the potential increase in expected revenue that the firm can hope to gain by performing experimentation is given by the optimal objective value of an ‘optimistic’ optimization problem of the form

$$\max_{S \in \mathcal{S}} \max_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S). \quad (\text{OO})$$

The above optimization problem can be viewed as an ‘best-case’ alternative to the robust optimization problem (RO), in the sense that (OO) will yield the assortment that has the maximum increase in the firm’s predicted expected revenue on a ranking-based choice model that is consistent with the historical sales data generated by the firm’s past assortments. The optimization problem (OO) is attractive because it delivers an upper bound on the increase in expected revenue that a firm can hope to gain by performing experimentation. If the gap between the optimal objective values of (OO) and (RO) is small, then the firm can conclude that the potential increase in expected revenue that a firm can hope to gain by experimentation with new assortments is small, thereby guiding the firm to avoid performing experimentation.

The main technical contribution of the present subsection is showing that our polynomial-time algorithms and mixed-integer optimization reformulation from §5 that were designed for the robust optimization problem (RO) can be extended to the optimistic optimization problem (OO). These technical contributions are summarized as follows.

THEOREM 6. (OO) can be solved in weakly $\mathcal{O}(\text{poly}(n))$ time for every fixed M .

THEOREM 7. If Assumption 1 holds and $S_1, \dots, S_M \in \mathcal{S}$, then (OO) can be reformulated as a mixed-integer linear optimization problem with $\mathcal{O}(n)$ binary decision variables, $\mathcal{O}(\text{poly}(n, M))$ continuous decision variables, and $\mathcal{O}(\text{poly}(n, M))$ constraints.

Theorems 6 and 7, which are analogues of Theorems 4 and 5 from §5, establish that the optimistic optimization problem (OO) enjoys similar tractability as the robust optimization problem (RO). The proofs of Theorems 6 and 7, which are found in Appendix J, follow using similar arguments as those from Theorems 4 and 5. In Appendix A.4, we show using synthetic data with nested past assortments that the optimistic optimization problem (OO) can yield practical upper bounds on the potential value of experimentation that can help a firm determine that the benefits of conducting experimentation will not outweigh the costs.

6.2. Low-Risk Approaches to Experimentation

Suppose that the gap between the upper bound given by (OO) and the expected revenue of the firm’s best past assortment is deemed by the firm to be large enough to warrant experimentation.

The firm, in this case, must decide which new assortments it will offer to its customers. Here, we take the perspective of a firm that wishes to perform experimentation in such a way that mitigates the worst-case short-run declines in expected revenue that come from experimenting with new assortments. That is, we assume that the firm is interested in experimenting with new assortments $S \in \mathcal{S}$ for offering to their customers that can increase the firm’s expected revenue while ensuring the worst-case expected revenue of the new assortment, $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$, remains close to the expected revenue of the firm’s best past assortment, $\max_{m \in \mathcal{M}} r^\top v_m$. We offer two approaches below for conducting low-risk experimentation that utilize our algorithms for solving (RO) and (OO).

6.2.1. Exact Algorithm for Pareto Optimization. Our first proposed approach to finding new assortments for experimentation is to solve the following ‘Pareto’ optimization problem:

$$\begin{aligned} & \underset{S \in \mathcal{S}}{\text{maximize}} && \max_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) \\ & \text{subject to} && \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) \geq \theta. \end{aligned} \tag{PO}$$

In contrast to the robust optimization problem (RO) and the optimistic optimization problem (OO), the Pareto optimization problem (PO) aims to find an assortment that maximizes the best-case expected revenue while imposing a lower bound on the worst-case expected revenue. By varying the lower bound θ , the above optimization problem thus leads to a *Pareto frontier* of new assortments that a firm can offer to its customers that offer varying degrees of possible improvement to the firm’s expected revenue while enjoying guarantees on the worst-case decrease in expected revenue that the firm might incur by experimenting with the new assortment. Such a Pareto frontier of new assortments provided by (PO) can be attractive to risk-averse firms in situations in which there do not exist any assortments that satisfy (IMP), as there may exist optimal assortments for the optimization problem (PO) that are different from the firm’s past assortments when choosing the parameter θ to be slightly less than the expected revenue of the firm’s best past assortment.

In the case of nested past assortments, we can use the same reformulation techniques from Theorems 5 and 7 to reformulate (PO) as a compact mixed-integer optimization problem:

THEOREM 8. *Let Assumption 1 hold and $S_1, \dots, S_M \in \mathcal{S}$. Then the optimization problem (PO) can be reformulated as a mixed-integer linear optimization problem with $\mathcal{O}(n)$ binary decision variables, $\mathcal{O}(\text{poly}(n, M))$ continuous decision variables, and $\mathcal{O}(\text{poly}(n, M))$ constraints.*

The description of the mixed-integer optimization reformulation from Theorem 8 is found at the end of Appendix J.3. In contrast to the polynomial-time algorithms that were developed for (RO) and (OO), we do not have an algorithm for solving (PO) that runs in polynomial-time for any fixed number of past assortments (see §6.2.2 for discussion). In Appendix A.5, we show using synthetic data with various collections of nested past assortments that (PO) can find new assortments that offer best-case increases in expected revenue of 20%-60% while enjoying worst-case decreases in expected revenues as low as 1%-10%.

6.2.2. Heuristic for Pareto Optimization. It is straightforward to reformulate (PO) as a compact mixed-integer optimization problem in the case of nested past assortments (see Theorem 8) by combining the compact mixed-integer optimization reformulations for (RO) and (OO) (see Theorems 5 and 7). The same reasoning, however, does not extend to solving (PO) in polynomial-time for any fixed number of past assortments. The issue stems from the fact that our polynomial-time algorithms for solving (RO) and (OO) (see Theorems 4 and 6) rely on performing an exhaustive search over polynomial-size collections of assortments that are guaranteed to contain optimal assortments for (RO) and (OO). In contrast to (RO) and (OO), we do not have a characterization

of the structural of optimal assortments for (PO), and thus cannot perform a polynomial-time exhaustive search that is guaranteed to deliver an optimal assortment for (PO).

To get around this, we propose a simple heuristic for finding new assortments for risk-averse experimentation. The heuristic runs in polynomial-time for any fixed number of past assortments and consists of three steps. In the first step, we construct the collection of assortments $\widehat{\mathcal{S}}$ from §4 that is guaranteed to contain an optimal solution to (RO). In the second step, we compute the worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ and the best-case expected revenue $\max_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ for each assortment $S \in \widehat{\mathcal{S}}$. In the third step, we compute the following *approximate Pareto frontier*:

$$\left\{ S \in \widehat{\mathcal{S}} : \left\{ S' \in \widehat{\mathcal{S}} : \begin{array}{l} \left[\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S') \geq \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) \text{ and } \max_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S') > \max_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) \right] \\ \text{or} \\ \left[\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S') > \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) \text{ and } \max_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S') \geq \max_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) \right] \end{array} \right\} = \emptyset \right\}. \quad (7)$$

The collection of assortments from line (7) can be interpreted as an approximation of the Pareto frontier that would be obtained by solving (PO) for all $\theta \geq 0$. In Appendix A.6, we show using real data from a conjoint dataset (Toubia et al. 2003) that the approximate Pareto frontier from line (7) can contain new assortments that offer increases in expected revenue with respect to the true ranking-based choice model, while simultaneously enjoying guarantees that implementing the new assortments will not lead to meaningful worst-case decreases in the firm’s expected revenue.

7. Conclusion and Future Research

In this work, we investigated a popular class of high-dimensional discrete choice models, known as ranking-based choice models, in the context of assortment planning problems. Motivated by the fact that many ranking-based choice models can be consistent with a firm’s historical sales data, we considered a class of robust optimization problems in which the goal is find an assortment that has high expected revenue across all of the ranking-based choice models that are consistent with the firm’s historical sales data. By developing the first structural results and practical algorithms for solving these robust optimization problems, our paper showed in a variety of applications that these robust optimization problems can be solved in reasonable computation times and offer value to firms in comparison to estimate-then-optimize in the context of data-driven assortment optimization with ranking-based choice models.

We believe this work opens up a number of promising directions for future research. First, our work showed for the first time that the question of whether it can be possible to identify assortments that satisfy condition (IMP) can be answered for one popular class of high-dimensional discrete choice models. Yet there are many other high-dimensional discrete choice models beyond the ranking-based choice model for which this identification question can be asked, such as models for capturing irrational customer choice (Berbeglia 2018, Chen and Mišić 2022, Jena et al. 2022). Second, our work showed that polynomial-time algorithms can be developed for solving the robust optimization problem (RO) for ranking-based choice models. However, it may be possible in certain settings that existing algorithms for estimating high-dimensional discrete choice models such as expectation-maximization (Talluri and Van Ryzin 2004, van Ryzin and Vulcano 2017, Şimşek and Topaloglu 2018), in combination with algorithms for finding assortments that maximize the predicted expected revenue, can lead to assortments with provable performance guarantees with respect to the identification question. Establishing such guarantees would provide new assurances to firms for using estimate-then-optimize in high-stakes assortment planning problems. Finally, we

believe that the present and related work (*e.g.*, Kallus and Zhou (2018), Sturt (2023)) provide a starting point for using robust optimization to develop efficient algorithms that are valuable for operations management problems in which good *average* performance is paramount. In particular, future work may extend the algorithms developed in the present paper to numerous other revenue management problems, ranging from multi-product pricing to dynamic assortment planning.

References

- Ahuja RK, Magnanti TL, Orlin JB (1988) Network flows .
- Aouad A, Farias V, Levi R (2021) Assortment optimization under consider-then-choose choice models. *Management Science* 67(6):3368–3386.
- Aouad A, Farias V, Levi R, Segev D (2018) The approximability of assortment optimization under ranking preferences. *Operations Research* 66(6):1661–1669.
- Ariely D (2010) Column: Why businesses don’t experiment. *Harvard Business Review* 88(4), ISSN 0017-8012.
- Belloni A, Freund R, Selove M, Simester D (2008) Optimizing product line designs: Efficient methods and comparisons. *Management Science* 54(9):1544–1552.
- Berbeglia G (2018) The generalized stochastic preference choice model. *arXiv preprint arXiv:1803.04244* .
- 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, Mišić VV (2017) Robust product line design. *Operations Research* 65(1):19–37.
- Bertsimas D, Mišić VV (2019) Exact first-choice product line optimization. *Operations Research* 67(3):651–670.
- Block H, Marschak J (1959) Random orderings and stochastic theories of response. Cowles Foundation Discussion Papers 66, Cowles Foundation for Research in Economics, Yale University, URL <https://EconPapers.repec.org/RePEc:cwl:cwldpp:66>.
- Chen YC, Mišić VV (2022) Decision forest: A nonparametric approach to modeling irrational choice. *Management Science* 68(10):7090–7111.
- Davis JM, Topaloglu H, Williamson DP (2015) Assortment optimization over time. *Operations Research Letters* 43(6):608–611.
- Désir A, Goyal V, Jagabathula S, Segev D (2021) Mallows-smoothed distribution over rankings approach for modeling choice. *Operations Research* 69(4):1206–1227.
- Désir A, Goyal V, Jiang B, Xie T, Zhang J (2023) Robust assortment optimization under the markov chain choice model. *Operations Research* .
- Farias VF, Jagabathula S, Shah D (2013) A nonparametric approach to modeling choice with limited data. *Management Science* 59(2):305–322.
- Feldman J, Paul A, Topaloglu H (2019) Assortment optimization with small consideration sets. *Operations Research* 67(5):1283–1299.
- Gallego G, Iyengar G, Phillips R, Dubey A (2004) Managing flexible products on a network. *Available at SSRN 3567371* .
- Honhon D, Jonnalagedda S, Pan XA (2012) Optimal algorithms for assortment selection under ranking-based consumer choice models. *Manufacturing & Service Operations Management* 14(2):279–289.
- Ignat A (2022) Saving the high street: Using experimentation in-store to improve customer experience. *Conversion* URL <https://conversion.com/blog/saving-the-high-street-using-experimentation-in-store-to-improve-customer-experience/>.
- Jagabathula S (2014) Assortment optimization under general choice. *Available at SSRN 2512831* .
- Jagabathula S, Rusmevichientong P (2019) The limit of rationality in choice modeling: Formulation, computation, and implications. *Management Science* 65(5):2196–2215.

- Jena SD, Lodi A, Sole C (2022) On the estimation of discrete choice models to capture irrational customer behaviors. *INFORMS Journal on Computing* 34(3):1606–1625.
- Kallus N, Zhou A (2018) Confounding-robust policy improvement. *Advances in Neural Information Processing Systems* 31.
- Mišić VV (2016) *Data, models and decisions for large-scale stochastic optimization problems*. Ph.D. thesis, Massachusetts Institute of Technology.
- Orlin JB (1997) A polynomial time primal network simplex algorithm for minimum cost flows. *Mathematical Programming* 78(2):109–129.
- 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.
- Sher I, Fox JT, il Kim K, Bajari P (2011) Partial identification of heterogeneity in preference orderings over discrete choices. Technical report, National Bureau of Economic Research.
- Şimşek AS, Topaloglu H (2018) An expectation-maximization algorithm to estimate the parameters of the markov chain choice model. *Operations Research* 66(3):748–760.
- Sturt B (2023) A nonparametric algorithm for optimal stopping based on robust optimization. *Operations Research* 71(5):1530–1557.
- Susan F, Golrezaei N, Emamjomeh-Zadeh E, Kempe D (2022) Active learning for non-parametric choice models. *arXiv preprint arXiv:2208.03346* .
- Talluri K, Van Ryzin G (2004) Revenue management under a general discrete choice model of consumer behavior. *Management Science* 50(1):15–33.
- Tarjan RE (1997) Dynamic trees as search trees via euler tours, applied to the network simplex algorithm. *Mathematical Programming* 78(2):169–177.
- Toubia O, Simester DI, Hauser JR, Dahan E (2003) Fast polyhedral adaptive conjoint estimation. *Marketing Science* 22(3):273–303.
- van Ryzin G, Vulcano G (2015) A market discovery algorithm to estimate a general class of nonparametric choice models. *Management Science* 61(2):281–300.
- van Ryzin G, Vulcano G (2017) An expectation-maximization method to estimate a rank-based choice model of demand. *Operations Research* 65(2):396–407.
- Wang Z, Peura H, Wiesemann W (2020) Randomized assortment optimization. *Available at SSRN 3685695* .

Technical Proofs and Additional Results

Appendix A: Numerical Experiments

This appendix contains the numerical experiments for the paper.

- Appendix A.1 shows that estimate-then-optimize with ranking-based choice models can risk leading to significant declines in a firm’s expected revenue when the firm’s past assortments are revenue-ordered. The motivation for Appendix A.1 is found at the end of §4.1.
- Appendix A.2 shows the practical efficiency of the polynomial-time algorithm from §5.1 for solving the robust optimization problem (RO) in the case of two past assortments.
- Appendix A.3 shows the practical tractability of the compact mixed-integer optimization reformulation of the robust optimization problem (RO) from §5.3 for the case of nested past assortments.
- Appendix A.4 shows the managerial value of using the optimistic optimization problem (OO) from §6.1 to compute upper bounds on the maximum increase in expected revenue that a firm can hope to gain by experimenting with new assortments.
- Appendix A.5 shows the managerial value of the exact algorithm from §6.2.1 for solving the Pareto optimization problem (PO) for finding new assortments that can be used by firms for low-risk experimentation.
- Appendix A.6 uses real data to show the managerial value of the heuristic from §6.2.2 for the Pareto optimization problem (PO) for finding new assortments that can be used by firms for low-risk experimentation.

A.1. Numerical Illustration of the Risks of Estimate-Then-Optimize

In this appendix, we perform numerical experiments to assess the performance of assortments obtained using estimate-then-optimize when the historical sales data is randomly generated from revenue-ordered assortments. To motivate our experiment setup, we recall the details of estimate-then-optimize from §1. Indeed, suppose that one estimates a ranking-based choice model from the historical sales data generated by the revenue-ordered assortments and then recommends that the firm implement a new assortment which maximizes the predicted expected revenue under the estimated ranking-based choice model. For this setting, Corollary 1 from §4.1 guarantees that this estimate-then-optimize technique will never offer *fidelity* to the firm, in the sense that there will always exist a ranking-based choice model which is consistent with the historical sales data for which the expected revenue for the new assortment will be less than or equal to the expected revenue of the best past assortment. Moreover, as we will further see through the following numerical experiment, the expected revenue from the assortment recommended by estimate-then-optimize can be strictly and significantly less than the expected revenue of the best past assortment.

To perform our numerical experiment, we begin by constructing randomly-generated problem instances. In each problem instance, the revenues for the products are drawn from the distribution $r_1, \dots, r_n \sim \text{Uniform}[0, 1]$, and a base choice for the parameters λ^* of a ranking-based choice model is drawn uniformly over the $(n + 1)$ -dimensional probability simplex.⁹ Using this base choice for the parameters, we generate historical sales data of the form v_1, \dots, v_n , where each v_m is the historical sales data generated by the revenue-ordered assortment $\{0, m, m + 1, \dots, n - 1, n\}$ under the ranking-based choice model with the base parameter λ^* .¹⁰ After we compute the historical sales data, we *forget* the base parameters λ^* of the ranking-based choice model and apply the estimate-then-optimize technique to obtain a new assortment. Specifically,

⁹ If the rankings in Σ are indexed by $\{\sigma_1, \dots, \sigma_{(n+1)!}\}$, then uniform sampling over the probability simplex is obtained by drawing $u_1, \dots, u_{(n+1)!} \sim \text{Uniform}[0, 1]$ and setting $\lambda_{\sigma_k} \leftarrow \log(u_k) / \sum_{k'=1}^{(n+1)!} \log(u_{k'})$ for each $k = 1, \dots, (n + 1)!$.

¹⁰ We sort the products in ascending order by revenue before performing our analysis, which ensures that $\{0, m, m + 1, \dots, n\}$ for each $m \in \{1, \dots, n\}$ is a revenue-ordered assortment.

we first estimate the parameters $\hat{\lambda}$ of a ranking-based choice model using the historical sales data; since many selections of the parameters may be consistent with the historical sales data, we choose our estimate $\hat{\lambda}$ as the optimal solution to the linear optimization problem $\min_{\lambda \in \mathcal{U}} c^\top \lambda$, where the cost vector c is drawn uniformly over $[0, 1]^{(n+1)!}$.^{11,12} We then obtain a new assortment S' as any optimal solution to the combinatorial optimization problem $\max_{S \in \mathcal{S}} \mathcal{R}^{\hat{\lambda}}(S)$ which maximizes the predicted expected revenue under the estimated ranking-based choice model.¹³ Finally, we evaluate the new assortment obtained using estimate-then-optimize by computing the worst-case expected revenue of the new assortment under all ranking-based choice models that are consistent with the historical sales data, $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S')$, the best-case expected revenue of the new assortment under all ranking-based choice models that are consistent with the historical sales data, $\max_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S')$, and the expected revenue of the best past assortment, $\max_{m \in \mathcal{M}} r^\top v_m$.

In Figures EC.1 and EC.2, we present the results of these numerical experiments for the case of $n = 4$ products over 1000 randomly-generated problem instances. In Figure EC.1, we compare the expected revenue under the best past assortment, $\max_{m \in \mathcal{M}} r^\top v_m$, to the expected revenue of the new assortments obtained using estimate-then-optimize under the worst-case ranking-based choice model that is consistent with the historical sales data, $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S')$. We observe that the results in Figure EC.1 are consistent with the impossibility result from Corollary 1; indeed, the worst-case expected revenues of the new assortments obtained using estimate-then-optimize never exceed the expected revenues of the best past assortments. Furthermore, we observe for many of the problem instances that there are ranking-based choice models $\lambda \in \mathcal{U}$ that are consistent with the historical sales data for which the resulting expected revenue of the new assortment $\mathcal{R}^\lambda(S')$ is strictly less than the expected revenue under the best past assortment, $\max_{m \in \mathcal{M}} r^\top v_m$.

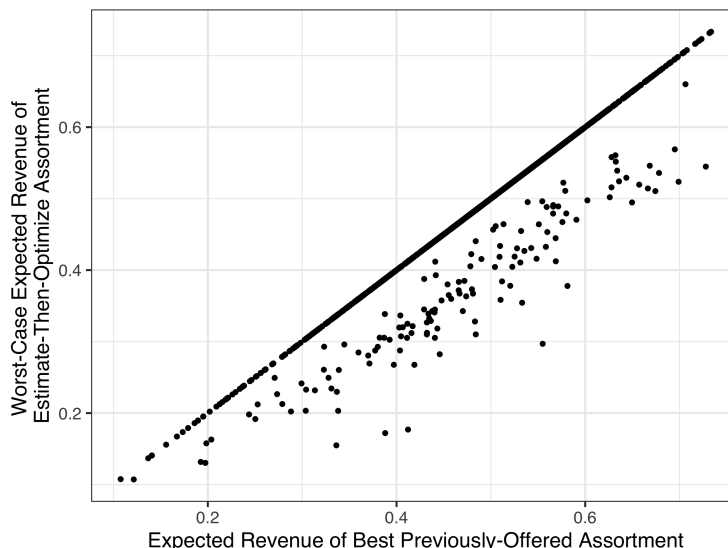
To assess whether the above findings are overly conservative from a practical standpoint, we turn to a detailed analysis of the 137 problem instances for which the expected revenue of the best past assortment, $\max_{m \in \mathcal{M}} r^\top v_m$, is strictly greater than the worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S')$ for the new assortment S' obtained using estimate-then-optimize. In Figure EC.2 we present, for each of these 137 problem instances, a visualization of the range of relative percentage improvements of the predicted expected revenue of the new assortment, $\mathcal{R}^\lambda(S')$, over the expected revenue of the firm's best past assortment, $\max_{m \in \mathcal{M}} r^\top v_m$, which are possible to obtain under a ranking-based choice model which is consistent with the historical sales data, $\lambda \in \mathcal{U}$. Stated more precisely, each of the x values in Figure EC.2 corresponds to one of these 137 problem instances, and the corresponding interval of y -values formed by the red and blue bars is the interval $[\Pi_\lambda : \lambda \in \mathcal{U}]$, where $\Pi_\lambda \triangleq 100\% \times (\mathcal{R}^\lambda(S') - \max_{m \in \mathcal{M}} r^\top v_m) / (\max_{m \in \mathcal{M}} r^\top v_m)$ is the relative percentage improvement of the predicted expected revenue of the new assortment S' over the expected revenue of the firm's best past assortment for a given ranking-based choice model $\lambda \in \mathcal{U}$. Hence, the red bars are the negative values of Π_λ that can be attained under the ranking-based choice models $\lambda \in \mathcal{U}$, and the dotted line shows the reflection of the endpoints of red bars over the horizontal line at zero. We do not assign any likelihood to the values in each interval $[\Pi_\lambda : \lambda \in \mathcal{U}]$, as there is no information available for inferring which of the ranking-based choice models are more likely to be the 'truth' among the ranking-based choice models $\lambda \in \mathcal{U}$ that are consistent with the historical sales data.

¹¹ Alternative approaches for estimating the parameters of a ranking-based choice model from historical sales data are provided in Mišić (2016), van Ryzin and Vulcano (2015, 2017), Désir et al. (2021). Our approach of estimating the parameters as $\hat{\lambda} \in \arg \min_{\lambda \in \mathcal{U}} c^\top \lambda$ for a randomly-chosen cost vector c is motivated by our desire to decouple any potential biases associated with any particular estimation procedure from an empirical assessment of the estimate-then-optimize technique. In particular, our approach is viewed as a simple way of randomly sampling the parameters from the set of all parameters of ranking-based choice models which are consistent with the historical sales data.

¹² An obvious downside of our simple estimation procedure for the parameters of the ranking-based choice model is that it requires solving a linear optimization problem with $\mathcal{O}(n!)$ decision variables, and, thus, our simple estimation procedure does not scale efficiently to problem instances with many products. Nonetheless, this estimation procedure is sufficiently fast for the purposes of this numerical study, where the aim is simply to assess the performance of assortments obtained by the estimate-then-optimize technique over revenue-ordered assortments. In particular, we believe it is a reasonable assumption that similar findings from our numerical experiment with $n = 4$ (see Figures EC.1 and EC.2) would be found in experiments with larger values of n .

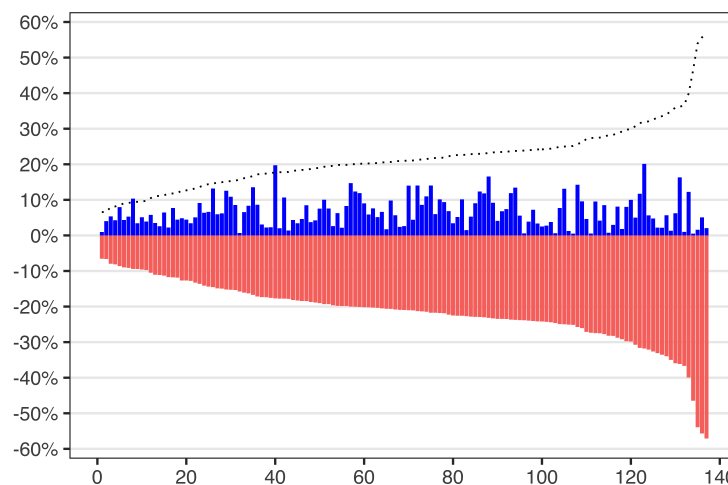
¹³ We solve this optimization problem using the mixed-integer optimization reformulation given by Bertsimas and Mišić (2019, §3.2), which is implemented using the Julia programming language with JuMP and solved using Gurobi.

Figure EC.1 Performance of new assortments obtained using estimate-then-optimize when $\mathcal{M} = \bar{\mathcal{S}}$.



Note. Each point corresponds to a randomly-generated problem instance. The x -axis shows the expected revenue of the best past assortment, $\max_{m \in \mathcal{M}} r^T v_m$. The y -axis shows the worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S')$ for the new assortment S' obtained using the estimate-then-optimize technique.

Figure EC.2 Relative improvement in expected revenue of new assortments obtained using estimate-then-optimize when $\mathcal{M} = \bar{\mathcal{S}}$.



Note. Each of the x -values corresponds to one of the 137 randomly-generated problem instances from Figure EC.1 for which the expected revenue of the best past assortment, $\max_{m \in \mathcal{M}} r^T v_m$, was strictly greater than the worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S')$ for the new assortment S' obtained using the estimate-then-optimize technique. The corresponding interval of y -values formed by the red and blue bars is the interval $[\Pi_\lambda : \lambda \in \mathcal{U}]$, where $\Pi_\lambda \triangleq 100\% \times (\mathcal{R}^\lambda(S') - \max_{m \in \mathcal{M}} r^T v_m) / (\max_{m \in \mathcal{M}} r^T v_m)$ is the relative percentage improvement of the expected revenue of the new assortment obtained using estimate-then-optimize over the expected revenue of the firm's best past assortment for a given ranking-based choice model $\lambda \in \mathcal{U}$. For clarity, the problem instances are sorted along the x -axis by the endpoints of the red bars, and the dotted line shows the reflection of the endpoints of red bars over the horizontal line at zero.

The results in Figure EC.2 reveal a striking asymmetry between the downside and upside of implementing a new assortment found by estimate-then-optimize. In all but two of the 137 instances, the worst-case decline in expected revenue from implementing the new assortment exceeded in magnitude the best-case increase in expected revenue. The difference in magnitude between the downside and upside is also found to be significant: the average best-case improvement of the new assortment over the best past assortment (i.e., the average of the blue endpoints) is 6.56%, while the average worst-case improvement of the new assortment over the best past assortment (i.e., the average of the red endpoints) is -21.71%. These numerical findings demonstrate that the downside risks to a firm from implementing a new assortment found by estimate-then-optimize can significantly exceed the potential upsides.

In conclusion, our theoretical and numerical analysis in this subsection lead to three main takeaways. The first takeaway is that there exist collections of past assortments in which it can be impossible to identify assortments that satisfy (IMP). The second takeaway is that estimate-then-optimize can lead to a strictly worse expected revenue than those of the past assortments. In fact, the numerical results in Figures EC.1 and EC.2 show that this decline in expected revenue can be significant and outweigh the potential upside for implementing the new assortment. The third takeaway is that all of the aforementioned issues arise when the past assortments are comprised of one of the most celebrated and widely-used classes of assortments from the literature, namely, the revenue-ordered assortments. All together, these takeaways raise concerns about whether the estimate-then-optimize technique with ranking-based choice models should be trusted to “first, do no harm” in assortment planning problems.

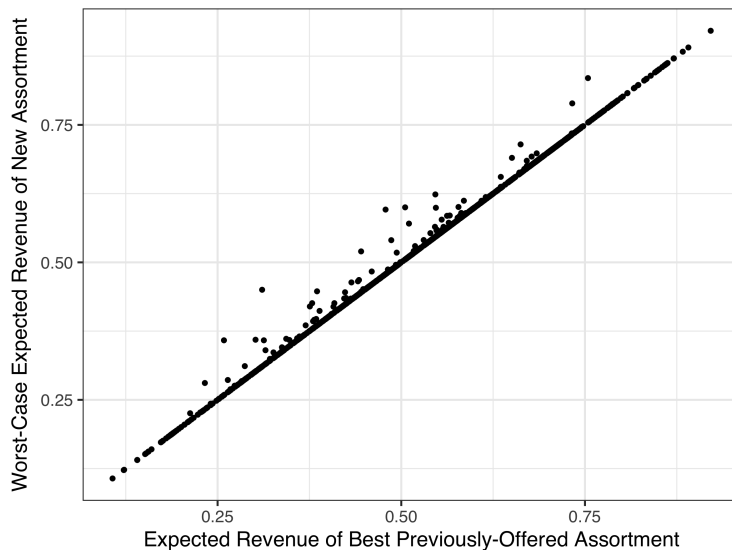
A.2. Numerical Experiments for Two Past Assortments

In this appendix, we perform numerical experiments to show the practical efficiency of our polynomial-time algorithm from §5.1 for solving the robust optimization problem (RO) in the case of $M = 2$ past assortments.

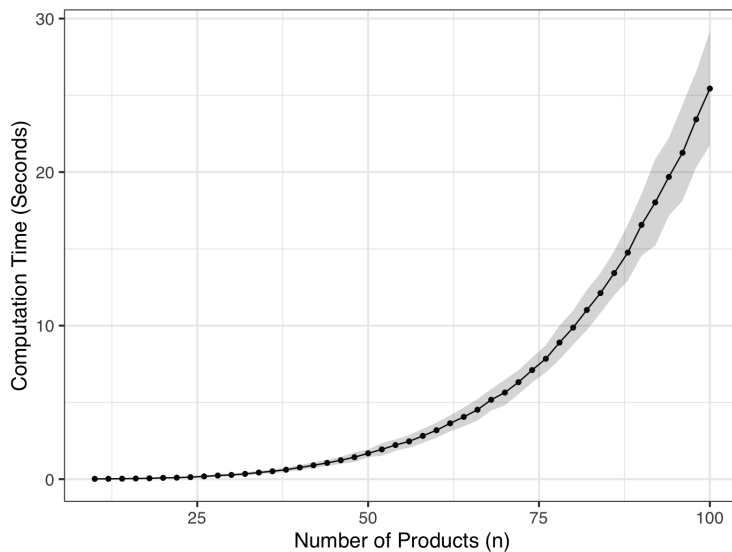
To perform our numerical experiments, we begin by constructing randomly-generated problem instances in a manner that is similar to that taken in Appendix A.1. In each randomly-generated problem instance, the revenues are drawn from the distribution $r_1, \dots, r_n \sim \text{Uniform}[0, 1]$, and we sort the products such that $r_1 < \dots < r_n$. The two past assortments $S_1, S_2 \in \mathcal{S}$ are also constructed randomly, whereby the two assortments satisfy $\{0, n\} \subseteq S_1 \cap S_2$ and, for each of the remaining products $j \in \{1, \dots, n-1\}$, we randomly assign the product to the assortments with distribution given by $P(j \in S_1 \cap S_2) = \frac{1}{3}$, $P(j \in S_1 \setminus S_2) = \frac{1}{3}$, and $P(j \in S_2 \setminus S_1) = \frac{1}{3}$. In order to generate historical sales data for these two assortments, we generate a base choice for the parameters λ^* of the ranking-based choice model. Because we will be performing numerical experiments on problem instances with larger values of n than were considered in Appendix A.1, it will not be viable from a computational tractability standpoint to generate base parameters λ^* that have nonzero values for each of the $(n+1)!$ parameters of a ranking-based choice model. To get around this, we restrict the numerical experiments to generating base parameters λ^* which are *sparse*. Specifically, we generate the base parameters in each problem instance by first randomly selecting a subset of rankings $\Sigma' \subseteq \Sigma$ of length $|\Sigma'| = K$ ¹⁴; we then assign $\lambda_\sigma^* \leftarrow 0$ for each ranking $\sigma \notin \Sigma'$, and we choose the remaining parameters $\{\lambda_\sigma : \sigma \in \Sigma'\}$ by drawing uniformly over the K -dimensional probability simplex. Using this base choice for the parameters, we generate historical sales data of the form v_1 and v_2 corresponding to the two assortments S_1 and S_2 under the ranking-based choice model with the base parameters λ^* . After we compute the historical sales data, we *forget* the base parameters λ^* of the ranking-based choice model as well as the choice of K , and we apply the algorithm from §5.1 to obtain a new assortment, denoted by S' .

In Figures EC.3 and EC.4, we present the results of the numerical experiments conducted as described above. In Figure EC.3, we present the results of these numerical experiments in the case of $n = 10$ products over 1000 randomly-generated problem instances and for $K = 10$. This figure compares the expected revenue under the best past assortment, $\max_{m \in \mathcal{M}} r^\top v_m$, to the predicted expected revenue of the new assortment under the worst-case ranking-based choice model that is consistent with the historical sales data, $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S')$. These results, in addition to the example from §3, establish that there can exist assortments that satisfy line (IMP) in the case of $M = 2$; indeed, we observe from Figure EC.3 that there are problem instances for which the predicted expected revenue of the new assortment obtained by our algorithm from

¹⁴ We use rejection sampling to ensure that each of the $\binom{(n+1)!}{K}$ subsets of rankings is selected with equal probability.

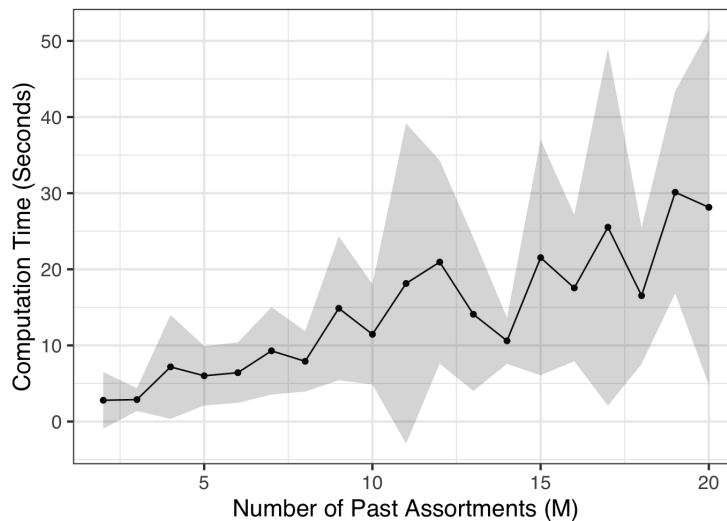
Figure EC.3 Performance of assortments obtained using algorithm from §5.1 when $M = 2$, $K = 10$, and $n = 10$.

Note. Each point corresponds to a randomly-generated problem instance. The x -axis shows the expected revenue of the best past assortment, $\max_{m \in \mathcal{M}} r^\top v_m$. The y -axis shows the worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S')$ for the new assortment S' obtained by using the algorithm from §5.1.

Figure EC.4 Computation time for the algorithm from §5.1 when $M = 2$ and $K = 1000$.

Note. Results are averaged over 100 replications for each $n \in \{10, 12, 14, \dots, 100\}$.

Figure EC.5 Computation time for solving mixed-integer optimization problem from Theorem 5 with $n = 20$ products and $M \in \{2, \dots, 20\}$ nested past assortments.



Note. Results are averaged over 10 replications for each $M \in \{2, \dots, 20\}$.

§5.1 is strictly greater than the expected revenue of the best past assortment under all ranking-based choice model that are consistent with the historical sales data. In Figure EC.4, we show the average computation times for our algorithm from §5.1 on problem instances in which the number of products is varied across $n \in \{10, 12, 14, \dots, 100\}$ and $K = 1000$. Here, we use the larger $K = 1000$ to ensure that the sparsity of the randomly-generated base parameters λ^* does not introduce any biases on the resulting computation times of our algorithm. The results in Figure EC.4 show that the computation time of the algorithm remains under 30 seconds even when there are one hundred products. This is viewed as promising from a practical perspective, as it shows that a general algorithm for solving the robust optimization problem (RO) in the case of $M = 2$ can scale to problem instances with numbers of products that are realistic in applications such as brick-and-mortar retail.

A.3. Numerical Experiments for Nested Past Assortments

In this appendix, we show that the compact mixed-integer optimization problem from Theorem 5 can be solved in reasonable computation times for robust optimization problems (RO) with up to $M = 20$ nested past assortments.

To show this, we conduct numerical experiments on randomly-generated problem instances with nested past assortments. Specifically, in each problem instance, we generate a sparse base parameter vector λ^* by drawing a subset of rankings $\Sigma' \subseteq \Sigma$ of length $|\Sigma'| = 80$ uniformly at random, and then choosing the values of the nonzero base parameters $\{\lambda_\sigma^* : \sigma \in \Sigma'\}$ by drawing uniformly over the 80-dimensional probability simplex. The revenues of the products are drawn from the distribution $r_1, \dots, r_n \sim \text{Uniform}\{1, \dots, 10000\}$, and we sort the products such that $r_1 < \dots < r_n$. To construct a random collection of nested past assortments of length M , we draw a permutation of products $\sigma : \mathcal{N} \rightarrow \mathcal{N}$ uniformly at random, draw a subset of distinct integers $1 \leq q_1 < \dots < q_{M-1} \leq n - 1$ uniformly at random, and define the past assortments as $S_m \triangleq \{0\} \cup \{\sigma(i) : i \in \{1, \dots, q_m\}\}$ for each $m \in \{1, \dots, M - 1\}$ and $S_M \triangleq \mathcal{N}_0$. We generate the historical sales data for each of the past assortments under the ranking-based choice model with the base parameters λ^* , and we then use the algorithm from Theorem 5 to solve the robust optimization problem (RO) in the case of $\eta = 0$.

In Figure EC.5, we present the computation times of our numerical experiments averaged over ten replications for each selection of $M \in \{2, \dots, 20\}$ and $n = 20$. To appreciate the computation times in Figure EC.5, let us recall the algorithm from §5.2 for solving the robust optimization problem (RO). In the case of $M = 20$ past assortments, it follows from Lemma 5 that our algorithm from §5.2 would require iterating over as many as $2^{20-2} \approx 260000$ assortments, and each iteration would have required solving a linear optimization

problem with more than $2^{20-1} \approx 520000$ decision variables (see Corollary 2 and Proposition 1). In contrast, we observe from Figure EC.5 that the computation time of our mixed-integer optimization problem remained below 30 seconds on average. More generally, we conclude from Figure EC.5 that our mixed-integer optimization reformulation of the robust optimization problem (RO) from Theorem 5 can scale to applications with numbers of products and past assortments that are realistic in applications such as brick-and-mortar retail.

A.4. Numerical Experiments for Optimistic Optimization

In this appendix, we present numerical experiments that showcase the managerial value of using the optimistic optimization problem (OO) from §6.1 to compute upper bounds on the maximum increase in expected revenue that a firm can hope to gain by experimenting with new assortments.

To perform our numerical experiments, we begin by constructing randomly-generated problem instances with nested past assortments. Specifically, in each problem instance, we generate a sparse base parameter vector λ^* by randomly selecting a subset of rankings $\Sigma' \subseteq \Sigma$ of length $|\Sigma'| = 100$ and assigning the parameters $\{\lambda_\sigma : \sigma \in \Sigma'\}$ by drawing uniformly over the 100-dimensional probability simplex. Using this base choice for the parameters, we then generate historical sales data for each of the past assortments under the ranking-based choice model with the base parameters λ^* , after which we forget the base parameters λ^* . In each randomly-generated problem instance, the revenues are drawn from the distribution $r_1, \dots, r_n \sim \text{Uniform}\{1, \dots, 10000\}$, and we sort the products such that $r_1 < \dots < r_n$. Finally, we use the algorithm from Theorem 7 to compute the optimal objective value of the optimistic optimization problem (OO).

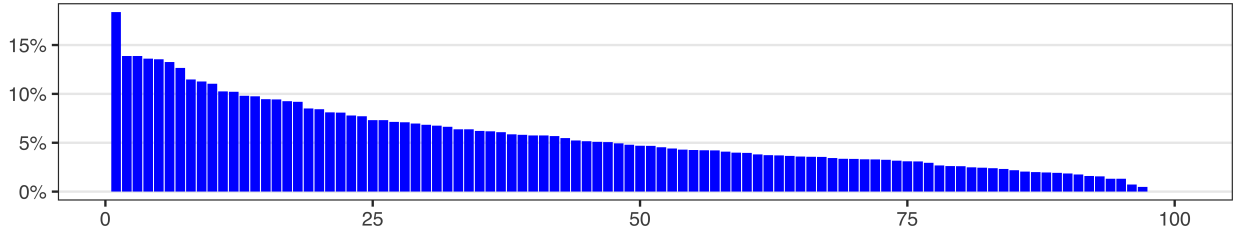
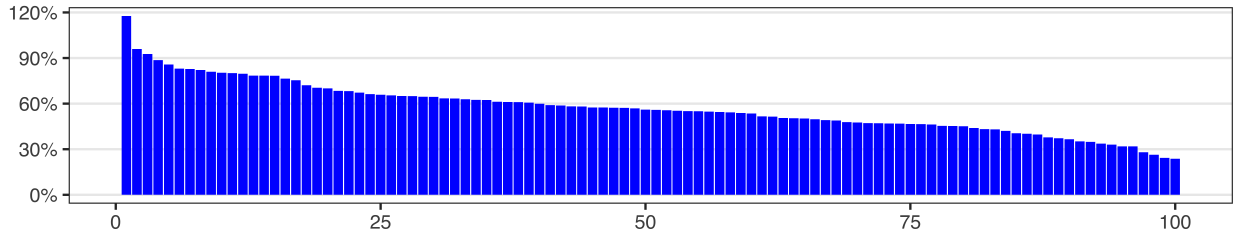
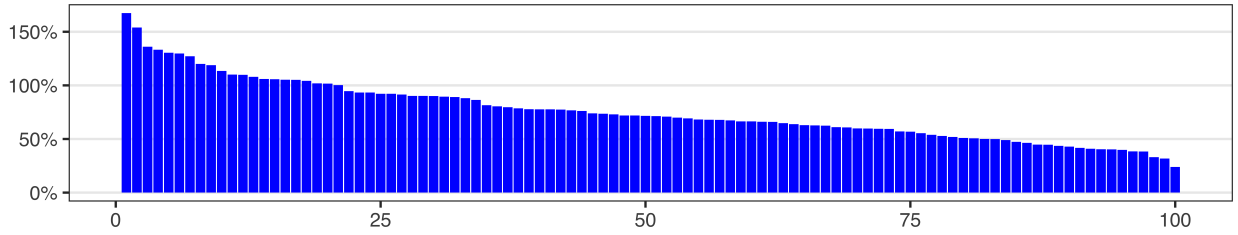
In Figure EC.6, we show the results of our experiments in cases where the collection of past assortments is the revenue-ordered assortments, the reverse revenue-ordered assortments, and neither revenue-ordered nor reverse revenue-ordered. In each of these three experiments, we randomly generate 100 problem instances and compute the optimal objective value of the optimistic optimization problem (OO). Figure EC.6 shows the value of $100\% \times (\max_{S \in \mathcal{S}} \max_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) - \max_{m \in \mathcal{M}} r^\top v_m) / (\max_{m \in \mathcal{M}} r^\top v_m)$ for each problem instance and each experiment, which represents the upper bound on the potential increase in expected revenue (relative to the firm's best past assortment) that can be gained by experimentation.

The results from Figure EC.6 show that the upper bound provided by (OO) can provide practical value in helping firms decide whether to perform experimentation. For example, each of the three experiments shown in Figure EC.6 contain problem instances in which the optimal objective value of (OO) is within 30% of the expected revenue of the firm's best past assortment. We further observe from Figure EC.7a that the optimal objective value of (OO) is within 15% of the expected revenue of the firm's best past assortment for nearly every problem instance when the past assortments are the revenue-ordered assortments. Such upper bounds observed for the problem instances in Figure EC.7a and many of the problem instances from Figure EC.7b and EC.7c can be viewed as compelling evidence to firms that the benefits of experimentation can be limited.

At first glance, the upper bounds shown in Figures EC.7b and EC.7c appear to be worse than the upper bounds shown in Figure EC.7a. However, the assessment of the quality of the upper bounds in Figures EC.7b and EC.7c should be calibrated by the quality of the collections of past assortments used in those two subfigures. Indeed, we observe that the collections of past assortments used in Figures EC.7b and EC.7c are not revenue-ordered assortments, and we have no theoretical or intuitive reasons to believe that the collections of past assortments used in Figures EC.7b and EC.7c will typically contain assortments that are near-optimal with respect to the true ranking-based choice models λ^* that generated the historical sales data. With these points in mind, the observations that the optimal objective value of (OO) in Figures EC.7b and EC.7c are sometimes tight (say, within 50% or 30% of the expected revenue of the firm's best past assortment) provides compelling evidence that (OO) is capable of giving actionable bounds even when the firm's past assortments are not chosen strategically.

A.5. Numerical Experiments for Pareto Optimization: The Exact Case

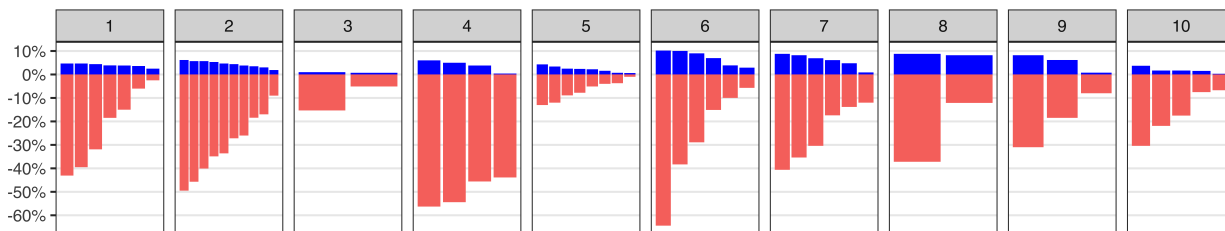
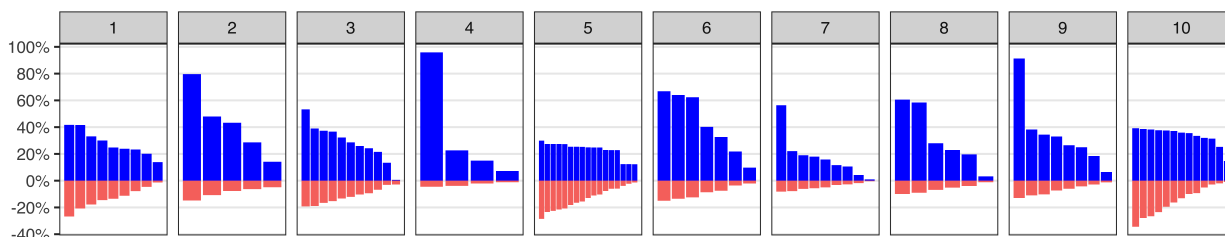
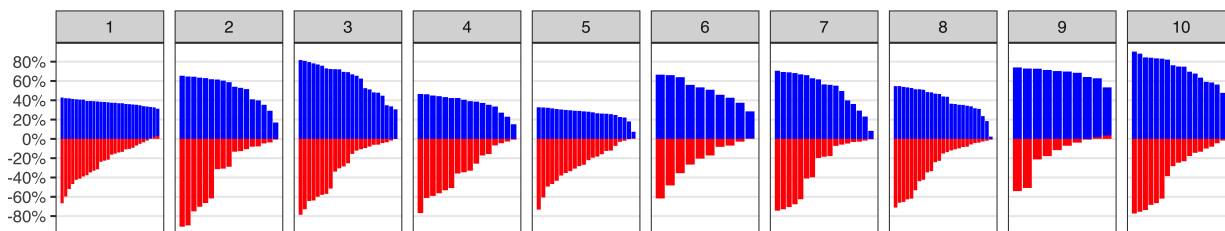
In this appendix, we showcase the managerial value of the exact algorithm from §6.2.1 for solving the Pareto optimization problem (PO) to find new assortments that can be used by firms for low-risk experimentation.

Figure EC.6 Best-case improvement for different collections of past assortments(a) $\mathcal{M} = \{\{0, 10\}, \{0, 9, 10\}, \{0, 8, 9, 10\}, \dots, \{0, 1, \dots, 9, 10\}\}$ (b) $\mathcal{M} = \{\{0, 10\}, \{0, 1, 10\}, \{0, 1, 2, 10\}, \dots, \{0, 1, \dots, 9, 10\}\}$ (c) $\mathcal{M} = \{\{0, 3, 8, 13\}, \{0, 3, 5, 8, 10, 13, 15\}, \{0, 1, 3, 5, 6, 8, 10, 11, 13, 15\}, \{0, 1, 3, 4, 5, 6, 8, 9, 10, 11, 13, 14, 15\}, \{0, \dots, 15\}\}$

Note. Each of the x -values in each experiment corresponds to one of the 100 randomly-generated problem instances. For each randomly-generated problem instance, the corresponding y -value shown by the blue bar is $100\% \times (\max_{S \in \mathcal{S}} \max_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) - \max_{m \in \mathcal{M}} r^T v_m) / (\max_{m \in \mathcal{M}} r^T v_m)$, which is the percentage increase of the optimal objective value of (OO) compared to the expected revenue of the firm's best past assortment. The top figure shows problem instances in which the past assortments are the revenue-ordered assortments with $n = 10$ products; the middle figure shows problem instances in which the past assortments are the reverse revenue-ordered assortments with $n = 10$ products; the bottom figure shows problem instances with a collection of nested past assortments with a complicated and more realistic structure, $n = 15$ products, and $M = 5$ past assortments.

To perform our numerical experiments, we begin by constructing randomly-generated problem instances with nested past assortments. Specifically, in each problem instance, we generate a sparse base parameter vector λ^* by randomly selecting a subset of rankings $\Sigma' \subseteq \Sigma$ of length $|\Sigma'| = 80$ and assigning the parameters $\{\lambda_\sigma : \sigma \in \Sigma'\}$ by drawing uniformly over the 80-dimensional probability simplex. Using this base choice for the parameters, we then generate historical sales data for each of the past assortments under the ranking-based choice model with the base parameters λ^* , after which we forget the base parameters λ^* . In each randomly-generated problem instance, the revenues are drawn from the distribution $r_1, \dots, r_n \sim \text{Uniform}\{1, \dots, 10000\}$, and we sort the products such that $r_1 < \dots < r_n$. Finally, we use the algorithm from Theorem 5 to compute the optimal objective value of the robust optimization problem (RO), and we then use the algorithm from Theorem 8 to solve the optimization problem (PO) for each parameter $\theta \in \{q \times (\text{RO}) : q \in \{0, 0.01, \dots, 0.99, 1.00\}\}$ to obtain a new assortment, denoted by S' .

In Figure EC.7, we show the results of our experiments in cases where the collection of past assortments is the revenue-ordered assortments, the reverse revenue-ordered assortments, and neither revenue-ordered nor reverse revenue-ordered. In each of these experiments, we visualize the Pareto frontier of new assortments generated under ten randomly-generated instances of the specified collections of past assortments. Compared

Figure EC.7 Pareto frontiers of new assortments from different collections of past assortments(a) $\mathcal{M} = \{\{0, 10\}, \{0, 9, 10\}, \{0, 8, 9, 10\}, \dots, \{0, 1, \dots, 9, 10\}\}$ (b) $\mathcal{M} = \{\{0, 10\}, \{0, 1, 10\}, \{0, 1, 2, 10\}, \dots, \{0, 1, \dots, 9, 10\}\}$ (c) $\mathcal{M} = \{\{0, 3, 8, 13\}, \{0, 3, 5, 8, 10, 13, 15\}, \{0, 1, 3, 5, 6, 8, 10, 11, 13, 15\}, \{0, 1, 3, 4, 5, 6, 8, 9, 10, 11, 13, 14, 15\}, \{0, \dots, 15\}\}$

Note. Each of the x -values in each problem instance corresponds to one of the *unique* assortments obtained from solving the optimization problem (PO) across the choices of parameters $\theta \in \{q \times (\text{RO}) : q \in \{0, 0.01, \dots, 0.99, 1.00\}\}$. The corresponding interval of y -values formed by the red and blue bars is the interval $[\Pi_\lambda : \lambda \in \mathcal{U}]$, where $\Pi_\lambda \triangleq 100\% \times (\mathcal{R}^\lambda(S') - \max_{m \in \mathcal{M}} r^\top v_m) / (\max_{m \in \mathcal{M}} r^\top v_m)$ is the relative percentage improvement of the expected revenue of the new assortment over the expected revenue of the firm's best past assortment for a given ranking-based choice model $\lambda \in \mathcal{U}$. The top figure shows problem instances in which the past assortments are the revenue-ordered assortments with $n = 10$ products; the middle figure shows problem instances in which the past assortments are the reverse revenue-ordered assortments with $n = 10$ products; the bottom figure shows problem instances with a collection of nested past assortments with a complicated and more realistic structure, $n = 15$ products, and $M = 5$ past assortments.

to new assortments found using estimate-then-optimize, the Pareto frontier of new assortments found from solving (PO) enables firms to select a new assortment that maximizes their best-case improvement in expected revenue for any guaranteed level of risk that a firm is willing to incur. This distinction is particularly salient in the case where the past assortments are the revenue-ordered assortments (top of Figure EC.7). Indeed, we recall from §3 and Appendix A.1 that the assortments found using estimate-then-optimize can have significant declines in expected revenue compared to the best expected revenue from the firm's past assortments. In contrast, we observe from Figure EC.7a that nine out of the ten problem instances yielded new assortments that firms can offer to their customers that, under any ranking-based choice model that is consistent with the firm's historical sales data, will cause the firm's expected revenue to decline by no more than 12%. Hence, the results from Figure EC.7 show that our algorithms can offer firms a low-risk way to experiment with the assortments that they offer to their customers.

A.6. Numerical Experiments for Pareto Optimization: The Approximate Case

In this appendix, we showcase the managerial value of the heuristic from §6.2.2 for the Pareto optimization problem (PO) to find new assortments that can be used by firms for low-risk experimentation.

Our experiments in this appendix are based on a real conjoint dataset involving laptop bags sold by the firm Timbuk2. This conjoint dataset was first described by Toubia et al. (2003) and has subsequently been studied in the context of assortment optimization by Belloni et al. (2008) and Bertsimas and Mišić (2019). The dataset from Timbuk2¹⁵ consists of 3584 variations of laptop bags, and the goal of the firm in the assortment optimization context to select a subset of the laptop bag variations to that maximizes the firm’s expected revenue. The conjoint study performed by Toubia et al. (2003) included 330 respondents, each of whom conveyed their preferences for different variations of the products. Based on their responses, Toubia et al. (2003) estimated the utilities for each of the variations of laptop bags for each respondent, which was then converted into rankings at various price points by Belloni et al. (2008). The resulting problem faced in this example is an assortment optimization problem in which the number of products is $n = 3594$, the ranking-based choice model satisfies $\|\lambda^*\|_0 = 330$, and the weights of the rankings corresponding to each of the respondents satisfies $\lambda_\sigma^* = \frac{1}{330}$.

We take several steps to convert the real data from the conjoint study into the setting studied in this paper. We begin by assuming that true ranking-based choice model λ^* is unknown, and that the only information available to the firm about the true choice model comes from historical sales data generated by $M \in \{3, 4, 5\}$ past assortments. We also restrict the universe of products from 3594 products to 15 products by drawing the subset uniformly at random and rejecting any subset that contains products with the identical revenues. We then randomly generate the past assortments $S_1, \dots, S_M \in \mathcal{S}$, whereby the assortments satisfy $\{0, n\} \subseteq S_1 \cap \dots \cap S_M$ and, for each of the remaining products $j \in \{1, \dots, n - 1\}$, we randomly assign the product to the past assortments $I \subseteq \mathcal{M}$ with distribution given by

$$P\left(j \in \bigcap_{m \in I} S_m \setminus \bigcup_{m \notin I} S_m\right) = \begin{cases} \frac{1}{2^M - 1}, & \text{if } |I| \geq 1, \\ 0, & \text{otherwise.} \end{cases}$$

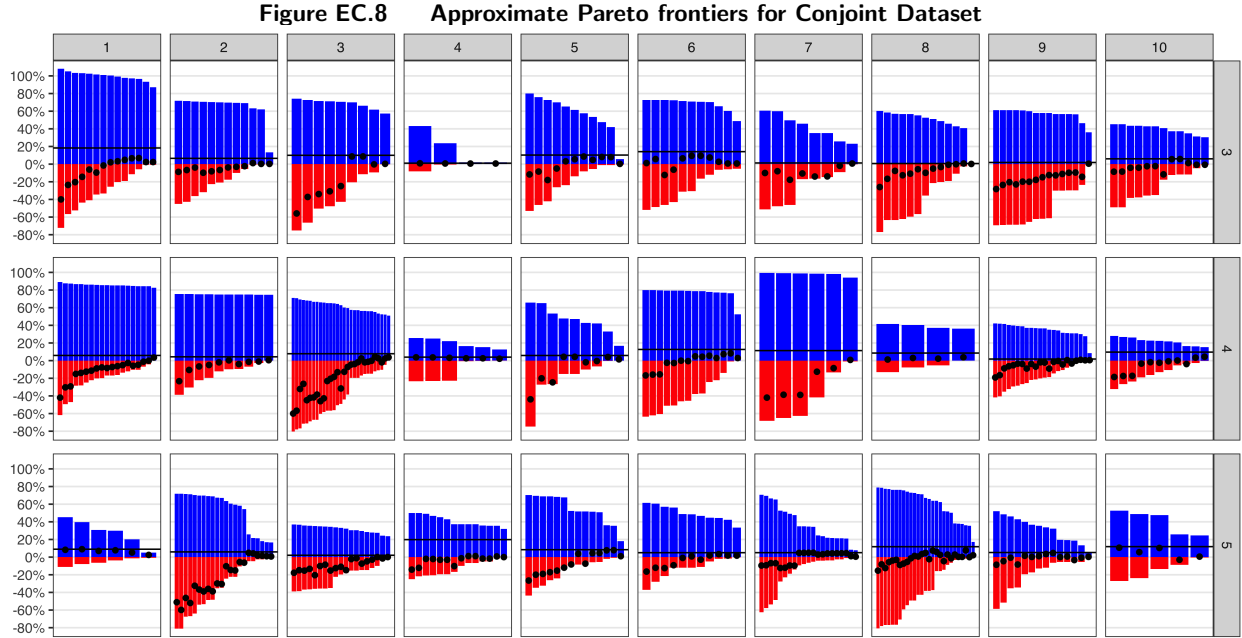
The above approach to constructing past assortments is a generalization of the approach used in Appendix A.2, and our approach for constructing past assortments ensures that each product appears in at least one past assortment. Finally, we use the heuristic from §6.2.2 to compute the approximate Pareto frontier of new assortments defined on line (7). We repeat this entire procedure ten times for each $M \in \{3, 4, 5\}$.

In Figure EC.8, we show the results of our experiments with $M \in \{3, 4, 5\}$ past assortments. The results of Figure EC.8 show that the approximate Pareto frontier often contains new assortments that have high best-case expected revenue and high true expected revenue relative to the expected revenue of the firm’s best past assortment. Moreover, the results from Figure EC.8 indicate that the assortments from the approximate Pareto frontier with the highest true expected revenue (i.e., the assortments with the highest black dots) are often the assortments from the approximate Pareto frontier with worst-case expected revenue that is rather close (e.g., within 10%) of the expected revenue of the firm’s best past assortment. These results suggest that the approximate Pareto frontier obtained using the heuristic from §6.2.2 can contain new assortments that lead to increases in the firm’s expected revenue with respect to a true but unknown ranking-based choice model λ^* , while simultaneously enjoying guarantees that implementing the new assortments will not lead to meaningful worst-case decreases in the firm’s expected revenue.

Appendix B: Proofs of Theorem 1 from §4 and Technical Results from §4.4

This appendix contains the proof of Theorem 1 by following the proof outline described in §4.4. We begin in Appendix B.1 by presenting the proof of Proposition 1 from §4.4. In Appendix B.2, we present the graphical interpretation of Definition 4 from §4.4. In Appendix B.3, we use this graphical interpretation to prove Proposition 2 from §4.4. We conclude in Appendix B.4 by using Propositions 1 and 2 to prove Theorem 1.

¹⁵ We accessed the publicly available dataset from <https://github.com/vvmisic/optimalPLD>.



Note. Each row corresponds to the number of past assortments, $M \in \{3, 4, 5\}$. For each number of past assortments M , we randomly generate ten problem instances (where each problem instance consists of a randomly-generated collection of past assortments and a random selection of the $n = 15$ products). For each problem instance, the distinct x -values have a one-to-one correspondence with the approximate Pareto frontier of assortments defined on line (7) in §6.2.2. For each assortment S' in the approximate Pareto frontier, the corresponding interval of y -values formed by the red and blue bars is the interval $[\Pi_\lambda : \lambda \in \mathcal{U}]$, where $\Pi_\lambda \triangleq 100\% \times (\mathcal{R}^\lambda(S') - \max_{m \in \mathcal{M}} r^\top v_m) / (\max_{m \in \mathcal{M}} r^\top v_m)$ is the relative percentage improvement of the expected revenue of the new assortment over the expected revenue of the firm's best past assortment for a given ranking-based choice model $\lambda \in \mathcal{U}$. The black dot for each assortment in the approximate Pareto frontier shows $100\% \times (\mathcal{R}^{\lambda^*}(S') - \max_{m \in \mathcal{M}} r^\top v_m) / (\max_{m \in \mathcal{M}} r^\top v_m)$, i.e., the percentage improvement of the true expected revenue of assortment S' with respect to the true ranking-based choice model λ^* from the conjoint dataset over the expected revenue of the firm's best past assortment. The solid horizontal line in each problem instance shows $100\% \times (\max_{S \in \mathcal{S}} \mathcal{R}^{\lambda^*}(S) - \max_{m \in \mathcal{M}} r^\top v_m) / (\max_{m \in \mathcal{M}} r^\top v_m)$, i.e., the percentage improvement of the expected revenue of the optimal assortment with respect to the true ranking-based choice model λ^* from the conjoint dataset over the expected revenue of the firm's best past assortment.

B.1. Proof of Proposition 1

Following the notation from §2, we readily observe that the worst-case expected revenue for the fixed assortment $S \in \mathcal{S}$ is equal to the optimal objective value of the following linear optimization problem:

$$\begin{aligned}
 & \underset{\lambda, \epsilon}{\text{minimize}} && \sum_{\sigma \in \Sigma} \sum_{i \in S} r_i \mathbb{I} \left\{ i = \arg \min_{j \in S} \sigma(j) \right\} \lambda_\sigma \\
 & \text{subject to} && \sum_{\sigma \in \Sigma} \mathbb{I} \left\{ i = \arg \min_{j \in S_m} \sigma(j) \right\} \lambda_\sigma - \epsilon_{m,i} = v_{m,i} \quad \forall m \in \mathcal{M} \text{ and } i \in S_m \\
 & && \sum_{\sigma \in \Sigma} \lambda_\sigma = 1 \\
 & && \|\epsilon\| \leq \eta \\
 & && \lambda_\sigma \geq 0 \quad \forall \sigma \in \Sigma.
 \end{aligned} \tag{WC-S}$$

For every assortment $S' \in \mathcal{S}$, product in the assortment $i \in S'$, and ranking $\sigma \in \Sigma$, it follows immediately from Definition 2 that the equality $\mathbb{I} \{ i = \arg \min_{j \in S'} \sigma(j) \} = 1$ is satisfied if and only if $\sigma \in \mathcal{D}_i(S')$. Therefore,

for every assortment $S \in \mathcal{S}$, we observe that the linear optimization problem (WC- S) is equivalent to

$$\begin{aligned}
& \underset{\lambda, \epsilon}{\text{minimize}} && \sum_{i \in S} \left(\sum_{\sigma \in \mathcal{D}_i(S)} \lambda_\sigma \right) r_i \\
& \text{subject to} && \sum_{\sigma \in \mathcal{D}_i(S_m)} \lambda_\sigma - \epsilon_{m,i} = v_{m,i} \quad \forall m \in \mathcal{M} \text{ and } i \in S_m \\
& && \sum_{\sigma \in \Sigma} \lambda_\sigma = 1 \\
& && \|\epsilon\| \leq \eta \\
& && \lambda_\sigma \geq 0 \quad \forall \sigma \in \Sigma.
\end{aligned} \tag{WC-S-1}$$

Moreover, we recall from Definitions 2 and 3 that there is a unique tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$ corresponding to each ranking $\sigma \in \Sigma$ that satisfies $\sigma \in \cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m)$. Therefore, it follows immediately from Definitions 2 and 3 that the linear optimization problem (WC- S -1) is equivalent to

$$\begin{aligned}
& \underset{\lambda, \epsilon}{\text{minimize}} && \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \sum_{i \in S} r_i \left(\sum_{\sigma \in \cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m) \cap \mathcal{D}_i(S)} \lambda_\sigma \right) \\
& \text{subject to} && \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i} \left(\sum_{\sigma \in \cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m)} \lambda_\sigma \right) - \epsilon_{m,i} = v_{m,i} \quad \forall m \in \mathcal{M} \text{ and } i \in S_m \\
& && \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \left(\sum_{\sigma \in \cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m)} \lambda_\sigma \right) = 1 \\
& && \|\epsilon\| \leq \eta \\
& && \lambda_\sigma \geq 0 \quad \forall \sigma \in \Sigma.
\end{aligned} \tag{WC-S-2}$$

We now simplify the linear optimization problem (WC- S -2) by performing a transformation on its decision variables. Specifically, we transform (WC- S -2) by creating the following new decision variables for each tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$ and each product $i \in S$:

$$\lambda_{i_1 \dots i_M} \leftarrow \sum_{\sigma \in \cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m)} \lambda_\sigma; \quad \omega_{i, i_1 \dots i_M} \leftarrow \sum_{\sigma \in \cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m) \cap \mathcal{D}_i(S)} \lambda_\sigma.$$

Let us make three observations about the new decision variables defined above. First, we observe immediately from the definition of the new decision variables $\omega_{i, i_1 \dots i_M}$ that the equality $\omega_{i, i_1 \dots i_M} = 0$ must hold for each tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$ and product $i \in S$ that satisfy $\cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m) \cap \mathcal{D}_i(S) = \emptyset$. Second, we observe that the nonnegativity constraints $\lambda_\sigma \geq 0$ in the linear optimization problem (WC- S -2) imply that the new decision variables must also satisfy the nonnegativity constraints $\lambda_{i_1 \dots i_M} \geq 0$ and $\omega_{i, i_1 \dots i_M} \geq 0$. Third, we observe from the definitions of the new decision variables that the equality $\sum_{i \in S} \omega_{i, i_1 \dots i_M} = \lambda_{i_1 \dots i_M}$ must hold for every tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$. Based on the definitions of the new decision variables and by applying the aforementioned three observations, we have shown that the linear optimization

problem (WC-S-2) is equivalent to

$$\begin{aligned}
& \underset{\lambda, \epsilon, \omega}{\text{minimize}} && \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \sum_{i \in S} r_i \omega_{i, i_1 \dots i_M} \\
& \text{subject to} && \sum_{i \in S} \omega_{i, i_1 \dots i_M} = \lambda_{i_1 \dots i_M} \quad \forall (i_1, \dots, i_M) \in \mathcal{L} \\
& && \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i} \lambda_{i_1 \dots i_M} - \epsilon_{m, i} = v_{m, i} \quad \forall m \in \mathcal{M} \text{ and } i \in S_m \\
& && \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \lambda_{i_1 \dots i_M} = 1 \\
& && \|\epsilon\| \leq \eta \\
& && \lambda_{i_1 \dots i_M} \geq 0 \quad \forall (i_1, \dots, i_M) \in \mathcal{L} \\
& && \omega_{i, i_1 \dots i_M} \geq 0 \quad \forall i \in S \text{ and } (i_1, \dots, i_M) \in \mathcal{L} \\
& && \omega_{i, i_1 \dots i_M} = 0 \quad \forall i \in S \text{ and } (i_1, \dots, i_M) \in \mathcal{L} \text{ such that } \bigcap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m) \cap \mathcal{D}_i(S) = \emptyset.
\end{aligned} \tag{WC-S-3}$$

After rearranging terms, we observe that the linear optimization problem (WC-S-3) is equivalent to

$$\begin{aligned}
& \underset{\lambda, \epsilon}{\text{minimize}} && \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \left[\begin{array}{l} \underset{\omega}{\text{minimize}} \quad \sum_{i \in S} r_i \omega_{i, i_1 \dots i_M} \\ \text{subject to} \quad \sum_{i \in S} \omega_{i, i_1 \dots i_M} = \lambda_{i_1 \dots i_M} \\ \omega_{i, i_1 \dots i_M} \geq 0 \quad \forall i \in S \\ \omega_{i, i_1 \dots i_M} = 0 \quad \forall i \in S \text{ such that } \bigcap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m) \cap \mathcal{D}_i(S) = \emptyset \end{array} \right] \\
& \text{subject to} && \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i} \lambda_{i_1 \dots i_M} - \epsilon_{m, i} = v_{m, i} \quad \forall m \in \mathcal{M} \text{ and } i \in S_m \\
& && \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \lambda_{i_1 \dots i_M} = 1 \\
& && \|\epsilon\| \leq \eta \\
& && \lambda_{i_1 \dots i_M} \geq 0 \quad \forall (i_1, \dots, i_M) \in \mathcal{L}.
\end{aligned} \tag{WC-S-4}$$

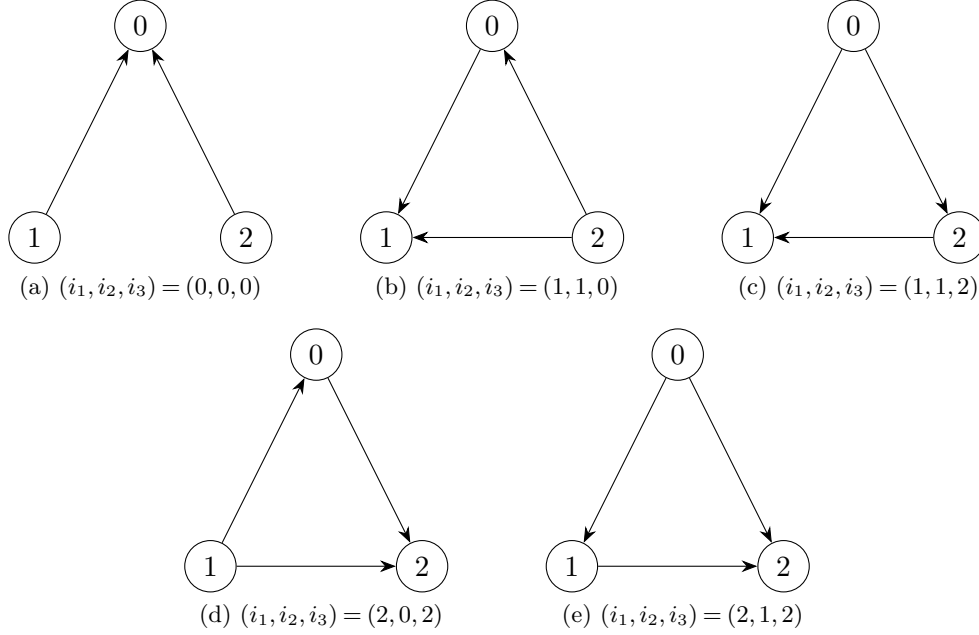
Finally, we readily observe from Definition 4 that the following equality holds for every $(i_1, \dots, i_M) \in \mathcal{L}$:

$$\left[\begin{array}{l} \underset{\omega}{\text{minimize}} \quad \sum_{i \in S} r_i \omega_{i, i_1 \dots i_M} \\ \text{subject to} \quad \sum_{i \in S} \omega_{i, i_1 \dots i_M} = \lambda_{i_1 \dots i_M} \\ \omega_{i, i_1 \dots i_M} \geq 0 \quad \forall i \in S \\ \omega_{i, i_1 \dots i_M} = 0 \quad \forall i \in S \text{ such that } \bigcap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m) \cap \mathcal{D}_i(S) = \emptyset \end{array} \right] = \rho_{i_1 \dots i_M}(S) \lambda_{i_1 \dots i_M}.$$

It follows from the above equality that (WC-S-4) is equivalent to the linear optimization problem (2), which concludes our proof of Proposition 1. \square

B.2. A Graphical Interpretation of Definition 4

In view of our overarching strategy for the proof of Proposition 2 that is described at the end of §4, we now proceed to analyze the behavior of the functions $S \mapsto \rho_{i_1 \dots i_M}(S)$. In particular, we will show in the rest of Appendix B.2 that $\rho_{i_1 \dots i_M}(S)$ can be computed by analyzing the reachability of vertices in a directed acyclic graph. In the subsequent Appendix B.3, we will use this graphical interpretation of Definition 4 to show for every arbitrary assortment $S \in \mathcal{S}$ that we can construct an assortment $S' \in \hat{\mathcal{S}}$ that satisfies $\rho_{i_1 \dots i_M}(S) \leq \rho_{i_1 \dots i_M}(S')$ for all tuples of products $(i_1, \dots, i_M) \in \mathcal{L}$, thereby completing the proof of Theorem 1.

Figure EC.9 Visualizations of directed graphs $\mathcal{G}_{i_1 \dots i_M}$ corresponding to tuples of products $(i_1, \dots, i_M) \in \mathcal{L}$.

Note. Each of the five figures presents a visualization of the directed graph $\mathcal{G}_{i_1 i_2 i_3}$ corresponding to a tuple of products $(i_1, i_2, i_3) \in \mathcal{L}$ in the case where there are $M = 3$ past assortments of the form $S_1 = \{0, 1, 2\}$, $S_2 = \{0, 1\}$, and $S_3 = \{0, 2\}$. We observe that there exists an incoming edge to vertex i if and only if there exists a past assortment $m \in \{1, 2, 3\}$ that satisfies $i = i_m$.

To develop our alternative representation of $\rho_{i_1 \dots i_M}(S)$, we begin by showing that the set of tuples of products \mathcal{L} can be interpreted as a set of directed acyclic graphs. Indeed, consider any selection of products from each of the past assortments, $(i_1, \dots, i_M) \in S_1 \times \dots \times S_M$. From this tuple of products, we will construct a directed graph, denoted by $\mathcal{G}_{i_1 \dots i_M}$, in which the set of vertices in the graph is equal to \mathcal{N}_0 , and the graph has a directed edge (i, i_m) from vertex i to vertex i_m for each past assortment $m \in \mathcal{M}$ and each product $i \in S_m \setminus \{i_m\}$. In Figure EC.9, we present visualizations of the directed graphs generated by this construction procedure. In the first intermediary result of this subsection, presented below as Lemma EC.1, we show that the tuple of products (i_1, \dots, i_M) is an element of \mathcal{L} if and only if the directed graph $\mathcal{G}_{i_1 \dots i_M}$ is acyclic.

LEMMA EC.1. $(i_1, \dots, i_M) \in \mathcal{L}$ if and only if $\mathcal{G}_{i_1 \dots i_M}$ is acyclic.

Proof. Consider any tuple of products $(i_1, \dots, i_M) \in S_1 \times \dots \times S_M$. First, we recall the fundamental result in graph theory that a directed graph is acyclic if and only if there exists a ranking $\sigma \in \Sigma$ that satisfies $\sigma(i) < \sigma(j)$ for every directed edge (j, i) in the graph (Ahuja et al. 1988, p. 77). Second, we recall from our construction of $\mathcal{G}_{i_1 \dots i_M}$ that the directed edges in $\mathcal{G}_{i_1 \dots i_M}$ are exactly those of the form (j, i_m) for each past assortment $m \in \mathcal{M}$ and product $j \in S_m \setminus \{i_m\}$. Combining these two recollections, we see that the directed graph $\mathcal{G}_{i_1 \dots i_M}$ is acyclic if and only if there exists a ranking $\sigma \in \Sigma$ that satisfies $\sigma(i_m) < \sigma(j)$ for each $m \in \mathcal{M}$ and $j \in S_m \setminus \{i_m\}$. Since

$$\begin{aligned} [\sigma(i_m) < \sigma(j) \text{ for all } m \in \mathcal{M} \text{ and } j \in S_m \setminus \{i_m\}] &\iff \left[i_m = \arg \min_{j \in S_m} \sigma(j) \text{ for all } m \in \mathcal{M} \right] \\ &\iff [\sigma \in \cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m)], \end{aligned}$$

we conclude that $\mathcal{G}_{i_1 \dots i_M}$ is acyclic if and only if $(i_1, \dots, i_M) \in \mathcal{L}$. \square

We next introduce the definition of *reachability* for vertices in the directed graph $\mathcal{G}_{i_1 \dots i_M}$. In particular, the following definition is standard in the study of directed graphs, and we will make extensive use of Definition EC.1 throughout the appendices of this paper.

DEFINITION EC.1. Let $(i_1, \dots, i_M) \in S_1 \times \dots \times S_M$ and $i, j \in \mathcal{N}_0$. We say that vertex j is *reachable* from vertex i , denoted by $j \prec_{i_1 \dots i_M} i$, if there exists a directed path in $\mathcal{G}_{i_1 \dots i_M}$ from i to j .

We adopt the convention throughout this paper that a directed path must contain at least one directed edge. Hence, we observe that if the directed graph $\mathcal{G}_{i_1 \dots i_M}$ is acyclic, then it must be the case that a vertex may never be reachable from itself, that is, $i \not\prec_{i_1 \dots i_M} i$ for all $i \in \mathcal{N}_0$.

We now use Definition EC.1 to develop several intermediary results regarding the structure of the directed acyclic graph $\mathcal{G}_{i_1 \dots i_M}$ for each tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$. We begin with a simple intermediary result, denoted below by Lemma EC.2, in which we characterize the vertices in the directed acyclic graph $\mathcal{G}_{i_1 \dots i_M}$ that can be reachable from other vertices.

LEMMA EC.2. *Let $(i_1, \dots, i_M) \in \mathcal{L}$ and $i, j \in \mathcal{N}_0$. If $j \prec_{i_1 \dots i_M} i$, then there exists a past assortment $m \in \mathcal{M}$ that satisfies $j = i_m$.*

Proof. Consider any tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$. We recall from our construction of $\mathcal{G}_{i_1 \dots i_M}$ that i_1, \dots, i_M are the only vertices in $\mathcal{G}_{i_1 \dots i_M}$ that have incoming edges. Since Lemma EC.1 implies that $\mathcal{G}_{i_1 \dots i_M}$ is acyclic, we conclude that i_1, \dots, i_M are the only vertices in $\mathcal{G}_{i_1 \dots i_M}$ that can be reachable from other vertices. \square

In our next intermediary result, denoted by Lemma EC.3, we relate the reachability of vertices in a directed acyclic graph $\mathcal{G}_{i_1 \dots i_M}$ to the set of rankings that correspond to $(i_1, \dots, i_M) \in \mathcal{L}$.

LEMMA EC.3. *Let $(i_1, \dots, i_M) \in \mathcal{L}$ and $i, j \in \mathcal{N}_0$. Then, there exists a ranking $\sigma \in \cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m)$ that satisfies $\sigma(i) < \sigma(j)$ if and only if $j \not\prec_{i_1 \dots i_M} i$.*

Proof. Let $i, j \in \mathcal{N}_0$ and $(i_1, \dots, i_M) \in \mathcal{L}$, in which case it follows from Lemma EC.1 that the directed graph $\mathcal{G}_{i_1 \dots i_M}$ is acyclic.

To show the first direction of the desired result, let us suppose that $j \not\prec_{i_1 \dots i_M} i$. In this case, we observe that the directed edge (j, i) can be added to $\mathcal{G}_{i_1 \dots i_M}$ without inducing any cycles. For notational convenience, let this augmented graph with the directed edge (j, i) be denoted by $\tilde{\mathcal{G}}_{i_1 \dots i_M}$. Since the augmented graph $\tilde{\mathcal{G}}_{i_1 \dots i_M}$ is acyclic, it follows from Ahuja et al. (1988, p. 77) that there exists a ranking $\sigma \in \Sigma$ that satisfies $\sigma(j') < \sigma(i')$ for all directed edges (i', j') in the augmented graph. Moreover, since the set of directed edges in the augmented graph $\tilde{\mathcal{G}}_{i_1 \dots i_M}$ is a superset of the set of directed edges in the original graph $\mathcal{G}_{i_1 \dots i_M}$, it follows from the construction of the original graph that this ranking satisfies $\sigma(i_m) < \sigma(i')$ for all $m \in \mathcal{M}$ and all $i' \in S_m \setminus \{i_m\}$, which implies that $\sigma \in \cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m)$. Since the augmented graph has the directed edge (j, i) , we observe that this ranking satisfies $\sigma(i) < \sigma(j)$. In summary, we have shown that if $j \not\prec_{i_1 \dots i_M} i$, then there exists a ranking $\sigma \in \cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m)$ that satisfies $\sigma(i) < \sigma(j)$. This concludes our proof of the first direction of Lemma EC.3.

To show the other direction, let us suppose that $j \prec_{i_1 \dots i_M} i$. In this case, consider any arbitrary ranking $\sigma \in \Sigma$ that satisfies $\sigma \in \cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m)$. It follows from the fact that $\sigma \in \cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m)$ and from the construction of $\mathcal{G}_{i_1 \dots i_M}$ that the inequality $\sigma(j') < \sigma(i')$ holds for each directed edge (i', j') in this graph. Moreover, without loss of generality, let the directed path from vertex i to vertex j be denoted by the sequence of vertices i, j_1, \dots, j_ν, j which satisfies the property that $(i, j_1), (j_1, j_2), \dots, (j_{\nu-1}, j_\nu), (j_\nu, j)$ are directed edges in $\mathcal{G}_{i_1 \dots i_M}$. It then follows from our earlier reasoning that $\sigma(i) > \sigma(j_1) > \dots > \sigma(j_\nu) > \sigma(j)$. Since the ranking $\sigma \in \cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m)$ was chosen arbitrarily, our proof of the other direction is complete. \square

Intuitively, Lemma EC.3 shows that the reachability of vertices in a directed acyclic graph $\mathcal{G}_{i_1 \dots i_M}$ provides an encoding of the rankings that correspond to a tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$. Said another way, Lemma EC.3 implies that if vertex i has a directed path to vertex j in a directed acyclic graph $\mathcal{G}_{i_1 \dots i_M}$, then product j is always preferred to product i under all rankings that correspond to the tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$. In our final intermediary result in Appendix B.2, denoted below by Lemma EC.4, we develop a generalization of Lemma EC.3 that relates the reachability of vertices in a directed acyclic graph $\mathcal{G}_{i_1 \dots i_M}$ to the most preferred products in an assortment $S \in \mathcal{S}$.

LEMMA EC.4. *Let $(i_1, \dots, i_M) \in \mathcal{L}$, $S \in \mathcal{S}$, and $i \in S$. Then, there exists a ranking $\sigma \in \cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m)$ that satisfies $i = \arg \min_{j \in S} \sigma(j)$ if and only if $j \not\prec_{i_1 \dots i_M} i$ for all $j \in S$.*

Proof. Consider any tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$, assortment $S \in \mathcal{S}$, and product $i \in S$.

To prove the first direction of Lemma EC.4, suppose there exists a product $j \in S$ that satisfies $j \prec_{i_1 \dots i_M} i$. In this case, it follows immediately from Lemma EC.2 that there exists a past assortment $m \in \mathcal{M}$ that

satisfies $i_m \in S$ and $i_m \prec_{i_1 \dots i_M} i$. Therefore, for each ranking $\sigma \in \cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m)$, it follows from Lemma EC.3 that $\sigma(i_m) < \sigma(i)$, which combined with the fact that $i_m \in S$ implies that $i \neq \arg \min_{j \in S} \sigma(j)$. That concludes our proof of the first direction.

To prove the other direction, suppose that $i \neq \arg \min_{j \in S} \sigma(j)$ for all rankings $\sigma \in \cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m)$. In this case, it follows immediately from Definition 2 that $\cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m) \cap \mathcal{D}_i(S) = \emptyset$. Moreover, it follows from the fact that $(i_1, \dots, i_M) \in \mathcal{L}$ and from Lemma EC.1 that the directed graph $\mathcal{G}_{i_1 \dots i_M}$ is acyclic. Therefore, we observe from the fact that $\cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m) \cap \mathcal{D}_i(S) = \emptyset$ and from Lemma EC.1 that adding the directed edges $\{(j, i) : j \in S \setminus \{i\}\}$ to the directed acyclic graph $\mathcal{G}_{i_1 \dots i_M}$ would result in a directed cycle. Since a cycle visits every vertex at most once, there must exist a single vertex $j \in S \setminus \{i\}$ such that adding only the directed edge (j, i) to $\mathcal{G}_{i_1 \dots i_M}$ would result in a directed cycle, which implies that $j \prec_{i_1 \dots i_M} i$. This concludes our proof of the other direction. \square

Lemma EC.4 establishes that a product $i \in S$ is the most preferred product from assortment $S \in \mathcal{S}$ under a ranking that corresponds to the tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$ if and only if there does not exist a directed path in the directed acyclic graph $\mathcal{G}_{i_1 \dots i_M}$ from vertex i to any vertex j that satisfies $j \in S$. In other words, Lemma EC.4 implies that the set $\cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m) \cap \mathcal{D}_i(S)$ is nonempty if and only if there is no vertex $j \in S$ which is reachable from vertex i .

In view of the above, we are now ready to develop our graphical interpretation of $\rho_{i_1 \dots i_M}(S)$. This interpretation of $\rho_{i_1 \dots i_M}(S)$, which is presented below in Proposition EC.1, will be instrumental to our analysis in the subsequent Appendix B.3, where we will use this interpretation to analyze the behavior of the functions $S \mapsto \rho_{i_1 \dots i_M}(S)$ for each tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$. Our graphical representation of $\rho_{i_1 \dots i_M}(S)$ requires the following definition of the set $\mathcal{I}_{i_1 \dots i_M}(S)$, which can be interpreted as the set of all vertices $i \in \mathcal{N}_0$ in the directed graph $\mathcal{G}_{i_1 \dots i_M}$ that do not have a directed path to any of the vertices i_1, \dots, i_M that are elements of the assortment S .

DEFINITION EC.2. $\mathcal{I}_{i_1 \dots i_M}(S) \triangleq \{i \in \mathcal{N}_0 : \text{for all } m \in \mathcal{M}, \text{ if } i_m \in S, \text{ then } i_m \not\prec_{i_1 \dots i_M} i\}$.

To make sense of Definition EC.2, we recall from Lemma EC.2 that a vertex j in a directed acyclic graph $\mathcal{G}_{i_1 \dots i_M}$ can be reachable from another vertex only if $j = i_m$ for some past assortment $m \in \mathcal{M}$. Therefore, it follows immediately from Lemma EC.4 that $S \cap \mathcal{I}_{i_1 \dots i_M}(S)$ is the set of products i for which the set of rankings $\cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m) \cap \mathcal{D}_i(S)$ is nonempty. Combining this with Definition 4, we have concluded the proof of the following Proposition EC.1, which establishes our graphical interpretation of $\rho_{i_1 \dots i_M}(S)$.

PROPOSITION EC.1. *For all $S \in \mathcal{S}$ and $(i_1, \dots, i_M) \in \mathcal{L}$, $\rho_{i_1 \dots i_M}(S) = \min_{i \in S \cap \mathcal{I}_{i_1 \dots i_M}(S)} r_i$.*

B.3. Proof of Proposition 2

Equipped with Proposition EC.1 from Appendix B.2, we are now ready to present our proof of Proposition 2. To show this, we begin by developing an intermediary result, denoted below by Claim EC.1, that will allow us to compare the values of $\rho_{i_1 \dots i_M}(S)$ and $\rho_{i_1 \dots i_M}(S \cup \{i\})$ for every assortment $S \in \mathcal{S}$ and every product i which is not in the assortment.

CLAIM EC.1. *For all $S \in \mathcal{S}$, $(i_1, \dots, i_M) \in \mathcal{L}$, and $i \notin S$,*

$$\rho_{i_1 \dots i_M}(S \cup \{i\}) = \begin{cases} \rho_{i_1 \dots i_M}(S), & \text{if } i \notin \mathcal{I}_{i_1 \dots i_M}(S), \\ \min \left\{ \min_{j \in S \cap \mathcal{I}_{i_1 \dots i_M}(S) \cap \{j' \in \mathcal{N}_0 : i \not\prec_{i_1 \dots i_M} j'\}} r_j, r_i \right\}, & \text{if } i \in \mathcal{I}_{i_1 \dots i_M}(S). \end{cases}$$

Proof of Claim EC.1. Consider any assortment $S \in \mathcal{S}$, tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$, and product $i \notin S$. We first observe that

$$\begin{aligned} \mathcal{I}_{i_1 \dots i_M}(S \cup \{i\}) &= \{j \in \mathcal{N}_0 : \text{for all } m \in \mathcal{M}, \text{ if } i_m \in S \cup \{i\}, \text{ then } i_m \not\prec_{i_1 \dots i_M} j\} \\ &= \{j \in \mathcal{N}_0 : \text{for all } m \in \mathcal{M}, \text{ if } i_m \in S, \text{ then } i_m \not\prec_{i_1 \dots i_M} j\} \\ &\quad \cap \{j \in \mathcal{N}_0 : \text{if there exists } m \in \mathcal{M} \text{ such that } i = i_m, \text{ then } i \not\prec_{i_1 \dots i_M} j\} \\ &= \mathcal{I}_{i_1 \dots i_M}(S) \cap \{j \in \mathcal{N}_0 : \text{if there exists } m \in \mathcal{M} \text{ such that } i = i_m, \text{ then } i \not\prec_{i_1 \dots i_M} j\} \end{aligned}$$

$$\begin{aligned}
&= \begin{cases} \mathcal{I}_{i_1 \dots i_M}(S), & \text{if } i \notin \{i_1, \dots, i_M\}, \\ \mathcal{I}_{i_1 \dots i_M}(S) \cap \{j \in \mathcal{N}_0 : i \not\prec_{i_1 \dots i_M} j\}, & \text{if } i \in \{i_1, \dots, i_M\} \end{cases} \\
&= \mathcal{I}_{i_1 \dots i_M}(S) \cap \{j \in \mathcal{N}_0 : i \not\prec_{i_1 \dots i_M} j\}.
\end{aligned} \tag{EC.1}$$

Indeed, the first and third equalities follow from Definition EC.2. The second and fourth equalities follow from algebra. The final equality follows from the fact that if $i \notin \{i_1, \dots, i_M\}$, then it follows from the construction of $\mathcal{G}_{i_1 \dots i_M}$ that there are no incoming edges to vertex i , which implies that $\{j \in \mathcal{N}_0 : i \not\prec_{i_1 \dots i_M} j\} = \mathcal{N}_0$.

Therefore, it follows from line (EC.1) and Proposition EC.1 that

$$\begin{aligned}
\rho_{i_1 \dots i_M}(S \cup \{i\}) &= \min_{j \in (S \cup \{i\}) \cap \mathcal{I}_{i_1 \dots i_M}(S \cup \{i\})} r_j \\
&= \min_{j \in (S \cup \{i\}) \cap \mathcal{I}_{i_1 \dots i_M}(S) \cap \{j' \in \mathcal{N}_0 : i \not\prec_{i_1 \dots i_M} j'\}} r_j \\
&= \begin{cases} \min_{j \in S \cap \mathcal{I}_{i_1 \dots i_M}(S) \cap \{j' \in \mathcal{N}_0 : i \not\prec_{i_1 \dots i_M} j'\}} r_j, & \text{if } i \notin \mathcal{I}_{i_1 \dots i_M}(S), \\ \min \left\{ \min_{j \in S \cap \mathcal{I}_{i_1 \dots i_M}(S) \cap \{j' \in \mathcal{N}_0 : i \not\prec_{i_1 \dots i_M} j'\}} r_j, r_i \right\}, & \text{if } i \in \mathcal{I}_{i_1 \dots i_M}(S), \end{cases}
\end{aligned} \tag{EC.2}$$

where the first equality follows from Proposition EC.1, the second equality follows from line (EC.1), and the last equality follows from algebra and the fact that $i \not\prec_{i_1 \dots i_M} i$.

We conclude the proof of Claim EC.1 by considering the case where $i \notin \mathcal{I}_{i_1 \dots i_M}(S)$. For this case, suppose for the sake of developing a contradiction that there exists $j \in S \cap \mathcal{I}_{i_1 \dots i_M}(S)$ which satisfies $i \prec_{i_1 \dots i_M} j$. Under that supposition, it would follow from the fact that $i \notin \mathcal{I}_{i_1 \dots i_M}(S)$ and Definition EC.2 that there would exist a past assortment $m \in \mathcal{M}$ that satisfies $i_m \in S$ and $i_m \prec_{i_1 \dots i_M} i$. Thus, by the transitive property, we have $i_m \prec_{i_1 \dots i_M} i \prec_{i_1 \dots i_M} j$, which contradicts the supposition that $j \in \mathcal{I}_{i_1 \dots i_M}(S)$. Because we have a contradiction, we have shown that $i \not\prec_{i_1 \dots i_M} j$ for all $j \in S \cap \mathcal{I}_{i_1 \dots i_M}(S)$, and so the desired result follows immediately from line (EC.2) and Definition 4. \square

Using the above intermediary result, we now complete the proof of Proposition 2. Indeed, consider any assortment $S \in \mathcal{S}$ and any product $i \notin S$. Suppose that there exists a product $i^* \in S$ which satisfies $r_{i^*} < r_i$ and $\mathcal{M}_{i^*} \subseteq \mathcal{M}_i$. For each tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$, we have two cases to consider:

- **Case 1:** Suppose that $i \notin \mathcal{I}_{i_1 \dots i_M}(S)$. In this case, it follows immediately from Claim EC.1 that

$$\rho_{i_1 \dots i_M}(S \cup \{i\}) = \rho_{i_1 \dots i_M}(S),$$

and so the inequality $\rho_{i_1 \dots i_M}(S) \leq \rho_{i_1 \dots i_M}(S \cup \{i\})$ holds when $i \notin \mathcal{I}_{i_1 \dots i_M}(S)$.

- **Case 2:** Suppose that $i \in \mathcal{I}_{i_1 \dots i_M}(S)$.

In this case, we begin by showing that $i_m \notin S$ for each $m \in \mathcal{M}_i$. Indeed, consider any past assortment $m \in \mathcal{M}_i$. On one hand, if $i = i_m$, then it follows immediately from the fact that $i \notin S$ that $i_m \notin S$. On the other hand, if $i \neq i_m$, then it also must be the case that $i_m \notin S$, else we would have a contradiction with the fact that $i \in \mathcal{I}_{i_1 \dots i_M}(S)$ and the fact that there is, by our construction of $\mathcal{G}_{i_1 \dots i_M}$, a directed edge from vertex i to vertex i_m . We have thus shown that $i_m \notin S$ for all $m \in \mathcal{M}_i$.

We next show that $i_m \in \mathcal{I}_{i_1 \dots i_M}(S)$ for all $m \in \mathcal{M}_i$. Indeed, consider any arbitrary $m \in \mathcal{M}_i$. On one hand, if $i_m = i$, then the statement $i_m \in \mathcal{I}_{i_1 \dots i_M}(S)$ follows immediately from the fact that $i \in \mathcal{I}_{i_1 \dots i_M}(S)$. On the other hand, if $i_m \neq i$, then it follows from the fact that $i \in \mathcal{I}_{i_1 \dots i_M}(S)$ and the fact that there is a directed edge from vertex i to vertex i_m that there must not be a directed path from vertex i_m to a vertex $i_{m'}$ that satisfies $i_{m'} \in S$ for any $m' \in \mathcal{M}$. We have thus shown that $i_m \in \mathcal{I}_{i_1 \dots i_M}(S)$ for all $m \in \mathcal{M}_i$.

Using the above results, we now prove that $i^* \in \mathcal{I}_{i_1 \dots i_M}(S)$. Indeed, we have shown in the above results that $i_m \notin S$ and $i_m \in \mathcal{I}_{i_1 \dots i_M}(S)$ for all $m \in \mathcal{M}_i$. Therefore, it follows from the supposition that $\mathcal{M}_{i^*} \subseteq \mathcal{M}_i$ that $i_m \notin S$ and $i_m \in \mathcal{I}_{i_1 \dots i_M}(S)$ for all $m \in \mathcal{M}_{i^*}$. Since we have supposed that $i^* \in S$, it follows from the fact that $i_m \notin S$ for all $m \in \mathcal{M}_{i^*}$ that $i^* \neq i_m$ for all $m \in \mathcal{M}_{i^*}$. Moreover, it follows from the construction of $\mathcal{G}_{i_1 \dots i_M}$ that all of the outgoing edges from vertex i^* are incoming edges to vertices i_m for $m \in \mathcal{M}_{i^*}$. Therefore, it follows from the fact that $i_m \in \mathcal{I}_{i_1 \dots i_M}(S)$ for all $m \in \mathcal{M}_{i^*}$ that $i^* \in \mathcal{I}_{i_1 \dots i_M}(S)$.

We now prove the desired result for Case 2. First, we observe that

$$\rho_{i_1 \dots i_M}(S) = \min_{j \in S \cap \mathcal{I}_{i_1 \dots i_M}(S)} r_j \leq r_{i^*}, \quad (\text{EC.3})$$

where the equality follows from Proposition EC.1 and the inequality follows from our supposition that $i^* \in S$ and because we have shown that $i^* \in \mathcal{I}_{i_1 \dots i_M}(S)$. Therefore, we have that

$$\begin{aligned} \rho_{i_1 \dots i_M}(S \cup \{i\}) &= \min \left\{ \min_{j \in S \cap \mathcal{I}_{i_1 \dots i_M}(S) \cap \{j' \in \mathcal{N}_0 : i \neq_{i_1 \dots i_M} j'\}} r_j, r_i \right\} \\ &\geq \min \left\{ \min_{j \in S \cap \mathcal{I}_{i_1 \dots i_M}(S)} r_j, r_i \right\} \\ &= \min \{ \rho_{i_1 \dots i_M}(S), r_i \} \\ &= \rho_{i_1 \dots i_M}(S), \end{aligned}$$

where the first equality follows from Claim EC.1 and the supposition of Case 2 that $i \in \mathcal{I}_{i_1 \dots i_M}(S)$, the inequality follows algebra, the second equality follows from Proposition EC.1, and the final equality holds because of our supposition that $r_{i^*} < r_i$ and because of line (EC.3), which showed that $\rho_{i_1 \dots i_M}(S) \leq r_{i^*}$. This concludes the proof of Case 2.

In both of the above two cases, we showed that $\rho_{i_1 \dots i_M}(S) \leq \rho_{i_1 \dots i_M}(S \cup \{i\})$, and so our proof of Proposition 2 is complete. \square

B.4. Proof of Theorem 1

Consider any arbitrary assortment $S \in \mathcal{S}$. For this assortment, we define a new assortment as

$$S' \triangleq S \cup \{i \in \mathcal{N}_0 : \text{there exists } i^* \in S \text{ such that } \mathcal{M}_{i^*} \subseteq \mathcal{M}_i \text{ and } r_{i^*} < r_i\}.$$

It follows immediately from the definition of the collection $\widehat{\mathcal{S}}$ that this new assortment satisfies $S' \in \widehat{\mathcal{S}}$. Moreover, let $\{j_1, \dots, j_\nu\} \triangleq S' \setminus S$ denote the new products that have been added into the assortment. Then we observe for each tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$ that

$$\rho_{i_1 \dots i_M}(S') = \rho_{i_1 \dots i_M}(S) + \sum_{\iota=1}^{\nu} (\rho_{i_1 \dots i_M}(S \cup \{j_1, \dots, j_\iota\}) - \rho_{i_1 \dots i_M}(S \cup \{j_1, \dots, j_{\iota-1}\})) \geq \rho_{i_1 \dots i_M}(S).$$

Indeed, the equality follows from algebra. The inequality follows from Proposition 2, which implies that $\rho_{i_1 \dots i_M}(S \cup \{j_1, \dots, j_\iota\}) \geq \rho_{i_1 \dots i_M}(S \cup \{j_1, \dots, j_{\iota-1}\})$ for each $\iota \in \{1, \dots, \nu\}$. Since the assortment $S \in \mathcal{S}$ was chosen arbitrarily, our proof of Theorem 1 follows from Proposition 1. \square

Appendix C: Proofs of Technical Results from §4.1

C.1. Proof of Lemma 1

Let $\mathcal{M} = \bar{\mathcal{S}}$, and let the past assortments be indexed by $\mathcal{M} = \{\bar{S}_1, \dots, \bar{S}_n\}$, whereby the i -th past assortment is $\bar{S}_i \triangleq \{0, i, i+1, \dots, n-1, n\}$. Equipped with the above notation, we first prove that the equality $\bar{\mathcal{S}} = \widehat{\mathcal{S}}$ holds. Indeed, choose any arbitrary assortment $S \in \widehat{\mathcal{S}}$, and let $i^* \triangleq \arg \min_{j \in S : r_j > 0} r_j$ denote the product in the chosen assortment that has the smallest nonzero revenue. It readily follows from the facts that $\mathcal{M} = \bar{\mathcal{S}}$ and $r_1 < \dots < r_n$ that the equalities $\mathcal{M}_{i^*} = \{m \in \mathcal{M} : m \leq i^*\}$ and $\mathcal{M}_i = \{m \in \mathcal{M} : m \leq i\}$ hold for each $i \in \{i^* + 1, \dots, n\}$. Therefore, for each $i \in \{i^* + 1, \dots, n\}$, it follows from the definition of the collection $\widehat{\mathcal{S}}$, from the fact that $r_{i^*} < r_i$, and from the fact that $\mathcal{M}_{i^*} \subseteq \mathcal{M}_i$ that $i \in S$. We have thus shown that $S = \{0, i^*, i^* + 1, \dots, n-1, n\} = \bar{S}_{i^*}$, which implies that $S \in \bar{\mathcal{S}}$. Since the assortment $S \in \widehat{\mathcal{S}}$ was chosen arbitrarily, we have shown that $\widehat{\mathcal{S}} \subseteq \bar{\mathcal{S}}$. The other direction of the proof that $\bar{\mathcal{S}} = \widehat{\mathcal{S}}$ follows from the fact that the inclusion $\mathcal{M} \subseteq \widehat{\mathcal{S}}$ always holds¹⁶ and from the fact that $\mathcal{M} = \bar{\mathcal{S}}$. Our proof that $\bar{\mathcal{S}} = \widehat{\mathcal{S}}$ is thus complete. \square

¹⁶ To see why the inclusion $\mathcal{M} \subseteq \widehat{\mathcal{S}}$ always holds, consider any past assortment $S \in \mathcal{M}$. For each product $i^* \in S$, suppose that there exists another product i which satisfies $r_{i^*} < r_i$ and $\mathcal{M}_{i^*} \subseteq \mathcal{M}_i$. Since $S \in \mathcal{M}_{i^*} \subseteq \mathcal{M}_i$, we conclude that $i \in S$ must hold, which proves that $S \in \widehat{\mathcal{S}}$.

C.2. Proof of Corollary 1

Using Lemma 1, we have

$$\max_{S \in \mathcal{S}} \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) = \max_{S \in \widehat{\mathcal{S}}} \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) = \max_{S \in \mathcal{S}} \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) = \max_{S \in \mathcal{M}} \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) = \max_{m \in \mathcal{M}} r^\top v_m,$$

where the first equality follows from Theorem 1, the second equality holds because $\widehat{\mathcal{S}} = \bar{\mathcal{S}}$, the third equality holds because $\mathcal{M} = \bar{\mathcal{S}}$, and the final equality follows from the construction of the set of ranking-based choice models \mathcal{U} (see §2) and from the fact that $\eta = 0$. \square

Appendix D: Proofs of Technical Results from §4.2

D.1. Proof of Lemma 2

For notational convenience, let the collection on the right side of line (1) in Lemma 2 be denoted by $\widehat{\mathcal{S}}'$. It follows immediately from the definition of $\widehat{\mathcal{S}}$ and from the fact that $r_0 < r_1 < \dots < r_n$ that each assortment $S \in \widehat{\mathcal{S}}'$ is also an element of $\widehat{\mathcal{S}}$. This proves that $\widehat{\mathcal{S}} \supseteq \widehat{\mathcal{S}}'$. To show the other direction, consider any arbitrary assortment $S \in \widehat{\mathcal{S}}$. We first show that $S_1 \cap S_2 \subseteq S$. Indeed, it follows from the fact that $S, S_1, S_2 \in \mathcal{S}$ that $0 \in S \cap S_1 \cap S_2$. Moreover, for each product $j \in (S_1 \cap S_2) \setminus \{0\}$, we readily observe that the inequality $r_j > r_0$ and the equality $\mathcal{M}_j = \mathcal{M}_0$ both hold. Therefore, it follows from the definition of $\widehat{\mathcal{S}}$ and the fact that $S \in \widehat{\mathcal{S}}$ that each product $j \in S_1 \cap S_2$ is also an element of S . We have thus shown that $S_1 \cap S_2 \subseteq S$ for all $S \in \widehat{\mathcal{S}}$. Next, we define the following integers:

$$i_1 \triangleq \min \left\{ n, \min_{j \in S_1 \setminus S_2} j \right\}, \quad i_2 \triangleq \min \left\{ n, \min_{j \in S_2 \setminus S_1} j \right\},$$

where any minimization problem over an empty feasible set is defined equal to ∞ . It follows from the fact that $S \in \widehat{\mathcal{S}}$, from the definition of $\widehat{\mathcal{S}}$, and from the assumption of $r_0 < r_1 < \dots < r_n$ that

$$\{j \in S_1 \setminus S_2 : j \geq i_1\} \subseteq S \text{ and } \{j \in S_2 \setminus S_1 : j \geq i_2\} \subseteq S.$$

Hence, it follows from the assumption that $S_1 \cup S_2 = \mathcal{N}_0$ that

$$S = (S_1 \cap S_2) \cup \{j \in S_1 \setminus S_2 : j \geq i_1\} \cup \{j \in S_2 \setminus S_1 : j \geq i_2\}.$$

We have thus shown that the assortment S is an element of the collection of assortments $\widehat{\mathcal{S}}'$. Since $S \in \widehat{\mathcal{S}}$ was chosen arbitrarily, we have shown that $\widehat{\mathcal{S}} \subseteq \widehat{\mathcal{S}}'$, which concludes our proof of Lemma 2. \square

D.2. Proof of Lemma 3

We readily observe that the collection of assortments $\widehat{\mathcal{S}}$ is a subset of

$$\widetilde{\mathcal{S}} \triangleq \{S \in \mathcal{S} : \text{if } i^* \in S, r_{i^*} < r_i, \text{ and } \mathcal{M}_{i^*} = \mathcal{M}_i, \text{ then } i \in S\},$$

where we recall from the beginning of §4 that \mathcal{M}_i is defined as the subset of the past assortments $\mathcal{M} \equiv \{1, \dots, M\}$ that offered product i . For each subset of past assortments $\mathcal{C} \subseteq \mathcal{M}$, let the products that are offered only in the assortments in \mathcal{C} be denoted by

$$\mathcal{N}_0(\mathcal{C}) \triangleq \{i \in \mathcal{N}_0 : \mathcal{M}_i = \mathcal{C}\}.$$

Equipped with the above definitions, we observe that $\{\mathcal{N}_0(\mathcal{C}) : \mathcal{C} \subseteq \mathcal{M}\}$ is the collection of subsets of products that always appear together in the past assortments, and we readily observe that $|\{\mathcal{N}_0(\mathcal{C}) : \mathcal{C} \subseteq \mathcal{M}\}| = 2^M$. Therefore,

$$|\widehat{\mathcal{S}}| \leq |\widetilde{\mathcal{S}}| \leq \prod_{\mathcal{C} \subseteq \mathcal{M}} (|\mathcal{N}_0(\mathcal{C})| + 1) \leq (n + 2)^{2^M}. \quad (\text{EC.4})$$

Indeed, the first inequality on line (EC.4) holds because the collection of assortments $\widehat{\mathcal{S}}$ is a subset of the collection of assortments $\widetilde{\mathcal{S}}$. To see why the second inequality on line (EC.4) holds, consider any arbitrary

subset of past assortments $\mathcal{C} \subseteq \mathcal{M}$, and let the products in $\mathcal{N}_0(\mathcal{C})$ be indexed in ascending order by revenue; that is, let the products that comprise $\mathcal{N}_0(\mathcal{C})$ be denoted by $i_1^{\mathcal{C}}, \dots, i_{|\mathcal{N}_0(\mathcal{C})|}^{\mathcal{C}}$, where $r_{i_1^{\mathcal{C}}} < \dots < r_{i_{|\mathcal{N}_0(\mathcal{C})|}^{\mathcal{C}}}$. Then, we observe from the definition of the collection $\tilde{\mathcal{S}}$ that every assortment $S \in \tilde{\mathcal{S}}$ must satisfy the condition $[\{i_{j+1}^{\mathcal{C}}, \dots, i_{|\mathcal{N}_0(\mathcal{C})|}^{\mathcal{C}}\} \subseteq S \text{ and } i_j^{\mathcal{C}} \notin S]$ for some $j \in \{0, \dots, |\mathcal{N}_0(\mathcal{C})|\}$. Since $\mathcal{C} \subseteq \mathcal{M}$ was chosen arbitrarily, we have shown that

$$\tilde{\mathcal{S}} \subseteq \left\{ \bigcup_{\mathcal{C} \subseteq \mathcal{M}} \mathcal{F}^{\mathcal{C}} : \mathcal{F}^{\mathcal{C}} \in \left\{ \emptyset, \{i_{|\mathcal{N}_0(\mathcal{C})|}^{\mathcal{C}}\}, \{i_{|\mathcal{N}_0(\mathcal{C})|-1}^{\mathcal{C}}, i_{|\mathcal{N}_0(\mathcal{C})|}^{\mathcal{C}}\}, \dots, \{i_1^{\mathcal{C}}, \dots, i_{|\mathcal{N}_0(\mathcal{C})|}^{\mathcal{C}}\} \right\} \right\},$$

which proves that the second inequality on line (EC.4) holds. The third inequality on line (EC.4) follows from the fact that $|\{\mathcal{N}_0(\mathcal{C}) : \mathcal{C} \subseteq \mathcal{M}\}| = 2^M$ and from the fact that $|\mathcal{N}_0(\mathcal{C})| \leq n + 1$ for every $\mathcal{C} \subseteq \mathcal{M}$. We have thus proven that $|\tilde{\mathcal{S}}|$ is at most $(n + 2)^{2^M}$, which concludes our proof of Lemma 3. \square

D.3. Proof of Lemma 4

As the first step in our proof of Lemma 4, we develop an algorithm for constructing a directed graph, denoted by $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$, in which the set of vertices in the directed graph is defined as $\mathcal{V} \triangleq \mathcal{N}_0$ and the set of directed edges is defined as $\mathcal{E} \triangleq \{(i^*, i) \in \mathcal{N}_0 \times \mathcal{N}_0 : r_{i^*} < r_i \text{ and } \mathcal{M}_{i^*} \subseteq \mathcal{M}_i\}$. This directed graph has a natural correspondence with the collection of assortments $\hat{\mathcal{S}}$, as we readily observe that an assortment satisfies $S \in \hat{\mathcal{S}}$ if and only if $[0 \in S]$ and $[i \in S \text{ whenever there exists a product } i^* \in S \text{ and a directed edge } (i^*, i) \in \mathcal{E}]$. It is easy to see that this directed graph is acyclic and has the property that a vertex $j \in \mathcal{V}$ is reachable from a vertex $i \in \mathcal{V}$ if and only if there is a directed edge $(i, j) \in \mathcal{E}$ from vertex i to vertex j . A directed acyclic graph which has the aforementioned property for each pair of vertices is referred to as a *transitive closure* (Ahuja et al. 1988, p.90).

Our algorithm for constructing the directed acyclic graph \mathcal{G} that is a transitive closure is presented in Algorithm 1. In the algorithm, we first iterate over each product $i \in \mathcal{N}_0$ and construct the corresponding set \mathcal{M}_i of past assortments which offered that product. It is easy to see that each of the sets \mathcal{M}_i can be constructed in $\mathcal{O}(M)$ computation time by checking whether the product satisfies $i \in S_m$ for each past assortment $m \in \mathcal{M}$; hence, we observe that all of the sets $\mathcal{M}_0, \dots, \mathcal{M}_n$ can be constructed in a total of $\mathcal{O}(Mn)$ computation time. Assume that we store the sets $\mathcal{M}_0, \dots, \mathcal{M}_n$ as unsorted arrays as well as hash tables. Given these data structures for $\mathcal{M}_0, \dots, \mathcal{M}_n$, we then iterate over each pair of products $(i, i^*) \in \mathcal{N}_0 \times \mathcal{N}_0$ and check in $\mathcal{O}(M)$ computation time whether $r_{i^*} < r_i$ and $\mathcal{M}_{i^*} \subseteq \mathcal{M}_i$. Since there are $\mathcal{O}(n^2)$ pairs of products in $\mathcal{N}_0 \times \mathcal{N}_0$, we conclude that the set of directed edges \mathcal{E} can be constructed in a total of $\mathcal{O}(n^2M)$ computation time. Combining all of the steps, and since it takes $\mathcal{O}(n)$ computation time to construct the set of vertices \mathcal{V} , we have shown that the total computation time for Algorithm 1 is $\mathcal{O}(n + Mn + n^2M) = \mathcal{O}(n^2M)$.

We next describe our algorithm for constructing the collection of assortments $\hat{\mathcal{S}}$ from the directed graph \mathcal{G} . This algorithm is denoted by $\text{CONSTRUCT-}\hat{\mathcal{S}}(\mathcal{M}, r)$ and is found in Algorithm 2. In this algorithm, we first use Algorithm 1 to construct the directed acyclic graph $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ which is a transitive closure, and then we invoke a recursive subroutine denoted by $\text{RECURSIVESTEP}(\mathcal{G})$ in Algorithm 3. The goal of the recursive subroutine is to take as an input a generic directed acyclic graph $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ which is a transitive closure, and for that graph, output the collection of subsets of vertices $\mathcal{A} \equiv \{S \subseteq \mathcal{V} : \text{if } i \in S \text{ and } (i, j) \in \mathcal{E}, \text{ then } j \in S\}$. Algorithm 2 concludes by removing the subsets of vertices from \mathcal{A} which do not include the no-purchase option 0, and then outputs the remaining subsets of vertices from \mathcal{A} . The correctness of Algorithm 2 follows immediately from our earlier observation that an assortment satisfies $S \in \hat{\mathcal{S}}$ if and only if $[0 \in S]$ and $[i \in S \text{ whenever there exists a product } i^* \in S \text{ and a directed edge } (i^*, i) \in \mathcal{E}]$.

At a high level, the recursive subroutine in Algorithm 3 is comprised of two cases. The base case of the subroutine is when the graph has no vertices, in which case it is clear that $\mathcal{A} = \{\emptyset\}$. If we are not in the base case, then the aim of the recursive subroutine is to construct the collections $\{S \in \mathcal{A} : i \notin S\}$ and $\{S \in \mathcal{A} : i \in S\}$ for a chosen vertex $i \in \mathcal{V}$ and output the union of these two collections. The construction of the collection $\{S \in \mathcal{A} : i \notin S\}$ takes place on lines (2b)-(2c) of Algorithm 3, and the construction of the collection $\{S \in \mathcal{A} : i \in S\}$ takes place on lines (2d)-(2f) of Algorithm 3.

Up to this point, we have established that Algorithm 1 is correct (that is, it delivers the desired output for any valid input), and we have established that Algorithm 2 is correct under the assumption that Algorithm 3

CONSTRUCT- $\mathcal{G}(\mathcal{M}, r)$ **Inputs:**

- The collection of past assortments, $\mathcal{M} \equiv \{S_1, \dots, S_M\}$.
- The revenues of the products, $r \equiv (r_0, r_1, \dots, r_n)$.

Output:

- $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} \equiv \mathcal{N}_0$ and $\mathcal{E} \equiv \{(i^*, i) \in \mathcal{N}_0 \times \mathcal{N}_0 : r_{i^*} < r_i \text{ and } \mathcal{M}_{i^*} \subseteq \mathcal{M}_i\}$.

Procedure:

1. Initialize the vertex set $\mathcal{V} \leftarrow \emptyset$ and edge set $\mathcal{E} \leftarrow \emptyset$.
2. For each product $i \in \mathcal{N}_0$:
 - (a) Update $\mathcal{V} \leftarrow \mathcal{V} \cup \{i\}$.
 - (b) Construct the set \mathcal{M}_i of past assortments which offered product i .
3. For each pair of products $(i^*, i) \in \mathcal{N}_0 \times \mathcal{N}_0$:
 - (a) If $r_{i^*} < r_i$ and $\mathcal{M}_{i^*} \subseteq \mathcal{M}_i$:
 - i. Update $\mathcal{E} \leftarrow \mathcal{E} \cup \{(i^*, i)\}$
4. Output $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ and terminate.

Algorithm 1: A procedure for constructing the directed acyclic graph \mathcal{G} .

CONSTRUCT- $\widehat{\mathcal{S}}(\mathcal{M}, r)$ **Inputs:**

- The collection of past assortments, $\mathcal{M} \equiv \{S_1, \dots, S_M\}$.
- The revenues of the products, $r \equiv (r_0, r_1, \dots, r_n)$.

Output:

- The collection of assortments $\widehat{\mathcal{S}}$ corresponding to the collection of past assortments \mathcal{M} and the revenues r .

Procedure:

1. Construct the directed acyclic graph $\mathcal{G} \leftarrow \text{CONSTRUCT-}\mathcal{G}(\mathcal{M}, r)$.
2. Compute the collection of assortments $\widehat{\mathcal{S}} \leftarrow \text{RECURSIVESTEP}(\mathcal{G})$.
3. For each $S \in \widehat{\mathcal{S}}$:
 - (a) If $0 \notin S$:
 - i. Update $\widehat{\mathcal{S}} \leftarrow \widehat{\mathcal{S}} \setminus \{S\}$.
4. Output $\widehat{\mathcal{S}}$ and terminate.

Algorithm 2: A procedure for constructing $\widehat{\mathcal{S}}$.

is correct. Therefore, it remains for us to prove that the recursive subroutine in Algorithm 3 is correct. To prove the correctness of the recursive subroutine, we will make use of four intermediary claims, which are denoted below by Claims EC.2-EC.5. The purpose of the first two intermediary claims, denoted by Claims EC.2 and EC.3, is to show that the graphs $\mathcal{G}' \equiv (\mathcal{V}', \mathcal{E}')$ and $\mathcal{G}'' \equiv (\mathcal{V}'', \mathcal{E}'')$ constructed on lines (2b) and (2d) of Algorithm 3 are directed acyclic graphs that are transitive closures, which implies that \mathcal{G}' and \mathcal{G}'' are valid inputs on lines (2c) and (2e) of Algorithm 3. The purpose of the second two intermediary claims, denoted by Claims EC.4 and EC.5, is to show that the union of the two collections \mathcal{A}' and \mathcal{A}'' constructed on lines (2c) and (2f) of Algorithm 3 provides the desired output on line (2g) of Algorithm 3.

RECURSIVESTEP(\mathcal{G})**Inputs:**

- A directed acyclic graph $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ which is a transitive closure.

Output:

- The collection $\mathcal{A} \equiv \{S \subseteq \mathcal{V} : \text{if } i \in S \text{ and } (i, j) \in \mathcal{E}, \text{ then } j \in S\}$.

Procedure:

1. If $\mathcal{V} = \emptyset$:
 - (a) Output the collection $\mathcal{A} \equiv \{\emptyset\}$ and terminate.
2. Otherwise:
 - (a) Choose any vertex $i \in \mathcal{V}$.
 - (b) Create a copy of $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ in which the vertices $\{i\} \cup \{\ell : (\ell, i) \in \mathcal{E}\}$ and the incoming and outgoing edges of these vertices are removed. Denote this new graph by $\mathcal{G}' \equiv (\mathcal{V}', \mathcal{E}')$.
 - (c) Compute the collection $\mathcal{A}' \leftarrow \text{RECURSIVESTEP}(\mathcal{G}')$.
 - (d) Create a copy of $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ in which the vertices $\{i\} \cup \{\ell : (i, \ell) \in \mathcal{E}\}$ and the incoming and outgoing edges of these vertices are removed. Denote this new graph by $\mathcal{G}'' \equiv (\mathcal{V}'', \mathcal{E}'')$.
 - (e) Compute the collection $\mathcal{A}'' \leftarrow \text{RECURSIVESTEP}(\mathcal{G}'')$.
 - (f) Compute the collection $\mathcal{A}''' \leftarrow \{S \cup \{i\} \cup \{\ell : (i, \ell) \in \mathcal{E}\} : S \in \mathcal{A}''\}$.
 - (g) Output the collection $\mathcal{A} \equiv \mathcal{A}' \cup \mathcal{A}'''$ and terminate.

Algorithm 3: A recursive subroutine which is used in Algorithm 2.

CLAIM EC.2. Let $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ be a directed acyclic graph that is a transitive closure. For notational convenience, let \mathcal{E}_i denote the set of incoming and outgoing edges from each vertex $i \in \mathcal{V}$. Then for each vertex $i \in \mathcal{V}$, we have that $\tilde{\mathcal{G}} \equiv (\mathcal{V} \setminus \{i\}, \mathcal{E} \setminus \mathcal{E}_i)$ is a directed acyclic graph that is a transitive closure.

Proof of Claim EC.2. Let $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ be a directed acyclic graph that is a transitive closure, and let $i \in \mathcal{V}$ be any chosen vertex. It is clear that removing a vertex and its associated incoming and outgoing edges from a directed acyclic graph will not induce any cycles. Therefore, it follows from the fact that the original graph $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ is a directed acyclic graph that the new graph $\tilde{\mathcal{G}} \equiv (\mathcal{V} \setminus \{i\}, \mathcal{E} \setminus \mathcal{E}_i)$ is also a directed acyclic graph. Moreover, consider any two arbitrary vertices $j, k \in \mathcal{V} \setminus \{i\}$ which satisfy the property that vertex k is reachable from vertex j in the new graph $\tilde{\mathcal{G}} \equiv (\mathcal{V} \setminus \{i\}, \mathcal{E} \setminus \mathcal{E}_i)$. Then it follows immediately from the construction of the new graph that vertex k is reachable from vertex j in the original graph $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$. Since the original graph $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ is a transitive closure, there must exist a directed edge from vertex j to vertex k in the original graph. Since neither j nor k are equal to i , we have thus shown that there is a directed edge $(j, k) \in \mathcal{E} \setminus \mathcal{E}_i$. Since the two vertices $j, k \in \mathcal{V} \setminus \{i\}$ were chosen arbitrarily, we have shown that the new graph is also a transitive closure. Our proof of Claim EC.2 is thus complete. \square

CLAIM EC.3. Let $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ be a directed acyclic graph that is a transitive closure. For notational convenience, let \mathcal{E}_i denote the set of incoming and outgoing edges from each vertex $i \in \mathcal{V}$. Then for each subset of vertices $\mathcal{B} \subseteq \mathcal{V}$, we have that $\tilde{\mathcal{G}} \equiv (\mathcal{V} \setminus \mathcal{B}, \mathcal{E} \setminus (\cup_{i \in \mathcal{B}} \mathcal{E}_i))$ is a directed acyclic graph that is a transitive closure.

Proof of Claim EC.3. Let $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ be a directed acyclic graph that is a transitive closure, let $\mathcal{B} \subseteq \mathcal{V}$ be a subset of vertices, and define the directed graph $\tilde{\mathcal{G}} \equiv (\mathcal{V} \setminus \mathcal{B}, \mathcal{E} \setminus (\cup_{i \in \mathcal{B}} \mathcal{E}_i))$. The rest of the proof follows a straightforward induction argument. Indeed, let the vertices of \mathcal{B} be indexed by $\mathcal{B} \equiv \{i_1^{\mathcal{B}}, \dots, i_{|\mathcal{B}|}^{\mathcal{B}}\}$. For each $j \in \{0, 1, \dots, |\mathcal{B}|\}$, we define the following directed graph:

$$\mathcal{G}_j^{\mathcal{B}} \equiv (\underbrace{\mathcal{V} \setminus \{i_1^{\mathcal{B}}, \dots, i_j^{\mathcal{B}}\}}_{\mathcal{V}_j^{\mathcal{B}}}, \underbrace{\mathcal{E} \setminus \{(k, \ell) \in \mathcal{E} : k \in \{i_1^{\mathcal{B}}, \dots, i_j^{\mathcal{B}}\} \text{ or } \ell \in \{i_1^{\mathcal{B}}, \dots, i_j^{\mathcal{B}}\}\}}_{\mathcal{E}_j^{\mathcal{B}}}).$$

We readily observe from the above definition that $\mathcal{G}_0^{\mathcal{B}} = \mathcal{G}$ and that $\mathcal{G}_{|\mathcal{B}|}^{\mathcal{B}} = \tilde{\mathcal{G}}$. In the remainder, we will prove by induction that each $\mathcal{G}_0^{\mathcal{B}}, \dots, \mathcal{G}_{|\mathcal{B}|}^{\mathcal{B}}$ is a directed acyclic graph that is a transitive closure. Our induction

proof proceeds as follows. In the base case, it follows from the equality $\mathcal{G}_0^{\mathcal{B}} = \mathcal{G}$ that $\mathcal{G}_0^{\mathcal{B}}$ is a directed acyclic graph that is a transitive closure. Next, assume by induction that $\mathcal{G}_0^{\mathcal{B}}, \dots, \mathcal{G}_{j-1}^{\mathcal{B}}$ are directed acyclic graphs that are transitive closures for any $j \in \{1, \dots, |\mathcal{B}|\}$. Then we readily observe from the definition of the directed graph $\mathcal{G}_j^{\mathcal{B}} \equiv (\mathcal{V}_j^{\mathcal{B}}, \mathcal{E}_j^{\mathcal{B}})$ that the following equalities hold:

$$\mathcal{V}_j^{\mathcal{B}} = \mathcal{V}_{j-1}^{\mathcal{B}} \setminus \{i_j^{\mathcal{B}}\}; \quad \mathcal{E}_j^{\mathcal{B}} = \mathcal{E}_{j-1}^{\mathcal{B}} \setminus \{(k, \ell) \in \mathcal{E}_{j-1}^{\mathcal{B}} : k = i_j^{\mathcal{B}} \text{ or } \ell = i_j^{\mathcal{B}}\}.$$

Using the above equalities, it follows immediately from Claim EC.2 and the induction hypothesis that $\mathcal{G}_j^{\mathcal{B}}$ is a directed acyclic graph that is a transitive closure. This concludes our induction proof, and since the equality $\mathcal{G}_{|\mathcal{B}|}^{\mathcal{B}} = \tilde{\mathcal{G}}$ holds, our proof of Claim EC.3 is complete. \square

CLAIM EC.4. *Let $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ be a directed acyclic graph that is a transitive closure. For notational convenience, let \mathcal{E}_i denote the set of incoming and outgoing edges from each vertex $i \in \mathcal{V}$. For each vertex $i \in \mathcal{V}$,*

$$\begin{aligned} & \{S \subseteq \mathcal{V} \setminus \{i\} : \text{if } k \in S \text{ and } (k, j) \in \mathcal{E}, \text{ then } j \in S\} \\ &= \left\{ S \subseteq \mathcal{V} \setminus (\{i\} \cup \{\ell : (\ell, i) \in \mathcal{E}\}) : \text{if } k \in S \text{ and } (k, j) \in \mathcal{E} \setminus \left(\mathcal{E}_i \cup \bigcup_{\ell: (\ell, i) \in \mathcal{E}} \mathcal{E}_\ell \right), \text{ then } j \in S \right\}. \end{aligned} \quad (\text{EC.5})$$

Proof of Claim EC.4. Let $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ be a directed acyclic graph that is a transitive closure, and let $i \in \mathcal{V}$ be any chosen vertex from this graph. We first observe that

$$\{S \subseteq \mathcal{V} \setminus \{i\} : \text{if } k \in S \text{ and } (k, j) \in \mathcal{E}, \text{ then } j \in S\} \quad (\text{EC.6})$$

$$= \{S \subseteq \mathcal{V} \setminus (\{i\} \cup \{\ell : (\ell, i) \in \mathcal{E}\}) : \text{if } k \in S \text{ and } (k, j) \in \mathcal{E}, \text{ then } j \in S\}, \quad (\text{EC.7})$$

where the above equality follows from the fact that any subset of vertices S from the collection on line (EC.6) must not contain any vertices in the graph that have an outgoing edge to vertex i . For notational convenience, let the collection of subsets of vertices on line (EC.5) be denoted by $\tilde{\mathcal{A}}$, and let the collection of subsets of vertices on line (EC.7) be denoted by $\tilde{\mathcal{A}}'$. In other words, let

$$\begin{aligned} \tilde{\mathcal{A}}' &\triangleq \{S \subseteq \mathcal{V} \setminus (\{i\} \cup \{\ell : (\ell, i) \in \mathcal{E}\}) : \text{if } k \in S \text{ and } (k, j) \in \mathcal{E}, \text{ then } j \in S\} \\ \tilde{\mathcal{A}} &\triangleq \left\{ S \subseteq \mathcal{V} \setminus (\{i\} \cup \{\ell : (\ell, i) \in \mathcal{E}\}) : \text{if } k \in S \text{ and } (k, j) \in \mathcal{E} \setminus \left(\mathcal{E}_i \cup \bigcup_{\ell: (\ell, i) \in \mathcal{E}} \mathcal{E}_\ell \right), \text{ then } j \in S \right\}. \end{aligned}$$

In the remainder of the proof, we will show that $\tilde{\mathcal{A}} = \tilde{\mathcal{A}}'$.

To show the first direction, consider any arbitrary subset of vertices $S \in \tilde{\mathcal{A}}'$. For this subset of vertices, consider any two arbitrary vertices $k, j \in \mathcal{V} \setminus (\{i\} \cup \{\ell : (\ell, i) \in \mathcal{E}\})$ which satisfy the conditions $[k \in S]$ and $[(k, j) \in \mathcal{E} \setminus (\mathcal{E}_i \cup \bigcup_{\ell: (\ell, i) \in \mathcal{E}} \mathcal{E}_\ell)]$. In this case, it follows immediately from the fact that $\mathcal{E} \setminus (\mathcal{E}_i \cup \bigcup_{\ell: (\ell, i) \in \mathcal{E}} \mathcal{E}_\ell) \subseteq \mathcal{E}$ that the condition $[(k, j) \in \mathcal{E}]$ holds. Therefore, it follows immediately from the definition of $\tilde{\mathcal{A}}'$ that $j \in S$. Since the two vertices $k, j \in \mathcal{V} \setminus (\{i\} \cup \{\ell : (\ell, i) \in \mathcal{E}\})$ which satisfy the conditions $[k \in S]$ and $[(k, j) \in \mathcal{E} \setminus (\mathcal{E}_i \cup \bigcup_{\ell: (\ell, i) \in \mathcal{E}} \mathcal{E}_\ell)]$ were chosen arbitrarily, we have shown that $S \in \tilde{\mathcal{A}}$, and since the subset of vertices $S \in \tilde{\mathcal{A}}'$ was chosen arbitrarily, we have shown that $\tilde{\mathcal{A}}' \subseteq \tilde{\mathcal{A}}$. Our proof of the first direction is thus complete.

To show the other direction, consider any arbitrary subset of vertices $S \in \tilde{\mathcal{A}}$. For this subset of vertices, consider any two arbitrary vertices $k, j \in \mathcal{V} \setminus (\{i\} \cup \{\ell : (\ell, i) \in \mathcal{E}\})$ which satisfy the conditions $[k \in S]$ and $[(k, j) \in \mathcal{E}]$. We now suppose, for the sake of developing a contradiction, that the directed edge from vertex k to vertex j satisfies $(k, j) \in \mathcal{E}_i \cup \bigcup_{\ell: (\ell, i) \in \mathcal{E}} \mathcal{E}_\ell$. Under this supposition, we have two cases to consider. In the first case, the directed edge satisfies $(k, j) \in \mathcal{E}_i$. However, that would imply that either $j = i$ or $k = i$, and those equalities contradict the facts that $k \in S$ and $S \subseteq \mathcal{V} \setminus (\{i\} \cup \{\ell : (\ell, i) \in \mathcal{E}\})$. In the second case, the directed edge satisfies $(k, j) \in \bigcup_{\ell: (\ell, i) \in \mathcal{E}} \mathcal{E}_\ell$, which implies that there are directed edges $(k, j) \in \mathcal{E}$ and $(j, i) \in \mathcal{E}$ in the graph \mathcal{G} . Since the graph is a transitive closure, the existence of directed edges $(k, j) \in \mathcal{E}$ and $(j, i) \in \mathcal{E}$ implies that there must exist a directed edge $(k, i) \in \mathcal{E}$. However, the existence of a directed edge $(k, i) \in \mathcal{E}$ contradicts the facts that $k \in S$ and $S \subseteq \mathcal{V} \setminus (\{i\} \cup \{\ell : (\ell, i) \in \mathcal{E}\})$. In all cases, we have proved by contradiction that the vertices k and j must satisfy the condition $[(k, j) \in \mathcal{E} \setminus (\mathcal{E}_i \cup \bigcup_{\ell: (\ell, i) \in \mathcal{E}} \mathcal{E}_\ell)]$. Since the two vertices $k, j \in \mathcal{V} \setminus (\{i\} \cup \{\ell : (\ell, i) \in \mathcal{E}\})$ which satisfy the conditions $[k \in S]$ and $[(k, j) \in \mathcal{E}]$ were chosen arbitrarily, we have shown that $S \in \tilde{\mathcal{A}}'$, and since the subset of vertices $S \in \tilde{\mathcal{A}}$ was chosen arbitrarily, we have shown that $\tilde{\mathcal{A}} \subseteq \tilde{\mathcal{A}}'$. Our proof of the second direction is thus complete.

We have thus shown that $\tilde{\mathcal{A}} = \tilde{\mathcal{A}}'$, which concludes our proof of Claim EC.4. \square

CLAIM EC.5. Let $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ be a directed acyclic graph that is a transitive closure. For notational convenience, let \mathcal{E}_i denote the set of incoming and outgoing edges from each vertex $i \in \mathcal{V}$. For each vertex $i \in \mathcal{V}$,

$$\begin{aligned} & \{S \subseteq \mathcal{V} : [i \in S] \text{ and } [\text{if } k \in S \text{ and } (k, j) \in \mathcal{E}, \text{ then } j \in S]\} \\ &= \left\{ \begin{array}{l} S' \cup \{i\} \cup \{\ell : (i, \ell) \in \mathcal{E}\} : \\ S' \in \left\{ S \subseteq \mathcal{V} \setminus (\{i\} \cup \{\ell : (i, \ell) \in \mathcal{E}\}) : \text{if } k \in S \text{ and } (k, j) \in \mathcal{E} \setminus \left(\mathcal{E}_i \cup \bigcup_{\ell: (i, \ell) \in \mathcal{E}} \mathcal{E}_\ell \right), \text{ then } j \in S \right\} \end{array} \right\}. \end{aligned} \quad (\text{EC.8})$$

Proof of Claim EC.5. Let $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ be a directed acyclic graph that is a transitive closure, and let $i \in \mathcal{V}$ be any chosen vertex from this graph. Then we observe that

$$\{S \subseteq \mathcal{V} : [i \in S] \text{ and } [\text{if } k \in S \text{ and } (k, j) \in \mathcal{E}, \text{ then } j \in S]\} \quad (\text{EC.9})$$

$$= \{S \subseteq \mathcal{V} : [\{i\} \cup \{\ell : (i, \ell) \in \mathcal{E}\} \subseteq S] \text{ and } [\text{if } k \in S \text{ and } (k, j) \in \mathcal{E}, \text{ then } j \in S]\} \quad (\text{EC.10})$$

$$= \left\{ S \subseteq \mathcal{V} : [\{i\} \cup \{\ell : (i, \ell) \in \mathcal{E}\} \subseteq S] \text{ and } \left[\text{if } k \in S \text{ and } (k, j) \in \mathcal{E} \setminus \left(\mathcal{E}_i \cup \bigcup_{\ell: (i, \ell) \in \mathcal{E}} \mathcal{E}_\ell \right), \text{ then } j \in S \right] \right\}. \quad (\text{EC.11})$$

Indeed, the equality of the collections of subsets of vertices on lines (EC.9) and (EC.10) follows from the fact that any subset of vertices S in the collection on line (EC.9) satisfies $i \in S$, and so it must contain all of the vertices in the graph which have an incoming edge from vertex i . For notational convenience, let the collections of subsets of vertices on lines (EC.10) and (EC.11) be denoted by \mathcal{A}' and \mathcal{A} , respectively. To show that line (EC.10) is equal to line (EC.11), we first observe that the inclusion $\mathcal{A}' \subseteq \mathcal{A}$ follows immediately from the definitions of the collections \mathcal{A}' and \mathcal{A} . To show the other direction, consider any subset of vertices $S \in \mathcal{A}$ and consider any vertex $k \in S$. If there exists a directed edge $(k, j) \in \mathcal{E}$ which satisfies $(k, j) \in \mathcal{E}_i \cup \bigcup_{\ell: (i, \ell) \in \mathcal{E}} \mathcal{E}_\ell$, then we observe that the vertex j must satisfy $j \in \{i\} \cup \{\ell : (i, \ell) \in \mathcal{E}\}$. Since it follows from $S \in \mathcal{A}$ that the inclusion $\{i\} \cup \{\ell : (i, \ell) \in \mathcal{E}\} \subseteq S$ holds, we conclude that $S \in \mathcal{A}'$. We have thus shown that line (EC.10) is equal to line (EC.11). Since the equivalence of lines (EC.11) and (EC.8) follows readily from algebra, our proof of Claim EC.5 is complete. \square

Using the above intermediary claims, we now prove the correctness of the recursive subroutine in Algorithm 3. Indeed, it is clear that the recursive subroutine yields the correct output in the base case where $\mathcal{V} = \emptyset$. Next, let us assume by induction that the recursive subroutine yields the correct output for all valid input graphs with up to $p - 1$ vertices, and consider any valid input $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ for which the number of vertices is $|\mathcal{V}| = p$. Let $i \in \mathcal{V}$ denote the vertex from this graph which is chosen in line (2a) of Algorithm 3, where the existence of such a vertex follows from the fact that we are in the case where $\mathcal{V} \neq \emptyset$. It follows immediately from Claim EC.3 that the graphs $\mathcal{G}' \equiv (\mathcal{V}', \mathcal{E}')$ and $\mathcal{G}'' \equiv (\mathcal{V}'', \mathcal{E}'')$ constructed on lines (2b) and (2d) of Algorithm 3 are directed acyclic graphs which are transitive closures, which implies that these graphs are valid inputs to Algorithm 3 in lines (2c) and (2e). Therefore, it follows from the induction hypothesis and lines (2c), (2e), and (2f) of Algorithm 3 that

$$\begin{aligned} \mathcal{A}' &= \left\{ S \subseteq \mathcal{V} \setminus (\{i\} \cup \{\ell : (\ell, i) \in \mathcal{E}\}) : \text{if } k \in S \text{ and } (k, j) \in \mathcal{E} \setminus \left(\mathcal{E}_i \cup \bigcup_{\ell: (\ell, i) \in \mathcal{E}} \mathcal{E}_\ell \right), \text{ then } j \in S \right\} \\ \mathcal{A}'' &= \left\{ S \subseteq \mathcal{V} \setminus (\{i\} \cup \{\ell : (i, \ell) \in \mathcal{E}\}) : \text{if } k \in S \text{ and } (k, j) \in \mathcal{E} \setminus \left(\mathcal{E}_i \cup \bigcup_{\ell: (i, \ell) \in \mathcal{E}} \mathcal{E}_\ell \right), \text{ then } j \in S \right\} \\ \mathcal{A}''' &= \{S \cup \{i\} \cup \{\ell : (i, \ell) \in \mathcal{E}\} : S \in \mathcal{A}''\}, \end{aligned}$$

where the induction hypothesis can be applied because $\mathcal{G}' \equiv (\mathcal{V}', \mathcal{E}')$ and $\mathcal{G}'' \equiv (\mathcal{V}'', \mathcal{E}'')$ are valid inputs to Algorithm 3 and because $|\mathcal{V}'| \leq p - 1$ and $|\mathcal{V}''| \leq p - 1$. Therefore, it follows from Claims EC.4 and EC.5 that

$$\begin{aligned} \mathcal{A}' &= \{S \subseteq \mathcal{V} \setminus \{i\} : \text{if } k \in S \text{ and } (k, j) \in \mathcal{E}, \text{ then } j \in S\} \\ \mathcal{A}''' &= \{S \subseteq \mathcal{V} : [i \in S] \text{ and } [\text{if } k \in S \text{ and } (k, j) \in \mathcal{E}, \text{ then } j \in S]\}, \end{aligned}$$

which proves that the output of Algorithm 3 is

$$\mathcal{A}' \cup \mathcal{A}''' = \{S \subseteq \mathcal{V} : \text{if } k \in S \text{ and } (k, j) \in \mathcal{E}, \text{ then } j \in S\}.$$

This completes our proof of the correctness of Algorithm 3.

To conclude our proof of Lemma 4, we analyze the computation time of Algorithm 2. Indeed, we recall that the computation time required for line (1) in Algorithm 2 is $\mathcal{O}(n^2M)$. In what follows, we assume that all directed graphs are stored as adjacency lists. Under this assumption, our analysis of the computation time for line (2) in Algorithm 3 is split into the following two intermediary claims, denoted by Claim EC.6 and EC.7. In our first intermediary claim, presented below as Claim EC.6, we establish the computation time required for lines (2b), (2d), (2f), and (2g) of Algorithm 3.

CLAIM EC.6. *If $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ is a directed acyclic graph that is a transitive closure with $|\mathcal{V}| \geq 1$, then lines (2b), (2d), (2f), and (2g) of Algorithm 3 can be performed in $\mathcal{O}(|\mathcal{V}| \times |\text{RECURSIVESTEP}(\mathcal{G})|)$ time.*

Proof. We observe that the set of vertices $\{i\} \cup \{\ell : (\ell, i) \in \mathcal{E}\}$ in line (2b) of Algorithm 3 can be queried and stored as a hash table in $\mathcal{O}(|\{i\} \cup \{\ell : (\ell, i) \in \mathcal{E}\}|) = \mathcal{O}(|\mathcal{V}|)$ computation time. Therefore, the directed graph $\mathcal{G}' \equiv (\mathcal{V}', \mathcal{E}')$ in line (2b) of Algorithm 3 can be constructed from scratch in $\mathcal{O}(|\mathcal{V}'| + |\mathcal{E}'|)$ computation time. By identical reasoning, we observe that the directed graph $\mathcal{G}'' \equiv (\mathcal{V}'', \mathcal{E}'')$ in line (2d) of Algorithm 3 can be constructed in $\mathcal{O}(|\mathcal{V}''| + |\mathcal{E}''|)$ computation time. It is easy to see for each $S \in \mathcal{A}''$ that $S \cup \{i\} \cup \{\ell : (i, \ell) \in \mathcal{E}\}$ is the union of three disjoint sets, which implies that the collection on line (2f) of Algorithm 3 can be constructed in a total of $\mathcal{O}(\sum_{S \in \mathcal{A}''} (|S| + |\{\ell : (i, \ell) \in \mathcal{E}\}|)) = \mathcal{O}(\sum_{S \in \mathcal{A}''} |S|)$ computation time. It is similarly easy to see that $\mathcal{A} \equiv \mathcal{A}' \cup \mathcal{A}'''$ is the union of two disjoint collections, which implies that line (2g) of Algorithm 3 can be performed in $\mathcal{O}(\sum_{S \in \mathcal{A}'} |S| + \sum_{S \in \mathcal{A}'''} |S|) = \mathcal{O}(\sum_{S \in \mathcal{A}} |S|)$ computation time. All combined, we have shown that the computation time required for lines (2b), (2d), (2f), and (2g) of Algorithm 3 is

$$\underbrace{\mathcal{O}(|\mathcal{V}| + |\mathcal{E}|)}_{(2b)} + \underbrace{\mathcal{O}(|\mathcal{V}| + |\mathcal{E}|)}_{(2d)} + \underbrace{\mathcal{O}\left(\sum_{S \in \mathcal{A}''} |S|\right)}_{(2f)} + \underbrace{\mathcal{O}\left(\sum_{S \in \mathcal{A}} |S|\right)}_{(2g)} = \mathcal{O}\left(|\mathcal{V}| + |\mathcal{E}| + \sum_{S \in \mathcal{A}} |S|\right).$$

Using the fact that graphs always satisfy the inequality $|\mathcal{E}| \leq |\mathcal{V}|^2$ and the fact that $|S| \leq |\mathcal{V}|$ for all $S \in \mathcal{A}$, the above computation time simplifies to $\mathcal{O}(|\mathcal{V}|^2 + |\mathcal{V}| \times |\mathcal{A}|)$. Since $\mathcal{A} \equiv \text{RECURSIVESTEP}(\mathcal{G})$ is the output of Algorithm 3, and since it is easy to see that $|\mathcal{A}| \geq |\mathcal{V}|$, our proof of Claim EC.6 is complete. \square

In our second intermediary claim, presented below as Claim EC.7, we use Claim EC.6 to establish the computation time for Algorithm 3 for any valid input.

CLAIM EC.7. *If $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ is a directed acyclic graph that is a transitive closure with $|\mathcal{V}| \geq 1$, then the computation time for Algorithm 3 is $\mathcal{O}(|\mathcal{V}|^2 \times |\text{RECURSIVESTEP}(\mathcal{G})|)$.*

Proof. Let $T(\mathcal{G})$ denote the computation time required to run Algorithm 3 for any valid input $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$. It follows from Claim EC.6 that $T(\mathcal{G})$ can be represented by an asymptotic recurrence of the form

$$T(\mathcal{G}) = \begin{cases} |\mathcal{V}| \times |\text{RECURSIVESTEP}(\mathcal{G})| + T(\mathcal{G}') + T(\mathcal{G}''), & \text{if } \mathcal{V} \neq \emptyset, \\ 1, & \text{if } \mathcal{V} = \emptyset, \end{cases} \quad (\text{EC.12})$$

where \mathcal{G}' and \mathcal{G}'' are the subgraphs constructed in Algorithm 3 on lines (2b) and (2d). We will now prove that the above recurrence satisfies the following inequality:

$$T(\mathcal{G}) \leq (|\mathcal{V}| + 1)^2 \times |\text{RECURSIVESTEP}(\mathcal{G})|.$$

Indeed, the above inequality clearly holds in the base case. Now assume by induction that the above inequality holds for all valid inputs graphs to Algorithm 3 with up to $p - 1$ vertices, and consider any valid input $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ to Algorithm 3 for which the number of vertices is $|\mathcal{V}| = p$. For this input, we have

$$\begin{aligned} T(\mathcal{G}) &= |\mathcal{V}| \times |\text{RECURSIVESTEP}(\mathcal{G})| + T(\mathcal{G}') + T(\mathcal{G}'') \\ &\leq |\mathcal{V}| \times |\text{RECURSIVESTEP}(\mathcal{G})| + (|\mathcal{V}'| + 1)^2 \times |\text{RECURSIVESTEP}(\mathcal{G}')| + (|\mathcal{V}''| + 1)^2 \times |\text{RECURSIVESTEP}(\mathcal{G}'')| \\ &\leq |\mathcal{V}| \times |\text{RECURSIVESTEP}(\mathcal{G})| + |\mathcal{V}|^2 \times |\text{RECURSIVESTEP}(\mathcal{G}')| + |\mathcal{V}|^2 \times |\text{RECURSIVESTEP}(\mathcal{G}'')| \\ &= |\mathcal{V}| \times |\text{RECURSIVESTEP}(\mathcal{G})| + |\mathcal{V}|^2 \times |\text{RECURSIVESTEP}(\mathcal{G})| \\ &\leq (|\mathcal{V}| + 1)^2 \times |\text{RECURSIVESTEP}(\mathcal{G})|, \end{aligned}$$

where the first line follows from (EC.12) and from the fact that $\mathcal{V} \neq \emptyset$, the second line follows from the induction hypothesis, the third line follows from the facts that $|\mathcal{V}'| < |\mathcal{V}|$ and $|\mathcal{V}''| < |\mathcal{V}|$, the fourth line follows from the fact that $|\text{RECURSIVESTEP}(\mathcal{G})| = |\text{RECURSIVESTEP}(\mathcal{G}')| + |\text{RECURSIVESTEP}(\mathcal{G}'')|$, and the final line follows from algebra. This concludes our proof of Claim EC.7. \square

Our analysis of the computation time for line (2) of Algorithm 2 follows readily from Claim EC.7. Indeed, we observe that the graph $\mathcal{G} \equiv (\mathcal{V}, \mathcal{E})$ that was constructed on line (1) of Algorithm 2 satisfies $|\mathcal{V}| = n + 1$ and $|\mathcal{E}| \leq (n + 1)^2$. Therefore, it follows from Claim EC.7 that line (2) of Algorithm 2 requires $\mathcal{O}((n + 1)^2 \times |\text{RECURSIVESTEP}(\mathcal{G})|) = \mathcal{O}(n^2 |\widehat{\mathcal{S}}|)$ computation time. Since the loop on line (3) of Algorithm 2 can be performed in $\mathcal{O}(\sum_{S \in \widehat{\mathcal{S}}} |S|) = \mathcal{O}(n |\widehat{\mathcal{S}}|)$ computation time, we have shown that the total computation time for Algorithm 2 is

$$\underbrace{\mathcal{O}(n^2 M)}_{(1)} + \underbrace{\mathcal{O}(n^2 |\widehat{\mathcal{S}}|)}_{(2)} + \underbrace{\mathcal{O}(n |\widehat{\mathcal{S}}|)}_{(3)} = \mathcal{O}\left(n^2(M + |\widehat{\mathcal{S}}|)\right).$$

This concludes our proof of Lemma 4. \square

Appendix E: Proof of Technical Results from §4.3

There are two technical results in §4.3: Lemma 5 and Theorem 2. Lemma 5 follows immediately from the definition of the collection of assortments $\widehat{\mathcal{S}}$ from §4, and so the proof of Lemma 5 is omitted. The remainder of this appendix thus contains the proof of Theorem 2, which is split into several steps. We will begin by choosing any arbitrary $\widehat{S} \in \widehat{\mathcal{S}}$ and, for that assortment, we will define a particular realization of the historical data $v \equiv (v_{m,i} : m \in \mathcal{M}, i \in S_m)$. We will then prove in Lemma EC.5 in Appendix E.2 for that realization of the historical data that the corresponding set of feasible solutions for the linear optimization problem (2) is a singleton. By showing that the realization of the historical data generates a unique feasible solution for the linear optimization problem (2), we conclude in Appendix E.3 by showing that \widehat{S} is the unique optimal solution of the robust optimization problem (RO).

E.1. Assumptions and Notation

Following the statement of Theorem 2, we will assume throughout Appendix E that $\mathcal{M} = \widehat{\mathcal{S}}$, that $\eta = 0$, and that the revenues $r_1 < \dots < r_n$ are fixed. We observe that Theorem 2 trivially holds in the case where $n = 1$; therefore, we assume throughout the rest of Appendix E that the number of products satisfies $n \geq 2$.¹⁷

We will make use of the following notation and preliminary results throughout Appendix E. First, it follows from the fact that $\mathcal{M} = \widehat{\mathcal{S}}$ and from Definition 3 that the set of tuples \mathcal{L} can be represented compactly as¹⁸

$$\mathcal{L} = \{(i_1, \dots, i_n) : i_1 \in \{0, n\} \text{ and } i_{j+1} \in \{j, i_j\} \text{ for all } j \in \{1, \dots, n-1\}\}. \quad (\text{EC.13})$$

Second, for each $j \in \{1, \dots, n-1\}$, we define the indices

$$\begin{aligned} \underline{i}_1^j &= \dots = \underline{i}_j^j = 0, & \underline{i}_{j+1}^j &= \dots = \underline{i}_n^j = j, \\ \bar{i}_1^j &= \dots = \bar{i}_j^j = n, & \bar{i}_{j+1}^j &= \dots = \bar{i}_n^j = j, \end{aligned}$$

and we denote the tuples constructed from those indices by

$$(\underline{i}_1^j, \dots, \underline{i}_n^j) \triangleq (0, \dots, 0, j, \dots, j) \text{ and } (\bar{i}_1^j, \dots, \bar{i}_n^j) \triangleq (n, \dots, n, j, \dots, j).$$

¹⁷ Suppose the number of products satisfies $n = 1$. In that case, we observe that $\widehat{\mathcal{S}} = \{\{0, 1\}\}$. Since $\widehat{\mathcal{S}}$ is a singleton, Theorem 2 follows immediately from Theorem 1.

¹⁸ A formal proof of the equality on line (EC.13) can be found as Lemma 10 in §5.3.

Finally, it follows from the fact that $\eta = 0$ that the constraints of the linear optimization problem (2) can be written compactly as

$$\begin{aligned} \sum_{(i_1, \dots, i_n) \in \mathcal{L}: i_m = i} \lambda_{i_1 \dots i_n} &= v_{m,i} \quad \forall m \in \mathcal{M}, i \in S_m \\ \sum_{(i_1, \dots, i_n) \in \mathcal{L}} \lambda_{i_1 \dots i_n} &= 1 \\ \lambda_{i_1 \dots i_n} &\geq 0 \quad \forall (i_1, \dots, i_n) \in \mathcal{L}, \end{aligned} \tag{EC.14}$$

where we restate for the sake of convenience that $S_m \equiv \{0, 1, \dots, m-1, n\}$ and $\mathcal{M} \equiv \{1, \dots, n\}$.

E.2. Construction of historical data for any given $\bar{S} \in \widehat{\mathcal{S}}$, and proof that the corresponding system (EC.14) has a unique solution

In view of the strategy outlined at the beginning of Appendix E, we now present the details of the proof of Theorem 2. Indeed, consider any arbitrary assortment $\bar{S} \in \widehat{\mathcal{S}}$. For that assortment, we define the realization of the historical data $v_{m,i}$ for each past assortment $m \in \mathcal{M} \equiv \{1, \dots, n\}$ and each product $i \in S_m$ as follows:

$$v_{m,i} \triangleq \begin{cases} \frac{1}{n} (1 + |\{m, \dots, n-1\} \setminus \bar{S}|), & \text{if } i = n, \\ \frac{1}{n} |\{m, \dots, n-1\} \cap \bar{S}|, & \text{if } i = 0, \\ \frac{1}{n}, & \text{otherwise.} \end{cases} \tag{EC.15}$$

With the realization of the historical data from line (EC.15) corresponding to the assortment \bar{S} , the remainder of Appendix E.2 is dedicated to proving the following lemma:

LEMMA EC.5. *The system (EC.14) is nonempty and has a unique solution λ which satisfies the following equality for each $(i_1, \dots, i_n) \in \mathcal{L}$:*

$$\lambda_{i_1 \dots i_n} = \begin{cases} \frac{1}{n}, & \text{if there exists } j \in \bar{S} \setminus \{0, n\} \text{ such that } (i_1, \dots, i_n) = (\underline{i}_1^j, \dots, \underline{i}_n^j), \\ \frac{1}{n}, & \text{if there exists } j \in \mathcal{N} \setminus \bar{S} \text{ such that } (i_1, \dots, i_n) = (\bar{i}_1^j, \dots, \bar{i}_n^j), \\ \frac{1}{n}, & \text{if } (i_1, \dots, i_n) = (n, \dots, n), \\ 0, & \text{otherwise.} \end{cases} \tag{EC.16}$$

Stated in words, the above Lemma EC.5 establishes that the vector λ defined by line (EC.16) is the unique feasible solution for the linear optimization problem (2) under the assumptions that the past assortments are $\mathcal{M} = \mathcal{S}$, that $\eta = 0$, and that the realization of the historical data satisfies (EC.15).

Let us provide an interpretation of the vector λ defined by line (EC.16). Speaking informally, the vector λ is comprised of three types of customers. The first type of customers are denoted by the tuples $(\underline{i}_1^j, \dots, \underline{i}_n^j)$ for each $j \in \bar{S} \setminus \{0, n\}$, and these are customers that prefer product j and, if that product is unavailable, will not make a purchase. The second type of customers are denoted by the tuples $(\bar{i}_1^j, \dots, \bar{i}_n^j)$ for each $j \in \mathcal{N} \setminus \bar{S}$, and these are customers whose most preferred product is j , whose second-most preferred product is n , and will not purchase a product if neither j nor n are available. The third type of customers are denoted by the tuple (n, \dots, n) , and these are customers that will purchase product n if it is available, and will otherwise not make a purchase. In Appendix E.3, we will use Lemma EC.5 to prove Theorem 2.

Our proof of Lemma EC.5 makes use of two intermediary claims, denoted below as Claims EC.8 and EC.9. These intermediary claims establish that any feasible solution of the system (EC.14) must satisfy certain structural properties. We first present these two intermediary claims, followed by the proof of Lemma EC.5.

CLAIM EC.8. *Suppose that λ is a feasible solution of the system (EC.14), and let $(i_1, \dots, i_n) \in \mathcal{L}$. If $\lambda_{i_1 \dots i_n} > 0$, then either $(i_1, \dots, i_n) = (n, \dots, n)$ or there exists $j \in \{1, \dots, n-1\}$ such that $i_{j+1} = j$.*

Proof of Claim EC.8. Suppose that λ is a solution of the system (EC.14), and let $(i_1, \dots, i_n) \in \mathcal{L}$.

We begin by showing that if $\lambda_{i_1 \dots i_n} > 0$, then either $(i_1, \dots, i_n) = (n, \dots, n)$ or there exists $m \in \{1, \dots, n\}$ such that $i_m \notin \{0, n\}$. We will show this claim by proving its contraposition. Indeed, suppose that $i_m \in \{0, n\}$ for all $m \in \{1, \dots, n\}$ and that there exists $\bar{m} \in \{1, \dots, n\}$ such that $i_{\bar{m}} = 0$. In this case, it follows from

(EC.13) and from the fact that $i_m \in \{0, n\}$ for all $m \in \{1, \dots, n\}$ that $i_1 = \dots = i_n = 0$. If $i_1 = \dots = i_n = 0$, then

$$\lambda_{i_1 \dots i_n} = \sum_{(i'_1, \dots, i'_n) \in \mathcal{L}: i'_n = 0} \lambda_{i'_1 \dots i'_n} = v_{n,0} = 0,$$

where the first equality follows from (EC.13), the second equality follows from the fact that λ is a feasible solution of the system (EC.14), and the third equality follows from (EC.15). We have thus shown that if $i_m \in \{0, n\}$ for all $m \in \{1, \dots, n\}$ and if there exists $m \in \{1, \dots, n\}$ such that $i_m = 0$, then $\lambda_{i_1 \dots i_n} = 0$. Note that the contrapositive of this statement is that if $\lambda_{i_1 \dots i_n} > 0$, then either $(i_1, \dots, i_n) = (n, \dots, n)$ or there exists $m \in \{1, \dots, n\}$ such that $i_m \notin \{0, n\}$.

We conclude the proof of Claim EC.8 by showing that if there exists $m \in \{1, \dots, n\}$ such that $i_m \notin \{0, n\}$, then there exists $j \in \{1, \dots, n-1\}$ such that $i_{j+1} = j$. Indeed, suppose there exists $m \in \{1, \dots, n\}$ such that $i_m \notin \{0, n\}$. Since the inclusion $i_1 \in \{0, n\}$ always holds, the existence of $m \in \{1, \dots, n\}$ satisfying $i_m \notin \{0, n\}$ implies that there must exist $j \in \{1, \dots, n-1\}$ such that $i_j \neq i_{j+1}$. Since we have assumed that the inclusion $(i_1, \dots, i_n) \in \mathcal{L}$ holds, it follows from (EC.13) and from the fact that $i_j \neq i_{j+1}$ that i_{j+1} must be equal to j . Our proof of Claim EC.8 is thus complete. \square

CLAIM EC.9. *Suppose that λ is a feasible solution of the system (EC.14), and let $(i_1, \dots, i_n) \in \mathcal{L}$. If $\lambda_{i_1 \dots i_n} > 0$ and if $(i_1, \dots, i_n) \neq (n, \dots, n)$, then there exists $j \in \{1, \dots, n-1\}$ and $\dagger \in \{0, n\}$ such that $i_1 = \dots = i_j = \dagger$ and $i_{j+1} = \dots = i_n = j$.*

Proof of Claim EC.9. Suppose that λ is a feasible solution of the system (EC.14), and suppose that the inequality $\lambda_{i_1 \dots i_n} > 0$ holds for a given tuple $(i_1, \dots, i_n) \in \mathcal{L}$ with $(i_1, \dots, i_n) \neq (n, \dots, n)$. Then it follows from Claim EC.8 that there exists a product $j' \in \{1, \dots, n-1\}$ such that $i_{j'+1} = j'$. From this point onward, we let the product $j' \in \{1, \dots, n-1\}$ with the smallest index that satisfies $i_{j'+1} = j'$ be denoted by

$$j \triangleq \min \{j' \in \{1, \dots, n-1\} : i_{j'+1} = j'\} = \min \{m \in \{1, \dots, n-1\} : i_{m+1} \notin \{0, n\}\}, \quad (\text{EC.17})$$

where the last equality follows from (EC.13).

For the product j defined in (EC.17), it follows from the fact that λ is a feasible solution of the system (EC.14) that the following two equalities must hold:

$$\sum_{(i'_1, \dots, i'_n) \in \mathcal{L}: i'_{j+1} = j} \lambda_{i'_1 \dots i'_n} = v_{j+1, j}; \quad \sum_{(i'_1, \dots, i'_n) \in \mathcal{L}: i'_n = j} \lambda_{i'_1 \dots i'_n} = v_{n, j}. \quad (\text{EC.18})$$

We observe from (EC.15) and from the inclusion $j \in \{1, \dots, n-1\}$ that the equality $v_{j+1, j} = v_{n, j}$ also holds. Therefore, (EC.18) implies that

$$\sum_{(i'_1, \dots, i'_n) \in \mathcal{L}: i'_{j+1} = j} \lambda_{i'_1 \dots i'_n} - \sum_{(i'_1, \dots, i'_n) \in \mathcal{L}: i'_n = j} \lambda_{i'_1 \dots i'_n} = 0. \quad (\text{EC.19})$$

Moreover, we observe from (EC.13) and from the inclusion $j \in \{1, \dots, n-1\}$ that a tuple $(i'_1, \dots, i'_n) \in \mathcal{L}$ satisfies the equality $i'_n = j$ if and only if $i'_{j+1} = \dots = i'_n = j$. Therefore, (EC.19) implies that

$$\sum_{(i'_1, \dots, i'_n) \in \mathcal{L}: \substack{i'_{j+1} = j \text{ and } i'_k \neq i'_{k+1} \\ \text{for some } k \in \{j+1, \dots, n-1\}}} \lambda_{i'_1 \dots i'_n} = 0. \quad (\text{EC.20})$$

Since λ is a feasible solution of the system (EC.14), and since (EC.14) implies that the inequality $\lambda_{i'_1 \dots i'_n} \geq 0$ holds for all $(i'_1, \dots, i'_n) \in \mathcal{L}$, it follows from (EC.20) that

$$\lambda_{i'_1 \dots i'_n} = 0 \quad \forall (i'_1, \dots, i'_n) \in \mathcal{L} \text{ such that } i'_{j+1} = j \text{ and } i'_k \neq i'_{k+1} \text{ for some } k \in \{j+1, \dots, n-1\}. \quad (\text{EC.21})$$

Since we supposed that $\lambda_{i_1 \dots i_n} > 0$, and since the tuple $(i_1, \dots, i_n) \in \mathcal{L}$ satisfied the equality $i_{j+1} = j$, (EC.21) proves that the equalities $i_{j+1} = \dots = i_n = j$ must hold. Moreover, it follows from the right-most equality in (EC.17) that $i_1, \dots, i_j \in \{0, n\}$, which together with (EC.13) implies that $i_1 = \dots = i_j$. Since (EC.13) implies that $i_1 \in \{0, n\}$, our proof of Claim EC.9 is complete. \square

In view of Claims EC.8 and EC.9, we now present the proof of Lemma EC.5.

Proof of Lemma EC.5. We begin by showing that if λ is a feasible solution of the system (EC.14), then λ must satisfy the equality on line (EC.16) for all $(i_1, \dots, i_n) \in \mathcal{L}$. Indeed, suppose that λ is a feasible solution of the system (EC.14). In Claim EC.9, we showed that if λ is a feasible solution of the system (EC.14) and if $\lambda_{i_1 \dots i_n} > 0$ for some tuple $(i_1, \dots, i_n) \in \mathcal{L}$ where $(i_1, \dots, i_n) \neq (n, \dots, n)$, then the tuple (i_1, \dots, i_n) must be contained in

$$\begin{aligned} \mathcal{L}^* &\triangleq \{(i'_1, \dots, i'_n) : \text{there exists } j \in \{1, \dots, n-1\} \text{ such that } i'_1 = \dots = i'_j = 0 \text{ and } i'_{j+1} = \dots = i'_n = j\} \\ &\cup \{(i'_1, \dots, i'_n) : \text{there exists } j \in \{1, \dots, n-1\} \text{ such that } i'_1 = \dots = i'_j = n \text{ and } i'_{j+1} = \dots = i'_n = j\} \\ &= \{(\underline{i}_1^1, \dots, \underline{i}_n^1), \dots, (\underline{i}_1^{n-1}, \dots, \underline{i}_n^{n-1}), (\bar{i}_1^1, \dots, \bar{i}_n^1), \dots, (\bar{i}_1^{n-1}, \dots, \bar{i}_n^{n-1})\}, \end{aligned}$$

where the notation $(\underline{i}_1^j, \dots, \underline{i}_n^j) \triangleq (0, \dots, 0, j, \dots, j)$ and $(\bar{i}_1^j, \dots, \bar{i}_n^j) \triangleq (n, \dots, n, j, \dots, j)$ for each $j \in \{1, \dots, n-1\}$ is defined in Appendix E.1. In view of the above notation, we now consider any arbitrary $j \in \{1, \dots, n-1\}$ and $\dagger \in \{0, n\}$. We first observe that

$$v_{j, \dagger} = \sum_{(i'_1, \dots, i'_n) \in \mathcal{L} : i'_j = \dagger} \lambda_{i'_1 \dots i'_n} = \sum_{(i'_1, \dots, i'_n) \in \mathcal{L}^* : i'_j = \dagger} \lambda_{i'_1 \dots i'_n} = \begin{cases} \sum_{j'=j}^{n-1} \lambda_{\underline{i}_1^{j'} \dots \underline{i}_n^{j'}}, & \text{if } \dagger = 0, \\ \lambda_{n \dots n} + \sum_{j'=j}^{n-1} \lambda_{\bar{i}_1^{j'} \dots \bar{i}_n^{j'}}, & \text{if } \dagger = n, \end{cases} \quad (\text{EC.22})$$

where the first equality follows from the fact that λ is a feasible solution of the system (EC.14), the second equality holds because $\lambda_{i'_1 \dots i'_n} > 0$ only if $(i'_1, \dots, i'_n) \in \mathcal{L}^* \cup \{(n, \dots, n)\}$, and the third equality holds because a tuple $(i'_1, \dots, i'_n) \in \mathcal{L}^* \cup \{(n, \dots, n)\}$ satisfies the equality $i'_j = \dagger$ if and only if $[(i'_1, \dots, i'_n) = (n, \dots, n) \text{ and } \dagger = n]$ or $[\dagger = 0 \text{ and there exists } j' \in \{j, \dots, n\} \text{ such that } (i'_1, \dots, i'_n) = (\underline{i}_1^{j'}, \dots, \underline{i}_n^{j'})]$ or $[\dagger = n \text{ and there exists } j' \in \{j, \dots, n\} \text{ such that } (i'_1, \dots, i'_n) = (\bar{i}_1^{j'}, \dots, \bar{i}_n^{j'})]$. Using identical reasoning as above, we also observe that

$$v_{j+1, \dagger} = \sum_{(i'_1, \dots, i'_n) \in \mathcal{L} : i'_{j+1} = \dagger} \lambda_{i'_1 \dots i'_n} = \sum_{(i'_1, \dots, i'_n) \in \mathcal{L}^* : i'_{j+1} = \dagger} \lambda_{i'_1 \dots i'_n} = \begin{cases} \sum_{j'=j+1}^{n-1} \lambda_{\underline{i}_1^{j'} \dots \underline{i}_n^{j'}}, & \text{if } \dagger = 0, \\ \lambda_{n \dots n} + \sum_{j'=j+1}^{n-1} \lambda_{\bar{i}_1^{j'} \dots \bar{i}_n^{j'}}, & \text{if } \dagger = n. \end{cases} \quad (\text{EC.23})$$

Combining lines (EC.22) and (EC.23), we have shown that

$$v_{j, \dagger} - v_{j+1, \dagger} = \begin{cases} \lambda_{\underline{i}_1^j \dots \underline{i}_n^j}, & \text{if } \dagger = 0, \\ \lambda_{\bar{i}_1^j \dots \bar{i}_n^j}, & \text{if } \dagger = n. \end{cases} \quad (\text{EC.24})$$

We also recall from (EC.15) that

$$v_{j, \dagger} - v_{j+1, \dagger} = \begin{cases} \frac{1}{n}, & \text{if } \dagger = 0 \text{ and } j \in \bar{S}, \\ \frac{1}{n}, & \text{if } \dagger = n \text{ and } j \notin \bar{S}, \\ 0, & \text{otherwise.} \end{cases} \quad (\text{EC.25})$$

Therefore, combining lines (EC.24) and (EC.25), we have shown that the following equalities hold for all $j \in \{1, \dots, n-1\}$:

$$\lambda_{\underline{i}_1^j \dots \underline{i}_n^j} = \begin{cases} \frac{1}{n}, & \text{if } j \in \bar{S}, \\ 0, & \text{otherwise;} \end{cases} \quad \lambda_{\bar{i}_1^j \dots \bar{i}_n^j} = \begin{cases} \frac{1}{n}, & \text{if } j \notin \bar{S}, \\ 0, & \text{otherwise;} \end{cases} \quad (\text{EC.26})$$

Moreover, it follows from the fact that λ is a feasible solution of the system (EC.14) that

$$\begin{aligned} \lambda_{n \dots n} &= 1 - \sum_{(i'_1, \dots, i'_n) \in \mathcal{L} : (i'_1, \dots, i'_n) \neq (n, \dots, n)} \lambda_{i'_1 \dots i'_n} \\ &= 1 - \sum_{(i'_1, \dots, i'_n) \in \mathcal{L}^*} \lambda_{i'_1 \dots i'_n} \\ &= 1 - \sum_{j' \in \bar{S} \setminus \{0, n\}} \lambda_{\underline{i}_1^{j'} \dots \underline{i}_n^{j'}} - \sum_{j' \in \mathcal{N} \setminus \bar{S}} \lambda_{\bar{i}_1^{j'} \dots \bar{i}_n^{j'}} \\ &= 1 - \sum_{j' \in \bar{S} \setminus \{0, n\}} \frac{1}{n} - \sum_{j' \in \mathcal{N} \setminus \bar{S}} \frac{1}{n} \\ &= 1 - \frac{n-1}{n} \\ &= \frac{1}{n}, \end{aligned}$$

where the first equality follows from the fact that λ is a feasible solution of the system (EC.14), the second equality follows from the fact that $\lambda_{i'_1 \dots i'_n} > 0$ for a tuple $(i'_1, \dots, i'_n) \in \mathcal{L} \setminus \{(n, \dots, n)\}$ if and only if $(i'_1, \dots, i'_n) \in \mathcal{L}^*$, the third and fourth equalities follow from (EC.26), and the fourth and fifth equalities follow from algebra. Combining the above analysis, we conclude that any feasible solution λ of the system (EC.14) must satisfy the following equality must hold for all $(i_1, \dots, i_n) \in \mathcal{L}$:

$$\lambda_{i_1 \dots i_n} = \begin{cases} \frac{1}{n}, & \text{if there exists } j \in \bar{S} \setminus \{0, n\} \text{ such that } (i_1, \dots, i_n) = (\underline{i}_1^j, \dots, \underline{i}_n^j), \\ \frac{1}{n}, & \text{if there exists } j \in \mathcal{N} \setminus \bar{S} \text{ such that } (i_1, \dots, i_n) = (\bar{i}_1^j, \dots, \bar{i}_n^j), \\ \frac{1}{n}, & \text{if } (i_1, \dots, i_n) = (n, \dots, n), \\ 0, & \text{otherwise.} \end{cases} \quad (\text{EC.16})$$

We conclude the proof of Lemma EC.5 by showing, for the sake of completeness, that if λ satisfies the equality on line (EC.16) for all $(i_1, \dots, i_n) \in \mathcal{L}$, then λ is a feasible solution of the system (EC.14). Indeed, let λ satisfy the equality on line (EC.16) for all $(i_1, \dots, i_n) \in \mathcal{L}$. For each assortment $m \in \mathcal{M}$ and product $i \in S_m$, we observe that

$$\begin{aligned} \sum_{(i_1, \dots, i_n) \in \mathcal{L}: i_m = i} \lambda_{i_1 \dots i_n} &= \begin{cases} \frac{1}{n}, & \text{if } i \in \mathcal{N} \setminus \bar{S}, \\ \frac{1}{n}, & \text{if } i \in \bar{S} \setminus \{0, n\}, \\ \frac{1}{n} + \sum_{j \in \mathcal{N} \setminus \bar{S}: j \geq m} \frac{1}{n}, & \text{if } i = n, \\ \sum_{j \in \bar{S} \setminus \{0, n\}: j \geq m} \frac{1}{n}, & \text{if } i = 0 \end{cases} \\ &= \begin{cases} \frac{1}{n}, & \text{if } i \in \mathcal{N} \setminus \{0, n\}, \\ \frac{1}{n} (1 + |\{m, \dots, n-1\} \setminus \bar{S}|), & \text{if } i = n, \\ \frac{1}{n} |\{m, \dots, n-1\} \cap \bar{S}|, & \text{if } i = 0, \end{cases} \\ &= v_{m,i}, \end{aligned}$$

where the first equality follows from the fact that λ satisfies the equality on line (EC.16) for all $(i_1, \dots, i_n) \in \mathcal{L}$, the second equality follows from algebra, and the third equality follows from line (EC.15). Furthermore, we observe that

$$\sum_{(i_1, \dots, i_n) \in \mathcal{L}} \lambda_{i_1 \dots i_n} = \sum_{j \in \bar{S} \setminus \{0, n\}} \frac{1}{n} + \sum_{j \in \mathcal{N} \setminus \bar{S}} \frac{1}{n} + \frac{1}{n} = \frac{|\mathcal{N} \setminus \{0, n\}|}{n} + \frac{1}{n} = 1,$$

where the first equality follows from the fact that λ satisfies the equality on line (EC.16) for all $(i_1, \dots, i_n) \in \mathcal{L}$, and the second and third equalities follows from algebra. Finally, we observe from line (EC.16) that $\lambda_{i_1 \dots i_n} \geq 0$ for all $(i_1, \dots, i_n) \in \mathcal{L}$. We have thus proven that λ is a feasible solution of the system (EC.14), which concludes our proof of Lemma EC.5. \square

E.3. Proof of Theorem 2

We now use the above Lemma EC.5 to conclude our proof of Theorem 2. We begin by using Lemma EC.5 to reformulate the robust optimization problem (RO). Indeed,

$$\begin{aligned} (\text{RO}) &= \max_{S \in \mathcal{S}} \left\{ \begin{array}{l} \text{minimize}_{\lambda} \quad \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \rho_{i_1 \dots i_M}(S) \lambda_{i_1 \dots i_M} \\ \text{subject to} \quad \sum_{(i_1, \dots, i_n) \in \mathcal{L}: i_m = i} \lambda_{i_1 \dots i_n} = v_{m,i} \quad \forall m \in \mathcal{M}, i \in S_m \\ \sum_{(i_1, \dots, i_n) \in \mathcal{L}} \lambda_{i_1 \dots i_n} = 1 \\ \lambda_{i_1 \dots i_n} \geq 0 \quad \forall (i_1, \dots, i_n) \in \mathcal{L}, \end{array} \right\} \\ &= \max_{S \in \mathcal{S}} \left\{ \frac{1}{n} \sum_{j \in \bar{S} \setminus \{0, n\}} \rho_{\underline{i}_1^j, \dots, \underline{i}_n^j}(S) + \frac{1}{n} \sum_{j \in \mathcal{N} \setminus \bar{S}} \rho_{\bar{i}_1^j, \dots, \bar{i}_n^j}(S) + \frac{1}{n} \rho_{n \dots n}(S) \right\}, \quad (\text{EC.27}) \end{aligned}$$

where the first equality follows from Proposition 1 and the fact that $\eta = 0$, and the second equality follows from Lemma EC.5.

We next analyze the objective function in (EC.27). Indeed, it follows readily from Proposition EC.1 from Appendix B.2 that the following equalities hold for all assortments $S \in \mathcal{S}$:

$$\begin{aligned}\rho_{\underline{i}_1^j \dots \underline{i}_n^j}(S) &= r_j \mathbb{I}\{j \in S\} & \forall j \in \bar{S} \setminus \{0, n\}, \\ \rho_{\bar{i}_1^j \dots \bar{i}_n^j}(S) &= r_j \mathbb{I}\{j \in S\} + r_n \mathbb{I}\{j \notin S \text{ and } n \in S\} & \forall j \in \mathcal{N} \setminus \bar{S}, \\ \rho_{n \dots n}(S) &= r_n \mathbb{I}\{n \in S\}.\end{aligned}$$

Using the above equalities, we rewrite the optimization problem on line (EC.27) as

$$\max_{S \in \mathcal{S}} \left\{ \underbrace{\frac{1}{n} \sum_{j \in \bar{S} \setminus \{0, n\}} r_j \mathbb{I}\{j \in S\} + \frac{1}{n} \sum_{j \in \mathcal{N} \setminus \bar{S}} (r_j \mathbb{I}\{j \in S\} + r_n \mathbb{I}\{j \notin S \text{ and } n \in S\}) + \frac{1}{n} r_n \mathbb{I}\{n \in S\}}_{\mathcal{R}'(S)} \right\}. \quad (\text{EC.28})$$

We next show that every optimal solution S of (EC.28) must contain the product n . Indeed, for every arbitrary assortment $S \in \mathcal{S}$ that satisfies $n \notin S$, we observe from algebra that

$$\mathcal{R}'(S \cup \{n\}) - \mathcal{R}'(S) = \frac{1}{n} \sum_{j \in \mathcal{N} \setminus \bar{S}} r_n (\mathbb{I}\{j \notin S\} - 0) + \frac{1}{n} r_n > 0.$$

Since the assortment $S \in \mathcal{S}$ with $n \notin S$ was chosen arbitrarily, we have shown that every optimal solution S of (EC.28) must satisfy $n \in S$. We can thus rewrite (EC.28) as

$$\max_{S \in \mathcal{S}: n \in S} \left\{ \underbrace{\frac{1}{n} \sum_{j \in \bar{S} \setminus \{0, n\}} r_j \mathbb{I}\{j \in S\} + \frac{1}{n} \sum_{j \in \mathcal{N} \setminus \bar{S}} (r_j \mathbb{I}\{j \in S\} + r_n \mathbb{I}\{j \notin S\}) + \frac{1}{n} r_n}_{\mathcal{R}''(S)} \right\}. \quad (\text{EC.29})$$

We next show that every optimal solution S of the optimization problem (EC.29) must satisfy $j' \notin S$ for all $j' \in \mathcal{N} \setminus \bar{S}$. Indeed, for every arbitrary assortment $S \in \mathcal{S}$ that satisfies $n \in S$ and satisfies $j' \in S$ for some product $j' \in \mathcal{N} \setminus \bar{S}$, we observe from algebra that

$$\mathcal{R}''(S \setminus \{j'\}) - \mathcal{R}''(S) = \frac{1}{n} (0 + r_n) - \frac{1}{n} (r_{j'} + 0) > 0.$$

Since the assortment $S \in \mathcal{S}$ with $n \in S$ and $j' \in S$ for some $j' \in \mathcal{N} \setminus \bar{S}$ was chosen arbitrarily, we have shown that every optimal solution S of (EC.29) must satisfy $j' \notin S$ for all $j' \in \mathcal{N} \setminus \bar{S}$. We can thus rewrite (EC.29) as

$$\max_{S \in \mathcal{S}: n \in S, j' \notin S \forall j' \in \mathcal{N} \setminus \bar{S}} \left\{ \underbrace{\frac{1}{n} \sum_{j \in \bar{S} \setminus \{0, n\}} r_j \mathbb{I}\{j \in S\} + \frac{1}{n} \sum_{j \in \mathcal{N} \setminus \bar{S}} r_n + \frac{1}{n} r_n}_{\mathcal{R}''(S)} \right\}. \quad (\text{EC.30})$$

Finally, we conclude by inspection that every optimal solution S of the optimization problem (EC.30) must satisfy $j' \in S$ for all $j' \in \bar{S} \setminus \{0, n\}$. We have thus shown that the only optimal solution of the robust optimization problem (RO) is \bar{S} , which concludes our proof of Theorem 2. \square

Appendix F: Assumption that $\mathcal{N}_0 = \cup_{m \in \mathcal{M}} S_m$

In Remark 3 of §5, we state an assumption that the universe of products \mathcal{N}_0 is equal to $\cup_{m \in \mathcal{M}} S_m$, where $S_1, \dots, S_M \subseteq \mathcal{N}_0$ are the assortments that the firm has previously offered to its customers. In that remark, we claim that this assumption can always be satisfied by redefining \mathcal{N}_0 to be equal to $\cup_{m \in \mathcal{M}} S_m$. In this appendix, we argue that this assumption is without loss of generality from the perspective of the robust optimization problem (RO). Specifically, we formalize Remark 3 via the following proposition:

PROPOSITION EC.2. $\max_{S \in \mathcal{S}} \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) = \max_{S \in \mathcal{S}} \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda \left(S \cap \left(\bigcup_{m \in \mathcal{M}} S_m \right) \right)$.

The above proposition says that if $S^* \in \mathcal{S}$ is an optimal solution of the robust optimization problem (RO), then $S^* \cap \left(\bigcup_{m \in \mathcal{M}} S_m \right)$ is also an optimal solution of the robust optimization problem (RO). Hence, the above proposition implies that we can assume without loss of generality that the robust optimization problem (RO) will only optimize over assortments S in the set $\mathcal{S} \cap \left(\bigcup_{m \in \mathcal{M}} S_m \right)$. The following proof of Proposition EC.2 relies on the graphical interpretation of $\rho_{i_1 \dots i_M}(S)$ that is developed in Appendix B.2.

Proof of Proposition EC.2. We first observe that

$$\max_{S \in \mathcal{S}} \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda \left(S \cap \left(\bigcup_{m \in \mathcal{M}} S_m \right) \right) = \max_{S \in \mathcal{S} \cap \left(\bigcup_{m \in \mathcal{M}} S_m \right)} \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) \leq \max_{S \in \mathcal{S}} \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S),$$

where the equality follows from the fact that $\mathcal{S} \triangleq \{S \subseteq \mathcal{N}_0 : 0 \in S\}$, and the inequality follows from algebra. To show the other direction, consider any arbitrary assortment $S \in \mathcal{S}$. We recall from Proposition 1 that the worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda \left(S \cap \left(\bigcup_{m \in \mathcal{M}} S_m \right) \right)$ is equal to the optimal objective value of the following linear optimization problem:

$$\begin{aligned} & \underset{\lambda, \epsilon}{\text{minimize}} && \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \rho_{i_1 \dots i_M} \left(S \cap \left(\bigcup_{m \in \mathcal{M}} S_m \right) \right) \lambda_{i_1 \dots i_M} \\ & \text{subject to} && \text{same constraints as (2)} \end{aligned} \quad (\text{EC.31})$$

Moreover, we recall from Proposition EC.1 in Appendix B.2 that the following equality holds for each $(i_1, \dots, i_M) \in \mathcal{L}$:

$$\rho_{i_1 \dots i_M} \left(S \cap \left(\bigcup_{m \in \mathcal{M}} S_m \right) \right) = \min_{i \in S \cap \left(\bigcup_{m \in \mathcal{M}} S_m \right) \cap \mathcal{I}_{i_1 \dots i_M} \left(S \cap \left(\bigcup_{m \in \mathcal{M}} S_m \right) \right)} r_i. \quad (\text{EC.32})$$

In particular, we observe for each $(i_1, \dots, i_M) \in \mathcal{L}$ that

$$\begin{aligned} \mathcal{I}_{i_1 \dots i_M} \left(S \cap \left(\bigcup_{m \in \mathcal{M}} S_m \right) \right) &= \left\{ i \in \mathcal{N}_0 : \text{for all } m \in \mathcal{M}, \text{ if } i_m \in S \cap \left(\bigcup_{m \in \mathcal{M}} S_m \right), \text{ then } i_m \not\prec_{i_1 \dots i_M} i \right\} \\ &= \{i \in \mathcal{N}_0 : \text{for all } m \in \mathcal{M}, \text{ if } i_m \in S, \text{ then } i_m \not\prec_{i_1 \dots i_M} i\} \\ &= \mathcal{I}_{i_1 \dots i_M}(S), \end{aligned} \quad (\text{EC.33})$$

where the first equality follows from Definition EC.2, the second equality follows from the fact that $i_1, \dots, i_M \in \bigcup_{m \in \mathcal{M}} S_m$, and the third equality follows from Definition EC.2. Therefore, for every feasible solution (λ, ϵ) for the linear optimization problem (2), we observe that

$$\begin{aligned} \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \rho_{i_1 \dots i_M} \left(S \cap \left(\bigcup_{m \in \mathcal{M}} S_m \right) \right) \lambda_{i_1 \dots i_M} &= \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \left(\min_{i \in S \cap \left(\bigcup_{m \in \mathcal{M}} S_m \right) \cap \mathcal{I}_{i_1 \dots i_M} \left(S \cap \left(\bigcup_{m \in \mathcal{M}} S_m \right) \right)} r_i \right) \lambda_{i_1 \dots i_M} \\ &= \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \left(\min_{i \in S \cap \left(\bigcup_{m \in \mathcal{M}} S_m \right) \cap \mathcal{I}_{i_1 \dots i_M}(S)} r_i \right) \lambda_{i_1 \dots i_M} \\ &\geq \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \left(\min_{i \in S \cap \mathcal{I}_{i_1 \dots i_M}(S)} r_i \right) \lambda_{i_1 \dots i_M} \\ &= \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \rho_{i_1 \dots i_M}(S) \lambda_{i_1 \dots i_M}, \end{aligned}$$

where the first equality follows from (EC.32), the second equality follows from (EC.33), the inequality follows from algebra, and the third equality follows from Proposition EC.1. Since the above analysis holds for every feasible solution (λ, ϵ) for the linear optimization problem (2), and since the assortment $S \in \mathcal{S}$ was chosen arbitrarily, we conclude that

$$\max_{S \in \mathcal{S}} \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda \left(S \cap \left(\bigcup_{m \in \mathcal{M}} S_m \right) \right) \geq \max_{S \in \mathcal{S}} \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S).$$

The proof of Proposition EC.2 is thus complete. \square

Appendix G: Proofs of Technical Results from §5.1

G.1. Proof of Lemma 6

It follows from Definition 3 and the fact that $M = 2$ that

$$\mathcal{L} = \{(i_1, i_2) : \mathcal{D}_{i_1}(S_1) \cap \mathcal{D}_{i_2}(S_2) \neq \emptyset\}.$$

The rest of the proof of Lemma 6 is split into three cases.

In the first case, consider any arbitrary pair of products $(i_1, i_2) \in (S_1 \setminus S_2) \times S_2$. For this pair of products, consider any ranking $\sigma \in \Sigma$ which satisfies the equalities $\sigma(i_1) = 0$ and $\sigma(i_2) = 1$. It follows from the fact that $i_1 \in S_1$ that this ranking satisfies the equality $\arg \min_{j \in S_1} \sigma(j) = 1$. Moreover, it follows from the facts that $i_1 \notin S_2$ and $i_2 \in S_2$ that this ranking also satisfies the equality $\arg \min_{j \in S_2} \sigma(j) = 2$. We have thus shown that $\sigma \in \mathcal{D}_1(S_1) \cap \mathcal{D}_2(S_2)$, which proves that $\mathcal{D}_1(S_1) \cap \mathcal{D}_2(S_2) \neq \emptyset$. Since the pair of products $(i_1, i_2) \in (S_1 \setminus S_2) \times S_2$ was chosen arbitrarily, we have shown that $(S_1 \setminus S_2) \times S_2 \subseteq \mathcal{L}$.

In the second case, consider any arbitrary pair of products $(i_1, i_2) \in S_1 \times (S_2 \setminus S_1)$. Using identical reasoning as the first case, we observe that $\mathcal{D}_1(S_1) \cap \mathcal{D}_2(S_2) \neq \emptyset$, which shows that $S_1 \times (S_2 \setminus S_1) \subseteq \mathcal{L}$.

In the third case, consider any arbitrary pair of products $(i_1, i_2) \in (S_1 \cap S_2) \times (S_1 \cap S_2)$. For this pair of products, it follows from Definition 2 that the inequality $\sigma(i_1) < \sigma(j)$ must hold for all rankings $\sigma \in \mathcal{D}_{i_1}(S_1)$ and all products $j \in (S_1 \cap S_2) \setminus \{i_1\}$. It also follows from Definition 2 that the inequality $\sigma(i_2) < \sigma(j)$ must hold for all rankings $\sigma \in \mathcal{D}_{i_2}(S_2)$ and all products $j \in (S_1 \cap S_2) \setminus \{i_2\}$. It follows immediately from these inequalities that there exists a ranking that satisfies $\sigma \in \mathcal{D}_{i_1}(S_1) \cap \mathcal{D}_{i_2}(S_2)$ if and only if $i_1 = i_2$, and so it follows from Definition 3 that $(i_1, i_2) \in \mathcal{L}$ if and only if $i_1 = i_2$.

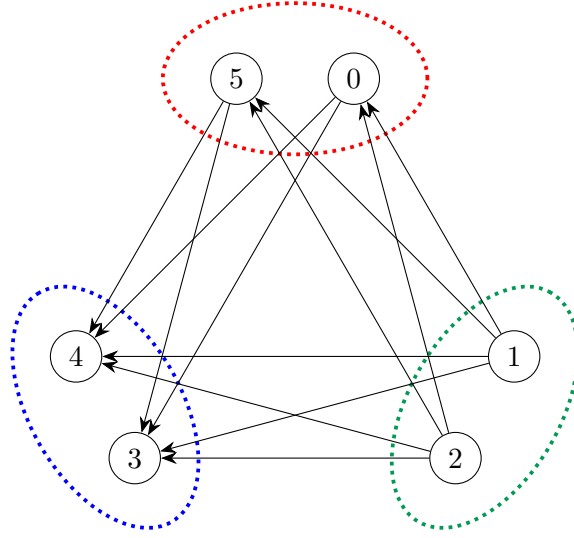
Since the three cases we have considered are exhaustive, our proof of Lemma 6 is complete. \square

G.2. Proof of Lemma 7

Let $M = 2$ and $\eta = 0$. In the remainder of the proof of Lemma 7, we will show that the following equality holds for all assortments $S \in \mathcal{S}$:

$$\begin{aligned} \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) &= \sum_{i \in S_1 \cap S_2} \rho_{ii}(S) v_{1,i} + \\ &\left[\begin{array}{l} \text{minimize}_{\lambda} \quad \sum_{i_1 \in S_1 \setminus S_2} \sum_{i_2 \in S_1 \cap S_2} \rho_{i_1 i_2}(S) \lambda_{i_1 i_2} \\ \quad + \sum_{i_1 \in S_1 \setminus S_2} \sum_{i_2 \in S_2 \setminus S_1} \rho_{i_1 i_2}(S) \lambda_{i_1 i_2} \\ \quad + \sum_{i_1 \in S_1 \cap S_2} \sum_{i_2 \in S_2 \setminus S_1} (\rho_{i_1 i_2}(S) - \rho_{i_1 i_1}(S)) \lambda_{i_1 i_2} \\ \text{subject to} \quad \sum_{i_2 \in S_2} \lambda_{i_1 i_2} = v_{1,i_1} \quad \forall i_1 \in S_1 \setminus S_2 \\ \quad \sum_{i_1 \in S_1} \lambda_{i_1 i_2} = v_{2,i_2} \quad \forall i_2 \in S_2 \setminus S_1 \\ \quad \sum_{i_2 \in S_2 \setminus S_1} \lambda_{i_1 i_2} - \sum_{i_1 \in S_1 \setminus S_2} \lambda_{i_1 i_2} = v_{1,i} - v_{2,i} \quad \forall i \in S_1 \cap S_2 \\ \quad \sum_{i_1 \in S_1 \setminus S_2} \lambda_{i_1 i} \leq v_{2,i} \quad \forall i \in S_1 \cap S_2 \\ \quad \lambda_{i_1 i_2} \geq 0 \quad \forall (i_1, i_2) \in \mathcal{L} \end{array} \right]. \end{aligned} \quad (\text{EC.34})$$

We readily observe that the linear optimization problem on line (EC.34) is a minimum-cost network flow problem, where each decision variable $\lambda_{i_1 i_2}$ corresponds to the flow on a directed edge from vertex i_1 to

Figure EC.10 Visualization of minimum-cost network flow problem from line (EC.34).

Note. The figure shows a visualization of the minimum-cost network flow problem corresponding to the linear optimization problem on line (EC.34) for the case where the past assortments are $S_1 = \{0, 1, 2, 5\}$ and $S_2 = \{0, 3, 4, 5\}$. We see that there is a vertex in the graph for each product $i \in \mathcal{N}_0 \equiv \{0, 1, 2, 3, 4, 5\}$. The graph is a tripartite directed graph, where the three partitions of vertices are denoted by the dotted ellipses and correspond to $S_1 \setminus S_2 = \{1, 2\}$, $S_1 \cap S_2 = \{0, 5\}$, and $S_2 \setminus S_1 = \{3, 4\}$. The flow demands at each of the vertices and the flow cost for each of the directed edges can be found on line (EC.34).

vertex i_2 (Ahuja et al. 1988, p. 296). In particular, we observe that the minimum-cost network flow problem on line (EC.34) takes place on a tripartite directed acyclic graph with $|S_1 \setminus S_2| + |S_2 \setminus S_1| + |S_1 \cap S_2| = n + 1$ vertices and $\mathcal{O}(n^2)$ directed edges. In Figure EC.10, we present a visualization of the network corresponding to the minimum-cost network flow problem from line (EC.34).

REMARK EC.1. The inequality of the form $\sum_{i_1 \in S_1 \setminus S_2} \lambda_{i_1 i_2} \leq v_{2,i}$ for each vertex $i \in S_1 \cap S_2$ can be interpreted as an upper bound on the incoming flow entering into vertex i . These inequalities can be converted into edge capacities by splitting each vertex $i \in S_1 \cap S_2$ into two vertices, i^{in} and i^{out} , where the incoming edges to vertex i^{in} are from the vertices in $S_1 \setminus S_2$, the outgoing edges from i^{out} are to the vertices in $S_2 \setminus S_1$, the only outgoing edge from vertex i^{in} is the only incoming edge to vertex i^{out} , the flow through the edge from vertex i^{in} to vertex i^{out} must satisfy the constraints $0 \leq g_{i^{\text{in}} i^{\text{out}}} \leq v_{2,i}$, and the supplies on the vertices i^{in} and i^{out} are 0 and $v_{1,i} - v_{2,i}$, respectively.

We now proceed to prove that line (EC.34) holds. Indeed, consider any arbitrary assortment $S \in \mathcal{S}$. It follows from Proposition 1 and our assumptions of $M = 2$ and $\eta = 0$ that

$$\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) = \left[\begin{array}{l} \text{minimize}_{\lambda} \quad \sum_{(i_1, i_2) \in \mathcal{L}} \rho_{i_1 i_2}(S) \lambda_{i_1 i_2} \\ \text{subject to} \quad \sum_{(i_1, i_2) \in \mathcal{L}: i_1 = i} \lambda_{i_1 i_2} = v_{1,i} \quad \forall i \in S_1 \\ \sum_{(i_1, i_2) \in \mathcal{L}: i_2 = i} \lambda_{i_1 i_2} = v_{2,i} \quad \forall i \in S_2 \\ \sum_{(i_1, i_2) \in \mathcal{L}} \lambda_{i_1 i_2} = 1 \\ \lambda_{i_1 i_2} \geq 0 \quad \forall (i_1, i_2) \in \mathcal{L} \end{array} \right].$$

After applying Lemma 6 to the above optimization problem and removing the redundant constraint $\sum_{(i_1, i_2) \in \mathcal{L}} \lambda_{i_1 i_2} = 1$, we observe that

$$\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) = \left[\begin{array}{l} \text{minimize}_{\lambda} \quad \sum_{i_1 \in S_1 \setminus S_2} \sum_{i_2 \in S_1 \cap S_2} \rho_{i_1 i_2}(S) \lambda_{i_1 i_2} + \sum_{i_1 \in S_1 \setminus S_2} \sum_{i_2 \in S_2 \setminus S_1} \rho_{i_1 i_2}(S) \lambda_{i_1 i_2} \\ \quad + \sum_{i_1 \in S_1 \cap S_2} \sum_{i_2 \in S_2 \setminus S_1} \rho_{i_1 i_2}(S) \lambda_{i_1 i_2} + \sum_{i \in S_1 \cap S_2} \rho_{ii}(S) \lambda_{ii} \\ \text{subject to} \quad \sum_{i_2 \in S_2} \lambda_{i_1 i_2} = v_{1, i_1} \quad \forall i_1 \in S_1 \setminus S_2 \\ \quad \sum_{i_1 \in S_1} \lambda_{i_1 i_2} = v_{2, i_2} \quad \forall i_2 \in S_2 \setminus S_1 \\ \quad \lambda_{ii} + \sum_{i_2 \in S_2 \setminus S_1} \lambda_{i i_2} = v_{1, i} \quad \forall i \in S_1 \cap S_2 \\ \quad \lambda_{ii} + \sum_{i_1 \in S_1 \setminus S_2} \lambda_{i_1 i} = v_{2, i} \quad \forall i \in S_1 \cap S_2 \\ \quad \lambda_{i_1 i_2} \geq 0 \quad \forall (i_1, i_2) \in \mathcal{L} \end{array} \right].$$

Finally, for each $i \in S_1 \cap S_2$, we eliminate the decision variable λ_{ii} from the above optimization problem by using the equality $\lambda_{ii} = v_{2, i} - \sum_{i_1 \in S_1 \setminus S_2} \lambda_{i_1 i}$. With this elimination, we obtain the desired result. \square

G.3. Proof of Theorem 3

We first describe our algorithm for solving the robust optimization problem (RO), which follows a brute-force strategy. Namely, our algorithm iterates over each of the assortments $S \in \widehat{\mathcal{S}}$, and, for each such assortment, the algorithm computes the corresponding worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$. The algorithm concludes by returning the maximum value of $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ across the assortments $S \in \widehat{\mathcal{S}}$. The correctness of this algorithm for solving the robust optimization problem (RO) follows immediately from Theorem 1.

We now analyze the running time of our algorithm by using our three intermediary results. We assume that the two assortments S_1 and S_2 are given as sorted arrays. Under this assumption, it is straightforward to see that the sets $S_1 \cap S_2$, $S_1 \setminus S_2$, and $S_2 \setminus S_1$ can be computed and stored as sorted arrays in $\mathcal{O}(n)$ computation time. We also require $\mathcal{O}(n)$ computation time to store copies of the sets $S_1, S_2, S_1 \cap S_2, S_1 \setminus S_2$, and $S_2 \setminus S_1$ in hash tables, which ensures that querying whether a given product is an element of any of these sets can be performed in $\mathcal{O}(1)$ time.

We next analyze the computation times for constructing the collection of assortments $\widehat{\mathcal{S}}$, constructing the set of pairs of products \mathcal{L} , and computing the quantities $\rho_{i_1 i_2}(S)$ for each assortment $S \in \widehat{\mathcal{S}}$ and each pair of products $(i_1, i_2) \in \mathcal{L}$. Indeed, using the aforementioned data structures, it follows readily from Lemma 2 that we can construct the collection of assortments $\widehat{\mathcal{S}}$ in $\mathcal{O}(n^3)$ computation time.¹⁹ Moreover, it follows from Lemma 6 that the set of pairs of products \mathcal{L} can be computed in $\mathcal{O}(n^2)$ time. Finally, we analyze the computation times for computing the quantities $\rho_{i_1 i_2}(S)$ for each assortment $S \in \widehat{\mathcal{S}}$ and each pair of products $(i_1, i_2) \in \mathcal{L}$. Indeed, we recall from Lemma 2 that $|\widehat{\mathcal{S}}| = \mathcal{O}(n^2)$, and we recall from Lemma 6 that $|\mathcal{L}| = \mathcal{O}(n^2)$. Therefore, there are $|\widehat{\mathcal{S}}| \times |\mathcal{L}| = \mathcal{O}(n^4)$ different ways of choosing an assortment $S \in \widehat{\mathcal{S}}$ and a pair of products

¹⁹ It follows from Lemma 2 that we can efficiently iterate over the assortments in $\widehat{\mathcal{S}}$ by iterating over the pairs of products in $S_1 \setminus S_2$ and $S_2 \setminus S_1$. Constructing the collection $\widehat{\mathcal{S}}$ thus requires iterating over the $|S_1 \setminus S_2| \times |S_2 \setminus S_1| = \mathcal{O}(n^2)$ assortments, and each of the assortments is comprised of at most $\mathcal{O}(n)$ products.

$(i_1, i_2) \in \mathcal{L}$. For each assortment $S \in \mathcal{S}$ and pair of products $(i_1, i_2) \in \mathcal{L}$, it follows readily from Definition 4 and Lemmas 2 and 6 that

$$\rho_{i_1 i_2}(S) = \begin{cases} r_{i_1}, & \text{if } i_1, i_2 \in S_1 \cap S_2, i_1 = i_2, \text{ and } i_1 \in S, \\ 0, & \text{if } i_1, i_2 \in S_1 \cap S_2, i_1 = i_2, \text{ and } i_1 \notin S, \\ r_{i_2}, & \text{if } i_1 \in S_1 \cap S_2, i_2 \in S_2 \setminus S_1, \text{ and } i_2 \in S, \\ \min\{r_{i_1}, \min_{j \in S \cap S_2 \setminus S_1} r_j\}, & \text{if } i_1 \in S_1 \cap S_2, i_2 \in S_2 \setminus S_1, i_2 \notin S, \text{ and } i_1 \in S, \\ 0, & \text{if } i_1 \in S_1 \cap S_2, i_2 \in S_2 \setminus S_1, i_2 \notin S, \text{ and } i_1 \notin S, \\ r_{i_1}, & \text{if } i_1 \in S_1 \setminus S_2, i_2 \in S_1 \cap S_2, \text{ and } i_1 \in S, \\ \min\{r_{i_2}, \min_{j \in S \cap S_1 \setminus S_2} r_j\}, & \text{if } i_1 \in S_1 \setminus S_2, i_2 \in S_1 \cap S_2, i_1 \notin S, \text{ and } i_2 \in S, \\ 0, & \text{if } i_1 \in S_1 \setminus S_2, i_2 \in S_1 \cap S_2, i_1 \notin S, \text{ and } i_2 \notin S, \\ \min\{r_{i_1}, r_{i_2}\}, & \text{if } i_1 \in S_1 \setminus S_2, i_2 \in S_2 \setminus S_1, i_1 \in S, \text{ and } i_2 \in S, \\ \min\{r_{i_1}, \min_{j \in S \cap S_2 \setminus S_1} r_j\}, & \text{if } i_1 \in S_1 \setminus S_2, i_2 \in S_2 \setminus S_1, i_1 \in S, \text{ and } i_2 \notin S, \\ \min\{r_{i_2}, \min_{j \in S \cap S_1 \setminus S_2} r_j\}, & \text{if } i_1 \in S_1 \setminus S_2, i_2 \in S_2 \setminus S_1, i_1 \notin S, \text{ and } i_2 \in S, \\ 0, & \text{if } i_1 \in S_1 \setminus S_2, i_2 \in S_2 \setminus S_1, i_1 \notin S, \text{ and } i_2 \notin S. \end{cases}$$

We observe that the quantities $\min_{j \in S \cap S_1 \setminus S_2} r_j$ and $\min_{j \in S \cap S_2 \setminus S_1} r_j$ appear in many of the above cases, and we see that these quantities can be precomputed for each of the assortments $S \in \widehat{\mathcal{S}}$ in a total of $|\widehat{\mathcal{S}}| \times \mathcal{O}(n) = \mathcal{O}(n^3)$ computation time. Given that we have precomputed these quantities, and given the fact that our data structures allow us to query whether any product is an element of the sets $S_1 \setminus S_2$, $S_2 \setminus S_1$, and $S_1 \cap S_2$ in $\mathcal{O}(1)$ time, we conclude that all of the $\rho_{i_1 i_2}(S)$ can be computed in a total of $\mathcal{O}(n^4 + n^3) = \mathcal{O}(n^4)$ time. In summary, we have established that constructing the collection of assortments $\widehat{\mathcal{S}}$, constructing the set of pairs of products \mathcal{L} , and computing the quantities $\rho_{i_1 i_2}(S)$ for each assortment $S \in \widehat{\mathcal{S}}$ and each pair of products $(i_1, i_2) \in \mathcal{L}$ can be performed in a total of $\mathcal{O}(n^4)$ computation time.

We conclude our proof of Theorem 3 by establishing the total computation time of our brute-force algorithm using the information computed above. In each iteration of our algorithm, we select an assortment $S \in \widehat{\mathcal{S}}$ and compute the worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$. As shown in Lemma 7, we can compute $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ by solving a minimum-cost network flow problem over a graph with $n + 1$ vertices and $\mathcal{O}(n^2)$ edges. Using the minimum-cost network flow algorithm of Orlin (1997) and Tarjan (1997), we observe that Problem (EC.34) can be solved in $\mathcal{O}(n^3 \log(nr_n))$ computation time.²⁰ Since the worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ must be computed for each assortment $S \in \widehat{\mathcal{S}}$, and since it follows readily from Lemma 2 that $|\widehat{\mathcal{S}}| = \mathcal{O}(n^2)$, we conclude that our algorithm requires a total of $\mathcal{O}(n^5 \log(nr_n))$ computation time. \square

Appendix H: Proofs of Technical Results from §5.2

H.1. Proof of Lemma 8

To construct the set of tuples of products \mathcal{L} , we iterate over each tuple of products $(i_1, \dots, i_M) \in S_1 \times \dots \times S_M$. For each such tuple of products, we can construct the corresponding directed graph $\mathcal{G}_{i_1 \dots i_M}$ that is described in the beginning of Appendix B.2. According to Lemma EC.1 in Appendix B.2, the tuple of products (i_1, \dots, i_M) satisfies $(i_1, \dots, i_M) \in \mathcal{L}$ if and only if the directed graph $\mathcal{G}_{i_1 \dots i_M}$ is acyclic. We can check whether a directed graph is acyclic by using the well-known topological sorting algorithm (Ahuja et al. 1988, p.79). Therefore, our algorithm for constructing the set of tuples of products \mathcal{L} is to iterate over each tuple of products $(i_1, \dots, i_M) \in S_1 \times \dots \times S_M$ and, for each such tuple of products, to check whether the tuple of

²⁰ The algorithm of Orlin (1997) and Tarjan (1997) computes the minimum-cost network flow on a directed graph in $\mathcal{O}((VE \log V) \min\{\log(VC), E \log V\})$ running time, where V is the number of vertices, E is the number of directed edges, and C is the maximum absolute value of any edge cost. The algorithm requires that C is integral; for more details, see Tarjan (1997, §3). In our case, Problem (EC.34) is a minimum-cost network flow problem in a directed graph where $V = n$, $E = \mathcal{O}(n^2)$, and $C = \max_{S \in \widehat{\mathcal{S}}, (i_1, i_2) \in \mathcal{L}} \rho_{i_1 i_2}(S) = r_n$.

products satisfies $(i_1, \dots, i_M) \in \mathcal{L}$ by performing the topological sorting algorithm on the directed graph $\mathcal{G}_{i_1 \dots i_M}$.

The computation time of the above algorithm for constructing \mathcal{L} can be analyzed as follows. The number of iterations in the algorithm is equal to $|S_1 \times \dots \times S_M| = \mathcal{O}(n^M)$, where we observe that it is trivial from a computation time analysis to enumerate and iterate over the tuples of products in $S_1 \times \dots \times S_M$. For each iteration, we must construct a directed graph $\mathcal{G}_{i_1 \dots i_M}$ that has $n+1 = \mathcal{O}(n)$ vertices and $\sum_{m \in \mathcal{M}} (|S_m| + 1) = \mathcal{O}(Mn)$ directed edges. Constructing this directed graph requires examining each product in each of the assortments S_1, \dots, S_M , and so we can construct $\mathcal{G}_{i_1 \dots i_M}$ in $\mathcal{O}(Mn)$ computation time. The computation time for the topological sort algorithm on a directed graph is equal to the number of vertices plus the number of directed edges in the directed graph, and so we can check whether $\mathcal{G}_{i_1 \dots i_M}$ is acyclic in $\mathcal{O}(n + Mn) = \mathcal{O}(Mn)$ computation time. Therefore, we have shown that our algorithm for constructing the set of tuples of products \mathcal{L} requires $\mathcal{O}(n^M \times (Mn + Mn)) = \mathcal{O}(Mn^{M+1})$ computation time. We also observe from the inequality $|\mathcal{L}| \leq |S_1 \times \dots \times S_M|$ that $|\mathcal{L}| = \mathcal{O}(n^M)$, which concludes our proof of Lemma 8. \square

H.2. Proof of Lemma 9

Consider any assortment $S \in \hat{\mathcal{S}}$ and any tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$. To motivate our algorithm for computing $\rho_{i_1 \dots i_M}(S)$, we begin by recalling the directed graph $\mathcal{G}_{i_1 \dots i_M}$ corresponding to the tuple of products (i_1, \dots, i_M) that is described in the beginning of Appendix B.2. According to Lemma EC.1 in Appendix B.2, it follows from the fact that $(i_1, \dots, i_M) \in \mathcal{L}$ that $\mathcal{G}_{i_1 \dots i_M}$ is a directed acyclic graph. Moreover, we recall that

$$\rho_{i_1 \dots i_M}(S) = \min_{i \in S \cap \mathcal{L}_{i_1 \dots i_M}(S)} r_i = \min_{i \in S: i_m \not\prec_{i_1 \dots i_M} i \text{ for all } i_m \in S} r_i, \quad (\text{EC.35})$$

where the first equality follows from Proposition EC.1 in Appendix B.2 and the second equality follows from Definition EC.2 in Appendix B.2. We recall from Definition EC.1 in Appendix B.2 that the notation $i \not\prec_{i_1 \dots i_M} j$ means that there is no directed path from vertex j to vertex i in the graph $\mathcal{G}_{i_1 \dots i_M}$. Stated in words, line (EC.35) shows that $\rho_{i_1 \dots i_M}(S)$ is equal to the minimum revenue r_i among all of the vertices i in the directed acyclic graph $\mathcal{G}_{i_1 \dots i_M}$ which do not have a directed path to a vertex i_m for any past assortment $m \in \mathcal{M}$ which satisfies $i_m \in S$.

Based on the above observations, we arrive at the following straightforward algorithm (presented as Algorithm 4) for computing $\rho_{i_1 \dots i_M}(S)$. The algorithm begins on line (1) of Algorithm 4 by constructing the directed acyclic graph $\mathcal{G}_{i_1 \dots i_M}$. In particular, it follows from identical reasoning as in the proof of Lemma 8 that $\mathcal{G}_{i_1 \dots i_M}$ can be constructed in $\mathcal{O}(Mn)$ computation time and that this graph is comprised of $\mathcal{O}(n)$ vertices and $\mathcal{O}(Mn)$ directed edges. In line (2) of Algorithm 4, we iterate over each of the previously offered assortments $m \in \mathcal{M}$ which satisfy $i_m \in S$. For each such past assortment m , we mark all of the vertices in the graph which have a directed path to vertex i_m . Since the graph has $\mathcal{O}(Mn)$ directed edges, and assuming henceforth that $\mathcal{G}_{i_1 \dots i_M}$ is stored in memory as an adjacency list, it is easy to see using a standard graph traversal algorithm like depth-first search that the loop in line (2) of Algorithm 4 can be performed in a total of $\mathcal{O}(M \times Mn) = \mathcal{O}(M^2n)$ computation time. Finally, line (3) of Algorithm 4 iterates over all of the vertices in at most $\mathcal{O}(Mn)$ computation time and outputs the minimum r_i among all of the vertices i which are unmarked. The correctness of Algorithm 4 follows immediately from our earlier reasoning on line (EC.35), and the total computation time required for Algorithm 4 is

$$\underbrace{\mathcal{O}(Mn)}_{(1)} + \underbrace{\mathcal{O}(M^2n)}_{(2)} + \underbrace{\mathcal{O}(n)}_{(3)} = \mathcal{O}(M^2n).$$

This concludes our proof of Lemma 9. \square

H.3. Proof of Theorem 4

As described at the beginning of §5.2, we consider a brute-force algorithm for solving the robust optimization problem (RO) that consists of computing the worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ for each candidate assortment $S \in \hat{\mathcal{S}}$ and outputting the assortment that has the maximum worst-case expected revenue. The computation time of this brute-force algorithm is analyzed as follows. In Lemmas 4 and 3, we established

CONSTRUCT- $\rho(S, (i_1, \dots, i_M), \mathcal{M}, r)$

Inputs:

- An assortment, $S \in \mathcal{S}$.
- A tuple of products, $(i_1, \dots, i_M) \in \mathcal{L}$.
- The collection of past assortments, $\mathcal{M} \equiv \{S_1, \dots, S_M\}$.
- The revenues of the products, $r \equiv (r_0, r_1, \dots, r_n)$.

Output:

- The quantity $\rho_{i_1 \dots i_M}(S)$.

Procedure:

1. Construct the directed acyclic graph $\mathcal{G}_{i_1 \dots i_M}$.
2. For each past assortment $m \in \mathcal{M}$:
 - (a) If $i_m \in S$:
 - i. Mark each unmarked vertex in $\mathcal{G}_{i_1 \dots i_M}$ that has a directed path to vertex i_m .
3. Output the minimum r_i among all unmarked vertices i in $\mathcal{G}_{i_1 \dots i_M}$ and terminate.

Algorithm 4: A procedure for computing $\rho_{i_1 \dots i_M}(S)$.

that the collection of assortments $\widehat{\mathcal{S}}$ can be constructed in $\mathcal{O}(n^2(M + |\widehat{\mathcal{S}}|)) = \mathcal{O}(n^2(M + (n + 2)^{2^M})) = \mathcal{O}(\text{poly}(n))$ computation time, where the last equality holds for any fixed M . In Lemma 8, we established that the set of tuples of products \mathcal{L} can be constructed in $\mathcal{O}(Mn^{M+1}) = \mathcal{O}(\text{poly}(n))$ computation time, where the equality holds for any fixed M . Our algorithm performs an iteration for each assortment $S \in \widehat{\mathcal{S}}$, and thus our algorithm will perform $|\widehat{\mathcal{S}}| \leq (n + 2)^{2^M} = \mathcal{O}(\text{poly}(n))$ iterations for any fixed M . Given any assortment $S \in \widehat{\mathcal{S}}$, we established in Lemma 9 that the quantities $\rho_{i_1 \dots i_M}(S)$ for each $(i_1, \dots, i_M) \in \mathcal{L}$ can all be computed in a total of $\mathcal{O}(|\mathcal{L}| \times M^2 n) = \mathcal{O}(n^M \times M^2 n) = \mathcal{O}(\text{poly}(n))$ computation time, where the last equality holds for any fixed M . Given the set of tuples of products \mathcal{L} and the quantities $\rho_{i_1 \dots i_M}(S)$ for each $(i_1, \dots, i_M) \in \mathcal{L}$, we established in Remark 2 that we can compute the worst-case expected revenue $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ by constructing and solving a linear optimization problem with $\mathcal{O}(n^M) = \mathcal{O}(\text{poly}(n))$ decision variables and $\mathcal{O}(nM) = \mathcal{O}(\text{poly}(n))$ constraints for any fixed M . Since linear optimization can be solved in weakly polynomial-time via the ellipsoid algorithm, we conclude that each iteration of our algorithm requires $\mathcal{O}(\text{poly}(n))$ time for any fixed M . Our proof of Theorem 4 is thus complete. \square

Appendix I: Proofs of Technical Results from §5.3

This appendix contains the proofs of the technical results from §5.3. Appendices I.1, I.2, and I.3 contain the proofs of Lemma 10, Corollary 2, and Lemma 11. Appendix I.4 contains two intermediary results (Lemmas EC.6 and EC.7) that are used in the proofs of Proposition 3 and Theorem 5. Appendices I.5 and I.6 contain the proofs of Proposition 3 and Theorem 5. We remark that the proofs in this appendix make extensive use of the notation and intermediary results from §4.4 and Appendix B.2.

I.1. Proof of Lemma 10

Let Assumption 1 hold, consider any tuple $(i_1, \dots, i_M) \in S_1 \times \dots \times S_M$, and consider the corresponding directed graph $\mathcal{G}_{i_1 \dots i_M}$. We recall that the set of vertices in the directed graph is equal to \mathcal{N}_0 , and the graph has a directed edge (i, i_m) from each vertex $i \in S_m \setminus \{i_m\}$ to vertex i_m for each past assortment $m \in \mathcal{M}$.

To show the first direction, suppose that there exists an $m \in \{1, \dots, M-1\}$ such that $i_{m+1} \notin \{i_m\} \cup \mathcal{B}_{m+1}$. In this case, we observe that

$$i_{m+1} \in S_{m+1} \setminus (\{i_m\} \cup \mathcal{B}_{m+1}) = S_{m+1} \setminus (\{i_m\} \cup (S_{m+1} \setminus S_m)) = S_m \setminus \{i_m\},$$

where the inclusion follows from the fact that $i_{m+1} \in S_{m+1}$ and from the supposition that $i_{m+1} \notin \{i_m\} \cup \mathcal{B}_{m+1}$, the first equality follows from line (3) and from the fact that $m+1 \in \{2, \dots, M\}$, and the second equality follows from algebra. Moreover, it follows from the fact that the past assortments are nested that $i_m \in S_{m+1}$ and from the supposition that $i_{m+1} \neq i_m$ that (i_m, i_{m+1}) is a directed edge in $\mathcal{G}_{i_1 \dots i_M}$. Since the directed graph $\mathcal{G}_{i_1 \dots i_M}$ has a directed edge from each vertex $i \in S_m \setminus \{i_m\}$ to vertex i_m , we conclude that the directed graph $\mathcal{G}_{i_1 \dots i_M}$ must have an edge from vertex i_{m+1} to vertex i_m , which implies that $\mathcal{G}_{i_1 \dots i_M}$ has a cycle. Therefore, it follows from Lemma EC.1 from Appendix B.2 that $(i_1, \dots, i_M) \notin \mathcal{L}$.

To show the other direction, suppose that the inclusion $i_{m+1} \in \{i_m\} \cup \mathcal{B}_{m+1}$ holds for all $m \in \{1, \dots, M-1\}$. For notational convenience, let i_{m_1}, \dots, i_{m_K} denote the elements of the set $\cup_{m=1}^M \{i_m\}$, where we assume without loss of generality that the indices m_1, \dots, m_K satisfy $m_1 < \dots < m_K$, satisfy $m_1 = 1$, and satisfy $i_{m_{k-1}} \neq i_{m_k}$ for all $k \in \{2, \dots, K\}$. It follows from the construction of the directed graph $\mathcal{G}_{i_1 \dots i_M}$ that i_{m_1}, \dots, i_{m_K} are the only vertices in the directed graph with incoming edges. Moreover, it follows from our choice of the indices m_1, \dots, m_K and from the supposition that $i_{m+1} \in \{i_m\} \cup \mathcal{B}_{m+1}$ for all $m \in \{1, \dots, M-1\}$ that $i_{m_k} \in \mathcal{B}_{m_k}$ for each $k \in \{1, \dots, K\}$, which implies that $i_{m_{k'}} \notin S_{m_k}$ for all $1 \leq k < k' \leq K$. Therefore, we conclude for all $1 \leq k < k' \leq K$ that there is no directed edge in $\mathcal{G}_{i_1 \dots i_M}$ from vertex $i_{m_{k'}}$ to vertex i_{m_k} , which implies that $\mathcal{G}_{i_1 \dots i_M}$ is acyclic. Therefore, it follows from Lemma EC.1 that $(i_1, \dots, i_M) \in \mathcal{L}$. \square

I.2. Proof of Corollary 2

Let Assumption 1 hold. For each $m \in \mathcal{M}$, let us define the following set of tuples:

$$\mathcal{L}_m \triangleq \begin{cases} \{(i_1) : i_1 \in \mathcal{B}_1\}, & \text{if } m = 1, \\ \{(i_1, \dots, i_m) : (i_1, \dots, i_{m-1}) \in \mathcal{L}_{m-1} \text{ and } i_m \in \{i_{m-1}\} \cup \mathcal{B}_m\}, & \text{if } m \in \{2, \dots, M\}. \end{cases} \quad (\text{EC.36})$$

It follows from a straightforward induction argument and from Lemma 10 that $\mathcal{L}_M = \mathcal{L}$. Moreover, it follows from a straightforward induction argument and from line (EC.36) that the following holds for each $m \in \mathcal{M}$:

$$|\mathcal{L}_m| = \begin{cases} |\mathcal{B}_1|, & \text{if } m = 1, \\ |\mathcal{B}_1| \times \prod_{m'=2}^m (1 + |\mathcal{B}_{m'}|), & \text{if } m \in \{2, \dots, M\}. \end{cases} \quad (\text{EC.37})$$

Therefore, we conclude that

$$|\mathcal{L}| = |\mathcal{B}_1| \times \prod_{m=2}^M (1 + |\mathcal{B}_m|) \geq 2^{M-1},$$

where the equality follows from line (EC.37) and from the fact that $\mathcal{L}_M = \mathcal{L}$, and the inequality follows from the fact that $|\mathcal{B}_m| \geq 1$ for all $m \in \mathcal{M}$. This completes our proof of Corollary 2. \square

I.3. Proof of Lemma 11

We assume throughout the proof of Lemma 11 that Assumption 1 holds, and that $S \subseteq \mathcal{N}_0$ is an assortment that satisfies $S \cap S_1 \neq \emptyset$.

We begin by considering any tuple $(i_1, \dots, i_M) \in \mathcal{L}$ and $m \in \mathcal{M}$ that satisfy $i_m \in S$. In this case, we recall from Appendix B.2 that the graph $\mathcal{G}_{i_1 \dots i_M}$ has a directed edge from each vertex $i \in S_m \setminus \{i_m\}$ to vertex i_m . In particular, that implies that there is a directed edge from each vertex $i \in (S \cap S_m) \setminus \{i_m\}$ to vertex i_m . Since $i_m \in S$, we have thus shown that $i_m \prec_{i_1 \dots i_M} i$ for all $i \in (S \cap S_m) \setminus \{i_m\}$. Therefore,

$$\begin{aligned} \rho_{i_1 \dots i_M}(S \cap S_m) &= \min_{i \in S \cap S_m \cap \mathcal{I}_{i_1 \dots i_M}(S \cap S_m)} r_i \\ &= \min_{i \in S \cap S_m \cap \{j \in \mathcal{N}_0 : \text{for all } m' \in \mathcal{M}, \text{ if } i_{m'} \in S \cap S_m, \text{ then } i_{m'} \not\prec_{i_1 \dots i_M} j\}} r_i \\ &= \min_{i \in S \cap S_m : \text{for all } m' \in \mathcal{M}, \text{ if } i_{m'} \in S \cap S_m, \text{ then } i_{m'} \not\prec_{i_1 \dots i_M} i} r_i \\ &= \min_{i \in S \cap S_m : i = i_m} r_i \\ &= r_{i_m}, \end{aligned} \quad (\text{EC.38})$$

where the first equality follows from Proposition EC.1 from Appendix B.2, the second equality follows from Definition EC.2 from Appendix B.2, the third equality follows from algebra, the fourth equality follows from the fact that $i_m \prec_{i_1 \dots i_M} i$ for all $i \in (S \cap S_m) \setminus \{i_m\}$, and the fifth equality follows from algebra. In summary, we have shown that if a tuple $(i_1, \dots, i_M) \in \mathcal{L}$ and $m \in \mathcal{M}$ satisfy $i_m \in S$, then $\rho_{i_1 \dots i_M}(S \cap S_m) = r_{i_m}$.

We next consider any tuple $(i_1, \dots, i_M) \in \mathcal{L}$ that satisfies $i_1 \notin S$. For this case, we observe that

$$\begin{aligned}
\rho_{i_1 \dots i_M}(S \cap S_1) &= \min_{i \in S \cap S_1 \cap \mathcal{I}_{i_1 \dots i_M}(S \cap S_1)} r_i \\
&= \min_{i \in S \cap S_1 \cap \{j \in \mathcal{N}_0: \text{for all } m' \in \mathcal{M}, \text{ if } i_{m'} \in S \cap S_1, \text{ then } i_{m'} \not\prec_{i_1 \dots i_M} j\}} r_i \\
&= \min_{i \in S \cap S_1: \text{for all } m' \in \mathcal{M}, \text{ if } i_{m'} \in S \cap S_1, \text{ then } i_{m'} \not\prec_{i_1 \dots i_M} i} r_i \\
&= \min_{i \in S \cap S_1: \text{for all } m' \in \{2, \dots, M\}, \text{ if } i_{m'} \in S \cap S_1, \text{ then } i_{m'} \not\prec_{i_1 \dots i_M} i} r_i \\
&= \min_{i \in S \cap S_1} r_i \\
&= \min_{i \in S \cap \mathcal{B}_1} r_i.
\end{aligned}$$

Indeed, the first equality follows from Proposition EC.1. The second equality follows from Definition EC.2. The third equality follows from algebra. The fourth equality follows from the fact that $i_1 \notin S$. The fifth equality follows from Lemma 10, which together with the facts that $(i_1, \dots, i_M) \in \mathcal{L}$ and $i_1 \notin S$ readily implies that $i_2, \dots, i_M \notin S \cap S_1$. The sixth equality follows from line (3). In summary, we have shown that if a tuple $(i_1, \dots, i_M) \in \mathcal{L}$ satisfies $i_1 \notin S$, then $\rho_{i_1 \dots i_M}(S \cap S_1) = \min_{i \in S \cap \mathcal{B}_1} r_i$.

Finally, we consider any tuple $(i_1, \dots, i_M) \in \mathcal{L}$ and $m \in \{2, \dots, M\}$ that satisfy $i_m \notin S$. In this case, we observe that

$$\begin{aligned}
\rho_{i_1 \dots i_M}(S \cap S_m) &= \min_{i \in S \cap S_m \cap \mathcal{I}_{i_1 \dots i_M}(S \cap S_m)} r_i \\
&= \min_{i \in S \cap (S_{m-1} \cup \mathcal{B}_m) \cap \mathcal{I}_{i_1 \dots i_M}(S \cap S_m)} r_i \\
&= \min \left\{ \min_{i \in S \cap S_{m-1} \cap \mathcal{I}_{i_1 \dots i_M}(S \cap S_m)} r_i, \min_{i \in S \cap \mathcal{B}_m \cap \mathcal{I}_{i_1 \dots i_M}(S \cap S_m)} r_i \right\}, \tag{EC.39}
\end{aligned}$$

where the first equality follows from Proposition EC.1, the second equality follows from Assumption 1, and the third equality follows from algebra. In the remainder of the proof, we focus on evaluating the two terms $\min_{i \in S \cap S_{m-1} \cap \mathcal{I}_{i_1 \dots i_M}(S \cap S_m)} r_i$ and $\min_{i \in S \cap \mathcal{B}_m \cap \mathcal{I}_{i_1 \dots i_M}(S \cap S_m)} r_i$ that appear in line (EC.39). First, we observe that

$$\begin{aligned}
\min_{i \in S \cap S_{m-1} \cap \mathcal{I}_{i_1 \dots i_M}(S \cap S_m)} r_i &= \min_{i \in S \cap S_{m-1} \cap \{j \in \mathcal{N}_0: \text{for all } m' \in \mathcal{M}, \text{ if } i_{m'} \in S \cap S_m, \text{ then } i_{m'} \not\prec_{i_1 \dots i_M} j\}} r_i \\
&= \min_{i \in S \cap S_{m-1} \cap \{j \in \mathcal{N}_0: \text{for all } m' \in \{1, \dots, m-1\}, \text{ if } i_{m'} \in S \cap S_m, \text{ then } i_{m'} \not\prec_{i_1 \dots i_M} j\}} r_i \\
&= \min_{i \in S \cap S_{m-1} \cap \{j \in \mathcal{N}_0: \text{for all } m' \in \{1, \dots, m-1\}, \text{ if } i_{m'} \in S \cap S_{m-1}, \text{ then } i_{m'} \not\prec_{i_1 \dots i_M} j\}} r_i \\
&= \min_{i \in S \cap S_{m-1} \cap \{j \in \mathcal{N}_0: \text{for all } m' \in \mathcal{M}, \text{ if } i_{m'} \in S \cap S_{m-1}, \text{ then } i_{m'} \not\prec_{i_1 \dots i_M} j\}} r_i \\
&= \min_{i \in S \cap S_{m-1} \cap \mathcal{I}_{i_1 \dots i_M}(S \cap S_{m-1})} r_i \\
&= \rho_{i_1 \dots i_M}(S \cap S_{m-1}). \tag{EC.40}
\end{aligned}$$

Indeed, the first equality follows from Definition EC.2. The second equality follows from Lemma 10, which together with the facts that $(i_1, \dots, i_M) \in \mathcal{L}$ and $i_m \notin S$ implies that $i_m, \dots, i_M \notin S \cap S_m$. The third equality follows from Assumption 1, which together with the fact that $i_m \notin S$ implies for each $m' \in \{1, \dots, m-1\}$ that $i_{m'} \in S \cap S_m$ if and only if $i_{m'} \in S \cap S_{m-1}$. The fourth equality follows from the fact that $i_m, \dots, i_M \notin S \cap S_m$

and the fact that $S \cap S_{m-1} \subseteq S \cap S_m$. The fifth equality follows from Definition EC.2, and the sixth equality follows from Proposition EC.1. Moreover, we observe that

$$\begin{aligned}
\min_{i \in S \cap \mathcal{B}_m \cap \mathcal{I}_{i_1 \dots i_M}} r_i &= \min_{i \in S \cap \mathcal{B}_m \cap \{j \in \mathcal{N}_0: \text{for all } m' \in \mathcal{M}, \text{ if } i_{m'} \in S \cap S_m, \text{ then } i_{m'} \not\prec_{i_1 \dots i_M} j\}} r_i \\
&= \min_{i \in S \cap \mathcal{B}_m \cap \{j \in \mathcal{N}_0: \text{for all } m' \in \{1, \dots, m-1\}, \text{ if } i_{m'} \in S \cap S_m, \text{ then } i_{m'} \not\prec_{i_1 \dots i_M} j\}} r_i \\
&= \min_{i \in S \cap \mathcal{B}_m \cap \{j \in \mathcal{N}_0: \text{for all } m' \in \{1, \dots, m-1\}, \text{ if } i_{m'} \in S \cap S_{m-1}, \text{ then } i_{m'} \not\prec_{i_1 \dots i_M} j\}} r_i \\
&= \min_{i \in S \cap \mathcal{B}_m: \text{for all } m' \in \{1, \dots, m-1\}, \text{ if } i_{m'} \in S \cap S_{m-1}, \text{ then } i_{m'} \not\prec_{i_1 \dots i_M} i} r_i. \tag{EC.41}
\end{aligned}$$

Indeed, the first equality follows from Definition EC.2. The second equality follows from Lemma 10, which together with the facts that $(i_1, \dots, i_M) \in \mathcal{L}$ and $i_m \notin S$ implies that $i_m, \dots, i_M \notin S \cap S_m$. The third equality follows from Assumption 1, which together with the fact that $i_m \notin S$ implies for each $m' \in \{1, \dots, m-1\}$ that $i_{m'} \in S \cap S_m$ if and only if $i_{m'} \in S \cap S_{m-1}$. To simplify the expression on line (EC.41), we make use of the following intermediary claim.

CLAIM EC.10. *Consider any $i \in S \cap \mathcal{B}_m$ and $m' \in \{1, \dots, m-1\}$. If $i_{m'} \in S$, then $i_{m'} \not\prec_{i_1 \dots i_M} i$.*

Proof of Claim EC.10. Consider any $i \in S \cap \mathcal{B}_m$, and consider any $m' \in \{1, \dots, m-1\}$ that satisfies the inclusion $i_{m'} \in S$.

We first show that $i_{m''} \prec_{i_1 \dots i_M} i_{m'}$ for all $m'' \in \{m, \dots, M\}$. Indeed, consider any arbitrary $m'' \in \{m, \dots, M\}$. We readily observe from Assumption 1 that $i_{m'} \in S_{m'} \subseteq S_{m''}$. Moreover, it follows from the fact that $i_{m'} \in S$ and from the fact that $i_m \notin S$ that the expression $i_{m'} \neq i_m$ must hold. Combining the expression $i_{m'} \neq i_m$ with Lemma 10, we observe that the expression $i_{m'} \neq i_{m''}$ must hold. Combining the expression $i_{m'} \neq i_{m''}$ with the fact that $i_{m'}, i_{m''} \in S_{m''}$, we conclude that the graph $\mathcal{G}_{i_1 \dots i_M}$ must have a directed edge from vertex $i_{m'}$ to vertex $i_{m''}$. Because $m'' \in \{m, \dots, M\}$ was chosen arbitrarily, we conclude that $i_{m''} \prec_{i_1 \dots i_M} i_{m'}$ for all $m'' \in \{m, \dots, M\}$.

We next show that the vertex i in the graph $\mathcal{G}_{i_1 \dots i_M}$ has an outgoing edge to vertex j only if there exists $m'' \in \{m, \dots, M\}$ such that $j = i_{m''}$. Indeed, we observe from the fact that $i \in S \cap \mathcal{B}_m \subseteq \mathcal{B}_m$ and from Assumption 1 that the inclusion $i \in S_{m''}$ holds if and only if $m'' \in \{m, \dots, M\}$. Therefore, it follows from the construction of the directed graph $\mathcal{G}_{i_1 \dots i_M}$ that (i, j) is a directed edge in $\mathcal{G}_{i_1 \dots i_M}$ only if $j = i_{m''}$ for some $m'' \in \{m, \dots, M\}$.

We now combine the above steps to conclude the proof of Claim EC.10. Indeed, suppose for the sake of developing a contradiction that $i_{m'} \prec_{i_1 \dots i_M} i$. We showed previously that the only outgoing edges from the vertex i in the directed graph $\mathcal{G}_{i_1 \dots i_M}$ are to vertices j that satisfy the equality $j = i_{m''}$ for some $m'' \in \{m, \dots, M\}$. Therefore, it follows from the supposition that $i_{m'} \prec_{i_1 \dots i_M} i$ that there must exist a $m'' \in \{m, \dots, M\}$ such that $i_{m'} \prec_{i_1 \dots i_M} i_{m''}$. Because we showed previously that $i_{m''} \prec_{i_1 \dots i_M} i_{m'}$, we conclude that the directed graph $\mathcal{G}_{i_1 \dots i_M}$ must have a cycle. However, we recall from Lemma EC.1 from Appendix B.2 and from the fact that $(i_1, \dots, i_M) \in \mathcal{L}$ that $\mathcal{G}_{i_1 \dots i_M}$ is acyclic. Because we have a contradiction, we conclude that $i_{m'} \not\prec_{i_1 \dots i_M} i$, which completes our proof of Claim EC.10. \square

Combining lines (EC.39), (EC.40), and (EC.41) with Claim EC.10, we conclude that

$$\begin{aligned}
\rho_{i_1 \dots i_M}(S \cap S_m) &= \min \left\{ \min_{i \in S \cap S_{m-1} \cap \mathcal{I}_{i_1 \dots i_M}} r_i, \min_{i \in S \cap \mathcal{B}_m \cap \mathcal{I}_{i_1 \dots i_M}} r_i \right\} \\
&= \min \left\{ \rho_{i_1 \dots i_M}(S \cap S_{m-1}), \min_{i \in S \cap \mathcal{B}_m \cap \mathcal{I}_{i_1 \dots i_M}} r_i \right\} \\
&= \min \left\{ \rho_{i_1 \dots i_M}(S \cap S_{m-1}), \min_{i \in S \cap \mathcal{B}_m: \text{for all } m' \in \{1, \dots, m-1\}, \text{ if } i_{m'} \in S \cap S_{m-1}, \text{ then } i_{m'} \not\prec_{i_1 \dots i_M} i} r_i \right\} \\
&= \min \left\{ \rho_{i_1 \dots i_M}(S \cap S_{m-1}), \min_{i \in S \cap \mathcal{B}_m} r_i \right\}.
\end{aligned}$$

Indeed, the first equality follows from (EC.39), the second equality follows from (EC.40), the third equality follows from (EC.41), and the fourth equality follows from Claim EC.10. Our proof of Lemma 11 is thus complete. \square

I.4. Intermediary Results for Proposition 3 and Theorem 5

This appendix contains two intermediary lemmas that will be used in the proofs of Proposition 3 in Appendix I.5 and the proof of Theorem 5 in Appendix I.6. The purpose of these intermediary lemmas, which are presented below as Lemmas EC.6 and EC.7, is to characterize the relationship between the the linear optimization problem (6) and the linear optimization problem (2).

For notational convenience, let the following set be defined for each assortment $S \subseteq \mathcal{N}_0$:

$$\mathfrak{N}(S) \triangleq \left\{ \begin{array}{l} [i \in S \text{ and } i \in \mathcal{B}_m \text{ and } \kappa \neq r_i] \\ (m, i, \kappa) \in \mathfrak{V}: \text{ or } [i \in S \text{ and } i \notin \mathcal{B}_m \text{ and } \kappa \in \{r_j : j \in \mathcal{B}_m\}] \\ \text{ or } [\kappa \in \{r_j : j \in \mathcal{B}_m \setminus S\}] \end{array} \right\}. \quad (\text{EC.42})$$

We remark that $\mathfrak{N}(S)$ is the set of vertices $(m, i, \kappa) \in \mathfrak{V}$ for which the linear optimization problem (6) has a constraint of the form $g_{m,i,\kappa} = 0$. Moreover, let the set of all paths of the form $(1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M) \in \mathfrak{V}$ in the directed acyclic graph $\mathfrak{G} \equiv (\mathfrak{V}, \mathfrak{E})$ be denoted by

$$\mathcal{P} \triangleq \{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) : ((m, i_m, \kappa_m), (m+1, i_{m+1}, \kappa_{m+1})) \in \mathfrak{E} \forall m \in \{1, \dots, M-1\}\}.$$

In our first intermediary lemma, denoted below by Lemma EC.6, we show that every feasible solution of the linear optimization problem (2) can be transformed into a feasible solution of the linear optimization problem (6) with the same objective value.

LEMMA EC.6. *Let Assumption 1 hold, let $S \subseteq \mathcal{N}_0$ satisfy $S \cap S_1 \neq \emptyset$, and let (λ, ϵ) be a feasible solution for the linear optimization problem (2). For each vertex $(m, i, \kappa) \in \mathfrak{V}$, let*

$$g_{m,i,\kappa} \triangleq \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i, \rho_{i_1 \dots i_M}(S \cap S_m) = \kappa} \lambda_{i_1 \dots i_M},$$

and for each directed edge $((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E}$, let

$$f_{m,i,\kappa,i',\kappa'} \triangleq \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i, \rho_{i_1 \dots i_M}(S \cap S_m) = \kappa, i_{m+1} = i', \rho_{i_1 \dots i_M}(S \cap S_{m+1}) = \kappa'} \lambda_{i_1 \dots i_M}.$$

Then (f, g, ϵ) is a feasible solution for the linear optimization problem (6), and the objective value associated with (f, g, ϵ) in the linear optimization problem (6) is equal to the objective value associated with (λ, ϵ) in the linear optimization problem (2).

Proof. Let Assumption 1 hold, let $S \subseteq \mathcal{N}_0$ satisfy $S \cap S_1 \neq \emptyset$, and let (λ, ϵ) be a feasible solution for the linear optimization problem (2). For each vertex $(m, i, \kappa) \in \mathfrak{V}$, let

$$g_{m,i,\kappa} \triangleq \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i, \rho_{i_1 \dots i_M}(S \cap S_m) = \kappa} \lambda_{i_1 \dots i_M},$$

and for each edge $((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E}$, let

$$f_{m,i,\kappa,i',\kappa'} \triangleq \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i, \rho_{i_1 \dots i_M}(S \cap S_m) = \kappa, i_{m+1} = i', \rho_{i_1 \dots i_M}(S \cap S_{m+1}) = \kappa'} \lambda_{i_1 \dots i_M}.$$

We begin by presenting an intermediary claim, denoted below by Claim EC.11, which will provide the key ingredient in the proof of Lemma EC.6. The purpose of this intermediary claim is to show that every tuple $(i_1, \dots, i_M) \in \mathcal{L}$ can be converted into a path $(1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)$ in which none of the vertices are elements of $\mathfrak{N}(S)$.

CLAIM EC.11. *For each $(i_1, \dots, i_M) \in \mathcal{L}$, we have that $((1, i_1, \rho_{i_1 \dots i_M}(S \cap S_1)), \dots, (M, i_M, \rho_{i_1 \dots i_M}(S \cap S_M))) \in \mathcal{P}$ and that $(m, i_m, \rho_{i_1, \dots, i_M}(S \cap S_m)) \notin \mathfrak{N}(S)$ for all $m \in \mathcal{M}$.*

Proof of Claim EC.11. Consider any tuple $(i_1, \dots, i_M) \in \mathcal{L}$. It follows readily from Definition 4 that the inclusion $\rho_{i_1 \dots i_M}(S \cap S_m) \in \{r_j : j \in S \cap S_m\}$ holds for each $m \in \mathcal{M}$, and so it follows from the definition of \mathfrak{V} that the inclusion $(m, i_m, \rho_{i_1 \dots i_M}(S \cap S_m)) \in \mathfrak{V}$ holds for each $m \in \mathcal{M}$. Moreover, for every arbitrary $m \in \{1, \dots, M-1\}$, we observe from Lemma 10 that $i_{m+1} \in \{i_m\} \cup \mathcal{B}_{m+1}$, and we observe from Lemma 11 and from the fact that $(i_1, \dots, i_M) \in \mathcal{L}$ that $\rho_{i_1 \dots i_M}(S \cap S_{m+1}) \in \{\rho_{i_1 \dots i_M}(S \cap S_m)\} \cup \{r_j : j \in \mathcal{B}_m\}$. Combining those two observations, we have shown that $((m, i_m, \rho_{i_1 \dots i_M}(S \cap S_m)), (m+1, i_{m+1}, \rho_{i_1 \dots i_M}(S \cap S_{m+1}))) \in \mathfrak{E}$. Since $m \in \{1, \dots, M-1\}$ was chosen arbitrarily, we conclude that $((1, i_1, \rho_{i_1 \dots i_M}(S \cap S_1)), \dots, (M, i_M, \rho_{i_1 \dots i_M}(S \cap S_M))) \in \mathcal{P}$.

To conclude the proof of Claim EC.11, it remains for us to show that $(m, i_m, \rho_{i_1, \dots, i_M}(S \cap S_m)) \notin \mathfrak{N}(S)$ for all $m \in \mathcal{M}$. Indeed, for each $m \in \mathcal{M}$, we consider the following three cases:

- If $i_m \in S$ and $i_m \in \mathcal{B}_m$, then it follows from Lemma 11 that $\rho_{i_1 \dots i_M}(S \cap S_m) = r_{i_m}$.
 - If $i_m \in S$ and $i_m \notin \mathcal{B}_m$, then it follows from Lemma 11 that $\rho_{i_1 \dots i_M}(S \cap S_m) = r_{i_m} \notin \{r_j : j \in \mathcal{B}_m\}$.
 - If $\rho_{i_1 \dots i_M}(S \cap S_m) \in \{r_j : j \in \mathcal{B}_m\}$, then it follows from Lemma 11 that $\rho_{i_1 \dots i_M}(S \cap S_m) \in \{r_j : j \in S \cap \mathcal{B}_m\}$.
- Combining the above cases with the definition of $\mathfrak{N}(S)$ on line (EC.42), we conclude that $(m, i_m, \rho_{i_1 \dots i_M}(S \cap S_m)) \notin \mathfrak{N}(S)$ for each $m \in \mathcal{M}$. Our proof of Claim EC.11 is thus complete. \square

Using Claim EC.11, we show in the following bullet points that the solution (f, g, ϵ) is a feasible solution for the linear optimization problem (6); that is, we show that (f, g, ϵ) satisfies each of the constraints in the linear optimization problem (6).

- For all $m \in \mathcal{M}$ and $i \in S_m$, we observe that

$$\begin{aligned}
\sum_{\kappa: (m, i, \kappa) \in \mathfrak{V}} g_{m, i, \kappa} - \epsilon_{m, i} &= \sum_{\kappa: (m, i, \kappa) \in \mathfrak{V}} \left(\sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i, \rho_{i_1 \dots i_M}(S \cap S_m) = \kappa} \lambda_{i_1 \dots i_M} \right) - \epsilon_{m, i} \\
&= \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i} \left(\sum_{\kappa: (m, i, \kappa) \in \mathfrak{V}, \rho_{i_1 \dots i_M}(S \cap S_m) = \kappa} \lambda_{i_1 \dots i_M} \right) - \epsilon_{m, i} \\
&= \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i} \mathbb{I}\{(m, i_m, \rho_{i_1 \dots i_M}(S \cap S_m)) \in \mathfrak{V}\} \lambda_{i_1 \dots i_M} - \epsilon_{m, i} \\
&= \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i} \lambda_{i_1 \dots i_M} - \epsilon_{m, i} \\
&= v_{m, i},
\end{aligned}$$

Indeed, the first equality follows from our construction of g . The second and third equalities follow from algebra. The fourth equality follows from Claim EC.11, which implies that $(m, i_m, \rho_{i_1 \dots i_M}(S \cap S_m)) \in \mathfrak{V}$ for all $(i_1, \dots, i_M) \in \mathcal{L}$. The fifth equality follows from the fact that (λ, ϵ) is a feasible solution for the linear optimization problem (2).

- For all $(m, i, \kappa) \in \mathfrak{V}$ such that $m \in \{1, \dots, M-1\}$, we observe that

$$\begin{aligned}
&\sum_{i', \kappa': ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E}} f_{m, i, \kappa, i', \kappa'} \\
&= \sum_{i', \kappa': ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E}} \left(\sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i, \rho_{i_1 \dots i_M}(S \cap S_m) = \kappa, i_{m+1} = i', \rho_{i_1 \dots i_M}(S \cap S_{m+1}) = \kappa'} \lambda_{i_1 \dots i_M} \right) \\
&= \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i, \rho_{i_1 \dots i_M}(S \cap S_m) = \kappa} \left(\sum_{i', \kappa': ((m, i, \rho_{i_1 \dots i_M}(S \cap S_m)), (m+1, i', \kappa')) \in \mathfrak{E}, i_{m+1} = i', \rho_{i_1 \dots i_M}(S \cap S_{m+1}) = \kappa'} \lambda_{i_1 \dots i_M} \right) \\
&= \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i, \rho_{i_1 \dots i_M}(S \cap S_m) = \kappa} \mathbb{I}\{((m, i_m, \rho_{i_1 \dots i_M}(S \cap S_m)), (m+1, i_{m+1}, \rho_{i_1 \dots i_M}(S \cap S_{m+1}))) \in \mathfrak{E}\} \lambda_{i_1 \dots i_M} \\
&= \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i, \rho_{i_1 \dots i_M}(S \cap S_m) = \kappa} \lambda_{i_1 \dots i_M} \\
&= g_{m, i, \kappa},
\end{aligned}$$

Indeed, the first equality follows from our construction of f . The second and third equalities follow from algebra. The fourth equality follows from Claim EC.11, which implies that $((m, i_m, \rho_{i_1 \dots i_M}(S \cap S_m)), (m+1, i_{m+1}, \rho_{i_1 \dots i_M}(S \cap S_{m+1}))) \in \mathfrak{E}$ for all $(i_1, \dots, i_M) \in \mathcal{L}$. The fifth equality follows from our construction of g .

- For all $(m, i, \kappa) \in \mathfrak{V}$ such that $m \in \{2, \dots, M\}$, we observe that

$$\begin{aligned}
& \sum_{i', \kappa': ((m-1, i'), (m, i, \kappa)) \in \mathfrak{E}} f_{m-1, i', \kappa', i, \kappa} \\
&= \sum_{i', \kappa': ((m-1, i'), (m, i, \kappa)) \in \mathfrak{E}} \left(\sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_{m-1} = i', \rho_{i_1 \dots i_M}(S \cap S_{m-1}) = \kappa', i_m = i, \rho_{i_1 \dots i_M}(S \cap S_m) = \kappa} \lambda_{i_1 \dots i_M} \right) \\
&= \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i, \rho_{i_1 \dots i_M}(S \cap S_m) = \kappa} \left(\sum_{i', \kappa': ((m-1, i'), (m, i, \rho_{i_1 \dots i_M}(S \cap S_m))) \in \mathfrak{E}, i_{m-1} = i', \rho_{i_1 \dots i_M}(S \cap S_{m-1}) = \kappa'} \lambda_{i_1 \dots i_M} \right) \\
&= \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i, \rho_{i_1 \dots i_M}(S \cap S_m) = \kappa} \mathbb{I}\{((m-1, i_{m-1}, \rho_{i_1 \dots i_M}(S \cap S_{m-1})), (m, i_m, \rho_{i_1 \dots i_M}(S \cap S_m))) \in \mathfrak{E}\} \lambda_{i_1 \dots i_M} \\
&= \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i, \rho_{i_1 \dots i_M}(S \cap S_m) = \kappa} \lambda_{i_1 \dots i_M} \\
&= g_{m, i, \kappa},
\end{aligned}$$

Indeed, the first equality follows from our construction of f . The second and third equalities follow from algebra. The fourth equality follows from Claim EC.11, which implies that $((m-1, i_{m-1}, \rho_{i_1 \dots i_M}(S \cap S_{m-1})), (m, i_m, \rho_{i_1 \dots i_M}(S \cap S_m))) \in \mathfrak{E}$ for all $(i_1, \dots, i_M) \in \mathcal{L}$. The fifth equality follows from our construction of g .

- Consider any vertex $(m, i, \kappa) \in \mathfrak{V}$ that satisfies $g_{m, i, \kappa} > 0$. Because $g_{m, i, \kappa} > 0$, it follows from our construction of g that there must exist a tuple $(i_1, \dots, i_M) \in \mathcal{L}$ that satisfies $i_m = i$, $\rho_{i_1 \dots i_M}(S \cap S_m) = \kappa$, and $\lambda_{i_1 \dots i_M} > 0$. Since Claim EC.11 implies that $(m, i_m, \rho_{i_1, \dots, i_M}(S \cap S_m)) \notin \mathfrak{N}(S)$ for all $m \in \mathcal{M}$ and all $(i_1, \dots, i_M) \in \mathcal{L}$, we conclude that

$$g_{m, i, \kappa} = 0 \quad \forall (m, i, \kappa) \in \mathfrak{N}(S).$$

- We observe that

$$\begin{aligned}
\sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} g_{M, i, \kappa} &= \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} \left(\sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_M = i, \rho_{i_1 \dots i_M}(S \cap S_M) = \kappa} \lambda_{i_1 \dots i_M} \right) \\
&= \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \lambda_{i_1 \dots i_M} \\
&= 1.
\end{aligned}$$

Indeed, the first equality follows from our construction of g . The second equality follows from Claim EC.11, which implies that $(M, i_M, \rho_{i_1 \dots i_M}(S \cap S_M)) \in \mathfrak{V}$ for all $(i_1, \dots, i_M) \in \mathcal{L}$. The third equality follows from the fact that (λ, ϵ) is a feasible solution for the linear optimization problem (2).

Since f, g are clearly nonnegative, the above bullet points show that (f, g, ϵ) satisfies each of the constraints in the linear optimization problem (6).

It remains for us to show that the objective value associated with (f, g, ϵ) in the linear optimization problem (6) is equal to the objective value associated with (λ, ϵ) in the linear optimization problem (2). Indeed, we observe that

$$\begin{aligned}
\sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} \kappa g_{M, i, \kappa} &= \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} \kappa \left(\sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_M = i, \rho_{i_1 \dots i_M}(S \cap S_M) = \kappa} \lambda_{i_1 \dots i_M} \right) \\
&= \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} \kappa \left(\sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_M = i, \rho_{i_1 \dots i_M}(S) = \kappa} \lambda_{i_1 \dots i_M} \right)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \left(\sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}, i_M = i, \rho_{i_1 \dots i_M}(S) = \kappa} \kappa \lambda_{i_1 \dots i_M} \right) \\
&= \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \mathbb{I}\{(M, i_M, \rho_{i_1 \dots i_M}(S)) \in \mathfrak{V}\} \rho_{i_1 \dots i_M}(S) \lambda_{i_1 \dots i_M} \\
&= \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \rho_{i_1 \dots i_M}(S) \lambda_{i_1 \dots i_M}.
\end{aligned}$$

The first equality follows from the construction of g . The second equality follows from Remark 3 and Assumption 1, which together imply that $S \cap S_M = S \cap \mathcal{N}_0 = S$. The third and fourth equalities follows from algebra. The fifth equality follows from Claim EC.11. Our proof of Lemma EC.6 is thus complete. \square

Our second intermediary lemma, denoted below by Lemma EC.7, will be used to show that the optimal objective value of the linear optimization problem (2) is less than or equal to the optimal objective value of the linear optimization problem (6).

LEMMA EC.7. *Let Assumption 1 hold, let $S \subseteq \mathcal{N}_0$ satisfy $S \cap S_1 \neq \emptyset$, and let (f, g, ϵ) satisfy the following constraints:*

$$\begin{aligned}
\sum_{\kappa: (m, i, \kappa) \in \mathfrak{V}} g_{m, i, \kappa} - \epsilon_{m, i} &= v_{m, i} & \forall m \in \mathcal{M}, i \in S_m \\
\sum_{i', \kappa': ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E}} f_{m, i, \kappa, i', \kappa'} &= g_{m, i, \kappa} & \forall (m, i, \kappa) \in \mathfrak{V}: m \in \{1, \dots, M-1\} \\
\sum_{i', \kappa': ((m-1, i', \kappa'), (m, i, \kappa)) \in \mathfrak{E}} f_{m-1, i', \kappa', i, \kappa} &= g_{m, i, \kappa} & \forall (m, i, \kappa) \in \mathfrak{V}: m \in \{2, \dots, M\} \\
\|\epsilon\| &\leq \eta \\
\sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} g_{M, i, \kappa} &= 1 \\
f_{m, i, \kappa, i', \kappa'} &\geq 0 & \forall ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E} \\
g_{m, i, \kappa} &\geq 0 & \forall (m, i, \kappa) \in \mathfrak{V}.
\end{aligned}$$

Then there exists a vector λ such that (λ, ϵ) is a feasible solution for the linear optimization problem (6) and

$$\sum_{(i_1, \dots, i_M) \in \mathcal{L}} \rho_{i_1 \dots i_M}(S) \lambda_{i_1 \dots i_M} \leq \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} \kappa g_{m, i, \kappa} + \sum_{(m, i, \kappa) \in \mathfrak{V}(S)} r_n g_{m, i, \kappa}.$$

Proof. Let Assumption 1 hold, let $S \subseteq \mathcal{N}_0$ satisfy $S \cap S_1 \neq \emptyset$, and let (f, g, ϵ) satisfy the following constraints:

$$\sum_{\kappa: (m, i, \kappa) \in \mathfrak{V}} g_{m, i, \kappa} - \epsilon_{m, i} = v_{m, i} \quad \forall m \in \mathcal{M}, i \in S_m \quad (\text{EC.43a})$$

$$\sum_{i', \kappa': ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E}} f_{m, i, \kappa, i', \kappa'} = g_{m, i, \kappa} \quad \forall (m, i, \kappa) \in \mathfrak{V}: m \in \{1, \dots, M-1\} \quad (\text{EC.43b})$$

$$\sum_{i', \kappa': ((m-1, i', \kappa'), (m, i, \kappa)) \in \mathfrak{E}} f_{m-1, i', \kappa', i, \kappa} = g_{m, i, \kappa} \quad \forall (m, i, \kappa) \in \mathfrak{V}: m \in \{2, \dots, M\} \quad (\text{EC.43c})$$

$$\|\epsilon\| \leq \eta \quad (\text{EC.43d})$$

$$\sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} g_{M, i, \kappa} = 1 \quad (\text{EC.43e})$$

$$f_{m, i, \kappa, i', \kappa'} \geq 0 \quad \forall ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E} \quad (\text{EC.43f})$$

$$g_{m, i, \kappa} \geq 0 \quad \forall (m, i, \kappa) \in \mathfrak{V}. \quad (\text{EC.43g})$$

We observe that f can be interpreted as a network flow through the directed acyclic graph $\mathfrak{G} \equiv (\mathfrak{V}, \mathfrak{E})$, in the sense that $f_{m, i, \kappa, i', \kappa'} \geq 0$ is the units of flow from vertex $(m, i, \kappa) \in \mathfrak{V}$ to vertex $(m+1, i', \kappa') \in \mathfrak{V}$, and the constraints of the form $\sum_{i', \kappa': ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E}} f_{m, i, \kappa, i', \kappa'} = g_{m, i, \kappa}$ and $\sum_{i', \kappa': ((m-1, i', \kappa'), (m, i, \kappa)) \in \mathfrak{E}} f_{m-1, i', \kappa', i, \kappa} =$

$g_{m,i,\kappa}$ ensure that flow is conserved through each vertex. Because f can be interpreted as a network flow in the directed acyclic graph $\mathfrak{G} \equiv (\mathfrak{V}, \mathfrak{E})$, it follows from the flow decomposition theorem (Ahuja et al. 1988, Theorem 3.5) that there exists a nonnegative vector $\pi \equiv (\pi_{i_1 \kappa_1 \dots i_M \kappa_M} : ((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}) \geq 0$ that satisfies the following equalities:

$$\begin{aligned} g_{m,i,\kappa} &= \sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}: i_m = i, \kappa_m = \kappa} \pi_{i_1 \kappa_1 \dots i_M \kappa_M} & \forall (m, i, \kappa) \in \mathfrak{V} & \quad (\text{EC.44}) \\ f_{m,i,\kappa,i',\kappa'} &= \sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}: i_m = i, \kappa_m = \kappa, i_{m+1} = i', \kappa_{m+1} = \kappa'} \pi_{i_1 \kappa_1 \dots i_M \kappa_M} & ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E}. \end{aligned}$$

From this point onward, consider any fixed choice of the nonnegative vector π that satisfies the above equalities. Given that choice of π , let the vector $\lambda \equiv (\pi_{i_1 \dots i_M} : (i_1, \dots, i_M) \in \mathcal{L})$ be defined as follows:

$$\lambda_{i_1 \dots i_M} \triangleq \sum_{\kappa_1, \dots, \kappa_M: ((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}} \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \quad \forall (i_1, \dots, i_M) \in \mathcal{L}. \quad (\text{EC.45})$$

Our first main step of the proof of Lemma EC.7 consists of showing that (λ, ϵ) is a feasible solution for the linear optimization problem (2). To show that (λ, ϵ) is a feasible solution for the linear optimization problem (2), we will utilize the following property of \mathcal{P} :

CLAIM EC.12. *If $((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}$, then $(i_1, \dots, i_M) \in \mathcal{L}$.*

Proof of Claim EC.12. Consider any path $((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}$. We readily observe from the construction of the graph $\mathfrak{G} \equiv (\mathfrak{V}, \mathfrak{E})$ and from the fact that $(1, i_1, \kappa_1) \in \mathfrak{V}$ that $i_1 \in S_1$. Next, choose any arbitrary $m \in \{2, \dots, M\}$. Because $(1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)$ is a path, we observe that there must exist a directed edge in the graph $\mathfrak{G} \equiv (\mathfrak{V}, \mathfrak{E})$ from vertex $(m-1, i_{m-1}, \kappa_{m-1}) \in \mathfrak{V}$ to vertex $(m, i_m, \kappa_m) \in \mathfrak{V}$. Therefore, it follows from the definition of \mathfrak{E} that the inclusion $i_m \in \{i_{m-1}\} \cup \mathcal{B}_m$ must hold. Because $m \in \{2, \dots, M\}$ was chosen arbitrarily, we conclude that $i_1 \in S_1$ and $i_m \in \{i_{m-1}\} \cup \mathcal{B}_m$ for all $m \in \{2, \dots, M\}$, and so Lemma 10 implies that $(i_1, \dots, i_M) \in \mathcal{L}$. Our proof of Claim EC.12 is thus complete. \square

Equipped with Claim EC.12, we now proceed to show that (λ, ϵ) is a feasible solution for the linear optimization problem (2). Indeed, for each $m \in \mathcal{M}$ and $i \in S_m$, we observe that

$$\begin{aligned} v_{m,i} &= \sum_{\kappa: (m,i,\kappa) \in \mathfrak{V}} g_{m,i,\kappa} - \epsilon_{m,i} \\ &= \sum_{\kappa: (m,i,\kappa) \in \mathfrak{V}} \left(\sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}: i_m = i, \kappa_m = \kappa} \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \right) - \epsilon_{m,i} \\ &= \sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}: i_m = i} \pi_{i_1 \kappa_1 \dots i_M \kappa_M} - \epsilon_{m,i} \\ &= \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i} \left(\sum_{\kappa_1, \dots, \kappa_M: ((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}} \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \right) - \epsilon_{m,i} \\ &= \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i} \lambda_{i_1 \dots i_M} - \epsilon_{m,i}, \end{aligned}$$

where the first equality follows from the fact that (f, g, ϵ) satisfies the constraints (EC.43), the second equality follows from line (EC.44), the third equality follows from algebra, the fourth equality follows from algebra and Claim EC.12, and the fifth equality follows from our construction of λ on line (EC.45). Moreover, we observe that

$$\begin{aligned} 1 &= \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} g_{M, i, \kappa} \\ &= \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} \sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}: i_M = i, \kappa_M = \kappa} \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \\ &= \sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}} \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \end{aligned}$$

$$\begin{aligned}
&= \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \left(\sum_{\kappa_1, \dots, \kappa_M: ((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}} \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \right) \\
&= \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \lambda_{i_1 \dots i_M},
\end{aligned}$$

where the first equality follows from the fact that (f, g, ϵ) satisfies the constraints (EC.43), the second equality follows from line (EC.44), the third equality follows from algebra, the fourth equality follows from algebra and Claim EC.12, and the fifth equality follows from our construction of λ on line (EC.45). Since λ is clearly nonnegative, our proof that (λ, ϵ) is a feasible solution for the linear optimization problem (2) is complete.

The remainder of the proof of Lemma EC.7 consists of bounding the objective value associated with (λ, ϵ) in the linear optimization problem (2). To develop our bound, we establish the following key property about directed acyclic graph $\mathfrak{G} \equiv (\mathfrak{V}, \mathfrak{E})$.

CLAIM EC.13. *If $((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}$ and $(m, i_m, \kappa_m) \notin \mathfrak{N}(S)$ for all $m \in \mathcal{M}$, then $\kappa_m \geq \rho_{i_1 \dots i_M}(S \cap S_m)$ for all $m \in \mathcal{M}$.*

Proof of Claim EC.13. Consider any path $((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}$ that satisfies $(m, i_m, \kappa_m) \notin \mathfrak{N}(S)$ for all $m \in \mathcal{M}$. Our proof of Claim EC.13 follows from an induction argument, and the induction argument proceeds as follows.

In the base case, we let $m = 1$. In this case, it follows from the fact that $((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}$ that the inclusion $(1, i_1, \kappa_1) \in \mathfrak{V}$ holds, and so it follows from the definition of \mathfrak{V} that $\kappa_1 \in \{r_j : j \in \mathcal{B}_1\}$. Thus, it follows from the facts that $\kappa_1 \in \{r_j : j \in \mathcal{B}_1\}$ and $(1, i_1, \kappa_1) \notin \mathfrak{N}(S)$ that

$$\kappa_1 \in \begin{cases} \{r_{i_1}\}, & \text{if } i_1 \in S, \\ \{r_j : j \in \mathcal{B}_1 \cap S\}, & \text{if } i_1 \notin S. \end{cases} \quad (\text{EC.46})$$

Therefore, we observe that

$$\begin{aligned}
\kappa_1 &\geq \begin{cases} r_{i_1}, & \text{if } i_1 \in S, \\ \min_{j \in \mathcal{B}_1 \cap S} r_j, & \text{if } i_1 \notin S \end{cases} \\
&= \rho_{i_1 \dots i_M}(S \cap S_1),
\end{aligned}$$

where the inequality follows from line (EC.46) and the equality follows from Claim EC.12 and Lemma 11. We have shown that $\kappa_1 \geq \rho_{i_1 \dots i_M}(S \cap S_1)$, which concludes our proof of the base case.

Now choose any arbitrary $m \in \{2, \dots, M\}$ and assume by induction that the inequality $\kappa_{m'} \geq \rho_{i_1 \dots i_M}(S \cap S_{m'})$ holds for all $m' \in \{1, \dots, m-1\}$. In this case, it follows from the fact that $((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}$ that the inclusion $((m-1, i_{m-1}, \kappa_{m-1}), (m, i_m, \kappa_m)) \in \mathfrak{E}$ holds, and so it follows from the definition of \mathfrak{E} that $\kappa_m \in \{\kappa_{m-1}\} \cup \{r_j : j \in \mathcal{B}_m\}$. Thus, it follows from the facts that $\kappa_m \in \{\kappa_{m-1}\} \cup \{r_j : j \in \mathcal{B}_m\}$ and $(m, i_m, \kappa_m) \notin \mathfrak{N}(S)$ that

$$\kappa_m \in \begin{cases} \{r_{i_m}\}, & \text{if } i_m \in S \text{ and } i_m \in \mathcal{B}_m, \\ \{\kappa_{m-1}\}, & \text{if } i_m \in S \text{ and } i_m \notin \mathcal{B}_m, \\ \{\kappa_{m-1}\} \cup \{r_j : j \in \mathcal{B}_m \cap S\}, & \text{if } i_m \notin S. \end{cases} \quad (\text{EC.47})$$

Therefore,

$$\begin{aligned}
\kappa_m &\geq \begin{cases} r_{i_m}, & \text{if } i_m \in S \text{ and } i_m \in \mathcal{B}_m, \\ \kappa_{m-1}, & \text{if } i_m \in S \text{ and } i_m \notin \mathcal{B}_m, \\ \min \left\{ \kappa_{m-1}, \min_{j \in \mathcal{B}_m \cap S} r_j \right\}, & \text{if } i_m \notin S \end{cases} \\
&\geq \begin{cases} r_{i_m}, & \text{if } i_m \in S \text{ and } i_m \in \mathcal{B}_m, \\ \rho_{i_1 \dots i_M}(S \cap S_{m-1}), & \text{if } i_m \in S \text{ and } i_m \notin \mathcal{B}_m, \\ \min \left\{ \rho_{i_1 \dots i_M}(S \cap S_{m-1}), \min_{j \in \mathcal{B}_m \cap S} r_j \right\}, & \text{if } i_m \notin S \end{cases} \\
&= \rho_{i_1 \dots i_M}(S \cap S_m),
\end{aligned}$$

where the first inequality follows from line (EC.47), the second inequality follows from the induction hypothesis, and the equality follows from Claim EC.12 and Lemma 11. We have thus shown that $\kappa_m \geq \rho_{i_1 \dots i_M}(S \cap S_m)$. This concludes our proof of the induction step, and thus our proof of Claim EC.13 is complete. \square

Equipped with Claims EC.12 and EC.13, we now develop the desired bound on the objective value associated with (λ, ϵ) in the linear optimization problem (2):

$$\begin{aligned} & \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \rho_{i_1 \dots i_M}(S) \lambda_{i_1 \dots i_M} \\ &= \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \rho_{i_1 \dots i_M}(S) \left(\sum_{(\kappa_1, \dots, \kappa_M) : ((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}} \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \right) \end{aligned} \quad (\text{EC.48})$$

$$= \sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}} \rho_{i_1 \dots i_M}(S) \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \quad (\text{EC.49})$$

$$= \sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}} \rho_{i_1 \dots i_M}(S \cap S_M) \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \quad (\text{EC.50})$$

$$\begin{aligned} &= \left(\sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P} : (m, i_m, \kappa_m) \notin \mathfrak{N}(S) \ \forall m \in \mathcal{M}} \rho_{i_1 \dots i_M}(S \cap S_M) \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \right) \\ &+ \left(\sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P} : \exists m \in \mathcal{M} \text{ such that } (m, i_m, \kappa_m) \in \mathfrak{N}(S)} \rho_{i_1 \dots i_M}(S \cap S_M) \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \right) \end{aligned} \quad (\text{EC.51})$$

$$\begin{aligned} &\leq \left(\sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P} : (m, i_m, \kappa_m) \notin \mathfrak{N}(S) \ \forall m \in \mathcal{M}} \kappa_M \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \right) \\ &+ \left(\sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P} : \exists m \in \mathcal{M} \text{ such that } (m, i_m, \kappa_m) \in \mathfrak{N}(S)} \rho_{i_1 \dots i_M}(S \cap S_M) \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \right) \end{aligned} \quad (\text{EC.52})$$

$$\begin{aligned} &\leq \left(\sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P} : (m, i_m, \kappa_m) \notin \mathfrak{N}(S) \ \forall m \in \mathcal{M}} \kappa_M \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \right) \\ &+ \left(\sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P} : \exists m \in \mathcal{M} \text{ such that } (m, i_m, \kappa_m) \in \mathfrak{N}(S)} r_n \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \right) \end{aligned} \quad (\text{EC.53})$$

$$\begin{aligned} &\leq \left(\sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P} : (m, i_m, \kappa_m) \notin \mathfrak{N}(S) \ \forall m \in \mathcal{M}} \kappa_M \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \right) \\ &+ \left(\sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}} \sum_{m \in \mathcal{M} : (m, i_m, \kappa_m) \in \mathfrak{N}(S)} r_n \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \right) \end{aligned} \quad (\text{EC.54})$$

$$\begin{aligned} &= \left(\sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P} : (m, i_m, \kappa_m) \notin \mathfrak{N}(S) \ \forall m \in \mathcal{M}} \kappa_M \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \right) \\ &+ \left(\sum_{(m, i, \kappa) \in \mathfrak{N}(S)} r_n \sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P} : i_m = i, \kappa_m = \kappa} \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \right) \end{aligned} \quad (\text{EC.55})$$

$$= \left(\sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P} : (m, i_m, \kappa_m) \notin \mathfrak{N}(S) \ \forall m \in \mathcal{M}} \kappa_M \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \right) + \sum_{(m, i, \kappa) \in \mathfrak{N}(S)} g_{m, i, \kappa} \quad (\text{EC.56})$$

$$\leq \left(\sum_{((1, i_1, \kappa_1), \dots, (M, i_M, \kappa_M)) \in \mathcal{P}} \kappa_M \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \right) + \sum_{(m, i, \kappa) \in \mathfrak{N}(S)} g_{m, i, \kappa} \quad (\text{EC.57})$$

$$= \left(\sum_{i,\kappa:(M,i,\kappa) \in \mathfrak{B}} \kappa \sum_{((1,i_1,\kappa_1), \dots, (M,i_M,\kappa_M)) \in \mathcal{P}: i_M=i, \kappa_M=\kappa} \pi_{i_1 \kappa_1 \dots i_M \kappa_M} \right) + \sum_{(m,i,\kappa) \in \mathfrak{N}(S)} g_{m,i,\kappa} \quad (\text{EC.58})$$

$$= \sum_{i,\kappa:(M,i,\kappa) \in \mathfrak{B}} \kappa g_{M,i,\kappa} + \sum_{(m,i,\kappa) \in \mathfrak{N}(S)} g_{m,i,\kappa}. \quad (\text{EC.59})$$

Indeed, line (EC.48) follows from our construction of λ on line (EC.45). Line (EC.49) follows from algebra and Claim EC.12. Line (EC.50) follows from Assumption 1 and Remark 3. Line (EC.51) follows from algebra. Line (EC.52) follows from Claim EC.13. Line (EC.53) follows from the fact that $\rho_{i_1 \dots i_M}(S \cap S_M) \leq r_n$ for all $(i_1, \dots, i_M) \in \mathcal{L}$. Lines (EC.54) and (EC.55) follows from algebra. Line (EC.56) follows from line (EC.44). Lines (EC.57) and (EC.58) follow from algebra. Line (EC.59) follows from line (EC.44). Our proof of Lemma EC.7 is thus complete. \square

I.5. Proof of Proposition 3

Let Assumption 1 hold, and consider any assortment $S \subseteq \mathcal{N}_0$ that satisfies $S \cap S_1 \neq \emptyset$. It follows immediately from Lemma EC.6 from Appendix I.4 that the optimal objective value of the linear optimization problem (2) is greater than or equal to the optimal objective value of the linear optimization problem (6). To show the other direction, consider any arbitrary feasible solution (f, g, ϵ) for the linear optimization problem (6). It follows from Lemma EC.7 from Appendix I.4 that there exists a vector λ such that (λ, ϵ) is a feasible solution for the linear optimization problem (2) and

$$\sum_{(i_1, \dots, i_M) \in \mathcal{L}} \rho_{i_1 \dots i_M}(S) \lambda_{i_1 \dots i_M} \leq \sum_{i,\kappa:(M,i,\kappa) \in \mathfrak{B}} \kappa g_{m,i,\kappa} + \sum_{(m,i,\kappa) \in \mathfrak{N}(S)} r_n g_{m,i,\kappa},$$

where the definition of the set $\mathfrak{N}(S)$ can be found at the beginning of Appendix I.4. In particular, it follows from the fact that (f, g, ϵ) is feasible for the linear optimization problem (6) that the equality $g_{m,i,\kappa} = 0$ holds for all $(m, i, \kappa) \in \mathfrak{N}(S)$. Therefore, we conclude that

$$\begin{aligned} \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \rho_{i_1 \dots i_M}(S) \lambda_{i_1 \dots i_M} &\leq \sum_{i,\kappa:(M,i,\kappa) \in \mathfrak{B}} \kappa g_{m,i,\kappa} + \sum_{(m,i,\kappa) \in \mathfrak{N}(S)} r_n 0 \\ &= \sum_{i,\kappa:(M,i,\kappa) \in \mathfrak{B}} \kappa g_{m,i,\kappa}. \end{aligned}$$

Since (f, g, ϵ) was chosen arbitrarily, we conclude that the optimal objective value of the linear optimization problem (6) is greater than or equal to the optimal objective value of the linear optimization problem (2). Our proof of Proposition 3 is thus complete. \square

I.6. Proof of Theorem 5

Let Assumption 1 hold and $S_1, \dots, S_M \in \mathcal{S}$. It follows from the fact that $S_1 \in \mathcal{S}$ that $S \cap S_1 \neq \emptyset$ for all $S \in \mathcal{S}$. Therefore, it follows from Proposition 3 for each assortment $S \in \mathcal{S}$ that the worst-case expected revenue

$\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ is equal to the optimal objective value of the following linear optimization problem:

$$\begin{aligned}
& \underset{f, g, \epsilon}{\text{minimize}} && \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} \kappa g_{M, i, \kappa} \\
& \text{subject to} && \sum_{\kappa: (m, i, \kappa) \in \mathfrak{V}} g_{m, i, \kappa} - \epsilon_{m, i} = v_{m, i} && \forall m \in \mathcal{M}, i \in S_m \\
& && \sum_{i', \kappa': ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E}} f_{m, i, \kappa, i', \kappa'} = g_{m, i, \kappa} && \forall (m, i, \kappa) \in \mathfrak{V}: m \in \{1, \dots, M-1\} \\
& && \sum_{i', \kappa': ((m-1, i', \kappa'), (m, i, \kappa)) \in \mathfrak{E}} f_{m-1, i', \kappa', i, \kappa} = g_{m, i, \kappa} && \forall (m, i, \kappa) \in \mathfrak{V}: m \in \{2, \dots, M\} \\
& && g_{m, i, \kappa} = 0 && \forall (m, i, \kappa) \in \mathfrak{N}(S) \\
& && \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} g_{M, i, \kappa} = 1 \\
& && \|\epsilon\| \leq \eta \\
& && f_{m, i, \kappa, i', \kappa'} \geq 0 && \forall ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E} \\
& && g_{m, i, \kappa} \geq 0 && \forall (m, i, \kappa) \in \mathfrak{V},
\end{aligned} \tag{6}$$

where the definition of the set $\mathfrak{N}(S)$ can be found in the beginning of Appendix I.4.

At a high level, our following analysis will consist of two main steps. First, we will use strong duality to transform the linear optimization problem (6) into a maximization problem. Second, we will introduce binary decision variables to represent the assortment $S \in \mathcal{S}$. The key challenge in performing these two main steps is ensuring that the resulting mixed-integer optimization problem does not have any nonlinear terms. To see why this is a challenge, suppose for the sake of discussion that we introduced a binary decision variable $x_i \in \{0, 1\}$ for each product $i \in \mathcal{N}_0$ that satisfies $x_i = 1$ if and only if product i is in the assortment S . Under this choice of binary decision variables, it is easy to verify using the definition of $\mathfrak{N}(S)$ from the beginning of Appendix I.4 that the constraints in the linear optimization problem (6) of the form

$$g_{m, i, \kappa} = 0 \quad \forall (m, i, \kappa) \in \mathfrak{N}(S)$$

will be satisfied if and only if

$$\begin{aligned}
g_{m, i, \kappa} &\leq 1 - x_i && \forall (m, i, \kappa) \in \mathfrak{V}: i \in \mathcal{B}_m \text{ and } \kappa \neq r_i \\
g_{m, i, \kappa} &\leq 1 - x_i && \forall (m, i, \kappa) \in \mathfrak{V}: i \notin \mathcal{B}_m \text{ and } \kappa \in \{r_j: j \in \mathcal{B}_m\} \\
g_{m, i, \kappa} &\leq x_j && \forall (m, i, \kappa) \in \mathfrak{V} \text{ and } j \in \mathcal{B}_m: \kappa = r_j.
\end{aligned}$$

However, we observe that the binary decision variables $x \in \{0, 1\}^{\mathcal{N}_0}$ appear on the right hand side of the above constraints. Therefore, if one added the above constraints into the linear optimization problem (6), then the dual of the linear optimization problem (6) would have an objective function in which dual decision variables are multiplied by the binary decision variables x .

To circumvent the aforementioned challenge and obtain a mixed-integer *linear* optimization reformulation of (RO), we will make use of the following Claim EC.14. Specifically, the following Claim EC.14 shows that the constraints of the form $g_{m, i, \kappa} = 0 \forall (m, i, \kappa) \in \mathfrak{N}(S)$ can be moved into the objective function of the linear optimization problem (6) without loss of generality. By using the following claim, we will subsequently be able to obtain a dual of the linear optimization problem (6) in which the binary decision variables $x \in \{0, 1\}^{\mathcal{N}_0}$ are not multiplied by dual decision variables. The proof of the following Claim EC.14 follows from similar reasoning as the proof of Proposition 3.

CLAIM EC.14. *Consider any assortment $S \in \mathcal{S}$, and let $\mathfrak{M}(S) \in \mathbb{R}^{\mathfrak{V}}$ be any vector that satisfies the inequality $\mathfrak{M}_{m, i, \kappa}(S) \geq r_n$ for all $(m, i, \kappa) \in \mathfrak{N}(S)$ and satisfies the equality $\mathfrak{M}_{m, i, \kappa}(S) = 0$ for all $(m, i, \kappa) \in \mathfrak{V} \setminus \mathfrak{N}(S)$.*

Then the optimal objective value of the linear optimization problem (6) is equal to the optimal objective value of the following linear optimization problem:

$$\begin{aligned}
& \underset{f, g, \epsilon}{\text{minimize}} && \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} \kappa g_{M, i, \kappa} + \sum_{(m, i, \kappa) \in \mathfrak{V}} \mathfrak{M}_{m, i, \kappa}(S) g_{m, i, \kappa} \\
& \text{subject to} && \sum_{\kappa: (m, i, \kappa) \in \mathfrak{V}} g_{m, i, \kappa} - \epsilon_{m, i} = v_{m, i} && \forall m \in \mathcal{M}, i \in S_m \\
& && \sum_{i', \kappa': ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E}} f_{m, i, \kappa, i', \kappa'} = g_{m, i, \kappa} && \forall (m, i, \kappa) \in \mathfrak{V} : m \in \{1, \dots, M-1\} \\
& && \sum_{i', \kappa': ((m-1, i', \kappa'), (m, i, \kappa)) \in \mathfrak{E}} f_{m-1, i', \kappa', i, \kappa} = g_{m, i, \kappa} && \forall (m, i, \kappa) \in \mathfrak{V} : m \in \{2, \dots, M\} \\
& && \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} g_{M, i, \kappa} = 1 && \\
& && \|\epsilon\| \leq \eta && \\
& && f_{m, i, \kappa, i', \kappa'} \geq 0 && \forall ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E} \\
& && g_{m, i, \kappa} \geq 0 && \forall (m, i, \kappa) \in \mathfrak{V}.
\end{aligned} \tag{EC.60}$$

Proof of Claim EC.14. Consider any assortment $S \in \mathcal{S}$, and let $\mathfrak{M}(S) \in \mathbb{R}^{\mathfrak{V}}$ be any vector that satisfies the inequality $\mathfrak{M}_{m, i, \kappa}(S) \geq r_n$ for all $(m, i, \kappa) \in \mathfrak{N}(S)$ and satisfies the equality $\mathfrak{M}_{m, i, \kappa}(S) = 0$ for all $(m, i, \kappa) \in \mathfrak{V} \setminus \mathfrak{N}(S)$. We readily observe that the optimal objective value of the linear optimization problem (EC.60) is a lower bound on the optimal objective value of the linear optimization problem (6), since the constraints of the form $g_{m, i, \kappa} = 0 \forall (m, i, \kappa) \in \mathfrak{N}(S)$ from the linear optimization problem (6) have been relaxed and since the equality $\mathfrak{M}_{m, i, \kappa}(S) = 0$ holds for all $(m, i, \kappa) \in \mathfrak{V} \setminus \mathfrak{N}(S)$. To show the other direction, consider any arbitrary feasible solution (f, g, ϵ) for the linear optimization problem (EC.60). Because Assumption 1 holds and $S_1 \cap S \neq \emptyset$, it follows from Lemma EC.7 from Appendix I.4 that there exists a vector λ such that (λ, ϵ) is a feasible solution for the linear optimization problem (2) and

$$\begin{aligned}
\sum_{(i_1, \dots, i_M) \in \mathcal{L}} \rho_{i_1 \dots i_M}(S) \lambda_{i_1 \dots i_M} &\leq \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} \kappa g_{m, i, \kappa} + \sum_{(m, i, \kappa) \in \mathfrak{N}(S)} r_n g_{m, i, \kappa} \\
&\leq \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} \kappa g_{m, i, \kappa} + \sum_{(m, i, \kappa) \in \mathfrak{V}} \mathfrak{M}(S) g_{m, i, \kappa},
\end{aligned}$$

where the first inequality follows from Lemma EC.7 and the second inequality follows from our construction of the vector $\mathfrak{M}(S)$. Since (f, g, ϵ) was chosen arbitrarily, we have shown that the optimal objective value of the linear optimization problem (2) is a lower bound on the optimal objective value of the linear optimization problem (EC.60). Since we showed in Proposition 3 that the optimal objective value of the linear optimization problem (EC.60) is equal to the optimal objective value of the linear optimization problem (6), we conclude that the optimal objective value of the linear optimization problem (6) is a lower bound on the optimal objective value of the linear optimization problem (EC.60). Our proof of Claim EC.14 is thus complete. \square

To simplify our exposition, the remainder of the proof of Theorem 5 focuses on the case where $\eta = 0$. Indeed, it is easy to see that our reformulations of the robust optimization problem (RO) for the case of $\eta = 0$ can be readily extended to the case of $\eta > 0$ by introducing $\mathcal{O}(\text{poly}(n, M))$ auxiliary decision variables and auxiliary linear constraints in the dual problem. In the case of $\eta = 0$, we observe that the linear optimization

problem (EC.60) can be rewritten equivalently as

$$\begin{aligned}
& \underset{f, g}{\text{minimize}} && \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} \kappa g_{M, i, \kappa} + \sum_{(m, i, \kappa) \in \mathfrak{V}} \mathfrak{M}_{m, i, \kappa}(S) g_{m, i, \kappa} \\
& \text{subject to} && \sum_{\kappa: (m, i, \kappa) \in \mathfrak{V}} g_{m, i, \kappa} = v_{m, i} && \forall m \in \mathcal{M}, i \in S_m \\
& && \sum_{i', \kappa': ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E}} f_{m, i, \kappa, i', \kappa'} = g_{m, i, \kappa} && \forall (m, i, \kappa) \in \mathfrak{V}: m \in \{1, \dots, M-1\} \\
& && \sum_{i', \kappa': ((m-1, i', \kappa'), (m, i, \kappa)) \in \mathfrak{E}} f_{m-1, i', \kappa', i, \kappa} = g_{m, i, \kappa} && \forall (m, i, \kappa) \in \mathfrak{V}: m \in \{2, \dots, M\} \\
& && f_{m, i, \kappa, i', \kappa'} \geq 0 && \forall ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E} \\
& && g_{m, i, \kappa} \geq 0 && \forall (m, i, \kappa) \in \mathfrak{V}.
\end{aligned} \tag{EC.60}$$

We observe from strong duality that the optimal objective value of the linear optimization problem (EC.60) is thus equal to the optimal objective value of the following linear optimization problem (EC.61) for every assortment $S \in \mathcal{S}$ and every choice of $\mathfrak{M}(S) \in \mathbb{R}^{\mathfrak{V}}$ that satisfies the inequality $\mathfrak{M}_{m, i, \kappa}(S) \geq r_n$ for all $(m, i, \kappa) \in \mathfrak{N}(S)$ and satisfies the equality $\mathfrak{M}_{m, i, \kappa}(S) = 0$ for all $(m, i, \kappa) \in \mathfrak{V} \setminus \mathfrak{N}(S)$:

$$\begin{aligned}
& \underset{\alpha, \beta, \gamma}{\text{maximize}} && \sum_{m \in \mathcal{M}, i \in S_m} v_{m, i} \alpha_{m, i} \\
& \text{subject to} && \alpha_{m, i} - \beta_{m, i, \kappa} \mathbb{I}\{m \in \{1, \dots, M-1\}\} - \gamma_{m, i, \kappa} \mathbb{I}\{m \in \{2, \dots, M\}\} \\
& && \leq \kappa \mathbb{I}\{m = M\} + \mathfrak{M}_{m, i, \kappa}(S) && \forall (m, i, \kappa) \in \mathfrak{V} \\
& && \beta_{m, i, \kappa} + \gamma_{m+1, i', \kappa'} \leq 0 && \forall ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E} \\
& && \alpha_{m, i} \in \mathbb{R} && \forall m \in \mathcal{M}, i \in S_m \\
& && \beta_{m, i, \kappa}, \gamma_{m, i, \kappa} \in \mathbb{R} && \forall (m, i, \kappa) \in \mathfrak{V}.
\end{aligned} \tag{EC.61}$$

Consequently, if $\mathfrak{M}(\cdot)$ satisfies the inequality $\mathfrak{M}_{m, i, \kappa}(S) \geq r_n$ for all $(m, i, \kappa) \in \mathfrak{N}(S)$ and satisfies the equality $\mathfrak{M}_{m, i, \kappa}(S) = 0$ for all $(m, i, \kappa) \in \mathfrak{V} \setminus \mathfrak{N}(S)$ and all assortments $S \in \mathcal{S}$, then

(RO)

$$\begin{aligned}
& = \max_{S \in \mathcal{S}} \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) \\
& = \max_{S \in \mathcal{S}} \left[\begin{array}{l} \underset{\alpha, \beta, \gamma}{\text{maximize}} \quad \sum_{m \in \mathcal{M}, i \in S_m} v_{m, i} \alpha_{m, i} \\ \text{subject to} \quad \alpha_{m, i} - \beta_{m, i, \kappa} \mathbb{I}\{m \in \{1, \dots, M-1\}\} - \gamma_{m, i, \kappa} \mathbb{I}\{m \in \{2, \dots, M\}\} \\ \quad \leq \kappa \mathbb{I}\{m = M\} + \mathfrak{M}_{m, i, \kappa}(S) && \forall (m, i, \kappa) \in \mathfrak{V} \\ \quad \beta_{m, i, \kappa} + \gamma_{m+1, i', \kappa'} \leq 0 && \forall ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E} \\ \quad \alpha_{m, i} \in \mathbb{R} && \forall m \in \mathcal{M}, i \in S_m \\ \quad \beta_{m, i, \kappa}, \gamma_{m, i, \kappa} \in \mathbb{R} && \forall (m, i, \kappa) \in \mathfrak{V} \end{array} \right] \\
& = \left[\begin{array}{l} \underset{S, \alpha, \beta, \gamma}{\text{maximize}} \quad \sum_{m \in \mathcal{M}, i \in S_m} v_{m, i} \alpha_{m, i} \\ \text{subject to} \quad \alpha_{m, i} - \beta_{m, i, \kappa} \mathbb{I}\{m \in \{1, \dots, M-1\}\} - \gamma_{m, i, \kappa} \mathbb{I}\{m \in \{2, \dots, M\}\} \\ \quad \leq \kappa \mathbb{I}\{m = M\} + \mathfrak{M}_{m, i, \kappa}(S) && \forall (m, i, \kappa) \in \mathfrak{V} \\ \quad \beta_{m, i, \kappa} + \gamma_{m+1, i', \kappa'} \leq 0 && \forall ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E} \\ \quad \alpha_{m, i} \in \mathbb{R} && \forall m \in \mathcal{M}, i \in S_m \\ \quad \beta_{m, i, \kappa}, \gamma_{m, i, \kappa} \in \mathbb{R} && \forall (m, i, \kappa) \in \mathfrak{V} \\ \quad S \in \mathcal{S} \end{array} \right],
\end{aligned} \tag{EC.62}$$

where the first equality follows from the definition of the robust optimization problem (RO), the second equality follows from Claim EC.14 (which implies that $\min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ is equal to the optimal objective value

of the linear optimization problem (EC.61) for all assortments $S \in \mathcal{S}$, and the third equality follows from algebra.

To transform the maximization problem in line (EC.62) into a mixed-integer optimization problem, we introduce a binary decision variable $x_i \in \{0, 1\}$ for each product $i \in \mathcal{N}_0$ that satisfies $x_i = 1$ if and only if product i is in the assortment S . Under this equivalence between an assortment $S \in \mathcal{S}$ and the binary decision variables $x \in \{0, 1\}^{\mathcal{N}_0}$, it follows from the definition of $\mathfrak{N}(\cdot)$ that the following equality holds for all assortments $S \in \mathcal{S}$:

$$\mathfrak{N}(S) \triangleq \left\{ \begin{array}{l} [x_i = 1 \text{ and } i \in \mathcal{B}_m \text{ and } \kappa \neq r_i] \\ (m, i, \kappa) \in \mathfrak{V}: \quad \text{or } [x_i = 1 \text{ and } i \notin \mathcal{B}_m \text{ and } \kappa \in \{r_j : j \in \mathcal{B}_m\}] \\ \text{or } [\text{there exists } j \in \mathcal{B}_m \text{ such that } \kappa = r_j \text{ and } x_j = 0] \end{array} \right\}. \quad (\text{EC.63})$$

Moreover, for each assortment $S \in \mathcal{S}$ and each $(m, i, \kappa) \in \mathfrak{V}$, let us define

$$\mathfrak{M}_{m,i,\kappa}(S) \triangleq r_n x_i \mathbb{I}\{i \in \mathcal{B}_m \text{ and } \kappa \neq r_i\} + r_n x_i \mathbb{I}\{i \notin \mathcal{B}_m \text{ and } \kappa \in \{r_j : j \in \mathcal{B}_m\}\} + \sum_{j \in \mathcal{B}_m} \mathbb{I}\{\kappa = r_j\} r_n (1 - x_j) \quad (\text{EC.64})$$

It follows immediately from line (EC.63) and from the equivalence between each assortment $S \in \mathcal{S}$ and the binary decision variables $x \in \{0, 1\}^{\mathcal{N}_0}$ that the function $\mathfrak{M}(\cdot)$ defined above satisfies the inequality $\mathfrak{M}_{m,i,\kappa}(S) \geq r_n$ for all $(m, i, \kappa) \in \mathfrak{N}(S)$ and satisfies the equality $\mathfrak{M}_{m,i,\kappa}(S) = 0$ for all $(m, i, \kappa) \in \mathfrak{V} \setminus \mathfrak{N}(S)$. Thus, it follows from (EC.62) that the robust optimization problem (RO) is equivalent to

$$\begin{array}{ll} \underset{x, \alpha, \beta, \gamma}{\text{maximize}} & \sum_{m \in \mathcal{M}, i \in S_m} v_{m,i} \alpha_{m,i} \\ \text{subject to} & \alpha_{m,i} - \beta_{m,i,\kappa} \mathbb{I}\{m \in \{1, \dots, M-1\}\} - \gamma_{m,i,\kappa} \mathbb{I}\{m \in \{2, \dots, M\}\} \\ & \leq \kappa \mathbb{I}\{m = M\} \\ & + r_n x_i \mathbb{I}\{i \in \mathcal{B}_m \text{ and } \kappa \neq r_i\} \\ & + r_n x_i \mathbb{I}\{i \notin \mathcal{B}_m \text{ and } \kappa \in \{r_j : j \in \mathcal{B}_m\}\} \\ & + \sum_{j \in \mathcal{B}_m} \mathbb{I}\{\kappa = r_j\} r_n (1 - x_j) & \forall (m, i, \kappa) \in \mathfrak{V} \\ & \beta_{m,i,\kappa} + \gamma_{m+1,i',\kappa'} \leq 0 & \forall ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E} \\ & x_0 = 1 \\ & \alpha_{m,i} \in \mathbb{R} & \forall m \in \mathcal{M}, i \in S_m \\ & \beta_{m,i,\kappa}, \gamma_{m,i,\kappa} \in \mathbb{R} & \forall (m, i, \kappa) \in \mathfrak{V} \\ & x_i \in \{0, 1\} & \forall i \in \mathcal{N}_0, \end{array} \quad (\text{EC.65})$$

where we observe that the mixed-integer optimization problem (EC.65) is a reformulation of (EC.62) in which the assortment $S \in \mathcal{S}$ is represented by binary decision variables $x \in \{0, 1\}^{\mathcal{N}_0}$ that satisfy $x_i = 1$ if and only if $i \in S$ and in which $\mathfrak{M}(\cdot)$ is implemented using the definition from line (EC.64). We note that the constraint $x_0 = 1$ is enforced in the mixed-integer optimization problem (EC.65) because $S \in \mathcal{S}$ if and only if $S \subseteq \mathcal{N}_0$ and $0 \in S$. This completes our mixed-integer optimization reformulation of the robust optimization problem (RO) in the case of $\eta = 0$. Since a similar mixed-integer optimization reformulation can be obtained for the case of $\eta > 0$, our proof of Theorem 5 is complete. \square

Appendix J: Proofs of Technical Results from §6.1

J.1. Proof of Theorem 6

Because the proof of Theorem 6 closely resembles the proof of Theorem 4, we focus in the present appendix on the main areas of distinction.

J.1.1. Characterization of Optimal Assortments. We begin our proof of Theorem 6 by characterizing the structure of optimal assortments for the optimistic optimization problem (OO). To do this, we introduce an optimistic version of Definition 4 from §4.4. Indeed, for each tuple $(i_1, \dots, i_M) \in \mathcal{L}$, let $\xi_{i_1 \dots i_M}(S)$ be defined as the maximum revenue among the products in the assortment S that can be the most preferred product in S under a ranking that corresponds to the tuple (i_1, \dots, i_M) .

$$\text{DEFINITION EC.3. } \xi_{i_1 \dots i_M}(S) \triangleq \max_{i \in S: \cap_{m \in \mathcal{M}} \mathcal{D}_{i_m}(S_m) \cap \mathcal{D}_i(S) \neq \emptyset} r_i.$$

Equipped with the above notation, the following result shows that the best-case expected revenue of any fixed assortment can be reformulated as a linear optimization problem. The proof of the following Proposition EC.3 follows from identical reasoning as Proposition 1 and is thus omitted.

PROPOSITION EC.3. *For each $S \in \mathcal{S}$, $\max_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ is equal to the optimal objective value of the following linear optimization problem:*

$$\begin{aligned} & \underset{\lambda, \epsilon}{\text{maximize}} && \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \xi_{i_1 \dots i_M}(S) \lambda_{i_1 \dots i_M} \\ & \text{subject to} && \sum_{(i_1, \dots, i_M) \in \mathcal{L}: i_m = i} \lambda_{i_1 \dots i_M} - \epsilon_{m,i} = v_{m,i} \quad \forall m \in \mathcal{M}, i \in S_m \\ & && \sum_{(i_1, \dots, i_M) \in \mathcal{L}} \lambda_{i_1 \dots i_M} = 1 \\ & && \|\epsilon\| \leq \eta \\ & && \lambda_{i_1 \dots i_M} \geq 0 \quad \forall (i_1, \dots, i_M) \in \mathcal{L}. \end{aligned} \tag{EC.66}$$

Following identical reasoning and using the same notation as developed in Appendix B.2, we obtain the following graphical interpretation of the quantity $\xi_{i_1 \dots i_M}(S)$. In particular, the proof of the following Proposition EC.4 follows from identical reasoning as Proposition EC.1 and is thus omitted.

PROPOSITION EC.4. *For all $S \in \mathcal{S}$ and $(i_1, \dots, i_M) \in \mathcal{L}$, $\xi_{i_1 \dots i_M}(S) = \max_{i \in S \cap \mathcal{I}_{i_1 \dots i_M}(S)} r_i$.*

Equipped with the graphical interpretation of $\xi_{i_1 \dots i_M}(S)$ provided by Proposition EC.4, we derive the following optimistic version of Proposition 2 from §4.4:

PROPOSITION EC.5. *Let $S \in \mathcal{S}$ and $i \notin S$. If there exists $i^* \in \mathcal{S}$ that satisfies $r_{i^*} > r_i$ and $\mathcal{M}_{i^*} \subseteq \mathcal{M}_i$, then $\xi_{i_1 \dots i_M}(S) \geq \xi_{i_1 \dots i_M}(S \cup \{i\})$ for each $(i_1, \dots, i_M) \in \mathcal{L}$.*

Proof. We begin our proof of Proposition EC.5 by developing an intermediary result, denoted below by Claim EC.15, that will allow us to compare the values of $\xi_{i_1 \dots i_M}(S)$ and $\xi_{i_1 \dots i_M}(S \cup \{i\})$ for every assortment $S \in \mathcal{S}$ and every product i which is not in the assortment. The proof of the following Claim EC.15, which uses the graphical interpretation of the quantity $\xi_{i_1 \dots i_M}(S)$ from Proposition EC.5, follows from identical reasoning as the proof of Claim EC.1 from Appendix B.3 and is thus omitted.

CLAIM EC.15. *For all $S \in \mathcal{S}$, $(i_1, \dots, i_M) \in \mathcal{L}$, and $i \notin S$,*

$$\xi_{i_1 \dots i_M}(S \cup \{i\}) = \begin{cases} \xi_{i_1 \dots i_M}(S), & \text{if } i \notin \mathcal{I}_{i_1 \dots i_M}(S), \\ \max \left\{ \max_{j \in S \cap \mathcal{I}_{i_1 \dots i_M}(S) \cap \{j' \in \mathcal{N}_0: i \neq j'\}} r_j, r_i \right\}, & \text{if } i \in \mathcal{I}_{i_1 \dots i_M}(S). \end{cases}$$

Using the above intermediary result, we now complete the proof of Proposition EC.5. Indeed, consider any assortment $S \in \mathcal{S}$ and any product $i \notin S$. Suppose that there exists a product $i^* \in \mathcal{S}$ which satisfies $r_{i^*} > r_i$ and $\mathcal{M}_{i^*} \subseteq \mathcal{M}_i$. For each tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$, we have two cases to consider:

- **Case 1:** Suppose that $i \notin \mathcal{I}_{i_1 \dots i_M}(S)$. In this case, it follows immediately from Claim EC.15 that

$$\xi_{i_1 \dots i_M}(S \cup \{i\}) = \xi_{i_1 \dots i_M}(S),$$

and so the inequality $\rho_{i_1 \dots i_M}(S) \leq \xi_{i_1 \dots i_M}(S \cup \{i\})$ holds when $i \notin \mathcal{I}_{i_1 \dots i_M}(S)$.

- Case 2: Suppose that $i \in \mathcal{I}_{i_1 \dots i_M}(S)$.

In this case, we begin by showing that $i_m \notin S$ for each $m \in \mathcal{M}_i$. Indeed, consider any past assortment $m \in \mathcal{M}_i$. On one hand, if $i = i_m$, then it follows immediately from the fact that $i \notin S$ that $i_m \notin S$. On the other hand, if $i \neq i_m$, then it also must be the case that $i_m \notin S$, else we would have a contradiction with the fact that $i \in \mathcal{I}_{i_1 \dots i_M}(S)$ and the fact that there is, by our construction of $\mathcal{G}_{i_1 \dots i_M}$, a directed edge from vertex i to vertex i_m . We have thus shown that $i_m \notin S$ for all $m \in \mathcal{M}_i$.

We next show that $i_m \in \mathcal{I}_{i_1 \dots i_M}(S)$ for all $m \in \mathcal{M}_i$. Indeed, consider any arbitrary $m \in \mathcal{M}_i$. On one hand, if $i_m = i$, then the statement $i_m \in \mathcal{I}_{i_1 \dots i_M}(S)$ follows immediately from the fact that $i \in \mathcal{I}_{i_1 \dots i_M}(S)$. On the other hand, if $i_m \neq i$, then it follows from the fact that $i \in \mathcal{I}_{i_1 \dots i_M}(S)$ and the fact that there is a directed edge from vertex i to vertex i_m that there must not be a directed path from vertex i_m to a vertex $i_{m'}$ that satisfies $i_{m'} \in S$ for any $m' \in \mathcal{M}$. We have thus shown that $i_m \in \mathcal{I}_{i_1 \dots i_M}(S)$ for all $m \in \mathcal{M}_i$.

Using the above results, we now prove that $i^* \in \mathcal{I}_{i_1 \dots i_M}(S)$. Indeed, we have shown in the above results that $i_m \notin S$ and $i_m \in \mathcal{I}_{i_1 \dots i_M}(S)$ for all $m \in \mathcal{M}_i$. Therefore, it follows from the supposition that $\mathcal{M}_{i^*} \subseteq \mathcal{M}_i$ that $i_m \notin S$ and $i_m \in \mathcal{I}_{i_1 \dots i_M}(S)$ for all $m \in \mathcal{M}_{i^*}$. Since we have supposed that $i^* \in S$, it follows from the fact that $i_m \notin S$ for all $m \in \mathcal{M}_{i^*}$ that $i^* \neq i_m$ for all $m \in \mathcal{M}_{i^*}$. Moreover, it follows from the construction of $\mathcal{G}_{i_1 \dots i_M}$ that all of the outgoing edges from vertex i^* are incoming edges to vertices i_m for $m \in \mathcal{M}_{i^*}$. Therefore, it follows from the fact that $i_m \in \mathcal{I}_{i_1 \dots i_M}(S)$ for all $m \in \mathcal{M}_{i^*}$ that $i^* \in \mathcal{I}_{i_1 \dots i_M}(S)$.

We now prove the desired result for Case 2. First, we observe that

$$\xi_{i_1 \dots i_M}(S) = \max_{j \in S \cap \mathcal{I}_{i_1 \dots i_M}(S)} r_j \geq r_{i^*}, \quad (\text{EC.67})$$

where the equality follows from Proposition EC.4 and the inequality follows from our supposition that $i^* \in S$ and because we have shown that $i^* \in \mathcal{I}_{i_1 \dots i_M}(S)$. Therefore, we have that

$$\begin{aligned} \xi_{i_1 \dots i_M}(S \cup \{i\}) &= \max \left\{ \max_{j \in S \cap \mathcal{I}_{i_1 \dots i_M}(S) \cap \{j' \in \mathcal{N}_0 : i \neq i_{1 \dots i_M} j'\}} r_j, r_i \right\} \\ &\leq \max \left\{ \max_{j \in S \cap \mathcal{I}_{i_1 \dots i_M}(S)} r_j, r_i \right\} \\ &= \max \{ \xi_{i_1 \dots i_M}(S), r_i \} \\ &= \xi_{i_1 \dots i_M}(S), \end{aligned}$$

where the first equality follows from Claim EC.1 and the supposition of Case 2 that $i \in \mathcal{I}_{i_1 \dots i_M}(S)$, the inequality follows algebra, the second equality follows from Proposition EC.4, and the final equality holds because of our supposition that $r_{i^*} > r_i$ and because of line (EC.67), which showed that $\xi_{i_1 \dots i_M}(S) \geq r_{i^*}$. This concludes the proof of Case 2.

In both of the above two cases, we showed that $\xi_{i_1 \dots i_M}(S) \geq \xi_{i_1 \dots i_M}(S \cup \{i\})$, and so our proof of Proposition EC.5 is complete. \square

In view of the above Proposition EC.5, we proceed to develop our main result of Appendix J.1.1, which is a characterization of the structure of optimal assortments for the optimistic optimization problem (OO). To this end, we introduce the following new collection of assortments:

$$\widehat{\mathcal{S}}^{\text{OO}} \triangleq \{S \in \mathcal{S} : \text{if } i^* \in S, 0 < r_i < r_{i^*}, \text{ and } \mathcal{M}_{i^*} \subseteq \mathcal{M}_i, \text{ then } i \notin S\}.$$

Our main result is the following:

THEOREM EC.1. *There exists an assortment $S \in \widehat{\mathcal{S}}^{\text{OO}}$ that is optimal for (OO).*

Proof. Consider any arbitrary assortment $S \in \mathcal{S}$. For this assortment, we define a new assortment as

$$S' \triangleq S \setminus \{i \in \mathcal{N} : \text{there exists } i^* \in S \text{ such that } \mathcal{M}_{i^*} \subseteq \mathcal{M}_i \text{ and } r_{i^*} > r_i\}.$$

We first show that this new assortment S' is an element of the collection $\widehat{\mathcal{S}}^{\text{OO}}$. Indeed, it follows from the fact that $S \in \mathcal{S}$ that $0 \in S'$, which implies that $S' \in \mathcal{S}$. Moreover, for every $i^* \in S'$, it follows from the construction of S' that there does not exist $i \in S \setminus \{0\}$ that satisfies $r_i < r_{i^*}$ and $\mathcal{M}_{i^*} \subseteq \mathcal{M}_i$. We thus conclude that $S' \in \widehat{\mathcal{S}}^{\text{OO}}$.

We next show for each product $i \in S' \setminus S$ that there must exist a product $i^* \in S'$ that satisfies $r_i < r_{i^*}$ and $\mathcal{M}_{i^*} \subseteq \mathcal{M}_i$. Indeed, consider any product $i \in S' \setminus S$. It follows from the construction of S' that there exists a product $i' \in S$ that satisfies $\mathcal{M}_{i'} \subseteq \mathcal{M}_i$ and $r_i < r_{i'}$. Let $i^* \triangleq \arg \max_{i' \in S: \mathcal{M}_{i'} \subseteq \mathcal{M}_i \text{ and } r_i < r_{i'}} r_{i'}$ be the product with the highest revenue among all products i' that satisfy $\mathcal{M}_{i'} \subseteq \mathcal{M}_i$ and $r_i < r_{i'}$. It follows from the definition of i^* that there does not exist any products $i'' \in S$ that satisfy $\mathcal{M}_{i''} \subseteq \mathcal{M}_{i^*}$ and $r_{i''} > r_{i^*}$. Therefore, we conclude that the product i^* is an element of S' .

Finally, let $\{j_1, \dots, j_\nu\} \triangleq S \setminus S'$ denote the products that have been removed from the assortment S . We observe for each tuple of products $(i_1, \dots, i_M) \in \mathcal{L}$ that

$$\xi_{i_1 \dots i_M}(S') = \xi_{i_1 \dots i_M}(S) + \sum_{\iota=1}^{\nu} (\xi_{i_1 \dots i_M}(S \setminus \{j_1, \dots, j_\iota\}) - \xi_{i_1 \dots i_M}(S \setminus \{j_1, \dots, j_{\iota-1}\})) \leq \xi_{i_1 \dots i_M}(S).$$

Indeed, the equality follows from algebra. The inequality follows from Proposition EC.5 together with the fact that for each $\iota \in \{1, \dots, \nu\}$, there exists a product $i^* \in S'$ that satisfies $r_{j_\iota} < r_{i^*}$ and $\mathcal{M}_{i^*} \subseteq \mathcal{M}_{j_\iota}$. Since the assortment $S \in \mathcal{S}$ was chosen arbitrarily, our proof of Theorem EC.1 follows from Proposition EC.3. \square

J.1.2. Upper Bound. As the second step of our proof of Theorem 6, we develop an upper bound on the cardinality of the collection of assortments $\widehat{\mathcal{S}}^{\text{OO}}$ that we showed in Theorem EC.1 contains an optimal assortment for the optimistic optimization problem (OO).

LEMMA EC.8. $|\widehat{\mathcal{S}}^{\text{OO}}| \leq (n+1)^{2^M}$.

Proof. We readily observe that the collection of assortments $\widehat{\mathcal{S}}^{\text{OO}}$ is a subset of

$$\widetilde{\mathcal{S}}^{\text{OO}} \triangleq \{S \in \mathcal{S} : \text{if } i^* \in S, 0 < r_i < r_{i^*}, \text{ and } \mathcal{M}_{i^*} = \mathcal{M}_i, \text{ then } i \notin S\},$$

where we recall from the beginning of §4 that \mathcal{M}_i is defined as the subset of the past assortments $\mathcal{M} \equiv \{1, \dots, M\}$ that offered product i . For each subset of past assortments $\mathcal{C} \subseteq \mathcal{M}$, let the products that are offered only in the assortments in \mathcal{C} (excluding the no-purchase option) be denoted by

$$\mathcal{N}(\mathcal{C}) \triangleq \{i \in \mathcal{N} : \mathcal{M}_i = \mathcal{C}\}.$$

Equipped with the above definitions, we observe that $\{\mathcal{N}(\mathcal{C}) : \mathcal{C} \subseteq \mathcal{M}\}$ is the collection of subsets of products that always appear together in the past assortments, and we readily observe that $|\{\mathcal{N}(\mathcal{C}) : \mathcal{C} \subseteq \mathcal{M}\}| = 2^M$. Therefore,

$$|\widehat{\mathcal{S}}^{\text{OO}}| \leq |\widetilde{\mathcal{S}}^{\text{OO}}| \leq \prod_{\mathcal{C} \subseteq \mathcal{M}} (|\mathcal{N}(\mathcal{C})| + 1) \leq (n+1)^{2^M}. \quad (\text{EC.68})$$

Indeed, the first inequality on line (EC.68) holds because the collection of assortments $\widehat{\mathcal{S}}^{\text{OO}}$ is a subset of the collection of assortments $\widetilde{\mathcal{S}}^{\text{OO}}$. To see why the second inequality on line (EC.68) holds, consider any arbitrary subset of past assortments $\mathcal{C} \subseteq \mathcal{M}$, and let the products in $\mathcal{N}(\mathcal{C})$ be indexed in ascending order by revenue; that is, let the products that comprise $\mathcal{N}(\mathcal{C})$ be denoted by $i_1^{\mathcal{C}}, \dots, i_{|\mathcal{N}(\mathcal{C})|}^{\mathcal{C}}$, where $r_{i_1^{\mathcal{C}}} < \dots < r_{i_{|\mathcal{N}(\mathcal{C})|}^{\mathcal{C}}}$. Then, it follows from the definition of the collection $\widetilde{\mathcal{S}}^{\text{OO}}$ that every assortment $S \in \widetilde{\mathcal{S}}^{\text{OO}}$ must satisfy the condition $\left| \left\{ i_1^{\mathcal{C}}, \dots, i_{|\mathcal{N}(\mathcal{C})|}^{\mathcal{C}} \right\} \cap S \right| \leq 1$. Since $\mathcal{C} \subseteq \mathcal{M}$ was chosen arbitrarily, we have shown that

$$\widetilde{\mathcal{S}}^{\text{OO}} \subseteq \left\{ \{0\} \cup \bigcup_{\mathcal{C} \subseteq \mathcal{M}} \mathcal{F}^{\mathcal{C}} : \mathcal{F}^{\mathcal{C}} \in \{\emptyset, \{i_1^{\mathcal{C}}\}, \dots, \{i_{|\mathcal{N}(\mathcal{C})|}^{\mathcal{C}}\}\} \right\},$$

which proves that the second inequality on line (EC.68) holds. The third inequality on line (EC.68) follows from the fact that $|\{\mathcal{N}(\mathcal{C}) : \mathcal{C} \subseteq \mathcal{M}\}| = 2^M$ and from the fact that $|\mathcal{N}(\mathcal{C})| \leq n$ for every $\mathcal{C} \subseteq \mathcal{M}$. We have thus proven that $|\widehat{\mathcal{S}}^{\text{OO}}|$ is at most $(n+1)^{2^M}$, which concludes our proof of Lemma EC.8. \square

CONSTRUCT- \mathcal{G}^{OO} (\mathcal{M}, r)**Inputs:**

- The collection of past assortments, $\mathcal{M} \equiv \{S_1, \dots, S_M\}$.
- The revenues of the products, $r \equiv (r_0, r_1, \dots, r_n)$.

Output:

- $\mathcal{G}^{\text{OO}} \equiv (\mathcal{V}^{\text{OO}}, \mathcal{E}^{\text{OO}})$, where $\mathcal{V}^{\text{OO}} \equiv \mathcal{N}$ and $\mathcal{E}^{\text{OO}} \equiv \{(i^*, i) \in \mathcal{N} \times \mathcal{N} : r_{i^*} > r_i \text{ and } \mathcal{M}_{i^*} \subseteq \mathcal{M}_i\}$.

Procedure:

1. Initialize the vertex set $\mathcal{V}^{\text{OO}} \leftarrow \emptyset$ and edge set $\mathcal{E}^{\text{OO}} \leftarrow \emptyset$.
2. For each product $i \in \mathcal{N}^{\text{OO}}$:
 - (a) Update $\mathcal{V}^{\text{OO}} \leftarrow \mathcal{V}^{\text{OO}} \cup \{i\}$.
 - (b) Construct the set \mathcal{M}_i of past assortments which offered product i .
3. For each pair of products $(i^*, i) \in \mathcal{N} \times \mathcal{N}$:
 - (a) If $r_{i^*} > r_i$ and $\mathcal{M}_{i^*} \subseteq \mathcal{M}_i$:
 - i. Update $\mathcal{E}^{\text{OO}} \leftarrow \mathcal{E}^{\text{OO}} \cup \{(i^*, i)\}$
4. Output $\mathcal{G}^{\text{OO}} \equiv (\mathcal{V}^{\text{OO}}, \mathcal{E}^{\text{OO}})$ and terminate.

Algorithm 5: A procedure for constructing the directed acyclic graph \mathcal{G}^{OO} .

J.1.3. Algorithm for Constructing $\widehat{\mathcal{S}}^{\text{OO}}$. As the third step of our proof of Theorem 6, we develop an efficient algorithm for constructing the collection of assortments $\widehat{\mathcal{S}}^{\text{OO}}$ that contains an optimal assortment for the optimistic optimization problem (OO). Specifically, our main result of Appendix J.1.3 is the following lemma:

LEMMA EC.9. *The collection of assortments $\widehat{\mathcal{S}}^{\text{OO}}$ can be constructed in $\mathcal{O}(n^2(M + |\widehat{\mathcal{S}}^{\text{OO}}|))$ time.*

Proof. We begin by developing an algorithm for constructing a directed graph, denoted by $\mathcal{G}^{\text{OO}} \equiv (\mathcal{V}^{\text{OO}}, \mathcal{E}^{\text{OO}})$, in which the set of vertices in the directed graph is defined as $\mathcal{V}^{\text{OO}} \triangleq \mathcal{N}$ and the set of directed edges is defined as $\mathcal{E}^{\text{OO}} \triangleq \{(i^*, i) \in \mathcal{N} \times \mathcal{N} : r_{i^*} > r_i \text{ and } \mathcal{M}_{i^*} \subseteq \mathcal{M}_i\}$. This directed graph has a natural correspondence with the collection of assortments $\widehat{\mathcal{S}}^{\text{OO}}$, as we readily observe that an assortment satisfies $S \in \widehat{\mathcal{S}}^{\text{OO}}$ if and only if $[0 \in S]$ and [we have that $i \notin S$ whenever there exists a product $i^* \in S \setminus \{0\}$ and a directed edge $(i^*, i) \in \mathcal{E}^{\text{OO}}$]. It is easy to see that this directed graph is acyclic and has the property that a vertex $j \in \mathcal{V}^{\text{OO}}$ is reachable from a vertex $i \in \mathcal{V}^{\text{OO}}$ if and only if there is a directed edge $(i, j) \in \mathcal{E}^{\text{OO}}$ from vertex i to vertex j . A directed acyclic graph which has the aforementioned property for each pair of vertices is referred to as a *transitive closure*.

Our algorithm for constructing the directed acyclic graph \mathcal{G}^{OO} that is a transitive closure is presented in Algorithm 5. In the algorithm, we first iterate over each product $i \in \mathcal{N}$ and construct the corresponding set \mathcal{M}_i of past assortments which offered that product. We then iterate over each pair of products $(i, i^*) \in \mathcal{N} \times \mathcal{N}$ and check whether $r_{i^*} > r_i$ and $\mathcal{M}_{i^*} \subseteq \mathcal{M}_i$. It follows from identical reasoning as that given for Algorithm 1 in Appendix D.3 that Algorithm 5 has a total computation time of $\mathcal{O}(n^2M)$.

We next describe our algorithm for constructing the collection of assortments $\widehat{\mathcal{S}}^{\text{OO}}$ from the directed graph \mathcal{G}^{OO} . This algorithm is denoted by $\text{CONSTRUCT-}\widehat{\mathcal{S}}^{\text{OO}}(\mathcal{M}, r)$ and is found in Algorithm 6. In this algorithm, we first use Algorithm 5 to construct the directed acyclic graph $\mathcal{G}^{\text{OO}} \equiv (\mathcal{V}^{\text{OO}}, \mathcal{E}^{\text{OO}})$, and then we invoke a recursive subroutine denoted by $\text{RECURSIVESTEP-OO}(\mathcal{G}^{\text{OO}})$ in Algorithm 7. The goal of the recursive subroutine is to take as an input a generic directed acyclic graph $\mathcal{G}^{\text{OO}} \equiv (\mathcal{V}^{\text{OO}}, \mathcal{E}^{\text{OO}})$, and for that graph, output the collection of subsets of vertices $\mathcal{A}^{\text{OO}} \equiv \{S \subseteq \mathcal{V}^{\text{OO}} : \text{if } i \in S \text{ and } (i, j) \in \mathcal{E}^{\text{OO}}, \text{ then } j \notin S\}$. Algorithm 6 concludes by adding the no-purchase option 0 into each of the subsets of vertices from \mathcal{A}^{OO} . The correctness of Algorithm 6 follows immediately from our earlier observation that an assortment satisfies $S \in \widehat{\mathcal{S}}^{\text{OO}}$ if and only if $[0 \in S]$ and [we have that $i \notin S$ whenever there exists a product $i^* \in S \setminus \{0\}$ and a directed edge $(i^*, i) \in \mathcal{E}^{\text{OO}}$].

At a high level, the recursive subroutine in Algorithm 7 is comprised of two cases. The base case of the subroutine is when the graph has no vertices, in which case it is clear that $\mathcal{A}^{\text{OO}} = \{\emptyset\}$. If we are not in the base case, then the aim of the recursive subroutine is to construct the collections $\{S \in \mathcal{A}^{\text{OO}} : i \notin S\}$ and

CONSTRUCT- $\widehat{\mathcal{S}}^{\text{OO}}(\mathcal{M}, r)$

Inputs:

- The collection of past assortments, $\mathcal{M} \equiv \{S_1, \dots, S_M\}$.
- The revenues of the products, $r \equiv (r_0, r_1, \dots, r_n)$.

Output:

- The collection of assortments $\widehat{\mathcal{S}}^{\text{OO}}$ corresponding to the collection of past assortments \mathcal{M} and the revenues r .

Procedure:

1. Construct the directed acyclic graph $\mathcal{G}^{\text{OO}} \leftarrow \text{CONSTRUCT-}\mathcal{G}^{\text{OO}}(\mathcal{M}, r)$.
2. Compute the collection of assortments $\widehat{\mathcal{S}}^{\text{OO}} \leftarrow \text{RECURSIVESTEP-OO}(\mathcal{G}^{\text{OO}})$.
3. Output $\{S \cup \{0\} : S \in \widehat{\mathcal{S}}^{\text{OO}}\}$ and terminate.

Algorithm 6: A procedure for constructing $\widehat{\mathcal{S}}^{\text{OO}}$.

$\{S \in \mathcal{A}^{\text{OO}} : i \in S\}$ for a chosen vertex $i \in \mathcal{V}^{\text{OO}}$ and output the union of these two collections. The construction of the collection $\{S \in \mathcal{A}^{\text{OO}} : i \notin S\}$ takes place on lines (2b)-(2c) of Algorithm 7, and the construction of the collection $\{S \in \mathcal{A}^{\text{OO}} : i \in S\}$ takes place on lines (2d)-(2f) of Algorithm 7.

Up to this point, we have established that Algorithm 5 is correct (that is, it delivers the desired output for any valid input), and we have established that Algorithm 6 is correct under the assumption that Algorithm 7 is correct. Therefore, it remains for us to prove that the recursive subroutine in Algorithm 7 is correct. To prove the correctness of the recursive subroutine, we will make use of four intermediary claims, which are denoted below by Claims EC.16-EC.19. The purpose of the first two intermediary claims, denoted by Claims EC.16 and EC.17, is to show that the graphs $\mathcal{G}^{\text{OO}'} \equiv (\mathcal{V}^{\text{OO}'}, \mathcal{E}^{\text{OO}'})$ and $\mathcal{G}^{\text{OO}''} \equiv (\mathcal{V}^{\text{OO}''}, \mathcal{E}^{\text{OO}''})$ constructed on lines (2b) and (2d) of Algorithm 7 are directed acyclic graphs that are transitive closures, which implies that $\mathcal{G}^{\text{OO}'}$ and $\mathcal{G}^{\text{OO}''}$ are valid inputs on lines (2c) and (2e) of Algorithm 7. The purpose of the second two intermediary claims, denoted by Claims EC.18 and EC.19, is to show that the union of the two collections $\mathcal{A}^{\text{OO}'}$ and $\mathcal{A}^{\text{OO}''}$ constructed on lines (2c) and (2f) of Algorithm 7 provides the desired output on line (2g) of Algorithm 7.

CLAIM EC.16. *Let $\mathcal{G}^{\text{OO}} \equiv (\mathcal{V}^{\text{OO}}, \mathcal{E}^{\text{OO}})$ be a directed acyclic graph that is a transitive closure. For notational convenience, let $\mathcal{E}_i^{\text{OO}}$ denote the set of incoming and outgoing edges from each vertex $i \in \mathcal{V}^{\text{OO}}$. Then for each vertex $i \in \mathcal{V}^{\text{OO}}$, we have that $\tilde{\mathcal{G}}^{\text{OO}} \equiv (\mathcal{V}^{\text{OO}} \setminus \{i\}, \mathcal{E}^{\text{OO}} \setminus \mathcal{E}_i^{\text{OO}})$ is a directed acyclic graph that is a transitive closure.*

CLAIM EC.17. *Let $\mathcal{G}^{\text{OO}} \equiv (\mathcal{V}^{\text{OO}}, \mathcal{E}^{\text{OO}})$ be a directed acyclic graph that is a transitive closure. For notational convenience, let $\mathcal{E}_i^{\text{OO}}$ denote the set of incoming and outgoing edges from each vertex $i \in \mathcal{V}^{\text{OO}}$. Then for each subset of vertices $\mathcal{B}^{\text{OO}} \subseteq \mathcal{V}^{\text{OO}}$, we have that $\mathcal{G}^{\text{OO}} \equiv (\mathcal{V}^{\text{OO}} \setminus \mathcal{B}^{\text{OO}}, \mathcal{E}^{\text{OO}} \setminus (\cup_{i \in \mathcal{B}^{\text{OO}}} \mathcal{E}_i^{\text{OO}}))$ is a directed acyclic graph that is a transitive closure.*

CLAIM EC.18. *Let $\mathcal{G}^{\text{OO}} \equiv (\mathcal{V}^{\text{OO}}, \mathcal{E}^{\text{OO}})$ be a directed acyclic graph that is a transitive closure. For notational convenience, let $\mathcal{E}_i^{\text{OO}}$ denote the set of incoming and outgoing edges from each vertex $i \in \mathcal{V}^{\text{OO}}$. For each vertex $i \in \mathcal{V}^{\text{OO}}$,*

$$\begin{aligned} & \{S \subseteq \mathcal{V}^{\text{OO}} \setminus \{i\} : \text{if } k \in S \text{ and } (k, j) \in \mathcal{E}^{\text{OO}}, \text{ then } j \notin S\} \\ & = \{S \subseteq \mathcal{V}^{\text{OO}} \setminus \{i\} : \text{if } k \in S \text{ and } (k, j) \in \mathcal{E}^{\text{OO}} \setminus \mathcal{E}_i^{\text{OO}}, \text{ then } j \notin S\}. \end{aligned}$$

CLAIM EC.19. *Let $\mathcal{G}^{\text{OO}} \equiv (\mathcal{V}^{\text{OO}}, \mathcal{E}^{\text{OO}})$ be a directed acyclic graph that is a transitive closure. For notational convenience, let $\mathcal{E}_i^{\text{OO}}$ denote the set of incoming and outgoing edges from each vertex $i \in \mathcal{V}^{\text{OO}}$. For each vertex $i \in \mathcal{V}^{\text{OO}}$,*

$$\{S \subseteq \mathcal{V}^{\text{OO}} : [i \in S] \text{ and } [\text{if } k \in S \text{ and } (k, j) \in \mathcal{E}^{\text{OO}}, \text{ then } j \notin S]\}$$

RECURSIVESTEP-OO(\mathcal{G}^{OO})**Inputs:**

- A directed acyclic graph $\mathcal{G}^{\text{OO}} \equiv (\mathcal{V}^{\text{OO}}, \mathcal{E}^{\text{OO}})$ that is a transitive closure.

Output:

- The collection $\mathcal{A}^{\text{OO}} \equiv \{S \subseteq \mathcal{V}^{\text{OO}} : \text{if } i \in S \text{ and } (i, j) \in \mathcal{E}^{\text{OO}}, \text{ then } j \notin S\}$.

Procedure:

1. If $\mathcal{V}^{\text{OO}} = \emptyset$:
 - (a) Output the collection $\mathcal{A}^{\text{OO}} \equiv \{\emptyset\}$ and terminate.
2. Otherwise:
 - (a) Choose any vertex $i \in \mathcal{V}^{\text{OO}}$.
 - (b) Create a copy of $\mathcal{G}^{\text{OO}} \equiv (\mathcal{V}^{\text{OO}}, \mathcal{E}^{\text{OO}})$ in which the vertex i and its incoming and outgoing edges are removed. Denote this new graph by $\mathcal{G}^{\text{OO}'} \equiv (\mathcal{V}^{\text{OO}'}, \mathcal{E}^{\text{OO}'})$.
 - (c) Compute the collection $\mathcal{A}^{\text{OO}'} \leftarrow \text{RECURSIVESTEP-OO}(\mathcal{G}^{\text{OO}'})$.
 - (d) Create a copy of $\mathcal{G}^{\text{OO}} \equiv (\mathcal{V}^{\text{OO}}, \mathcal{E}^{\text{OO}})$ in which the vertices $\{i\} \cup \{\ell : (i, \ell) \in \mathcal{E}^{\text{OO}}\} \cup \{\ell : (\ell, i) \in \mathcal{E}^{\text{OO}}\}$ and the incoming and outgoing edges of these vertices are removed. Denote this new graph by $\mathcal{G}^{\text{OO}''} \equiv (\mathcal{V}^{\text{OO}''}, \mathcal{E}^{\text{OO}''})$.
 - (e) Compute the collection $\mathcal{A}^{\text{OO}''} \leftarrow \text{RECURSIVESTEP}(\mathcal{G}^{\text{OO}''})$.
 - (f) Compute the collection $\mathcal{A}^{\text{OO}'''} \leftarrow \{S \cup \{i\} : S \in \mathcal{A}^{\text{OO}''}\}$.
 - (g) Output the collection $\mathcal{A}^{\text{OO}} \equiv \mathcal{A}^{\text{OO}'} \cup \mathcal{A}^{\text{OO}'''}$ and terminate.

Algorithm 7: A recursive subroutine which is used in Algorithm 6.

$$= \left\{ \begin{array}{l} S' \cup \{i\} : \\ S' \in \left\{ \begin{array}{l} S \subseteq \mathcal{V}^{\text{OO}} \setminus (\{i\} \cup \{\ell : (i, \ell) \in \mathcal{E}^{\text{OO}}\} \cup \{\ell : (\ell, i) \in \mathcal{E}^{\text{OO}}\}) : \\ \text{if } k \in S \text{ and } (k, j) \in \mathcal{E}^{\text{OO}} \setminus \left(\mathcal{E}_i^{\text{OO}} \cup \bigcup_{\ell: (i, \ell) \in \mathcal{E}^{\text{OO}}} \mathcal{E}_\ell^{\text{OO}} \cup \bigcup_{\ell: (\ell, i) \in \mathcal{E}^{\text{OO}}} \mathcal{E}_\ell^{\text{OO}} \right), \text{ then } j \in S \end{array} \right\} \end{array} \right\}.$$

Proof. The proofs of Claims EC.16, EC.17, EC.18, and EC.19 follows from identical reasoning as the proofs of Claims EC.2, EC.3, EC.4, and EC.5 from Appendix D.3 and are thus omitted. \square

Using the above intermediary claims, we now prove the correctness of the recursive subroutine in Algorithm 7. Indeed, it is clear that the recursive subroutine yields the correct output in the base case where $\mathcal{V}^{\text{OO}} = \emptyset$. Next, let us assume by induction that the recursive subroutine yields the correct output for all valid input graphs with up to $p-1$ vertices, and consider any valid input $\mathcal{G}^{\text{OO}} \equiv (\mathcal{V}^{\text{OO}}, \mathcal{E}^{\text{OO}})$ for which the number of vertices is $|\mathcal{V}^{\text{OO}}| = p$. Let $i \in \mathcal{V}^{\text{OO}}$ denote the vertex from this graph which is chosen in line (2a) of Algorithm 3, where the existence of such a vertex follows from the fact that we are in the case where $\mathcal{V}^{\text{OO}} \neq \emptyset$. It follows immediately from Claim EC.3 that the graphs $\mathcal{G}^{\text{OO}'} \equiv (\mathcal{V}^{\text{OO}'}, \mathcal{E}^{\text{OO}'})$ and $\mathcal{G}^{\text{OO}''} \equiv (\mathcal{V}^{\text{OO}''}, \mathcal{E}^{\text{OO}''})$ constructed on lines (2b) and (2d) of Algorithm 7 are directed acyclic graphs which are transitive closures, which implies that these graphs are valid inputs to Algorithm 7 in lines (2c) and (2e). Therefore, it follows from the induction hypothesis and lines (2c), (2e), and (2f) of Algorithm 7 that

$$\mathcal{A}^{\text{OO}'} = \{S \subseteq \mathcal{V}^{\text{OO}} \setminus \{i\} : \text{if } k \in S \text{ and } (k, j) \in \mathcal{E}^{\text{OO}} \setminus \mathcal{E}_i^{\text{OO}}, \text{ then } j \notin S\}$$

$$\mathcal{A}^{\text{OO}'''} = \left\{ \begin{array}{l} S' \cup \{i\} : \\ S' \in \left\{ \begin{array}{l} S \subseteq \mathcal{V}^{\text{OO}} \setminus (\{i\} \cup \{\ell : (i, \ell) \in \mathcal{E}^{\text{OO}}\} \cup \{\ell : (\ell, i) \in \mathcal{E}^{\text{OO}}\}) : \\ \text{if } k \in S \text{ and } (k, j) \in \mathcal{E}^{\text{OO}} \setminus \left(\mathcal{E}_i^{\text{OO}} \cup \bigcup_{\ell: (i, \ell) \in \mathcal{E}^{\text{OO}}} \mathcal{E}_\ell^{\text{OO}} \cup \bigcup_{\ell: (\ell, i) \in \mathcal{E}^{\text{OO}}} \mathcal{E}_\ell^{\text{OO}} \right), \text{ then } j \in S \end{array} \right\} \end{array} \right\}$$

where the induction hypothesis can be applied because $\mathcal{G}^{\text{OO}'} \equiv (\mathcal{V}^{\text{OO}'}, \mathcal{E}^{\text{OO}'})$ and $\mathcal{G}^{\text{OO}''} \equiv (\mathcal{V}^{\text{OO}''}, \mathcal{E}^{\text{OO}''})$ are valid inputs to Algorithm 7 and because $|\mathcal{V}^{\text{OO}'}| \leq p-1$ and $|\mathcal{V}^{\text{OO}''}| \leq p-1$. Therefore, it follows from Claims EC.18 and EC.19 that

$$\begin{aligned} \mathcal{A}^{\text{OO}'} &= \{S \subseteq \mathcal{V}^{\text{OO}} \setminus \{i\} : \text{if } k \in S \text{ and } (k, j) \in \mathcal{E}^{\text{OO}}, \text{ then } j \notin S\} \\ \mathcal{A}^{\text{OO}'''} &= \{S \subseteq \mathcal{V}^{\text{OO}} : [i \in S] \text{ and } [\text{if } k \in S \text{ and } (k, j) \in \mathcal{E}^{\text{OO}}, \text{ then } j \notin S]\}, \end{aligned}$$

which proves that the output of Algorithm 7 is

$$\mathcal{A}^{\text{OO}'} \cup \mathcal{A}^{\text{OO}'''} = \{S \subseteq \mathcal{V}^{\text{OO}} : \text{if } k \in S \text{ and } (k, j) \in \mathcal{E}^{\text{OO}}, \text{ then } j \notin S\}.$$

This completes our proof of the correctness of Algorithm 7.

We conclude Appendix J.1.3 by analyzing the computation time of Algorithm 6. Indeed, we recall that the computation time required for line (1) in Algorithm 6 is $\mathcal{O}(n^2M)$. In what follows, we assume that all directed graphs are stored as adjacency lists. Under this assumption, our analysis of the computation time for line (2) in Algorithm 7 is split into the following two intermediary claims, denoted by Claim EC.20 and EC.21. In our first intermediary claim, presented below as Claim EC.20, we establish the computation time required for lines (2b), (2d), (2f), and (2g) of Algorithm 7. The proof of the following Claim EC.20 follows from similar reasoning as the proof of Claim EC.6 from Appendix D.3 and is thus omitted.

CLAIM EC.20. *If $\mathcal{G}^{\text{OO}} \equiv (\mathcal{V}^{\text{OO}}, \mathcal{E}^{\text{OO}})$ is a directed acyclic graph that is a transitive closure with $|\mathcal{V}^{\text{OO}}| \geq 1$, then lines (2b), (2d), (2f), and (2g) of Algorithm 7 can be performed in $\mathcal{O}(|\mathcal{V}^{\text{OO}}| \times |\text{RECURSIVESTEP-OO}(\mathcal{G}^{\text{OO}})|)$ time.*

In our second intermediary claim, presented below as Claim EC.21, we use Claim EC.20 to establish the computation time for Algorithm 7 for any valid input. The proof of the following Claim EC.21 follows from similar reasoning as the proof of Claim EC.7 from Appendix D.3 and is thus omitted.

CLAIM EC.21. *If $\mathcal{G}^{\text{OO}} \equiv (\mathcal{V}^{\text{OO}}, \mathcal{E}^{\text{OO}})$ is a directed acyclic graph that is a transitive closure with $|\mathcal{V}^{\text{OO}}| \geq 1$, then the computation time for Algorithm 7 is $\mathcal{O}(|\mathcal{V}^{\text{OO}}|^2 \times |\text{RECURSIVESTEP-OO}(\mathcal{G}^{\text{OO}})|)$.*

Our analysis of the computation time for line (2) of Algorithm 6 follows readily from Claim EC.21. Indeed, we observe that the graph $\mathcal{G}^{\text{OO}} \equiv (\mathcal{V}^{\text{OO}}, \mathcal{E}^{\text{OO}})$ that was constructed on line (1) of Algorithm 6 satisfies $|\mathcal{V}^{\text{OO}}| = n$ and $|\mathcal{E}^{\text{OO}}| \leq n^2$. Therefore, it follows from Claim EC.21 that line (2) of Algorithm 6 requires $\mathcal{O}((n+1)^2 \times |\text{RECURSIVESTEP-OO}(\mathcal{G}^{\text{OO}})|) = \mathcal{O}(n^2|\widehat{\mathcal{S}}^{\text{OO}}|)$ computation time. We have shown that the total computation time for Algorithm 6 is

$$\underbrace{\mathcal{O}(n^2M)}_{(1)} + \underbrace{\mathcal{O}(n^2|\widehat{\mathcal{S}}^{\text{OO}}|)}_{(2)} = \mathcal{O}\left(n^2(M + |\widehat{\mathcal{S}}^{\text{OO}}|)\right).$$

□

J.1.4. Computation time for solving (OO). We conclude our proof of Theorem 6 by combining the intermediary results from Appendices J.1.1, J.1.2, and J.1.3. Indeed, we showed in Theorem EC.1 in Appendix J.1.1 that there exists an optimal solution for the optimistic optimization problem (OO) that is an element of the collection of assortments $\widehat{\mathcal{S}}^{\text{OO}}$. Moreover, we showed in Lemma EC.8 in Appendix J.1.2 that $|\widehat{\mathcal{S}}^{\text{OO}}|$ is upper bounded by a polynomial in the number of products n for any fixed number of past assortments M , and so it follows from Lemma EC.9 in Appendix J.1.3 that the collection $\widehat{\mathcal{S}}^{\text{OO}}$ can be constructed in computation time that is polynomial in the number of products n for any fixed number of past assortments M . Finally, for each assortment $S \in \widehat{\mathcal{S}}^{\text{OO}}$, it follows from Proposition EC.3 in Appendix J.1.1 that the best-case expected revenue $\max_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ can be computed by solving a linear optimization problem with numbers of decision variables and constraints that scale as a polynomial in the number of products n for any fixed number of past assortments M .²¹ We thus conclude that (OO) can be solved in time that is polynomial in the number of products n for any fixed number of past assortments M .

²¹ It follows identical reasoning as Lemmas 8 and 9 from §5.2 that the linear optimization reformulation of the best-case expected revenue $\max_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ can be constructed in time that is polynomial in the number of products n for any fixed number of past assortments M .

J.2. Proof of Theorem 7

Let Assumption 1 hold, and let $S_1, \dots, S_M \in \mathcal{S}$. To develop our reformulation of the objective function of the optimization problem (OO), we will utilize the following intermediary claim:

CLAIM EC.22. For each $S \subseteq \mathcal{N}_0$ that satisfies $S \cap S_1 \neq \emptyset$, $\max_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ is equal to the optimal objective value of the following linear optimization problem:

$$\begin{aligned} & \underset{f, g, \epsilon}{\text{maximize}} && \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{B}} \kappa g_{M, i, \kappa} \\ & \text{subject to} && \text{same constraints as (6)} \end{aligned} \tag{EC.69}$$

Proof of Claim EC.22. Consider any assortment $S \subseteq \mathcal{N}_0$ that satisfies $S \cap S_1 \neq \emptyset$. For each $i \in \mathcal{N}_0$, let $\bar{r}_i \triangleq r_n - r_i$. We observe that

$$\begin{aligned} \max_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) &= \max_{\lambda \in \mathcal{U}} \sum_{i \in \mathcal{N}_0} r_i \mathcal{D}_i^\lambda(S) \\ &= r_n + \max_{\lambda \in \mathcal{U}} \sum_{i \in \mathcal{N}_0} (r_i - r_n) \mathcal{D}_i^\lambda(S) \\ &= r_n - \min_{\lambda \in \mathcal{U}} \sum_{i \in \mathcal{N}_0} (r_n - r_i) \mathcal{D}_i^\lambda(S) \\ &= r_n - \min_{\lambda \in \mathcal{U}} \sum_{i \in \mathcal{N}_0} \bar{r}_i \mathcal{D}_i^\lambda(S). \end{aligned} \tag{EC.70}$$

Indeed, the first equality follows from the definition of the predicted expected revenue function $\mathcal{R}^\lambda(\cdot)$ from §2. The second equality follows from algebra and from the fact that $\sum_{i \in \mathcal{N}_0} \mathcal{D}_i^\lambda(S) = 1$ for all $\lambda \in \Delta_\Sigma$. The third equality follows from algebra. The fourth equality follows from the definition of the vector \bar{r} .

We observe that the optimization problem $\min_{\lambda \in \mathcal{U}} \sum_{i \in \mathcal{N}_0} \bar{r}_i \mathcal{D}_i^\lambda(S)$ is the same as the optimization problem $\min_{\lambda \in \mathcal{U}} \sum_{i \in \mathcal{N}_0} r_i \mathcal{D}_i^\lambda(S)$, except with the revenue r_i for each product $i \in \mathcal{N}_0$ having been changed to \bar{r}_i . In particular, we observe that the vector \bar{r} satisfies $0 = \bar{r}_n < \bar{r}_{n-1} < \dots < \bar{r}_1 < \bar{r}_0$. Thus, with a suitable reindexing of the products, it follows readily from Assumption 1, from the fact that $S \cap S_1 \neq \emptyset$, and from Proposition 3 that $\min_{\lambda \in \mathcal{U}} \sum_{i \in \mathcal{N}_0} \bar{r}_i \mathcal{D}_i^\lambda(S)$ is equal to the optimal objective value of the following linear optimization problem:

$$\begin{aligned} & \underset{f, g, \epsilon}{\text{minimize}} && \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{B}} (r_n - \kappa) g_{M, i, \kappa} \\ & \text{subject to} && \text{same constraints as (6)} \end{aligned} \tag{EC.71}$$

Combining the above analysis, we conclude that

$$\begin{aligned} \max_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) &= r_n - \left[\begin{array}{l} \underset{f, g, \epsilon}{\text{minimize}} \quad \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{B}} (r_n - \kappa) g_{M, i, \kappa} \\ \text{subject to} \quad \text{same constraints as (6)} \end{array} \right] \\ &= r_n + \left[\begin{array}{l} \underset{f, g, \epsilon}{\text{maximize}} \quad \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{B}} (\kappa - r_n) g_{M, i, \kappa} \\ \text{subject to} \quad \text{same constraints as (6)} \end{array} \right] \\ &= \left[\begin{array}{l} \underset{f, g, \epsilon}{\text{maximize}} \quad \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{B}} \kappa g_{M, i, \kappa} \\ \text{subject to} \quad \text{same constraints as (6)} \end{array} \right], \end{aligned}$$

where the first equality follows from line (EC.70) and from the fact that $\min_{\lambda \in \mathcal{U}} \sum_{i \in \mathcal{N}_0} \bar{r}_i \mathcal{D}_i^\lambda(S)$ is equal to the optimal objective value of the linear optimization problem (EC.71), the second equality follows from algebra, and the third equality follows from the fact that $\sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{B}} g_{M, i, \kappa} = 1$ is one of the constraints in the linear optimization problem (6). Our proof of Claim EC.22 is thus complete. \square

To simplify our exposition, the remainder of the proof of Theorem 7 focuses on the case where $\eta = 0$. Indeed, the reformulations of objective function of the optimization problem (OO) for the case of $\eta = 0$ can be readily extended to the case of $\eta > 0$ by introducing $\mathcal{O}(\text{poly}(n, M))$ auxiliary decision variables and auxiliary linear constraints in the dual problem. In the case of $\eta = 0$, we observe from Claim EC.22 that the linear optimization problem (EC.69) can be rewritten equivalently as

$$\begin{aligned}
& \underset{f, g}{\text{maximize}} && \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} \kappa g_{M, i, \kappa} \\
& \text{subject to} && \sum_{\kappa: (m, i, \kappa) \in \mathfrak{V}} g_{m, i, \kappa} = v_{m, i} && \forall m \in \mathcal{M}, i \in S_m \\
& && \sum_{i', \kappa': ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E}} f_{m, i, \kappa, i', \kappa'} = g_{m, i, \kappa} && \forall (m, i, \kappa) \in \mathfrak{V}: m \in \{1, \dots, M-1\} \\
& && \sum_{i', \kappa': ((m-1, i', \kappa'), (m, i, \kappa)) \in \mathfrak{E}} f_{m-1, i', \kappa', i, \kappa} = g_{m, i, \kappa} && \forall (m, i, \kappa) \in \mathfrak{V}: m \in \{2, \dots, M\} \\
& && g_{m, i, \kappa} = 0 && \forall (m, i, \kappa) \in \mathfrak{N}(S) \\
& && f_{m, i, \kappa, i', \kappa'} \geq 0 && \forall ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E} \\
& && g_{m, i, \kappa} \geq 0 && \forall (m, i, \kappa) \in \mathfrak{V},
\end{aligned} \tag{EC.69}$$

where the definition of the set $\mathfrak{N}(S)$ can be found in the beginning of Appendix I.4. It remains for us to show that the constraints of the linear optimization problem (EC.69) can be reformulated using linear constraints and mixed-integer decision variables. Indeed, let $x_i \in \{0, 1\}$ be a binary decision variable for each product $i \in \mathcal{N}_0$ that satisfies $x_i = 1$ if and only if product i is in the assortment S . Under this equivalence between an assortment $S \in \mathcal{S}$ and the binary decision variables $x \in \{0, 1\}^{\mathcal{N}_0}$, it follows from the definition of $\mathfrak{N}(\cdot)$ that the following equality holds for all assortments $S \in \mathcal{S}$:

$$\mathfrak{N}(S) \triangleq \left\{ \begin{array}{l} [x_i = 1 \text{ and } i \in \mathcal{B}_m \text{ and } \kappa \neq r_i] \\ (m, i, \kappa) \in \mathfrak{V}: \quad \text{or } [x_i = 1 \text{ and } i \notin \mathcal{B}_m \text{ and } \kappa \in \{r_j : j \in \mathcal{B}_m\}] \\ \quad \text{or } [\text{there exists } j \in \mathcal{B}_m \text{ such that } \kappa = r_j \text{ and } x_j = 0] \end{array} \right\}.$$

Hence, it is easy to verify that the constraints in the linear optimization problem (EC.69) of the form

$$g_{m, i, \kappa} = 0 \quad \forall (m, i, \kappa) \in \mathfrak{N}(S)$$

will be satisfied if and only if

$$\begin{aligned}
g_{m, i, \kappa} &\leq 1 - x_i && \forall (m, i, \kappa) \in \mathfrak{V}: i \in \mathcal{B}_m \text{ and } \kappa \neq r_i \\
g_{m, i, \kappa} &\leq 1 - x_i && \forall (m, i, \kappa) \in \mathfrak{V}: i \notin \mathcal{B}_m \text{ and } \kappa \in \{r_j : j \in \mathcal{B}_m\} \\
g_{m, i, \kappa} &\leq x_j && \forall (m, i, \kappa) \in \mathfrak{V} \text{ and } j \in \mathcal{B}_m : \kappa = r_j.
\end{aligned}$$

We thus conclude from Claim EC.22 that the best-case expected revenue $\max_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ for each assortment $S \in \mathcal{S}$ is equal to the optimal objective value of the following linear optimization problem:

$$\begin{aligned}
& \underset{f, g}{\text{maximize}} && \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} \kappa g_{M, i, \kappa} \\
& \text{subject to} && \sum_{\kappa: (m, i, \kappa) \in \mathfrak{V}} g_{m, i, \kappa} = v_{m, i} && \forall m \in \mathcal{M}, i \in S_m \\
& && \sum_{i', \kappa': ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E}} f_{m, i, \kappa, i', \kappa'} = g_{m, i, \kappa} && \forall (m, i, \kappa) \in \mathfrak{V}: m \in \{1, \dots, M-1\} \\
& && \sum_{i', \kappa': ((m-1, i', \kappa'), (m, i, \kappa)) \in \mathfrak{E}} f_{m-1, i', \kappa', i, \kappa} = g_{m, i, \kappa} && \forall (m, i, \kappa) \in \mathfrak{V}: m \in \{2, \dots, M\} \\
& && g_{m, i, \kappa} \leq 1 - x_i && \forall (m, i, \kappa) \in \mathfrak{V}: i \in \mathcal{B}_m \text{ and } \kappa \neq r_i \\
& && g_{m, i, \kappa} \leq 1 - x_i && \forall (m, i, \kappa) \in \mathfrak{V}: i \notin \mathcal{B}_m \text{ and } \kappa \in \{r_j : j \in \mathcal{B}_m\} \\
& && g_{m, i, \kappa} \leq x_j && \forall (m, i, \kappa) \in \mathfrak{V} \text{ and } j \in \mathcal{B}_m : \kappa = r_j \\
& && f_{m, i, \kappa, i', \kappa'} \geq 0 && \forall ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E} \\
& && g_{m, i, \kappa} \geq 0 && \forall (m, i, \kappa) \in \mathfrak{V}.
\end{aligned} \tag{EC.72}$$

Using the linear optimization problem (EC.72) to reformulate the objective function of (OO), we conclude for the case of $\eta = 0$ that the optimization problem (OO) is equivalent to the following mixed-integer optimization problem:

$$\begin{aligned}
& \underset{f, g, x}{\text{maximize}} && \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{B}} \kappa g_{M, i, \kappa} \\
& \text{subject to} && \sum_{\kappa: (m, i, \kappa) \in \mathfrak{B}} g_{m, i, \kappa} = v_{m, i} && \forall m \in \mathcal{M}, i \in S_m \\
& && \sum_{i', \kappa': ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E}} f_{m, i, \kappa, i', \kappa'} = g_{m, i, \kappa} && \forall (m, i, \kappa) \in \mathfrak{B}: m \in \{1, \dots, M-1\} \\
& && \sum_{i', \kappa': ((m-1, i', \kappa'), (m, i, \kappa)) \in \mathfrak{E}} f_{m-1, i', \kappa', i, \kappa} = g_{m, i, \kappa} && \forall (m, i, \kappa) \in \mathfrak{B}: m \in \{2, \dots, M\} \\
& && g_{m, i, \kappa} \leq 1 - x_i && \forall (m, i, \kappa) \in \mathfrak{B}: i \in \mathcal{B}_m \text{ and } \kappa \neq r_i \\
& && g_{m, i, \kappa} \leq 1 - x_i && \forall (m, i, \kappa) \in \mathfrak{B}: i \notin \mathcal{B}_m \text{ and } \kappa \in \{r_j: j \in \mathcal{B}_m\} \\
& && g_{m, i, \kappa} \leq x_j && \forall (m, i, \kappa) \in \mathfrak{B} \text{ and } j \in \mathcal{B}_m: \kappa = r_j \\
& && f_{m, i, \kappa, i', \kappa'} \geq 0 && \forall ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E} \\
& && g_{m, i, \kappa} \geq 0 && \forall (m, i, \kappa) \in \mathfrak{B} \\
& && x_0 = 1 \\
& && x_i \in \{0, 1\} && \forall i \in \mathcal{N}_0.
\end{aligned} \tag{EC.73}$$

Our proof of Theorem 7 is thus complete. \square

J.3. Proof of Theorem 8

Let Assumption 1 hold, let $S_1, \dots, S_M \in \mathcal{S}$, and let θ be any constant that satisfies $\theta \leq \max_{S \in \mathcal{S}} \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$.²² For any assortment $S \in \mathcal{S}$, let $x_i \in \{0, 1\}$ be a binary decision variable for each product $i \in \mathcal{N}_0$ that satisfies $x_i = 1$ if and only if product i is in the assortment S . Under this equivalence between an assortment $S \in \mathcal{S}$ and the binary decision variables $x \in \{0, 1\}^{\mathcal{N}_0}$, it follows immediately from the proof of Theorem 5 that the set $\{S \in \mathcal{S}: \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) \geq \theta\}$ can be reformulated using the binary decision variables $x \in \{0, 1\}^{\mathcal{N}_0}$ and an addition of $\mathcal{O}(\text{poly}(n, M))$ auxiliary decision variables and $\mathcal{O}(\text{poly}(n, M))$ linear constraints. Moreover, it follows immediately from the proof of Theorem 7 that the level set $\{(S, t) \in \mathcal{S} \times \mathbb{R}: \max_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S) \geq t\}$ can be reformulated using the binary decision variables $x \in \{0, 1\}^{\mathcal{N}_0}$ and an addition of $\mathcal{O}(\text{poly}(n, M))$ auxiliary decision variables and $\mathcal{O}(\text{poly}(n, M))$ linear constraints. We thus

²² It is easy to see that the requirement that $\theta \leq \max_{S \in \mathcal{S}} \min_{\lambda \in \mathcal{U}} \mathcal{R}^\lambda(S)$ is a necessary and sufficient condition for the optimization problem (PO) to have a feasible solution.

obtain the following mixed-integer optimization reformulation of (PO):

$$\begin{aligned}
& \underset{f, g, \alpha, \beta, \gamma, x}{\text{maximize}} && \sum_{i, \kappa: (M, i, \kappa) \in \mathfrak{V}} \kappa g_{M, i, \kappa} \\
& \text{subject to} && \sum_{\kappa: (m, i, \kappa) \in \mathfrak{V}} g_{m, i, \kappa} = v_{m, i} && \forall m \in \mathcal{M}, i \in S_m \\
& && \sum_{i', \kappa': ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E}} f_{m, i, \kappa, i', \kappa'} = g_{m, i, \kappa} && \forall (m, i, \kappa) \in \mathfrak{V} : m \in \{1, \dots, M-1\} \\
& && \sum_{i', \kappa': ((m-1, i', \kappa'), (m, i, \kappa)) \in \mathfrak{E}} f_{m-1, i', \kappa', i, \kappa} = g_{m, i, \kappa} && \forall (m, i, \kappa) \in \mathfrak{V} : m \in \{2, \dots, M\} \\
& && g_{m, i, \kappa} \leq 1 - x_i && \forall (m, i, \kappa) \in \mathfrak{V} : i \in \mathcal{B}_m \text{ and } \kappa \neq r_i \\
& && g_{m, i, \kappa} \leq 1 - x_i && \forall (m, i, \kappa) \in \mathfrak{V} : i \notin \mathcal{B}_m \text{ and } \kappa \in \{r_j : j \in \mathcal{B}_m\} \\
& && g_{m, i, \kappa} \leq x_j && \forall (m, i, \kappa) \in \mathfrak{V} \text{ and } j \in \mathcal{B}_m : \kappa = r_j \\
& && \sum_{m \in \mathcal{M}, i \in S_m} v_{m, i} \alpha_{m, i} \geq \theta \\
& && \alpha_{m, i} - \beta_{m, i, \kappa} \mathbb{I}\{m \in \{1, \dots, M-1\}\} - \gamma_{m, i, \kappa} \mathbb{I}\{m \in \{2, \dots, M\}\} \\
& && \leq \kappa \mathbb{I}\{m = M\} \\
& && \quad + r_n x_i \mathbb{I}\{i \in \mathcal{B}_m \text{ and } \kappa \neq r_i\} \\
& && \quad + r_n x_i \mathbb{I}\{i \notin \mathcal{B}_m \text{ and } \kappa \in \{r_j : j \in \mathcal{B}_m\}\} \\
& && \quad + \sum_{j \in \mathcal{B}_m} \mathbb{I}\{\kappa = r_j\} r_n (1 - x_j) && \forall (m, i, \kappa) \in \mathfrak{V} \\
& && \beta_{m, i, \kappa} + \gamma_{m+1, i', \kappa'} \leq 0 && \forall ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E} \\
& && x_0 = 1 \\
& && \alpha_{m, i} \in \mathbb{R} && \forall m \in \mathcal{M}, i \in S_m \\
& && \beta_{m, i, \kappa}, \gamma_{m, i, \kappa} \in \mathbb{R} && \forall (m, i, \kappa) \in \mathfrak{V} \\
& && x_i \in \{0, 1\} && \forall i \in \mathcal{N}_0 \\
& && f_{m, i, \kappa, i', \kappa'} \geq 0 && \forall ((m, i, \kappa), (m+1, i', \kappa')) \in \mathfrak{E} \\
& && g_{m, i, \kappa} \geq 0 && \forall (m, i, \kappa) \in \mathfrak{V}.
\end{aligned} \tag{EC.74}$$

Our proof of Theorem 8 is thus complete. \square