Online Assortment Optimization for Two-sided Matching Platforms

Ali Aouad
London Business School, aaouad@london.edu

Daniela Saban
Stanford Graduate School of Business, dsaban@stanford.edu

Motivated by online labor markets, we consider the online assortment optimization problem faced by a two-sided matching platform that hosts a set of suppliers waiting to match with a customer. Arriving customers are shown an assortment of suppliers, and may choose to issue a match request to one of them. After spending some time on the platform, each supplier reviews all the match requests she has received and, based on her preferences, she chooses whether to match with a customer or to leave unmatched. We study how platforms should design online assortment algorithms to maximize the expected number of matches in such two-sided settings. We establish that a simple greedy algorithm is 1/2-competitive against an optimal clairvoyant algorithm that knows in advance the full sequence of customers' arrivals. However, unlike related online assortment problems, no randomized algorithm can achieve a better competitive ratio, even in asymptotic regimes. To advance beyond this general impossibility, we consider structured settings where suppliers' preferences are described by the Multinomial Logit and Nested Logit choice models. We develop new forms of balancing algorithms, which we call *preference-aware*, that leverage structural information about suppliers' choice models to design the associated discount function. In certain settings, these algorithms attain competitive ratios provably larger than the standard "barrier" of $1 - 1/e$ in the adversarial arrival model. Our results suggest that the shape and timing of suppliers' choices play critical roles in designing online assortment algorithms for two-sided matching platforms.

Key words: online algorithms, two-sided matching, assortment optimization, choice models

1. Introduction

In recent years, online platforms that match customers with service providers have experienced sustained growth in a variety of two-sided markets. For example, travelers find accommodations on Airbnb, independent workers connect with relevant businesses using online marketplaces such as Upwork, Rover, and DesignCrowd, etc. In the majority of these settings, matches are not centrally dictated by the platform. For instance, in graphic design marketplaces such as DesignCrowd, designers browse through open projects (or design contests) and choose where to submit their designs. Generally speaking, the matches are formed through a decentralized search process initiated by one side of the market. When the corresponding users reach the platform (e.g., freelancers), they are presented with an assortment of alternatives listed on the other side of the market (e.g., job openings). Based on her preferences, each user chooses whether to request a match and with whom (e.g., applications for job openings). If a request is accepted by the other party, a match is formed. From an operational standpoint, the platform influences the matching process by selecting the assortment initially displayed to the users. This core decision is analogous to assortment planning, i.e., the question of how firms should optimize the variety of products offered to customers based on underlying assumptions on their preferences. While this issue is extensively studied in previous literature, researchers have, by and large, focused on applications to brick-and-mortar retailing as well as e-commerce, where the displayed assortment can be further personalized.

Naturally, one may consider leveraging the assortment algorithms designed for retail applications in the context of online matching platforms. However, one salient difference between retail and matching settings is the presence of one-sided versus two-sided preferences. In retailing, customers have preferences over products; a request to purchase an available product generally yields a successful transaction. By contrast, in many online matching platforms, a transaction (i.e., match) does not automatically occur after a match request is submitted; rather, a transaction occurs only after both parties agree on it. For example, for a match to happen on DesignCrowd, not only must the designer choose to submit a design to an open project, but also the client must select this design out of all the submissions she has received. Importantly, the decision of whether to accept a request and "complete the transaction" ultimately depends on the set of received match requests and the preferences of the client who reviews and evaluates them. Going forward, the notion of *two-sided preferences* refers to the fact that transactions occur only if they suit both parties.

The preceding discussion suggests a shift in perspective from one-sided marketplaces such as those for retail, where only one side has preferences and actively makes choices, to two-sided matching platforms where both parties sequentially express their preferences. The platform plays a crucial role in guiding the matching outcomes by controlling the assortment displayed to the users on the side of the market that initiates the match requests. Hence, it is natural to hypothesize that preferences on both sides of the market should be accounted for when designing online assortment algorithms in such settings. The goal of our research is twofold. First, we ask whether and when the presence of two-sided preferences has fundamental implications on the performance of online algorithms in comparison with one-sided preferences. Second, we study how to design near-optimal online assortment algorithms: what type of information is useful to optimize their design? Which characteristics of the underlying market affect their performance?

1.1. Preview of our main results

This paper studies, both conceptually and technically, how to design online assortment algorithms for two-sided matching platforms. To this end, we first introduce a model for the assortment optimization problem faced by these platforms. As formally described in Section 2.1, we take the perspective of a platform that hosts a set of *suppliers* (e.g., firms with open positions), who are waiting to match with a *customer* (e.g., freelancer). We use the terminology of customers and suppliers to simplify the comparison with models studied in the previous literature: the matching process is initiated from the customer. In practice, however, what we refer to as a customer and a supplier might not be natural in certain applications.

Our analysis leads to the following results:

1. There are significant conceptual and algorithmic differences between one-sided and two-sided assortment problems. Specifically, when considering our two-sided setting in a general form, it is not possible to design efficient algorithms achieving a competitive ratio higher than that of a simple greedy algorithm.

2. Structural information about the suppliers' choice models is useful for developing algorithms that achieve better performance guarantees. In particular, we devise so-called *preference-aware* balancing algorithms that exploit the substitution patterns of these choice models.

3. Depending on the structure of the suppliers' preferences, the online assortment optimization problem for two-sided matching markets can be easier or harder than for retail (one-sided) markets. For instance, assortment optimization is harder in markets where suppliers' preferences induce greater competition among customers.

In what follows, we discuss our contributions in more detail.

Modeling approach. Our first main contribution is to introduce a model for the online assortment optimization problem faced by a platform operating in the aforementioned markets. As formally described in Section 2.1, we take the perspective of a platform that hosts a set of suppliers who are waiting to match with a customer. Each arriving customer is shown an assortment of suppliers, and may choose to issue a match request to one of them. Before leaving the platform, each supplier reviews all the match requests she has received, and based on her preferences, she chooses whether to accept one these requests and match with a customer or to leave unmatched. The platform's objective is to devise an online assortment algorithm that maximizes the total expected number of matches over a finite time horizon. As the suppliers' choices are captured by a general set function, our formulation generalizes the online one-sided assortment optimization setting.

Differences with one-sided settings: Analysis of the greedy algorithm. As is common in the literature, we use the notion of *competitive ratio* as a performance criterion for online assortment algorithms; this notion is formally defined Section 2.1. Extending what is known with regard to special cases of the online two-sided assortment problem, we show in Section 3 that a simple greedy algorithm, which offers to each arriving customer the assortment that maximizes the expected increase in matches, is 1/2-competitive for this problem. By leveraging the hardness result for online submodular welfare maximization (Kapralov et al. 2013), we show that this performance guarantee is best-possible in the two-sided setting; i.e., there is no randomized polynomial-time algorithm that can achieve a better competitive ratio. Interestingly, the greedy algorithm remains optimal even in a suitably defined scaling of the problem instances, termed the *small probability regime*, for which improved asymptotic competitive ratios were established for related online matching and assortment settings. These results suggest that the probabilistic structure of customers' choices has no effect on the best-possible competitive ratio. By contrast, the hardness of the online two-sided assortment problem stems from the structure of suppliers' choice preferences.

Beyond greedy: Preference-aware algorithms. The impossibility of devising algorithms better than 1/2-competitive is established when suppliers' preferences are captured by any arbitrary rank-based choice model without any further (parametric) assumption. However, in real-world applications, practitioners often specify structured families of choice models, such as the Multinomial Logit (MNL) and Nested Logit (NL) models. Hence, to advance beyond this impossibility result, our main contribution is to introduce a family of simple deterministic algorithms, termed *preference-aware* balancing algorithms, that leverage this structure to achieve improved competitive ratios.

Balancing algorithms can be viewed as variants of the greedy algorithm that optimize a tradeoff between maximizing the immediate rewards from matching the current customer and leaving sufficient "capacity" to match future customers. This notion is implemented by applying a discount on the matching rewards, where the *discount function* maps the remaining capacity of each supplier to a multiplicative discount. Previous literature shows that balancing algorithms with exponentially shaped discount functions asymptotically achieve the optimal constant-factor competitive ratio of $1-1/e$ in many related settings, including the online one-sided assortment problem (Mehta et al. 2005, Golrezaei et al. 2014).

Using structural properties of the MNL and NL models, we show that carefully calibrated discount functions, which satisfy certain ordinary differential inequalities, yield balancing algorithms with better competitive ratios. Our analysis provides several interesting conceptual findings. When suppliers have MNL preferences, we construct algorithms with competitive ratios better than $1-1/e$ whereas this constant is a fundamental barrier for related online assortment problems under adversarial arrivals. Interestingly, the discount function used to attain this result significantly differs from exponentially shaped discount functions used in related settings. Moreover, our analysis of the NL choice model sheds light on structural properties of suppliers' preferences that affect the tractability of the online two-sided assortment problem. Specifically, we show that our performance guarantees improve as a function of the *nest dissimilarity coefficient* of the NL model, a parameter that intuitively controls the degree of substitution between comparable customers in a nest. Specifically, competitive ratios better than $1 - 1/e$ can be achieved for large dissimilarity coefficients; however, the performance guarantee achieved by our algorithms deteriorates to $1/2$ for small dissimilarity coefficients. This suggests that online assortment optimization is inherently more difficult in markets where suppliers' preferences capture starker competition effects between comparable customers.

Organization. In Section 1.2, we review the literature that is directly related to our work. In Section 2, we present our modeling approach and discuss its connection with practice. In Section 3, we analyze the performance of a greedy algorithm and present our general impossibility result. In Section 4, we introduce and analyze the class of preference-aware balancing algorithms. Finally, in Section 5, we summarize our main findings and discuss some modeling extensions.

1.2. Related Literature

In this section, we review three lines of research that are directly related to our paper: online matching, dynamic matching, and assortment optimization.

Online matching. Online matching problems have been widely studied in the computer science, operations research, and computational economics communities. The seminal paper Karp et al. (1990) introduces the online bipartite matching problem, where vertices on one side (online customers) are revealed sequentially and can be irrevocably matched to neighbor vertices on the other side of the graph (offline suppliers). The objective is to maximize the size of the matching. The authors propose a randomized algorithm that achieves a competitive ratio of $1 - 1/e$ and provide a matching lower bound. Numerous papers have extended the online bipartite matching model to capture more complex resource allocation settings (Mehta 2013), including the online b-matching problem (Kalyanasundaram and Pruhs 2000) and the Adwords problem (Mehta et al. 2005). Finally, there is a line of research that focuses on the design of online matching algorithms in Bayesian settings (Feldman et al. 2009, Manshadi et al. 2012, Jaillet and Lu 2013).

Closer to the present paper, Kapralov et al. (2013) study the online welfare maximization problem, a generalization of online matching where suppliers are assigned multiple customers and have submodular valuations over the set of customers. This paper establishes that no polynomial-time randomized algorithm can achieve a constant-factor competitive ratio better than $1/2$ under coverage valuation functions. As shown in Section 3, the online welfare maximization problem can be cast as a special case of our problem in its utmost generality, due to a direct analogy between coverage valuation functions and rank-based choice models. However, rather than viewing suppliers' valuations as arbitrary functions, we consider parametric choice models; these structural assumptions enable us to attain competitive ratios better than $1 - 1/e$ in certain settings. The online two-sided assortment problem is also connected to online fractional matching with concave returns (Devanur and Jain 2012), whereby competitive ratios better than $1-1/e$ are achievable. Nonetheless, the latter setting has two salient differences: items are fractionally matched rather than randomly assigned to customers' via probabilistic choice models, and rewards are captured by univariate functions as opposed to set functions.[1] Due to the latter, our algorithm and analytical approach are significantly different.

Dynamic matching. Motivated by the growth of online markets, several recent papers have studied dynamic variants of the online matching setting, where vertices arrive dynamically on the two sides of the graph; see, e.g., Gurvich and Ward (2015), Akbarpour et al. (2020), Anderson et al. (2017), Ashlagi et al. (2019), Afeche et al. (2019), and Aouad and Saritaç (2020) which propose various models in this spirit. This line of work focuses on capturing the market dynamics that describe the agents' arrivals and departures, while the matching decisions are controlled by a central platform. By contrast, we focus on a setting in which both sides of the market have complex choice preferences and the users choose whether and with whom to match.

Several recent papers study platform design questions in dynamic matching settings where users can choose among a set of potential partners; we refer the reader to the survey in Lobel (2021) that reviews recent research on marketplace management. These papers argue that various design choices can improve the efficiency of the market, including carefully selecting which side of the market should be sending the match requests and limiting the information that users see (Kanoria and Saban 2021), limiting the number of match requests a user can submit (Arnosti et al. 2021), optimizing the information disclosure policy (Bimpikis et al. 2020), and more. However, these papers do not take an algorithmic approach.

Assortment optimization. Since the seminal paper of Talluri and van Ryzin (2004), a great deal of attention in the revenue management literature has been paid to assortment optimization problems. In its standard form, the *static assortment problem* aims to determine an assortment of products that should be offered to customers in order to maximize the expected revenue with respect to a given choice model, i.e., a probabilistic prior on the customers' choice preferences. This problem has been studied under a variety of choice models and constrained formulations (Rusmevichientong et al. 2010, Sumida et al. 2020, Désir et al. 2014, Blanchet et al. 2016). Closer to our setting, Golrezaei et al. (2014) introduce what we call the *online one-sided assortment problem*, the problem of personalizing the assortment of products offered to each arriving customer given initial inventory levels for each product. This problem is a generalization of the online matching setting, where customers make random choices rather than being assigned to an item. Several extensions are

[1] While MNL preferences can be represented using similar univariate valuation functions, this is no longer true for the NL choice model.

considered in recent literature (Gong et al. 2019, Rusmevichientong et al. 2020, Ma and Simchi-Levi 2020). Our setting not only considers customers' choices but also introduces suppliers' choices, thereby further generalizing the online assortment problem in Golrezaei et al. (2014).

There are only a few papers that examine assortment optimization problems that are relevant to the design of matching markets. Motivated by applications to school choice, Shi (2019) studies the design of priority systems for matching agents to heterogeneous items, but accounts for one-sided preferences only. Ashlagi et al. (2020) study an assortment optimization problem for two-sided matching markets where customers and suppliers sequentially express their preferences. Nevertheless, this problem is formulated in a static setting, where all customers and suppliers are initially known. Rios et al. (2020) study a dynamic assortment optimization problem in collaboration with a dating app; however, they focus on practical applications, which result in significant differences in terms of modeling angle and algorithmic contributions.

2. Model

We consider a platform that facilitates the matching between customers and suppliers, both of which have exogenous preferences. The platform hosts a set of suppliers, who are waiting to match with a customer.

Customers arrive dynamically over time. Each time a customer arrives, the platform presents her with an assortment of suppliers, out of which the customer may choose at most one alternative, whereupon the corresponding supplier receives a match request. Before leaving the platform, each supplier chooses either to match with one customer among all the requests she has received or to leave unmatched; when the former happens, we say that a match occurs. The platform must choose an online assortment algorithm that maximizes the total expected number of matches over a finite time horizon. In Section 2.1, we describe each component of the model in more detail. Next, in Section 2.2, we discuss the connections with previous literature.

2.1. Problem formulation

Preliminaries. We consider a discrete-time problem over a finite horizon of T time periods, and use $[T]$ to denote the set $\{1, \dots, T\}$. Let $\mathcal{S}$ be the set of all potential suppliers that may join the platform. In line with the applications described in Section 1, suppliers need not be on the platform throughout the entire time horizon. For example, in design marketplaces, new projects arrive over time and have a deadline by which the supplier (client requesting designs) makes a decision. As a result, we allow suppliers to arrive and depart over time and denote by $\{\mathcal{S}_t\}_{t=1}^{T}$ an *exogenous* stochastic process that describes the evolution of the set of available suppliers on the platform.

Customers' choice process. In each time period $t \in [T]$, a new customer arrives. Each customer is associated with a type that specifies (i) a choice model that describes, for every given assortment, the probability that she requests a match with each supplier, and (ii) customer-specific features that determine how attractive this customer is to each supplier. For example, in labor markets such as Upwork and DesignCrowd, customers correspond to freelancers and the supply side is formed by all the jobs listed by clients on the platform. Each freelancer has preferences over the jobs available; her job applications can be viewed as a random experiment described by a choice model. In turn, due to freelancers' unique skills and each job's characteristics, a freelancer's application or design submission is more or less attractive to different clients. These customer-specific features will be implicitly captured by the suppliers' choice models, which we define subsequently.

Formally, we denote by $\mathcal{Z}$ the set of all possible customer types. Once the t-th customer arrives, her type $z_t \in \mathcal{Z}$ is revealed. Knowing the customer's type, the platform may offer her any feasible assortment $A \subseteq \mathcal{S}_t$ of suppliers, where $\mathcal{S}_t \subseteq \mathcal{S}$ denotes the set of suppliers available at the time of the customer's arrival. Then, the customer either requests a match with her preferred supplier within the assortment A, or leaves the platform unmatched. Specifically, customer t's choices are prescribed by a probabilistic choice model that depends only on her type z_t. We denote by $\phi_i^z(A)$ the probability that a customer of type $z \in \mathcal{Z}$ chooses supplier i when presented with assortment A. Consequently, we use $\xi_t(A)$ to denote the random variable that describes customer t's choice within assortment A, i.e., $\xi_t(A)$ takes values in $A \cup \{0\}$, where 0 refers to the no-match option in case customer t chooses to leave the platform and $\Pr[\xi_t(A) = i | z_t = z] = \phi_i^z(A)$ for every $i \in A$ and $z \in \mathcal{Z}$. By a slight abuse of notation, we sometimes make time implicit and use $\xi_z(A)$ to designate this multinomial random variable.

We do not impose any structural restriction on the customers' choice model. That said, we make the following computational assumption, which states that we have access to an efficient oracle to solve the corresponding class of static assortment optimization problems.

Assumption 1 **(Static oracle)**. *For all $z \in \mathcal{Z}$, there exists a polynomial-time algorithm to solve the static assortment optimization problem with respect to type z's choice model. Here, the static assortment optimization problem consists in finding $A \subseteq S$ so as to maximize $\sum_{i \in A} r_i \phi_i^z(A)$, for every subset of suppliers $S \subseteq \mathcal{S}$ and collection of non-negative per-unit revenues $\{r_i\}_{i \in S}$.*

Assumption 1 is not particularly restrictive for practical purposes since it is satisfied by a large class of choice models, including widely used logit-based models (Talluri and van Ryzin 2004, Rusmevichientong et al. 2006, Davis et al. 2014). A similar pre-requisite is common in the literature

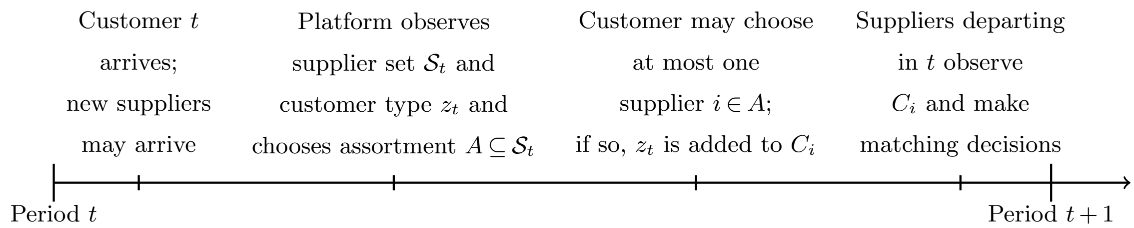

Figure 1 **Timing of events.**

on online assortment optimization; see, e.g., Golrezaei et al. (2014), Ma and Simchi-Levi (2020), and Goyal et al. (2020).[2]

Finally, we briefly discuss two natural extensions of our model. First, to ease the exposition, we assume that customers request to match with at most one supplier. That said, as long as Assumption 1 is satisfied, the case where customers submit several requests to match can be readily accommodated; we defer a detailed discussion on this topic to Section 5. Second, while we focus on assortment decisions for simplicity, all our results extend to settings in which the platform instead chooses a display ranking (Gallego et al. 2020, Aouad and Segev 2021, Sumida et al. 2020).

Suppliers' choice process. Before her permanent departure from the platform, a supplier examines the set of all customer requests she has received and decides whether to accept a match with one of these customers or to leave unmatched. If the supplier accepts a request, a match occurs.

In more detail, suppose that $C \in \mathbb{N}_0^{\mathcal{Z}}$ is the set of customer requests that supplier i has received. Here, a set $C \in \mathbb{N}_0^{\mathcal{Z}}$ is represented by a vector that counts, in coordinate z, how many customers of type z are in the set. We assume that each supplier's random choices depend only on the vector of customer requests C at the time of her departure. Specifically, the expected outcome from supplier i's choice decision is summarized through a set function $w_i : \mathbb{N}_0^{\mathcal{Z}} \to \mathbb{R}$, which we refer to as supplier i's *aggregate match probability* function (or aggregate match function, for short). Here, $w_i(C)$ is the total probability that supplier i chooses to match with any customer in C. That said, we note that all our subsequent results do not require the image of $w_i(\cdot)$ to be in the interval $[0,1]$.

In what follows, we impose very mild assumptions on the functions $w_i(\cdot)$ that are satisfied by all commonly used choice models including the class of random utility maximization models. First, for computational considerations, we assume that the functions $w_i(\cdot)$ are presented by means of a value oracle, which can answer queries of the form "*What is the value of $w_i(C)$?*"[3] Second, we require that these functions satisfy two natural properties, diminishing returns and monotonicity.

[2] Online assortment algorithms typically use a subroutine to solve the static assortment optimization problem in each time period. As such, their computational performance hinges on the ability to efficiently compute an optimal (or near-optimal) static assortment. Moreover, it is worth mentioning that Assumption 1 can be relaxed and our algorithmic results continue to hold in a weaker form: if there exists a polynomial-time β-approximation algorithm for the static assortment optimization problem, our performance guarantees degrade by a multiplicative factor of β.

[3] For most choice models, these queries can be answered in time polynomial in T, which is the maximum number of customer requests that a supplier can receive.

Definition 1 (Diminishing returns). A function $f:\mathbb{N}_0^{\mathcal{Z}} \to \mathbb{R}$ is said to satisfy the *diminishing returns* property if, for any $\mathbf{x} \le \mathbf{y}$ (coordinate-wise) and any unit basis vector $\mathbf{e}_z$ for $z \in \mathcal{Z}$, we have that $f(\mathbf{x}+\mathbf{e}_z) - f(\mathbf{x}) \ge f(\mathbf{y}+\mathbf{e}_z) - f(\mathbf{y})$.

We observe that, when restricted to the $\{0,1\}^{\mathcal{Z}}$ domain, this property is equivalent to submodularity. We also consider the following natural notion of monotonicity.

Definition 2 (Monotonicity). A function $f:\mathbb{N}_0^{\mathcal{Z}} \to \mathbb{R}$ is *monotone* if, for any $\mathbf{x} \le \mathbf{y}$, we have that $f(\mathbf{x}) \le f(\mathbf{y})$.

Consequently, we make the following assumption that we keep throughout the paper.

Assumption 2. *For every $i \in \mathcal{S}$, we assume that $w_i(\cdot)$ is monotone and satisfies the property of diminishing returns.*

To illustrate the above concepts, we introduce two examples of set functions $w_i(\cdot)$ arising from supplier-side choice models, that satisfy Assumption 2.

Definition 3 (MNL). Suppose that supplier i's preferences are governed by the Multinomial Logit (MNL) choice model. For every customer type $z \in \mathcal{Z}$, let $q_{z,i}$ be the preference weight that supplier i assigns to a customer of type z. Then, for any given customer set C,

$$w_i(C) = \frac{\sum_{z\in\mathcal{Z}} C_z \cdot q_{z,i}}{1+\sum_{z\in\mathcal{Z}} C_z \cdot q_{z,i}},$$

where C_z represents the number of customers of type z in set C and the value of the no-match option has been normalized to 1.

Definition 4 (Rank-based preferences). Suppose that supplier i's preferences are governed by a distribution $(\alpha_1,\ldots,\alpha_K)$ over ranked lists $L_1,\ldots,L_K$. Each list L_k is a subset of customer types together with a strict order on these types. For any given customer set C, supplier i samples a ranked list out of $L_1,\ldots,L_K$ according to the multinomial distribution $(\alpha_1,\ldots,\alpha_K)$ and chooses a customer of the topmost ranked type among those in C; if there is no such customer, then supplier i is unmatched. Here, we have $w_i(C) = \sum_{k\in\mathcal{K}} \alpha_k \cdot \min\{|L_k \cap C|, 1\}$, where we abuse notation and use L_k and C to denote the set of types contained in the list L_k and the customer set C, respectively.

Platform's algorithm. The platform's objective is to design an online assortment algorithm that maximizes the expected total number of matches in the T periods. An algorithm π for the *online two-sided assortment problem* describes a non-anticipating randomized sequence of assortments $\{A_t^\pi\}_{t=1}^T$ shown to the incoming customers. Formally, the stochastic process $\{A_t^\pi\}_{t=1}^T$ is $\mathcal{H}_t^\pi$-adapted, where $\{\mathcal{H}_t^\pi\}_{t=1}^T$ is the filtration describing algorithm π's history; i.e., letting s^π be π's random seed, we have $\mathcal{H}_1^\pi = \sigma(s^\pi, z_1, \mathcal{S}_1)$ and $\mathcal{H}_t^\pi = \sigma(\mathcal{H}_{t-1}^\pi, A_{t-1}^\pi, \xi_{t-1}(A_{t-1}^\pi), z_t, \mathcal{S}_t)$ for every $t \in \{2,\ldots,T\}$. We require that the assortment of suppliers be restricted to the subset of available suppliers in each time period, i.e., $A_t^\pi \subseteq \mathcal{S}_t$ for every $t \in [T]$. We denote by Π the set of all admissible online assortment algorithms. We mention in passing that our modeling approach can easily capture additional structural constraints on the assortments selected in each time period, such as constraints on the number of suppliers shown to each customer.

The platform's objective is to maximize the total expected number of successful matches, denoted by $\mathcal{M}^\pi = \mathbb{E}[\sum_{i\in\mathcal{S}} w_i(C_i^{T+1,\pi})]$, where $C_i^{t,\pi} \in \mathbb{N}_0^{\mathcal{Z}}$ is the set of customer requests that supplier i has received by the beginning of period t, i.e., the requests from the first $t-1$ arriving customers. As the platform may only show suppliers who are available, the set $C_i^{T+1,\pi}$ coincides with the set seen by supplier i immediately before departing. Thus, without loss of generality, we sometimes assume that suppliers' choices occur at the end of the time horizon.

To characterize the performance of an algorithm with respect to specific arrival patterns, we denote by $\mathcal{M}^\pi(\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T)$ the expected number of matches conditional on the sequences of customer types and of suppliers available given by $\{z_t\}_{t=1}^T$ and $\{\mathcal{S}_t\}_{t=1}^T$, respectively. Here, the expectation is taken with respect to the choices made by each customer, the choices made by the suppliers, and possibly random selections of the algorithm, if the latter is not deterministic.

Competitive ratio. In the remainder of the paper, we do not impose any assumption on the sequence of arrivals and suppliers, unless explicitly said otherwise. Thus, our goal is to devise online assortment algorithms that are robust to arbitrary arrival patterns $\{z_t\}_{t=1}^T$ and $\{\mathcal{S}_t\}_{t=1}^T$. To analyze the performance of a proposed algorithm $\pi \in \Pi$, we compare it to an optimal *clairvoyant algorithm* π^C that knows a priori the realization of the sequence of customers' arrivals $\{z_t\}_{t=1}^T$, and of suppliers' sets $\{\mathcal{S}_t\}_{t=1}^T$, but does not know the realization of customers' and suppliers' choices. We define the *competitive ratio* as the following quantity that measures the relative expected number of matches generated by π against π^C:

$$\inf_T \inf_{\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T} \frac{\mathcal{M}^\pi\left(\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T\right)}{\mathcal{M}^{\pi^C}\left(\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T\right)} .$$

Unless specified otherwise, this criterion is used to characterize the performance of online assortment algorithms. To conclude, we note that the same online optimization problem can be specified more succinctly by "folding" the information regarding the sequence of available suppliers $\{\mathcal{S}_t\}_{t=1}^T$ into the customer types[4] $\{z_t\}_{t=1}^T$. However, we include a time-varying set of available suppliers in our model description as it is more natural for the applications of interest.

Asymptotic analysis: Small probability regime. For purposes of tractability, the related literature (see Section 1.2) often focuses on asymptotic competitive analysis by considering a scaling of the problem instances. At a high level, the initial amount of resources is assumed to be large compared

[4] This means that we redefine the collection of customer types as $\mathcal{Z} \times 2^{\mathcal{S}}$. Here, the choice model of a new customer type $z' = (z, S)$ with $z \in \mathcal{Z}$ and $S \subseteq \mathcal{S}$ satisfies $\phi_i^{z'}(A) = \phi_i^z(A \cap S)$ for every $i \in \mathcal{S} \cup \{0\}$ and $A \subseteq \mathcal{S}$.

to the quantity depleted by each individual match. For example, the online one-sided assortment problem is analyzed under the large inventory regime, meaning that the initial inventory of each product tends to infinity (Golrezaei et al. 2014, Rusmevichientong et al. 2020, Goyal et al. 2020). Similarly, the small bid regime introduced in Mehta et al. (2005) for the AdWords problem assumes that bids are small relative to the total advertisers' budgets. Along similar lines, we introduce the *small probability regime* in the context of the online two-sided assortment problem.

Definition 5 (ϵ-small probability regime). Given a parameter $\epsilon \in (0,1)$, we define the ϵ-*small probability regime* as the family of instances for which $w_i(\boldsymbol{e}_z) \le \epsilon \max_{C \in \mathbb{N}_0^{\mathcal{Z}}} w_i(C)$ for all $i \in \mathcal{S}$ and $z \in \mathcal{Z}$.

Technically speaking, the small probability condition generalizes both the small bid and large inventory regimes discussed above. This connection is formally stated in Appendix EC.1.2. From a practical standpoint, this regime best applies to markets where the suppliers are "hard-to-please," meaning that each individual match request has a small probability of being successful. As a concrete example in which this regime would be meaningful, contests listed on DesignContest.UK often receive hundreds of submissions, whereas typically one submission is selected. As a result, the probability that a designer wins any given contest tends to be small.

2.2. Connection with related optimization settings

We conclude this section by describing in more detail the connections between the online two-sided assortment problem described in Section 2.1 and related online optimization problems.

Connection to online (one-sided) assortment. We first argue that our problem formulation subsumes the standard online one-sided assortment problem introduced by Golrezaei et al. (2014) as a special case. Here, an instance is defined by the inputs $(T, \mathcal{Z}, \{c_i\}_{i\in\mathcal{S}}, \{r_i\}_{i\in\mathcal{S}})$, where T is the time horizon, $\mathcal{Z}$ is the set of customer types, $\mathcal{S}$ is the collection of offline resources, c_i is the initial inventory for each offline item $i \in \mathcal{S}$, and r_i is the revenue associated with selling one unit of item i. Given an instance of the online one-sided assortment problem, for each offline item $i \in \mathcal{S}$, we specify a corresponding set function $w_i(\cdot)$ such that $w_i(C) = r_i \cdot \min\{|C|, c_i\}$ for every customer set C. Next, we construct an equivalent instance of the online two-sided assortment problem that has identical collections of customer types $\mathcal{Z}$ and suppliers $\mathcal{S}$; each supplier $i \in \mathcal{S}$ is endowed with the aggregate match function $w_i(\cdot)$. This instance clearly satisfies all the assumptions stated in Section 2.1. This example illustrates that, while for simplicity we interpret the functions $\{w_i(\cdot)\}_{i\in\mathcal{S}}$ as suppliers' choice probabilities over customers' requests, we do not restrict the image of these functions to be in the $[0,1]$ domain. Therefore, we can easily incorporate supplier-dependent weights, akin to vertex weights in online matching and per-unit revenues in assortment optimization problems.

To highlight conceptual differences between the online one-sided and two-sided assortment problems, as a brief detour, we first examine a variant of our problem formulation with *immediate supplier feedback.* In this context, immediately after receiving a match request, each supplier irrevocably decides whether to reject the request or to accept it and leave the platform. In Appendix EC.1.1, we show that the resulting online optimization problem is in fact reducible to the online one-sided assortment problem.[5] Hence, in two-sided markets with immediate supplier feedback, it is possible to leverage the online assortment algorithms developed by the previous literature for the one-sided setting. While two-sided markets with immediate supplier feedback need not be different from one-sided markets, many settings of interest do not fit the latter framework. For instance, in the online labor marketplaces mentioned in Section 1, job openings have deadlines and "suppliers" accumulate a number of match requests before making a decision. As shown by our subsequent analysis, under this form of delay, the online two-sided assortment problem can no longer be mapped to its one-sided counterpart, and fundamental differences between these settings arise.

Connection to online matching with submodular valuations. At the other end of the spectrum, our modeling approach for the two-sided assortment problem has close connections with the online welfare maximization problem with submodular valuations, which was studied by Kapralov et al. (2013) among others. Here, customers arrive online; each customer can be assigned to an offline resource within a subset of compatible suppliers, revealed upon her arrival. In turn, suppliers have submodular valuations over the subset of customers they have been assigned to, and the objective is to maximize the sum of suppliers' valuations. Technically speaking, this online optimization problem can be viewed as a special case of our problem formulation whereby suppliers can only be offered a singleton assortment containing a compatible supplier, which is then chosen with probability 1; i.e., letting $S \subseteq \mathcal{S}$ be the subset of compatible suppliers for customer type $z \in \mathcal{Z}$, we have $\phi_i^z(\{j\}) = 1$ for all $j \in S$ and $\phi_i^z(\{j\}) = 0$ for all $j \in \mathcal{S} \setminus S$. This reduction is made formal in Section 3.2 to establish our hardness result.

Despite this formal connection, the online two-sided assortment problem presents a significant modeling extension by capturing stochastic customers' choices. At this stage, it is important to note that the incorporation of customers' choices is precisely what differentiates the online (one-sided) assortment problem from the classical online matching problem. The implications of this generalization are non-trivial since the best-possible competitive ratio for the online assortment problem is still unknown (Golrezaei et al. 2014), even in special cases (Goyal and Udwani 2020).

[5] Intuitively, a match between a customer and a supplier occurs if and only if two (independent) events occur: (i) the customer, after being presented with an assortment, requests to match with that supplier and, (ii) conditional on receiving the request, the supplier accepts it. Thus, an equivalent instance of the one-sided problem is constructed by specifying customers' choice models formed by the multiplicative combination of the probabilities of these events.

The added complexity is closely related to the formulation of a suitable benchmark due to the post-allocation stochasticity of customers' choices. Moreover, in online welfare maximization problems, the valuation functions (the counterpart of the aggregate matching function in our setting) can be arbitrarily and adversarially chosen under Assumptions 1 and 2. By contrast, we aim to better understand how the structure of the suppliers' choice models affect the performance of online algorithms. Thus, in Section 4, we study special cases in which suppliers' preferences are governed by practical families of choice models.

3. The Optimality of the Greedy Algorithm

The goal of this section is to better understand the differences between the online one-sided and two-sided assortment problems. For this purpose, we establish a clear distinction between these settings: the best-possible constant-factor competitive ratio in the two-sided setting is worse than that achievable in the one-sided setting. In fact, we show that a simple greedy algorithm is generally optimal, and we discuss a number of conceptual implications of this result.

3.1. The greedy algorithm

We formally introduce the greedy algorithm. As before, we denote by $\mathcal{S}_t$ the set of suppliers present on the platform at time t. Recall that $C_i^{t,\pi}$ stands for the set of customer requests that supplier $i \in \mathcal{S}$ has received by the beginning of time t under algorithm π. Define $\bar{C}^{t,\pi} = (C_i^{t,\pi})_{i\in\mathcal{S}}$ to be the sets of customer requests over all suppliers in vector form.

Suppose that a customer of type $z \in \mathcal{Z}$ arrives at a time when the set of customer requests is $\bar{C} = (C_i)_{i\in\mathcal{S}}$. If the arriving customer is shown assortment $A \subseteq \mathcal{S}$, the total expected number of matches increases by the following quantity:

$$\begin{aligned} R_z\left(\bar{C}, A\right) &= \mathbb{E}_{\xi_z(A)}\left[\sum_{i\in A} \left(w_i(C_i + \mathbf{e}_z \cdot \mathbb{I}\left[\xi_z(A) = i\right]) - w_i(C_i)\right)\right] \\ &= \sum_{i\in A} \underbrace{\phi_i^z(A)\left(w_i(C_i + \mathbf{e}_z) - w_i(C_i)\right)}_{\text{expected marginal increase in i's matches}} \ , \end{aligned}$$

In what follows, the critical quantity $R_z(\bar{C}, A)$ is called the *marginal reward*, for short.

The *greedy algorithm* π^g is the algorithm that maximizes the marginal reward in each time period. That is, for every $t \in [T]$, the greedy algorithm presents the t-th arriving customer with an assortment A^{π^g} satisfying

$$A^{\pi^g} \in \arg\max_{A\subseteq\mathcal{S}_t} \ R_{z_t}\left(\bar{C}^{t,\pi^g}, A\right) \ . \tag{1}$$

We note that the maximization problem (1) is an instance of the static assortment optimization problem defined in Assumption 1, where the "revenue" associated with each supplier $i \in \mathcal{S}_t$ is equal to $r_i = w_i(C_i^{t,\pi^g} + \mathbf{e}_{z_t}) - w_i(C_i^{t,\pi^g})$. Hence, by Assumption 1, the selection of an assortment at each greedy step can be solved in polynomial time. It is worth noting that the greedy algorithm is adaptive. By the property of diminishing returns (Assumption 2), the contribution of each supplier to the marginal reward is non-increasing over time, and thus the greedy algorithm is less likely to pick those suppliers who are most likely to match given the customer requests they have received so far. Nonetheless, the greedy algorithm is myopic: it takes into account the type of the current customer and the set of requests received by the suppliers so far, but it ignores the impact of the current decision on the marginal rewards to be collected in the future. Moreover, the information about suppliers' preferences required by the algorithm is minimal: no specific (or distributional) knowledge about the suppliers' choice model is needed other than the information revealed up until the current customer.

3.2. Performance of the greedy algorithm

Here, we analyze the greedy algorithm in a general setting, where Assumption 2 is the only restriction on the suppliers' aggregate match probability functions. The proofs of the results presented in this section are deferred to Appendix EC.2.2. We begin by stating the performance guarantee for the greedy algorithm.

PROPOSITION 1. *The greedy algorithm achieves a competitive ratio of 1/2 for the online two-sided assortment problem.*

The proof follows the same line of argumentation as that of Chan and Farias (2009), which identifies sufficient conditions for the greedy algorithm to be 1/2-competitive. This result cannot be directly invoked in our setting since the depletion of resources is described by multiset state variables. Nonetheless, we establish Proposition 1 by showing that the sufficient conditions of Chan and Farias (2009) are easily extended to our setting.

It is well known that the greedy algorithm is 1/2-competitive in the special cases discussed in Section 2.2, including the online matching problem and the online one-sided assortment problem. However, there exist algorithms with better (constant-factor) competitive ratios in these settings. For instance, Karp et al. (1990) show that the best-possible competitive ratio for the online matching problem is $1-1/e$, thus substantially improving upon greedy's 1/2-competitiveness. Similarly, in the ϵ-small probability regime defined in Section 2, the previous literature has developed algorithms that asymptotically attain a competitive ratio of $1-1/e$ as ϵ tends to zero for the b-matching problem (Kalyanasundaram and Pruhs 2000), the Adwords problem (Mehta et al. 2005), and the online one-sided assortment problem (Golrezaei et al. 2014). By contrast, our next proposition shows that there is no efficient randomized online algorithm that achieves a competitive ratio better than 1/2, not even asymptotically.

PROPOSITION 2. *Unless*[6] *NP = RP, for any constants* $\delta, \epsilon > 0$*, there is no* $(1/2 + \delta)$*-competitive polynomial-time algorithm for the online two-sided assortment problem when suppliers have rank-based preferences in the* ϵ*-small probability regime.*

The above proposition establishes a fundamental difference between the one-sided and two-sided online assortment problems: the two-sided setting is not only a generalization of the one-sided one, but also harder to solve. This difference emerges from a connection to the online welfare maximization problem. In this context, Kapralov et al. (2013) establish that no online algorithm (even randomized, against an oblivious adversary) is better than 1/2-competitive for online welfare maximization with coverage valuations. Our proof of Proposition 2 exploits this result by providing a direct mapping between coverage valuations and rank-based preferences. We show that the restriction imposed by the ϵ-small probability assumption can be viewed as a cardinality constraint on the set system generating the valuation functions. This constraint can be imposed on the "hard" instances without any loss.

It is instructive to contrast the effects of incorporating customers' preferences versus suppliers' preferences on the performance of online algorithms. The celebrated result of Karp et al. (1990) shows that no online algorithm can be better than $(1 - 1/e)$-competitive for the online matching problem, even though the offline version of this problem (maximum cardinality matching) can be optimally solved in polynomial time. Naturally, this negative result also applies to the online one-sided assortment problem, which is a generalization of the online matching problem. However, in this setting as well, there exist polynomial-time algorithms attaining $(1 - 1/e)$-competitiveness in an asymptotic sense, as shown by Golrezaei et al. (2014). Hence, this hardness stems mainly from the online nature of the problem (i.e., not knowing the sequence of arrivals in advance), rather than from the complexity added by modeling customers' preferences. Focusing now on suppliers' preferences, the above negative result shows that incorporating such preferences limits the best-possible constant-factor competitive ratio to 1/2. To explain this phenomenon, note that, in addition to facing the challenges related to the online nature of the problem, even the offline version of the welfare maximization problem is hard to approximate. This inapproximability is crucially exploited in the proof of Kapralov et al. (2013).

We conclude this section by showing that a stochastic assumption on the arrival process leads to improved performance guarantees. Specifically, we tightly characterize the competitive ratio of the greedy algorithm under i.i.d. sequences of customers' arrivals.

[6] Randomized polynomial time (RP) is the complexity class of decision problems for which a polynomial-time randomized algorithm exists such that it always returns NO when the correct answer is NO, and returns YES with probability at least c for a fixed constant $c \in (0, 1]$ whenever the correct answer is YES (Motwani and Raghavan 1995).

Proposition 3. *Suppose that the set of suppliers is time invariant (i.e., $\mathcal{S}_t = \mathcal{S}$) and that in every period $t = 1, \ldots, T$, the type of an arriving customer, z_t, is drawn independently and identically from a common distribution $\mathcal{D}$ over the set of types $\mathcal{Z}$. Then, we have*

$$\inf_{\mathcal{D}} \frac{\mathbb{E}_{\{z_t\}_{t=1}^T}\left[\mathcal{M}^{\pi^g}\left(\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T\right)\right]}{\mathbb{E}_{\{z_t\}_{t=1}^T}\left[\mathcal{M}^{\pi^C}\left(\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T\right)\right]} \geq 1 - \frac{1}{e}.$$

Moreover, unless NP = RP, there is no polynomial-time algorithm for the online two-sided assortment problem with rank-based suppliers' preferences that achieves a competitive ratio of $1 - 1/e + \delta$, for fixed $\delta > 0$, in this setting.

Consequently, the performance of the greedy algorithm is optimal even under an i.i.d. arrival process. This result further highlights the differences between the one-sided and two-sided settings. Indeed, no algorithm is better than $(1 - 1/e)$-competitive in the two-sided setting, whereas Golrezaei et al. (2014) present an algorithm that asymptotically achieves a competitive ratio of at least 3/4 in the one-sided setting.

The preceding results suggest that the hardness of the online two-sided assortment problem is driven by the lack of immediate supplier feedback, which ultimately allows suppliers to choose from a set of customer requests according to rank-based preferences: as in the one-sided setting, the uncertainty associated with the customers' choice process does not make the problem harder.

Our analysis so far has focused on rank-based suppliers' choice models. However, this modeling approach is almost generic: rank-based models have the power to express the class of random utility maximization choice models within any desired degree of accuracy (Rusmevichientong et al. 2006, Farias et al. 2013). The optimality of the greedy algorithm in such a general setting raises the question of whether improved competitive algorithms can be developed by leveraging additional structural properties of suppliers' preferences. This important practical development is examined in the next section.

4. Beyond Greedy: Accounting for Suppliers' Preferences

In Section 3, we established that the greedy algorithm is essentially optimal under general rank-based preferences. Here, our objective is to demonstrate that structural properties of suppliers' choice models may be useful for constructing more competitive algorithms. From an algorithmic perspective, we develop new *preference-aware* balancing algorithms, which leverage general information about suppliers' choice models to achieve improved performance guarantees. From a conceptual standpoint, we show that the "shape" of these choice models (a notion we will make precise in the sequel) directly affects the performance of online assortment algorithms.

4.1. General framework: Preference-aware balancing algorithms

We start by describing the general framework within which our subsequent results unfold. As discussed in Section 3, the related literature shows that it is possible to develop competitive algorithms better than the greedy algorithm, with near-optimal asymptotic performance for several special cases of the online two-sided assortment problem. Results in this spirit include the b-matching problem (Kalyanasundaram and Pruhs 2000), the AdWords problem (Mehta et al. 2005), and the online one-sided assortment problem (Golrezaei et al. 2014). The approach followed by these papers is based on two ingredients, which we describe using the terminology of our model:

(i) *Asymptotic regime:* Instances are analyzed under a suitable problem scaling, whereby the amount of offline resources depleted by each individual assignment is small relative to their initial quantities. As explained in Section 2, this condition is captured in our setting by the ϵ-small probability assumption (Definition 5) for some parameter $\epsilon \in (0,1)$.

(ij) *Balancing algorithm:* A so-called balancing algorithm is introduced. Intuitively, these algorithms aim to balance between maximizing the increase in matches due to the current customer and leaving sufficient capacity to match future customers.

Similarly to previous literature, we define the family of balancing algorithms for the online two-sided assortment problem. Next, we conceptually explain how the design of these algorithms can be guided by underlying assumptions on suppliers' choice models. Throughout the rest of the section, to simplify the exposition, we assume that the aggregate match functions $w_i(\cdot)$ have an image in the $[0,1]$ domain.

The family of balancing algorithms. The central ingredient of balancing algorithms is a *discount function* $f(\cdot):[0,1]\to[0,1]$, which is usually assumed to be monotone non-decreasing. Each discount function $f(\cdot)$ induces a corresponding balancing algorithm, which we denote by π^f. Specifically, for every $t\in[T]$, the assortment $A_t^{\pi^f}$ offered to the t-th customer is picked according to the static optimization problem:

$$A_t^{\pi^f} \in \operatorname*{arg\,max}_{A\subseteq \mathcal{S}_t} \sum_{i\in A} \phi_i^{z_t}(A) \cdot \underbrace{\left(w(C_i^{t,\pi^f}+\boldsymbol{e}_{z_t}) - w(C_i^{t,\pi^f})\right)}_{\text{Marginal increase in matches}} \cdot \underbrace{\left(1 - f\left(w_i\left(C_i^{t,\pi^f}\right)\right)\right)}_{\text{Discount due to past requests}} . \tag{2}$$

Here, recall that $C_i^{t,\pi}$ stands for the (random) set of arriving customers who have requested to match with supplier i before period t (i.e., by the time customer t arrives) under algorithm π. To gain intuition on Equation (2), we note that the greedy algorithm is a special case of the family of balancing algorithms for which the discount is constant and equal to zero; i.e., by specifying $f(x)=0$ for all $x\in[0,1]$, we recover Equation (1). In general, at each arrival $t\in[T]$, balancing algorithms solve a modified static assortment problem, where the expected marginal increase in the aggregate match probability of each supplier $i\in\mathcal{S}$ is discounted by the multiplicative factor

$1 - f(w_i(C_i^{t,\pi^f}))$. Intuitively, this discount is related to the notion of opportunity cost: the match request sent by the current customer reduces the platform's ability to match the corresponding supplier with future customers. The discount level increases as a supplier is more likely to be satisfied with past customer requests; this likelihood is captured by the state variable $w_i(C_i^{t,\pi^f})$. We mention in passing that discount-based algorithms in this spirit were also employed in the literature on online matching with stochastic rewards (Mehta and Panigrahi 2012, Mehta et al. 2015, Goyal and Udwani 2020).

Preference-awareness. When designing such balancing algorithms, the intriguing question is: how should the discount function be specified for optimal performance? Quite surprisingly, most papers in the above-mentioned literature show that an exponentially shaped discount function, e.g., $f(x) = e^{-(1-x)}$, yields an asymptotically optimal balancing algorithm. However, as shown in Section 3, one challenge inherent in the online two-sided assortment problem is that no polynomial-time algorithm can achieve a competitive ratio better than 1/2, even asymptotically in the small probability regime (Proposition 2). Hence, regardless of how the discount function is specified, the resulting balancing algorithm is subject to the same fundamental limitation.

To advance beyond this impossibility result, we restrict attention to practical instances of the two-sided assortment problem. Key to our hardness reduction in the proof of Proposition 2 is the fact that suppliers' preferences are formed by arbitrary rank-based choice models. By contrast, we now turn our attention to widely used parametric choice models, namely, the Multinomial Logit (MNL) and Nested Logit (NL) models, which accurately describe preference rankings in real-world settings. In Section 4.2, we present a calibration method for the discount function that embeds the parametric structure of suppliers' preferences by using non-linear ordinary differential inequalities (ODI). We call this approach *preference-awareness*. We show that preference-aware balancing algorithms achieve constant-factor competitive ratios significantly better than 1/2.

Going forward, it is worth noting that our notion of information on the structure of the suppliers' choice models is minimal; we assume only that suppliers' choice models fall in a broad class of choice models (e.g., MNL), out of which any particular instance is feasible. Our competitive analysis continues to be under adversarial arrivals, meaning that the performance guarantees are robust to worst-case instances within the class of choice models that we consider.

4.2. Main results

In this section, we develop the preference-aware algorithmic approach described in Section 4.1. To present our approach in its simplest possible form, we first examine the special case of the MNL choice model, for which our ordinary differential inequalities are explicitly formulated. Next, we consider the NL model in full generality; here, our approach is more technically involved and the crucial ODI is only implicitly specified. To simplify our notation, without loss of generality, we assume that each arriving customer is of a different type, i.e., $\mathcal{Z} = [T]$.

Warm-up: MNL model. We assume that the suppliers' aggregate match probability functions $w_i(\cdot)$ are consistent with the MNL choice model. Recall from Definition 3 that the MNL model is described as follows. For every supplier $i \in \mathcal{S}$ and customer set $C \in \{0,1\}^T$, we have

$$w_i(C) = \frac{\sum_{t\in[T]} C_t \cdot q_{i,t}}{1 + \sum_{t\in[T]} C_t \cdot q_{i,t}} \ ,$$

where the weight parameter $q_{i,t} > 0$ measures the attractiveness of customer t to supplier i. Under such choice models, the ϵ-small probability assumption is equivalent to the condition $q_{i,t} \leq \frac{\epsilon}{1-\epsilon}$ for all $i \in \mathcal{S}$ and $t \in [T]$. We highlight that the weight parameters depend on both the customer type and the supplier. Hence, the MNL model already captures a rich spectrum of heterogeneous suppliers' preferences from a practical standpoint.

We proceed by introducing the family of preference-aware balancing algorithms for the MNL choice model. To this end, the discount function is constructed using the following ODI in the variable $x \in [0,1]$, parametrized by $\kappa \in (0,1)$:

$$\sqrt{\int_0^x f(u)du} + \sqrt{1 - f(x)} \cdot (1-x) \geq \sqrt{\kappa} \ . \tag{MNL-ODI}$$

Let κ^* be the supremum value of the parameter $\kappa \in (0,1)$ for which the inequality (MNL-ODI) admits a non-decreasing Lipschitz solution $f : [0,1] \to [0,1]$. For all $\kappa \in (0, \kappa^*)$, we denote by f_κ a non-decreasing Lipschitz function that satisfies the latter properties; let $\eta_\kappa > 0$ stand for a Lipschitz constant associated with f_κ, i.e., $|f_\kappa(x) - f_\kappa(y)| \leq \eta_\kappa \cdot |x-y|$ for all $x, y \in [0,1]$. Later, we will explain how these discount functions are computed, and how the value of κ^* is numerically approximated. For the time being, we note that $\kappa^* \geq 1/2$, since the inequality (MNL-ODI) is met when specifying the constant function $f(x) = 1/2$ with $\kappa = 1/2$.

The family of preference-aware balancing algorithms for the MNL choice model corresponds to the balancing algorithms π^{f_κ} obtained by specifying the discount functions f_κ where $\kappa \in (0, \kappa^*)$. The next theorem shows that the preference-aware balancing algorithm π^{f_κ} asymptotically achieves a competitive ratio of $\kappa - O(\epsilon)$ in the ϵ-small probability regime, as ϵ tends to zero.

THEOREM 1. *Fix $\kappa \in (0, \kappa^*)$ and let $f_\kappa(\cdot)$ be a corresponding Lipschitz non-decreasing discount function that satisfies the ODI in* (MNL-ODI) *with a corresponding Lipschitz constant η_κ. Then, for all $\epsilon \in (0, 1/\eta_\kappa)$, the preference-aware balancing algorithm π^{f_κ} achieves a competitive ratio of $\kappa - (\kappa + \kappa\eta_\kappa)\epsilon + \kappa\eta_\kappa\epsilon^2$ for the online two-sided assortment problem with MNL suppliers' preferences in the ϵ-small probability regime.*

A detailed outline of the proof is provided in Section 4.4, where we invoke the main technical ingredients of our analysis and give further intuition on the ODI. Rather than relying on a linear-programming benchmark, we construct an upper bound on the optimal expected number of matches by using a careful decomposition of the expected number of matches in the spirit of Goyal et al. (2020). Next, by revealing the connection with inequality (MNL-ODI), we show that this upper bound is matched by the performance of our preference-aware algorithm up to a constant factor; inequality (MNL-ODI) is closely related to the substitution patterns prescribed by the MNL choice model.

An important caveat is that Theorem 1 is non-constructive. This raises the question of whether it is possible to construct a discount function that satisfies inequality (MNL-ODI). We answer this question in the affirmative in Appendix EC.3.6 by formulating an efficient dynamic program that constructs a preference-aware discount function f_κ for every $\kappa \in (0, \kappa^*)$ and certifies all its required properties. Consequently, we numerically show that $\kappa^* \geq 0.67$ using a computer-aided approach. These findings are summarized by the following claim.

COROLLARY 1. *There exists an online assortment algorithm that achieves a competitive ratio of* $0.67 - O(\epsilon) > 1 - \frac{1}{e} - O(\epsilon)$ *for the online two-sided assortment problem with MNL-based suppliers' preferences in the* ϵ*-small probability regime.*

General case: $\boldsymbol{\gamma}$*-NL model.* We now describe our result in the most general setting. We start by formally defining the γ-Nested Logit choice model for any given parameter $\gamma \in (0, 1]$.

DEFINITION 6 (γ-NESTED LOGIT). Let $\gamma \in (0, 1]$. Suppose that supplier i's preferences are governed by the γ*-Nested Logit choice model* with partially captured nests (abbreviated γ-NL). Then, there exists a partition $N_1, \ldots, N_K$ of the set of customer types $[T]$ into nests, and non-negative weights $\{q_{i,t}\}_{t=1}^T$ such that, for every customer set $C \in \{0, 1\}^T$,

$$w_i(C) = \frac{\sum_{k=1}^K (\sum_{t \in N_k} C_t \cdot q_{i,t})^\gamma}{1 + \sum_{k=1}^K (\sum_{t \in N_k} C_t \cdot q_{i,t})^\gamma} .$$

The parameter γ is referred to as the *nest dissimilarity coefficient*; the MNL model corresponds to the case where $\gamma = 1$. By restricting attention to the regime $\gamma \in (0, 1]$, our modeling approach conforms to the standard form of the Nested Logit model (Börsch-Supan 1990), in which the corresponding aggregate match functions $\{w_i(\cdot)\}_{i \in \mathcal{S}}$ satisfy the property of diminishing returns (Assumption 2). Interestingly, the nest dissimilarity coefficient γ has concrete implications for suppliers' preferences; smaller values of γ capture starker competition effects between customers in each nest. Specifically, when $\gamma = 1$, customers within a nest are "dissimilar," and thus the benefit of an extra request is independent of how the previous customers are distributed over the nests. By contrast, as γ decreases, customers within a nest are more "similar" and the requests in the same nest cannibalize each other's acceptance probabilities. As γ approaches zero, the benefit of a second customer request in the same nest is only marginal. Technically speaking, these properties are related to the curvature of the aggregate match functions.

Next, we present our ODI in order to construct preference-aware discount functions under the γ-NL choice model. We assume that the platform knows the value of $\gamma \in (0,1]$ that is consistent with all suppliers' choice models. With this information at hand, we introduce the ODI in the variable $x \in [0,1]$, parametrized by $\kappa \in (0,1)$, where

$$\min_{\alpha \in (0,1)} \frac{1}{\alpha} \cdot \left(\int_0^x f(u)du + (1 - f(x)) \cdot x(1-x) \cdot \left(\left(1 + \left(\frac{\alpha \cdot (1-x)}{x \cdot (1-\alpha)} \right)^{\frac{1}{\gamma}} \right)^{\gamma} - 1 \right) \right) \geq \kappa \ . \qquad \text{(NL-ODI)}$$

It is not difficult to verify that, when $\gamma = 1$, inequalities (NL-ODI) and (MNL-ODI) are identical. In this case, the inner optimization problem with respect to $\alpha \in (0,1)$ can be easily solved in closed form. Extending our previous notation, we denote by $\kappa^*(\gamma)$ the supremum value of the parameter $\kappa \in (0,1)$ for which inequality (NL-ODI) admits a non-decreasing Lipschitz solution $f : [0,1] \to [0,1]$. We denote by $f_{\gamma,\kappa}$ a discount function with the latter properties. We are now ready to state our main analytical result.

Theorem 2. *Fix $\gamma \in (0,1]$ and $\kappa \in (0, \kappa^*(\gamma))$. Then, there exists $\epsilon(\kappa,\gamma) > 0$ such that, for all $\epsilon \in (0, \epsilon(\kappa,\gamma))$, the preference-aware balancing algorithm $\pi^{f_{\gamma,\kappa}}$ achieves a competitive ratio of $\kappa - O(\epsilon)$ for the online two-sided assortment problem with γ-NL-based suppliers' preferences in the ϵ-small probability regime.*

Due to lengthy technical details, the proof is presented in Appendix EC.3.4. Our analysis is similar to that of the MNL setting; in particular, we recover the statement of Theorem 1 when $\gamma = 1$. Nonetheless, the substitution patterns of the γ-NL choice model are far more complex to analyze than those of the MNL model. The γ-NL choice model adds a combinatorial dimension to the system state: for each supplier $i \in \mathcal{S}$, the state variable $w_i(C^{t,\pi^{f_{\gamma,\kappa}}})$ is jointly characterized by the nest-level quantities $\sum_{\tau \in N_k \cap [t-1]} C^{\tau,\pi^{f_{\gamma,\kappa}}} \cdot q_{i,\tau}$ for every $k \in [K]$, rather than the aggregate quantity $\sum_{\tau \leq t-1} C^{\tau,\pi^{f_{\gamma,\kappa}}} \cdot q_{i,\tau}$ for the MNL model. Hence, our competitive analysis is based on a non-convex mathematical program that reveals the worst-case arrival sequence under the γ-NL choice model; this mathematical program is closely related to inequality (NL-ODI).

4.3. Analytical and managerial implications

Here, we discuss various implications of our algorithmic results and analysis.

Beating the $(1 - 1/e)$ *barrier.* The significance of Theorems 1 and 2 is to provide rigorous evidence that online algorithms may leverage the structure of suppliers' preferences to achieve improved performance guarantees. This is an important deviation from the one-sided assortment problem, where the structure of the customers' choice models has no bearing on the design of competitive algorithms, other than the computational aspects of the static assortment problem.

To our knowledge, our result in the MNL setting is the first competitive algorithm that beats the $(1 - 1/e)$ barrier for an online assortment problem with adversarial (non-stochastic) arrivals.

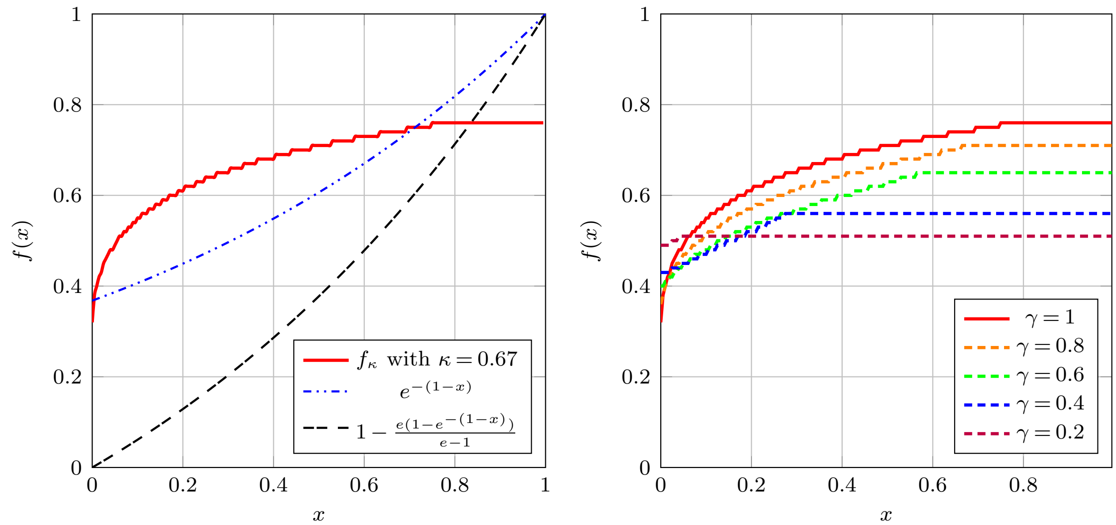

Figure 2 **Visualization of various discount functions. Left-hand plot: The solid (red) line represents the function f_κ for $\kappa = 0.67$ generated by our preference-aware approach in the MNL setting. The dashed (black) and the dotted-dashed (blue) functions represent the discount functions used in the one-sided assortment and AdWords problems, respectively. Right-hand plot: Plots of $f_{\gamma,\kappa^*(\gamma)}$ for different values of $\gamma \in (0,1]$ using close underestimates of $\kappa^*(\gamma)$.**

At first glance, this fact might be surprising in light of the negative result of Section 3.2. Indeed, Proposition 2 shows that no competitive algorithm can surpass the 1/2 upper bound even in the small probability regime. Additionally, it is easy to show that the MNL-based online two-sided assortment problem generalizes the one-sided problem. To see this, we adopt a similar reduction to that of Section 2.2, where each unit of inventory is represented as a distinct supplier $i \in \mathcal{S}$ with the aggregate match function $w_i(C) = r_i \cdot \min\{|C|, 1\}$. Observe that the aggregate match function $w_i(\cdot)$ can be approximated up to an $O(\delta)$-factor by specifying an MNL instance with weights $q_{i,t} = \frac{1}{\delta}$ for every compatible customer $t \in [T]$. Therefore, neither the MNL model assumption nor the small probability assumption in itself is sufficient to alleviate the $1 - 1/e$ upper bound on the achievable competitive ratios. The conjunction of these structural properties, *small probability regime* and *MNL-based suppliers' preferences*, is paramount to derive Theorem 1. Note that, using the classical upper triangular family of instances, we derive an upper bound of 0.8074 on the competitive ratio achievable by any online assortment algorithm (see Proposition EC.3 in Appendix EC.4).

Shape of the discount functions. To gain a better understanding of why it is possible to develop algorithms better than $(1-1/e)$-competitive, it is instructive to contrast the shape of preference-aware discount functions with previously used exponential discount functions. We provide an illustration of these differences in Figure 2. With respect to the AdWords problem, Mehta et al. (2005) proposed a balancing algorithm that assigns the ad to the bidder with the highest discounted bid, where bids are discounted using the function $f(x) = e^{-(1-x)}$ over the fraction x of unspent budget.

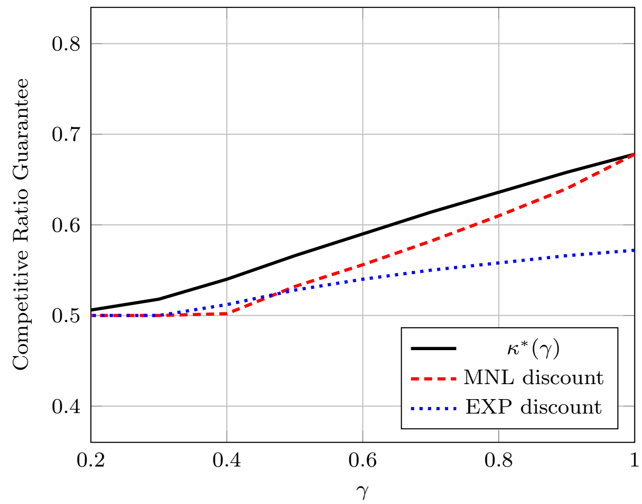

Figure 3 **The performance guarantees attained by different discount functions under the γ-NL choice model as a function of $\gamma \in [0,1]$.**

In the context of the online one-sided assortment problem, Golrezaei et al. (2014) analyze a balancing algorithm with the discount function $f(x) = 1 - \frac{e}{e-1}(1 - e^{-x})$, where x stands for the fraction of remaining inventory. In fact, Goyal et al. (2020) recently observed that a similar competitive ratio can be achieved by specifying the same discount function $f(x) = e^{-(1-x)}$ as in the paper of Mehta et al. (2005). In both settings, the constant-factor competitive ratio $(1 - 1/e)$ is optimal.

However, one fundamental property of these problems is that the marginal rewards associated with matching a supplier or selling a unit of product are constant; each unit of product sold generates the same unit of revenue for the platform regardless of how many units of the same product have been previously sold. By contrast, the online two-sided assortment problem exhibits diminishing returns; the marginal increases in the expected number of matches due to additional customer requests are strictly diminishing. Intuitively, this implies that, regardless of how the discount function is specified, the optimization problem (2) solved by balancing algorithms at each greedy step already promotes suppliers having received fewer requests. This observation might explain the concave shape of the discount functions $f_{\gamma,\kappa^*(\gamma)}$ generated by our preference-aware approach.

The effects of within-nest competition. Using a computer-aided approach described in Appendix EC.3.6, we numerically approximate the largest competitive ratio $\kappa^*(\gamma)$ that can be achieved by Theorem 2. The variations of the asymptotic competitive ratio $\kappa^*(\gamma)$ as a function of the nest dissimilarity coefficient $\gamma \in (0,1]$ are plotted in Figure 3. We observe that the performance of our online assortment algorithm increases as a function of γ. The highest competitive ratio is attained when suppliers' preferences follow the MNL choice model (i.e., $\gamma = 1$). That our performance guarantee is maximized under the MNL model (i.e., $\kappa^*(\gamma) \le \kappa^*(1)$) is not surprising. In fact, for any fixed value of $\gamma \in (0,1]$, the γ-NL setting subsumes the MNL model as a special case. This generalization property is due to the arbitrary structure of the partition into nests. Indeed, given an instance of the MNL model described by the weights $\{q_t\}_{t=1}^T$, an equivalent instance of the γ-NL model is constructed by specifying the nests $N_t = \{t\}$ for every $t \in [T]$ and the modified weights $\tilde{q}_{i,t} = q_t^{\frac{1}{\gamma}}$. We note in passing that the ϵ-small probability assumption is preserved by this reduction.

When γ tends to zero, the competitive ratio converges to $1/2$, which suggests that our γ-NL-balancing algorithm might only be marginally better than greedy in this regime. This observation suggests that assortment optimization is inherently harder in markets where there is greater competition among comparable customers. This is mirrored by the fact that the discount functions we have generated numerically converge to the constant function $f(x) = 1/2$, as shown by the right-hand plot of Figure 2. In fact, it is not difficult to formally establish that no deterministic algorithm is better than 1/2-competitive as γ tends to zero. To see this, we construct instances comprised of two suppliers $i \in \{1,2\}$ having γ-NL preferences, and two customers $t \in \{1,2\}$ forming distinct nests. Suppose that each customer $t \in \{1,2\}$ chooses to send a match request to the highest-numbered supplier $i \in A_t \cap S_t$, where A_t is the assortment displayed to customer t and $S_t \subseteq \mathcal{S}$ is her "compatibility set." For example, if $S_t = \{1,2\}$, then customer t sends a match request to supplier 1 when $A_t = \{1,2\}$, and to supplier 2 when $A_t = \{2\}$. To satisfy the ϵ-small probability condition, we specify uniform preference weights $q_{i,t} = w = (\epsilon/(1-\epsilon))^{1/\gamma}$ for all $i \in \{1,2\}$ and $t \in \{1,2\}$. Now, our analysis proceeds by fixing a deterministic algorithm π and by considering the following two instances. For Instance 1, we have $S_1 = \{1,2\}, S_2 = \{2\}$, whereas, for Instance 2, we have $S_1 = \{1,2\}, S_2 = \{1\}$. Due to the symmetry between suppliers 1 and 2, without loss of generality, we may assume that algorithm π assigns customer 1 to supplier 1. Hence, in Instance 2, the total expected number of matches is $(2w)^\gamma/(1+(2w)^\gamma)$ since supplier 1 receives two requests. By contrast, the clairvoyant algorithm obtains $2 \cdot w^\gamma/(1+w^\gamma)$ matches in expectation since each supplier receives one request. Consequently, we have just shown that, as γ tends to zero, the best-possible constant-factor competitive ratio is $1/2$.

The value of being preference-aware. Finally, we evaluate the theoretical gains in performance achieved by our preference-aware approach against other balancing algorithms. Specifically, in Figure 3, we plot the performance guarantees that can be established in the γ-NL setting for the MNL-optimal discount function $f_{1,\kappa^*(1)}$ as well as the exponential discount function $f : x \mapsto e^{-(1-x)}$, i.e., the maximum value of κ for which these functions satisfy inequality (NL-ODI). Here, the comparison with the MNL-optimal discount function informs about the loss of performance incurred by a misspecification of the nest dissimilarity coefficient. Additionally, the comparison with the exponential discount function underlines the value added of calibrating the discount function to the problem at hand, rather than employing existing ones. In both cases, we find that the gaps between the performance guarantees can be very significant. That said, when interpreting this plot, it is important to bear in mind that our ODI only gives a sufficient condition to guarantee a certain competitive ratio; it is unclear whether or not this analysis is tight.

4.4. Proof outline of Theorem 1

In this section, we provide a comprehensive outline of our proof for Theorem 1. To this end, we fix a sequence of customer-type arrivals $\{z_t\}_{t=1}^T$ and a sequence of available suppliers $\{\mathcal{S}_t\}_{t=1}^T$. Let $\epsilon \in (0, 1/\eta_\kappa)$, where $\eta_\kappa > 0$ is a Lipschitz constant for f_κ. For simplicity of notation, for every $i \in \mathcal{S}$ and $t \in [T]$, let $\Delta_i^t = w_i(C_i^{t,\pi^{f_\kappa}} + \boldsymbol{e}_t) - w_i(C_i^{t,\pi^{f_\kappa}})$ stand for the marginal increase of the aggregate match probability for supplier i upon receiving a request from customer t.

General approach. The objective of our proof is to bound the ratio between the expected number of matches generated by the preference-aware balancing algorithm π^{f_κ} and that generated by the optimal clairvoyant algorithm π^C as follows:

$$\frac{\mathcal{M}^{\pi^{f_\kappa}}(\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T)}{\mathcal{M}^{\pi^C}(\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T)} = \frac{\sum_{i\in\mathcal{S}} \mathbb{E}\left[w_i\left(C_i^{T+1,\pi^{f_\kappa}}\right)\right]}{\sum_{i\in\mathcal{S}} \mathbb{E}\left[w_i\left(C_i^{T+1,\pi^C}\right)\right]} \geq \kappa - (\kappa + \kappa\eta_\kappa)\,\epsilon + \kappa\eta_\kappa\epsilon^2 \ . \tag{3}$$

To establish this inequality, we will relate the random variables $w_i(C_i^{T+1,,\pi^{f_\kappa}})$ and $w_i(C_i^{T+1,\pi^C})$ for every $i \in \mathcal{S}$ using a sample path-based analysis. For this purpose, we introduce a coupling ω that describes a joint probabilistic space for the sequences of assortments shown and customers' match requests under the arrival patterns $(\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T)$. This probabilistic coupling is arbitrarily chosen: our subsequent sample path arguments hold for any valid coupling ω. Going forward, as all random variables are expressed with respect to this probabilistic space, we drop the reference to ω to ease readability.

Our proof proceeds in three steps, which are detailed next.

Step 1: Lower bound via a decomposition method. We begin by decomposing the expected number of matches generated by our algorithm π^{f_κ} to obtain a lower bound on $\mathcal{M}^{\pi^{f_\kappa}}(\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T)$. Importantly, this lower bound relates the matching decisions of our balancing algorithm π^{f_κ} to those of the clairvoyant benchmark π^C, as formally described by the next lemma.

LEMMA 1. *For every supplier $i \in \mathcal{S}$, we define the random variable R_i as follows:*

$$R_i = \left(1 - f_\kappa\left(w_i\left(C_i^{T+1,,\pi^{f_\kappa}}\right)\right)\right) \cdot \left(\sum_{t=1}^T \Delta_i^t \cdot \mathbb{I}\left[\xi_t\left(A_t^{\pi^C}\right) = i\right]\right) \tag{4}$$

$$+ \sum_{t=1}^T \Delta_i^t \cdot \mathbb{I}\left[\xi_t\left(A_t^{\pi^{f_\kappa}}\right) = i\right] \cdot f_\kappa\left(\min\{1, w_i\left(C_i^{t,\pi^{f_\kappa}}\right) + \epsilon\}\right) \ .$$

Then, the expected number of matches generated by π^{f_κ} is lower bounded by

$$\mathcal{M}^{\pi^{f_\kappa}}\left(\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T\right) \geq (1-\eta_\kappa \cdot \epsilon) \cdot \left(\sum_{i\in\mathcal{S}} \mathbb{E}_\omega\left[R_i\right]\right) .$$

The proof, presented in Appendix EC.3.1, is based on simple algebraic comparisons using the optimality criterion of the balancing algorithm in (2), together with the fact that $w_i(\boldsymbol{e}_t) \leq \epsilon$ for every $i \in \mathcal{S}$ and $t \in [T]$. However, we do not use any structural property of the MNL choice model. As such, this lower bounding method holds for any rank-based choice model.

Taking a closer look at Equation (4), we observe that, for every $i \in \mathcal{S}$, the random variable R_i decomposes into the sum of two terms, which we interpret as follows:

$$R_i = \underbrace{\left(1 - f_\kappa\left(w_i\left(C_i^{T+1,\pi^{f_\kappa}}\right)\right)\right) \cdot \left(\sum_{t=1}^T \Delta_i^t \cdot \mathbb{I}\left[\xi_t\left(A_t^{\pi^C}\right) = i\right]\right)}_{\text{Discounted rewards of } \pi^C \text{ with respect to the state variables } \{C_i^{t,\pi^{f_\kappa}}\}_{t\in[T]}} \tag{5}$$

$$+ \underbrace{\sum_{t=1}^T \Delta_i^t \cdot \mathbb{I}\left[\xi_t\left(A_t^{\pi^{f_\kappa}}\right) = i\right] \cdot f_\kappa\left(\min\{1, w_i\left(C_i^{t,\pi^{f_\kappa}}\right) + \epsilon\}\right)}_{\text{Sum of discounts accumulated by } \pi^{f_\kappa}} . \tag{6}$$

The second term (6) is simply the cumulative discounts that the balancing algorithm π^{f_κ} accrues over T time periods. The first term (5) has a more complex interpretation: it can be viewed as a "counterfactual" to the discounted rewards generated by the assortment decisions of the clairvoyant algorithm π^C with respect to the sequence of state variables $\{C_i^{t,\pi^{f_\kappa}}\}_{t\in[T]}$. That is, for every time period $t \in [T]$, we compute what would have been the marginal increase $\Delta_i^t \cdot \mathbb{I} \cdot [\xi_t(A_t^{\pi^{f_\kappa}}) = i]$ of supplier i's aggregate match probability if one were to implement the assortment decisions of π^C, given the set of customer requests $C_i^{t,\pi^{f_\kappa}}$ generated by π^{f_κ} thus far. Then, this quantity is multiplied by $(1 - f_\kappa(w_i(C_i^{T+1,\pi^{f_\kappa}}))$, the discount factor reached by π^{f_κ} at the end of the time horizon.

Now, why is this decomposition useful? Intuitively, we wish to show that the assortment decisions of π^{f_κ} are competitive against those of π^C for every supplier $i \in \mathcal{S}$. The first term of the decomposition, (5), precisely relates the assortment decisions of π^{f_κ} to those of π^C. However, the challenge is that the counterfactual rewards are computed with respect to the sequence of state variables $\{C_i^{t,\pi^{f_\kappa}}\}_{t\in[T]}$ created by the assortment decisions of π^{f_κ}, rather than the sequence $\{C_i^{t,\pi^C}\}_{t\in[T]}$ created by the assortment decisions of π^C. The subsequent analysis will show that the potential loss against the expected number of matches of π^C is "more than offset" by the cumulative discounts accrued by π^{f_κ}; we argue that the second term (6) compensates any loss in (5) relative to the sequence of state variables $\{C_i^{t,\pi^C}\}_{t\in[T]}$, up to a factor of $\kappa - O(\epsilon)$.

Step 2: Sample-path analysis using the MNL structure. Next, we separately analyze the random variables R_i for each supplier $i \in \mathcal{S}$ on every sample path. Namely, our next lemmas bound each of the two terms appearing in the decomposition (4).

LEMMA 2. *For every sample realization of ω and every supplier $i \in \mathcal{S}$, we have*

$$\sum_{t=1}^{T} \Delta_i^t \cdot \mathbb{I}\left[\xi_t\left(A_t^{\pi^C}\right) = i\right] \geq (1-\epsilon) \cdot \left(1 - w_i\left(C_i^{T+1,\pi^{f_\kappa}}\right)\right)^2 \cdot \frac{w_i(C_i^{T+1,\pi^C})}{1 - w_i(C_i^{T+1,\pi^C})} \ .$$

LEMMA 3. *For every sample realization of ω and every supplier $i \in \mathcal{S}$, we have*

$$\sum_{t=1}^{T} \Delta_i^t \cdot \mathbb{I}\left[\xi_t\left(A_t^{\pi^{f_\kappa}}\right) = i\right] \cdot f_\kappa\left(\min\left\{1, w_i\left(C_i^{t,\pi^{f_\kappa}}\right) + \epsilon\right\}\right) \geq \int_0^{w_i(C_i^{T+1,\pi^{f_\kappa}})} f_\kappa(u)\, du \ .$$

The proofs are presented in Appendices EC.3.2 and EC.3.3. Here, contrary to Step 1, our analysis exploits the specific structural properties of the MNL-based online two-sided assortment instances we consider. In particular, Lemma 2 follows from the substitution patterns captured by the MNL choice model, among other properties. Lemma 3 crucially relies on the ϵ-small probability regime and the Lipschitz property of the discount function. Intuitively, we show that the sum of discounts under policy π^{f_κ} is well approximated by an integral given the small marginal increments accrued in each time period.

At this point, we note that the main technical challenge to establish our general result for the Nested Logit choice model (Theorem 2) resides in deriving a lower bound on $\sum_{t=1}^{T} \Delta_i^t \cdot \mathbb{I}\left[\xi_t\left(A_t^{\pi^C}\right) = i\right]$ along the same lines as Lemma 2. In this context, we formulate a non-convex mathematical problem that implicitly determines the worst-possible sequence of arrivals under γ-NL preferences. Contrary to Lemma 2, the optimal solution of this mathematical problem cannot be fully expressed in closed form.

Step 3: Revealing the connection with the ODI. The final step of our proof combines together the lower bounds obtained in Step 2 and reveals why (MNL-ODI) gives a sufficient condition for a competitive ratio of $\kappa - O(\epsilon)$. That is, for every supplier $i \in \mathcal{S}$ and every sample realization of ω, Lemmas 2 and 3 imply that

$$\begin{aligned}
&\frac{R_i}{w_i(C_i^{T+1,\pi^C})} \\
&\quad \geq \frac{1-\epsilon}{1 - w_i(C_i^{T+1,\pi^C})} \cdot \left(1 - f_\kappa\left(w_i\left(C_i^{T+1,\pi^{f_\kappa}}\right)\right)\right) \cdot \left(1 - w_i\left(C_i^{T+1,\pi^{f_\kappa}}\right)\right)^2 \\
&\qquad + \frac{1}{w_i(C_i^{T+1,\pi^C})} \cdot \left(\int_0^{w_i(C_i^{T+1,\pi^{f_\kappa}})} f_\kappa(u)\, du\right) \\
&\quad \geq \left(\sqrt{\int_0^{w_i(C_i^{T+1,\pi^{f_\kappa}})} f_\kappa(u)\, du} + \sqrt{(1-\epsilon) \cdot \left(1 - f_\kappa\left(w_i\left(C_i^{T+1,\pi^{f_\kappa}}\right)\right)\right)} \cdot \left(1 - w_i\left(C_i^{T+1,\pi^{f_\kappa}}\right)\right)\right)^2
\end{aligned} \tag{7}$$

where the second inequality is obtained by minimizing over $w_i(C_i^{T+1,\pi^C})$ and observing that the minimum of the function $\alpha \mapsto \frac{c_1}{\alpha} + \frac{c_2}{1-\alpha}$ is attained for $\alpha = \frac{\sqrt{c_1}}{\sqrt{c_1}+\sqrt{c_2}}$. Consequently, we wish to bound the last expression in (7). This lower bound originates from inequality (MNL-ODI), which requires that, for all $x \in [0,1]$,

$$\sqrt{\int_0^x f(u)du} + \sqrt{1-f(x)} \cdot (1-x) \geq \sqrt{\kappa} \ .$$

Taking $x = w_i(C_i^{T+1,\pi^{f_\kappa}})$, since $f_\kappa(\cdot)$ is a solution of (MNL-ODI) with parameter κ, we have

$$\frac{R_i}{w_i(C_i^{T+1,\pi^C})} \ \geq \ \left(\sqrt{1-\epsilon} \cdot \sqrt{\kappa}\right)^2 = (1-\epsilon) \cdot \kappa \ . \tag{8}$$

By summing this inequality over all $i \in \mathcal{S}$ and by taking the expectation, we conclude that

$$\begin{aligned}
\mathcal{M}^{\pi^{f_\kappa}}\left(\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T\right) &= \sum_{i \in \mathcal{S}} \mathbb{E}_\omega\left[w_i\left(C_i^{T+1}\right)\right] \\
&\geq (1-\eta_\kappa \cdot \epsilon) \cdot \left(\sum_{i \in \mathcal{S}} \mathbb{E}_\omega\left[R_i\right]\right) \\
&\geq (1-\eta_\kappa \cdot \epsilon) \cdot \left(\sum_{i \in \mathcal{S}} \mathbb{E}_\omega\left[(1-\epsilon) \cdot \kappa \cdot w_i(C_i^{T+1,\pi^C})\right]\right) \\
&= \left(\kappa - \kappa(1+\eta_\kappa)\epsilon + \kappa\eta_\kappa\epsilon^2\right) \cdot \left(\sum_{i \in \mathcal{S}} \mathbb{E}_\omega\left[w_i(C_i^{T+1,\pi^C})\right]\right) \\
&= \left(\kappa - \kappa(1+\eta_\kappa)\epsilon + \kappa\eta_\kappa\epsilon^2\right) \cdot \mathcal{M}^{\pi^C}\left(\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T\right) \ ,
\end{aligned}$$

where the first inequality follows from Lemma 1 and the second one from (8).

5. Concluding Remarks

To our knowledge, our work is the first to formally study the implications that the presence of two-sided preferences has on the design of online assortment algorithms for matching platforms. We find that, in settings where suppliers do not immediately accept or reject the requests to match, the two-sided setting is a strict generalization of the one-sided problem. In fact, we show that no efficient algorithm is better than greedy under rank-based suppliers' preferences. That said, our work also demonstrates that the design and performance analysis of online assortment algorithms strongly depend on the functional form of suppliers' preferences. When restricting attention to preference models widely used by academics and practitioners, *online assortment optimization in a two-sided matching platform can be easier or harder than in retail platforms*. When suppliers have MNL preferences, we develop simple algorithms better than $(1-1/e)$-competitive, whereas, in the Nested Logit case, we establish that no deterministic algorithm has a constant-factor competitive ratio better than 1/2 in certain cases. Overall, our results showcase the importance of carefully accounting for suppliers' preferences in two-sided markets.

Model extensions. We briefly discuss the robustness of our modeling approach when considering practically relevant extensions. First, our model may account for time-varying constraints on the assortment of suppliers shown to each customer, such as constraints of the assortment size; our only requirement, stated by Assumption 1, is that the single-stage constrained assortment optimization problem needs to be computationally tractable. Second, in online marketplaces, it is not uncommon that the same worker simultaneously desires to be matched with multiple jobs/tasks. As discussed in Section 2, our modeling approach and algorithmic results can be directly leveraged when customers choose to send multiple matching requests on the condition that Assumption 1 still holds. Nonetheless, one implicit assumption of our model is that, if the firm agrees to match in return, the worker proceeds with the job in question. In reality, this assumption is reasonable when the worker has sufficient capacity to fulfil all the tasks she is matched with, or if the work is completed by the time the match request is sent. For example, the latter condition holds in contest design marketplaces; here, each match request corresponds to a full design submission, and there is often no further work to be done after the firms reveal which submissions are rewarded. Hence, in these settings, our approach and results are directly applicable. For completeness, in Appendix EC.5, we also examine a variant of our problem where workers can fulfil a single match, even if they submit multiple match requests. Through numerical simulations, we find that, in realistic market regimes, the *number of conflicts* is relatively small compared to the total expected number of matches; a conflict designates two match requests from the same customer being simultaneously accepted by suppliers. This finding suggests that our model may still provide a good first-order approximation of the matching outcomes.

Open questions. This paper is a first stride towards formulating and analyzing online assortment problems in two-sided matching markets where preferences are expressed by both parties. A number of interesting and challenging directions are left for future research. In particular, one may want to account for the interdependencies between the timing of suppliers' feedback and the process describing the customers' match requests. For example, suppliers might endogenously decide when to match based on the requests they have accumulated thus far. Moreover, the customers' arrivals and choice process might exhibit serial dependencies; e.g., an unmatched customer might return to the platform to seek additional matches if her previous requests are left unsatisfied. While practically relevant, all these extensions introduce complex information dynamics and feedback loops; making progress in analyzing such models likely requires introducing a higher degree of stylization in modeling the customers' and suppliers' preferences.

Acknowledgments

The authors gratefully acknowledge the feedback from the Associate Editor and reviewers, attendees at various seminars and conferences as well as the anonymous review team of the 22nd ACM Conference on Economics and Computation (EC'21).

References

Afeche, P., Caldentey, R. and Gupta, V. (2019), 'On the optimal design of a bipartite matching queueing system', *Available at SSRN 3345302* .

Akbarpour, M., Li, S. and Gharan, S. O. (2020), 'Thickness and information in dynamic matching markets', *Journal of Political Economy* **128**(3), 785–815.

Anderson, R., Ashlagi, I., Gamarnik, D. and Kanoria, Y. (2017), 'Efficient dynamic barter exchange', *Operations Research* **65**(6), 1446–1459.

Aouad, A. and Saritaç, Ö. (2020), Dynamic stochastic matching under limited time, *in* 'Proceedings of the 21st ACM Conference on Economics and Computation', pp. 789–790.

Aouad, A. and Segev, D. (2021), 'Display optimization for vertically differentiated locations under multinomial logit preferences', *Management Science* **67**(6), 3519–3550.

Arnosti, N., Johari, R. and Kanoria, Y. (2021), 'Managing congestion in matching markets', *Manufacturing & Service Operations Management* **23**(3), 620–636.

Ashlagi, I., Burq, M., Dutta, C., Jaillet, P., Saberi, A. and Sholley, C. (2019), Edge weighted online windowed matching, *in* 'Proceedings of the 2019 ACM Conference on Economics and Computation', ACM, pp. 729–742.

Ashlagi, I., Krishnaswamy, A. K., Makhijani, R., Saban, D. and Shiragur, K. (2020), 'Assortment planning for two-sided sequential matching markets', *arXiv:1907.04485* .

Bimpikis, K., Papanastasiou, Y. and Zhang, W. (2020), 'Information provision in two-sided platforms: Optimizing for supply', *Available at SSRN* .

Blanchet, J., Gallego, G. and Goyal, V. (2016), 'A Markov chain approximation to choice modeling', *Operations Research* **64**(4), 886–905.

Börsch-Supan, A. (1990), 'On the compatibility of nested logit models with utility maximization', *Journal of Econometrics* **43**(3), 373–388.

Chan, C. W. and Farias, V. F. (2009), 'Stochastic depletion problems: Effective myopic policies for a class of dynamic optimization problems', *Mathematics of Operations Research* **34**(2), 333–350.

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

Désir, A., Goyal, V. and Zhang, J. (2014), 'Near-optimal algorithms for capacity constrained assortment optimization'. Working paper. Available online as SSRN report #2543309.

Devanur, N. R. and Jain, K. (2012), Online matching with concave returns, *in* 'Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing', pp. 137–144.

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

Feige, U. (1998), 'A threshold of $\ln(n)$ for approximating set cover', *Journal of the ACM* **45**(4), 634–652.

Feldman, J., Mehta, A., Mirrokni, V. and Muthukrishnan, S. (2009), Online stochastic matching: Beating 1-1/e, *in* '2009 50th Annual IEEE Symposium on Foundations of Computer Science', IEEE, pp. 117–126.

Gallego, G., Li, A., Truong, V.-A. and Wang, X. (2020), 'Approximation algorithms for product framing and pricing', *Operations Research* **68**(1), 134–160.

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

Gong, X.-Y., Goyal, V., Iyengar, G., Simchi-Levi, D., Udwani, R. and Wang, S. (2019), 'Online assortment optimization with reusable resources', *Available at SSRN 3334789* .

Goyal, V., Iyengar, G. and Udwani, R. (2020), 'Online allocation of reusable resources: Achieving optimal competitive ratio', *arXiv preprint arXiv:2002.02430* .

Goyal, V. and Udwani, R. (2020), Online matching with stochastic rewards: Optimal competitive ratio via path based formulation, *in* 'Proceedings of the 21st ACM Conference on Economics and Computation', pp. 791–791.

Gurvich, I. and Ward, A. (2015), 'On the dynamic control of matching queues', *Stochastic Systems* **4**(2), 479–523.

Horton, J. J. and Vasserman, S. (2021), 'Job-seekers send too many applications: Experimental evidence and a partial solution', *Working paper* .

Jaillet, P. and Lu, X. (2013), 'Online stochastic matching: New algorithms with better bounds', *Mathematics of Operations Research* **39**(3), 624–646.

Kalyanasundaram, B. and Pruhs, K. R. (2000), 'An optimal deterministic algorithm for online b-matching', *Theoretical Computer Science* **233**(1-2), 319–325.

Kanoria, Y. and Saban, D. (2021), 'Facilitating the search for partners on matching platforms', *Management Science* .

Kapralov, M., Post, I. and Vondrák, J. (2013), Online submodular welfare maximization: Greedy is optimal, *in* 'Proceedings of the Twenty-fourth Annual ACM-SIAM Symposium on Discrete Algorithms', SIAM, pp. 1216–1225.

Karp, R. M., Vazirani, U. V. and Vazirani, V. V. (1990), An optimal algorithm for on-line bipartite matching, *in* 'Proceedings of the Twenty-second Annual ACM Symposium on Theory of Computing', ACM, pp. 352–358.

Lobel, I. (2021), 'Revenue management and the rise of the algorithmic economy', *Management Science* **67**(9), 5389–5398.

Ma, W. and Simchi-Levi, D. (2020), 'Algorithms for online matching, assortment, and pricing with tight weight-dependent competitive ratios', *Operations Research* **68**(6), 1787–1803.

Manshadi, V. H., Gharan, S. O. and Saberi, A. (2012), 'Online stochastic matching: Online actions based on offline statistics', *Mathematics of Operations Research* **37**(4), 559–573.

Mehta, A. (2013), 'Online matching and ad allocation', *Foundations and Trends in Theoretical Computer Science* **8**(4), 265–368.

Mehta, A. and Panigrahi, D. (2012), Online matching with stochastic rewards, *in* '2012 IEEE 53rd Annual Symposium on Foundations of Computer Science', IEEE, pp. 728–737.

Mehta, A., Saberi, A., Vazirani, U. and Vazirani, V. (2005), Adwords and generalized on-line matching, *in* '46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)', IEEE, pp. 264–273.

Mehta, A., Waggoner, B. and Zadimoghaddam, M. (2015), Online stochastic matching with unequal probabilities, *in* 'Proceedings of the Twenty-sixth Annual ACM-SIAM Symposium on Discrete Algorithms', SIAM, pp. 1388–1404.

Motwani, R. and Raghavan, P. (1995), *Randomized Algorithms*, Cambridge University Press.

Rios, I., Saban, D. and Zheng, F. (2020), 'Optimizing assortments in two-sided markets: An application to dating', *Working paper* .

Rusmevichientong, P., Shen, Z.-J. M. and Shmoys, D. B. (2010), 'Dynamic assortment optimization with a multinomial logit choice model and capacity constraint', *Operations Research* **58**(6), 1666–1680.

Rusmevichientong, P., Sumida, M. and Topaloglu, H. (2020), 'Dynamic assortment optimization for reusable products with random usage durations', *Management Science* **66**(7), 2820–2844.

Rusmevichientong, P., Van Roy, B. and Glynn, P. W. (2006), 'A nonparametric approach to multiproduct pricing', *Operations Research* **54**(1), 82–98.

Shi, P. (2019), 'Optimal priority-based allocation mechanisms', *Available at SSRN 3425348* .

Sumida, M., Gallego, G., Rusmevichientong, P., Topaloglu, H. and Davis, J. (2020), 'Revenue-utility tradeoff in assortment optimization under the multinomial logit model with totally unimodular constraints', *Management Science* **67**(5), 2845–2869.

Talluri, K. and van Ryzin, G. (2004), 'Revenue management under a general discrete choice model of consumer behavior', *Management Science* **50**(24), 15–33.

Tulabandhula, T., Sinha, D. and Patidar, P. (2020), 'Multi-purchase behavior: Modeling and optimization', *Available at SSRN 3626788* .

Yao, A. C.-C. (1977), Probabilistic computations: 'toward a unified measure of complexity', *in* '18th Annual Symposium on Foundations of Computer Science', IEEE Computer Society, pp. 222–227.

This page is intentionally blank. Proper e-companion title page, with INFORMS branding and exact metadata of the main paper, will be produced by the INFORMS office when the issue is being assembled.

Online Companion

EC.1. Additional material to Section 2

EC.1.1. Online two-sided assortment problem with immediate feedback

As described in Section 1, one of our main research objectives is to understand when the online two-sided assortment problem significantly differs from the one-sided setting. We next show that a setting where suppliers provide immediate feedback after each customer match request can be reduced to the online one-sided assortment problem. We formalize this idea by introducing the *online two-sided assortment problem with immediate feedback* and constructing a mapping to the one-sided setting.

Definition EC.1 (Immediate supplier feedback). An instance of the *online two-sided assortment problem with immediate feedback* is described by the set of customer types $\mathcal{Z}$, the choice probability $\phi_i^z(\cdot)$ that a customer of type $z \in \mathcal{Z}$ chooses supplier $i \in \mathcal{S}$, the probability $w_{i,z}$ that supplier $i \in \mathcal{S}$ immediately accepts a request from a customer of type $z \in \mathcal{Z}$, and the sequences of arriving customer types $\{z_t\}_{t=1}^T$ and supplier sets $\{\mathcal{S}_t\}_{t=1}^T$.

Proposition EC.1. *There exists a polynomial-time approximation-preserving reduction from the online two-sided assortment problem with immediate feedback to the online one-sided assortment problem.*

Proof. Building on the observation made in Section 2.2, without loss of generality, we focus on the setting where $\mathcal{S}_t = \mathcal{S}$ for all $t \in [T]$. Consequently, given an instance $\mathcal{I}$ of the online two-sided assortment problem with immediate feedback, we construct an instance $\tilde{\mathcal{I}}$ of the online one-sided assortment problem through the following mapping:

1. For every supplier $i \in \mathcal{S}$, we create a corresponding item i in $\tilde{\mathcal{I}}$ with revenue $r_i = 1$ and corresponding inventory $c_i = 1$.
2. We keep the number of time periods to be the same, meaning that T is the number of time periods in $\tilde{\mathcal{I}}$. For every customer type $z \in \mathcal{Z}$, we create a corresponding type z in $\tilde{\mathcal{I}}$.
3. Customers' choice models in $\tilde{\mathcal{I}}$ are specified as follows. For every customer type $z \in \mathcal{Z}$, assortment $A \subseteq \mathcal{S}$, and item $i \in A$, the probability $\tilde{\phi}_i^z(A)$ that a customer of type z chooses item i in assortment A is given by $\tilde{\phi}_i^z(A) = w_{i,z} \cdot \phi_i^z(A)$.

In what follows, we fix a sequence of customer types $\{z_t\}_{t=1}^T$. We show that there is a direct correspondence between algorithms for $\mathcal{I}$ and those for $\tilde{\mathcal{I}}$. Specifically, given an algorithm π for the two-sided problem $\mathcal{I}$, we devise an algorithm $\tilde{\pi}$ for the one-sided problem $\tilde{\mathcal{I}}$ such that the expected revenue $\mathcal{R}^{\tilde{\pi}}(\{z_t\}_{t=1}^T)$ generated by $\tilde{\pi}$ is equal to the expected number of matches $\mathcal{M}^{\pi}(\{z_t\}_{t=1}^T)$ generated by π.

The construction of $\tilde{\pi}$ proceeds from an inductive coupling argument. To this end, we use $S_t \subseteq \mathcal{S}$ to denote the subset of suppliers available at time $t \in [T]$ with respect to the algorithm π, and S_{T+1} is the set of unmatched suppliers after all customers' arrivals. Similarly, we denote by $\tilde{S}_t \subseteq \mathcal{S}$ the subset of items remaining at time $t \in [T+1]$ with respect to the algorithm $\tilde{\pi}$. Hence, S_t and $\tilde{S}_t$ are random variables with respect to our coupled probabilistic space. The invariant of our induction is the property that $\tilde{S}_t = S_t$ and $A^{\pi}_t = A^{\tilde{\pi}}_t$ for every $t \in [T]$.

Base case: $t = 1$. By item 1 of our mapping, we initially have $\tilde{S}_1 = S_1$. Consequently, it is clear that it is feasible to select the assortment $A^{\tilde{\pi}}_1 = A^{\pi}_1$.

Inductive case: $t \geq 2$. Suppose that $\tilde{S}_{t-1} = S_{t-1}$ and $A^{\tilde{\pi}}_{t-1} = A^{\pi}_{t-1}$. Define the Bernoulli random variable $\eta_{i,t-1}$ that indicates whether supplier i would accept a match request from the $t-1$-th arriving customer in the two-sided problem $\mathcal{I}$, i.e, $\Pr[\eta_{i,t-1} = 1 | \mathcal{H}^{\pi}_{t-1}] = w_{i,z_{t-1}}$. In addition, we define $\tilde{\xi}_{t-1}(A)$ as the random item chosen by the $t-1$-th arriving customer when presented with the assortment $A \subseteq \mathcal{S}$ in the one-sided problem $\mathcal{I}$. Note that

$$
\begin{aligned}
\Pr\left[\tilde{S}_t = \tilde{S}_{t-1} \setminus \{i\} \middle| \mathcal{H}^{\tilde{\pi}}_{t-1}\right] &= \Pr\left[\tilde{\xi}_{t-1}\left(A^{\tilde{\pi}}_{t-1}\right) = i \middle| A^{\tilde{\pi}}_{t-1}, z_{t-1}\right] \\
&= \tilde{\phi}^{z_{t-1}}_i\left(A^{\tilde{\pi}}_{t-1}\right) \\
&= w_{iz_t} \cdot \phi^{z_{t-1}}_i\left(A^{\tilde{\pi}}_{t-1}\right) \\
&= w_{iz_t} \cdot \phi^{z_{t-1}}_i\left(A^{\pi}_{t-1}\right) \\
&= \Pr\left[\eta_{i,t} = 1 | z_{t-1}\right] \cdot \Pr\left[\xi_{t-1}\left(A^{\pi}_{t-1}\right) = i \middle| A^{\pi}_{t-1}, z_{t-1}\right] \\
&= \Pr\left[S_t = S_{t-1} \setminus \{i\} | \mathcal{H}^{\pi}_{t-1}\right] \ ,
\end{aligned}
$$

where the third equality proceeds from item 3 of our mapping between $\mathcal{I}$ and $\tilde{\mathcal{I}}$, and the next equality holds since our induction hypothesis implies that $A^{\tilde{\pi}}_{t-1} = A^{\pi}_{t-1}$. Based on the above equality, it is clear that there exists a coupling of the random variables $\tilde{S}_t$ and S_t such that $\tilde{S}_t = S_t$. Consequently, for every $t \in [T]$, it is feasible to select the assortment $A^{\tilde{\pi}}_t = A^{\pi}_t$ with respect to our inductive coupling.

To conclude the proof of Proposition EC.1, observe that

$$
\mathcal{R}^{\tilde{\pi}}(\{z_t\}_{t=1}^T) = \mathbb{E}\left[\left|\mathcal{S} \setminus \tilde{S}_{T+1}\right|\right] = \mathbb{E}\left[\left|\mathcal{S} \setminus S_{T+1}\right|\right] = \mathcal{M}^{\pi}\left(\{z_t\}_{t=1}^T\right) \ ,
$$

where the second equality holds since our induction shows that $\tilde{S}_{T+1} = S_{T+1}$. $\square$

EC.1.2. Connection between small probability and large inventory regimes

In the one-sided setting, the platform is endowed with c_i units of inventory for each item $i \in \mathcal{S}$. The large inventory regime consists in analyzing the performance of algorithms as a function of the parameter $c_{\min} = \min_{i \in \mathcal{S}} c_i$ in the asymptotic regime $c_{\min} \to +\infty$. Here, we formally show that the

ϵ-small probability condition subsumes the class of online one-sided assortment problems where the inventory satisfies $c_{\min} \geq 1/\epsilon$. Thus, the small probability regime is indeed a generalization of the large inventory regime studied in the previous literature.

To show this, we slightly modify the reduction devised in Section 2.2, where units of inventory in the online one-sided assortment problem were treated as distinct suppliers $i \in \mathcal{S}$ in the two-sided setting. Instead, we develop an exact reduction to the two-sided assortment problem by representing each item i as a supplier $i \in \mathcal{S}$ endowed with the aggregate matching function $w_i(\cdot)$ such that, for every $C \in \mathbb{N}_0^{\mathcal{Z}}$,

$$w_i(C) = r_i c_i \cdot \min\left\{ \frac{\sum_{z \in \mathcal{Z}} C_z}{c_i}, 1 \right\} .$$

Observe that this reduction is equivalent to the one presented in Section 2.2. Now, the important observation is that, for every $t \in [T]$, we have $w_i(\boldsymbol{e}_t) \leq (1/c_{\min}) \cdot \max_{C \in \mathbb{N}_0^{\mathcal{Z}}} w_i(C)$. Consequently, the above-constructed instance of the online two-sided assortment problem satisfies the ϵ-small probability condition when $\epsilon \geq 1/c_{\min}$.

EC.2. Proofs of Section 3

EC.2.1. Proof of Proposition 1

To establish the result in Proposition 1, we start by highlighting the connection to the broad class of stochastic depletion problems introduced in Chan and Farias (2009). Informally, a *stochastic depletion problem* is specified by item types and activity sets. In each time period, an activity must be chosen, which results in (random) depletion of items of various types. The distribution of the number of items of each type that are depleted may depend on the activity employed, the number of items of that type available, and exogenous stochastic processes. Item depletions generate rewards. The objective is to design an adaptive activity selection policy that can use its knowledge of system dynamics to maximize the total expected reward. Many challenging dynamic optimization problems can be formulated as stochastic depletion problems, including optimal control of parallel-server queuing models, optimal AdWords allocation, and more. Surprisingly, Chan and Farias (2009) identify some simple sufficient conditions of the reward function under which a *myopic* policy generates an expected reward within a factor of 1/2 of the optimal adaptive policy for that problem.

To establish our result, it is useful to think of the platform's problem as a stochastic depletion problem. In each time period, the platform needs to choose an assortment for the arriving customer; choices made by customers ultimately result in *depletions* of the suppliers' aggregate matching quantity. We describe our problem in those terms next. In particular, we adopt the same terminology as in Chan and Farias (2009): the stochastic depletion problem is formulated as a dynamic program and online algorithms are referred to as policies.

Notation and preliminaries. Recall that C_i^t denotes the set of customer requests received by supplier i at time t, and $\bar{C}^t = (C_i^t)_{i\in\mathcal{S}}$ is the vector of sets of customer requests up to time t. In each time period t, the policy selects an assortment $A \subseteq \mathcal{S}_t$ that results in a random *depletion vector* $\mathbf{X}_t^A \in \{0,1\}^{\mathcal{S}}$, where $\mathbf{X}_{t,i}^A = 1$ if customer t requests a match with supplier i. Based on our model described in Section 2, the distribution of $\mathbf{X}_t^A$ is specified by customer t's choice model over the set of alternatives in A, i.e., $\mathbf{X}_{t,i}^A = 1$ if and only if $\xi_{z_t}(A) = i$. In particular, $\mathbf{X}_t^A$ is also independent of all past depletions. (For the sake of readability, from now on we omit the superscript A from $\mathbf{X}_t^A$ whenever this is clear from the context.)

Dynamic programming formulation. Based on the above observations, we cast the online two-sided assortment problem as a dynamic program. Recall that, to demonstrate the performance of the greedy policy, we compare it to a *clairvoyant policy* that knows the realization of the sequence of customer arrivals and of the suppliers' set. As a result, we assume that $\{z_t\}_{t=1}^T$ and $\{\mathcal{S}_t\}_{t=1}^T$ are fixed, and we work with a reduced state space where each state is characterised by the time period and the set of customer requests received so far. Formally, let Σ denote the state space where each state $\sigma \in \Sigma$ is described by $(t(\sigma), \bar{C}(\sigma))$, where $0 \le t(\sigma) \le T$ is the current time period and $\bar{C}(\sigma) = (C_1(\sigma), \ldots, C_m(\sigma)) \subseteq \mathbb{N}_0^{m\times\mathcal{Z}}$ is the vector of sets of customer requests at the beginning of period $t(\sigma)$.

In state σ, after the algorithm shows assortment A to customer $t(\sigma)$, the next state is $\left(t(\sigma)+1, \bar{C}^{\uparrow}(\sigma)\right)$, where $\bar{C}^{\uparrow}(\sigma)$ is the set of customer requests $\bar{C}(\sigma)$ shifted by at most one unit, i.e.,

$$\bar{C}_i^{\uparrow}(\sigma) = \bar{C}_i(\sigma) + \mathbf{e}_{z_{t(\sigma)}} \cdot \mathbf{X}_{t,i} \qquad \forall i \in \mathcal{S} \ . \tag{EC.1}$$

To ease exposition, we sometimes write $\bar{C}(\sigma) + \mathbf{e}_{z_{t(\sigma)}} \cdot \mathbf{X}_t$ or simply $\bar{C}(\sigma) + \mathbf{X}_t$ as a shortcut for the expression in (EC.1).

Let $g : \mathbb{N}_0^{m\times\mathcal{Z}} \times \mathbb{N}_0^{m\times\mathcal{Z}} \to \mathbb{R}$ be defined as

$$g(\bar{C}, \bar{C}') = \sum_{i\in\mathcal{S}} \left(w_i(C_i') - w_i(C_i)\right),$$

and define the random reward function $R : \Sigma \times 2^{|\mathcal{S}|} \to \mathbb{R}_+$ as

$$R(\sigma, A) = g(\bar{C}(\sigma), \bar{C}(\sigma')) = \sum_{i\in\mathcal{S}} w_i(C_i(\sigma')) - \sum_{i\in\mathcal{S}} w_i(C_i(\sigma)), \tag{EC.2}$$

where σ' is the random state reached by the system after assortment A is shown to the incoming customer in state σ. With this notation at hand, we define the reward-to-go under policy π starting at state σ as

$$J^{\pi}(\sigma) = \mathbb{E}\left[\sum_{u=t(\sigma)}^{T-1} R\left(\sigma_u, \pi\left(\sigma_u\right)\right)\right] \ . \tag{EC.3}$$

Consequently, we let $J^\star(\sigma) = \sup_{\pi\in\Pi} J^\pi(\sigma)$ denote the maximum expected number of matches that can be achieved starting from state σ, and let $\pi^\star$ denote the corresponding optimal policy. Let π^g denote the greedy policy, i.e.,

$$\pi^g(\sigma) \in \arg\max_A \mathbb{E}\left[R(\sigma, A)\right] \ ,$$

where R is defined as in (EC.2) and the expectation is taken with respect to customer t's choices. Note that this definition agrees with that of the greedy algorithm in Section 3. Therefore, the following claim directly implies Proposition 1.

PROPOSITION EC.2. *Fix $\{z_t\}_{t=1}^T$ and $\{\mathcal{S}_t\}_{t=1}^T$. For every $\sigma \in \Sigma$ we have $J^\star(\sigma) \le 2 \cdot J^{\pi^g}(\sigma)$.*

To establish the above result, we follow a line of argumentation similar to that of Chan and Farias (2009). Namely, we establish two properties, value function monotonicity and immediate rewards, that will allow us to establish the desired result. These properties are established in the subsequent lemmas.

LEMMA EC.1. *Consider states σ and σ' satisfying $t(\sigma) = t(\sigma')$ and $C_i(\sigma) \subseteq C_i(\sigma')$ for all $i \in \mathcal{S}$. Then, $J^\star(\sigma) \ge J^\star(\sigma')$.*

Proof. We consider a coupling of the systems starting at states σ and σ', with $C_i(\sigma_t) \subseteq C_i(\sigma'_t)$ for all $i \in \mathcal{S}$, such that both systems witness identical sample paths for the customers' decisions. In what follows, we use our standard notation for the system starting at state σ, while prime notation is used to designate random variables corresponding to the system starting at state σ'. Hence, if the same assortment is shown to the arriving customer at time t in both systems, the resulting depletion vectors are equal, i.e., $\mathbf{X}_t = \mathbf{X}'_t$.

Now, assume that the system starting at σ' uses the optimal policy $\pi^\star$ whereas the system starting at σ uses a policy $\bar\pi$ that mimics the actions of the optimal policy under state σ'. Note that $\bar\pi$ is an admissible policy. Under the coupling, it follows that

$$\begin{aligned}
R(\sigma, \bar\pi(\sigma)) &= \sum_{i\in\mathcal{S}} \Big(w_i(C_i(\sigma) + \mathbf{e}_{z_{t(\sigma)}} \mathbf{X}_{t,i}) - w_i(C_i(\sigma)) \Big) \\
&= \sum_{i\in\mathcal{S}} \Big(w_i(C_i(\sigma) + \mathbf{e}_{z_{t(\sigma)}} \mathbf{X}'_{t,i}) - w_i(C_i(\sigma)) \Big) \\
&\ge \sum_{i\in\mathcal{S}} \Big(w_i(C_i(\sigma') + \mathbf{e}_{z_{t(\sigma)}} \mathbf{X}'_{t,i}) - w_i(C_i(\sigma')) \Big) \\
&= R(\sigma', \pi^\star(\sigma')) \ ,
\end{aligned}$$

where the second equality follows from the coupling argument, the first inequality follows from the fact that $C_i(\sigma_t) \subseteq C_i(\sigma'_t)$ together with the diminishing returns assumption (see Assumption 2) and the fact that the arrival sequence $\{z_t\}_{t=1}^T$ and the suppliers' set sequences $\{\mathcal{S}_t\}_{t=1}^T$ are assumed to be a priori given.

Moreover, suppose that, under the above policies, both systems transition to states $\sigma_{t(\sigma)+1}$ and $\sigma'_{t(\sigma)+1}$, respectively. By the coupling argument, $C_i(\sigma_{t(\sigma)+1}) \subseteq C_i(\sigma_{t(\sigma)+1})$ for all $i \in \mathcal{S}$, and thus the argument above can be repeated for step $t(\sigma)+1$. Therefore, one can establish that, in every time set, the system starting at σ and controlled by policy $\bar{\pi}$ earns a reward at least as large as that of the system starting at σ' and controlled by the optimal policy $\pi^\star$. Taking expectation over the $\mathbf{X}_t$ and $\mathbf{X}'_t$ processes, we obtain $J^\star(\sigma') \leq J^{\bar{\pi}}(\sigma) \leq J^\star(\sigma)$. $\square$

LEMMA EC.2. *Define $\tilde{\sigma} : \Sigma \times \mathbb{N}_0^{m\times\mathcal{Z}} \to \Sigma$ as $\tilde{\sigma}(\sigma,\alpha) = \sigma'$ with $t(\sigma') = t(\sigma)$ and $C_i(\sigma') = C_i(\sigma') + \alpha_{i,*}$. That is, state $\tilde{\sigma}(\sigma,\alpha)$ is obtained if one is permitted to add $\alpha_{i,*} \geq \mathbf{0}$ customer request to each set without using a time step. Then, for all $\sigma \in \Sigma$ and all $\alpha \in \mathbb{N}_0^{m\times\mathcal{Z}}$ we have*

$$J^\star(\sigma) \ \leq \ g(\bar{C}(\sigma), \bar{C}(\sigma)+\alpha) + J^\star(\tilde{\sigma}(\sigma,\alpha));$$

i.e., it is advantageous to incur a depletion of α without advancing the system by any time step.

Proof. Fix a state $\sigma \in \Sigma$ and $\alpha \in \mathbb{N}_0^{m\times\mathcal{Z}}$. Let $\sigma_T^\star$ denote the random state under the optimal policy at the end of the horizon starting from σ. Similarly, let $\tilde{\sigma}_T$ denote the random state under the optimal policy at the end of the horizon starting from $\sigma' = \tilde{\sigma}(\sigma,\alpha)$. Note that

$$J^\star(\sigma) = \mathbb{E}\left[\sum_{i\in\mathcal{S}} w_i(C_i(\sigma_T^\star))\right] - \sum_{i\in\mathcal{S}} w_i(C_i(\sigma))$$

and

$$J^\star(\sigma') = \mathbb{E}\left[\sum_{i\in\mathcal{S}} w_i(C_i(\tilde{\sigma}_T^\star))\right] - \sum_{i\in\mathcal{S}} w_i(C_i(\sigma'))\ .$$

We consider a coupling of the systems starting at states σ and σ' such that both systems witness identical sample paths for the customers' decisions. If the same assortment is shown to the arriving customer at time t in both systems, the resulting depletion vectors are equal, i.e., $\mathbf{X}_t = \mathbf{X}'_t$. Now, assume that the system starting at σ uses the optimal policy $\pi^\star$ whereas the system starting at σ' uses a policy $\bar{\pi}$ that mimics the actions of the optimal policy under state σ, that is, $\bar{\pi}(\sigma') = \pi^\star(t(\sigma'), (\bar{C}(\sigma') - \alpha))$. It follows that

$$\begin{aligned}
J^\star(\sigma) &= \mathbb{E}\left[\sum_{i\in\mathcal{S}} w_i(C_i(\sigma_T^\star))\right] - \sum_{i\in\mathcal{S}} w_i(C_i(\sigma)) \\
&= \mathbb{E}\left[\sum_{i\in\mathcal{S}} w_i\left(C_i(\sigma) + \sum_{t=t(\sigma)}^{T} \mathbf{e}_{z_t}\mathbf{X}_{t,i}\right)\right] - \sum_{i\in\mathcal{S}} w_i(C_i(\sigma)) \\
&= \mathbb{E}\left[\sum_{i\in\mathcal{S}} w_i\left(C_i(\sigma) + \sum_{t=t(\sigma)}^{T} \mathbf{e}_{z_t}\mathbf{X}'_{t,i}\right)\right] - \sum_{i\in\mathcal{S}} w_i(C_i(\sigma))
\end{aligned}$$

$$
\begin{aligned}
&\leq \mathbb{E}\left[\sum_{i\in\mathcal{S}} w_i\left(C_i(\sigma)+\alpha_{i,*}+\sum_{t=t(\sigma)}^{T}\mathbf{e}_{z_t}\mathbf{X}'_{t,i}\right)\right]-\sum_{i\in\mathcal{S}} w_i(C_i(\sigma))\\
&=\mathbb{E}\left[\sum_{i\in\mathcal{S}} w_i\left(C_i(\sigma')+\sum_{t=t(\sigma)}^{T}\mathbf{e}_{z_t}\mathbf{X}'_{t,i}\right)\right]-\sum_{i\in\mathcal{S}} w_i(C_i(\sigma'))+\sum_{i\in\mathcal{S}} w_i(C_i(\sigma'))-\sum_{i\in\mathcal{S}} w_i(C_i(\sigma))\\
&=J^{\bar{\pi}}(\sigma')+\sum_{i\in\mathcal{S}} w_i(C_i(\sigma'))-\sum_{i\in\mathcal{S}} w_i(C_i(\sigma))\\
&\leq J^{\star}(\sigma')+\sum_{i\in\mathcal{S}} w_i(C_i(\sigma'))-\sum_{i\in\mathcal{S}} w_i(C_i(\sigma))\ ,
\end{aligned}
$$

where the first equality follows from the definition of the optimal policy, the second equality proceeds by conditioning on the $\mathbf{X}_t$ process, and the third equality follows from the coupling argument. The inequality in the fourth line is a consequence of the monotonicity of $w_i(\cdot)$, and the equality in the fifth line uses the fact that $C_i(\sigma') = C_i(\sigma) + \alpha_{i,*}$. □

Proof of Proposition EC.2. The proof is by induction in the number of remaining steps. If $t(\sigma) = T$, then the myopic and optimal policies agree. Consider a state with $t(\sigma) < T$ and assume that the claim is true for all σ' with $t(\sigma') > t$. If $\pi^\star(\sigma) = \pi^g(\sigma)$, then the next states are identically distributed and using the induction hypothesis immediately yields the desired result.

Suppose that $\pi^\star(\sigma) \neq \pi^g(\sigma)$. We denote by $\mathbf{X}^\star_{t(\sigma)}$ and $\mathbf{X}^g_{t(\sigma)}$ the random vector of depletions in period $t(\sigma)$ under the optimal and greedy policies, respectively. Let $\sigma^\uparrow : \Sigma \times \mathbb{N}_0^{m\times\mathcal{Z}} \to \Sigma$ be the mapping such that, for every $\alpha \in \mathbb{N}_0^{m\times\mathcal{Z}}$, we have $\sigma^\uparrow(\sigma,\alpha) = \sigma'$ with $t(\sigma') = t(\sigma)+1$ and $C_i(\sigma') = C_i(\sigma') + \alpha_{i,*}$. That is, $\sigma^\uparrow$ specifies the next state obtained by starting from σ and depleting the suppliers' matching sets by α. Consequently, we have

$$
\begin{aligned}
&\mathbb{E}\left[R(\sigma,\pi^\star(\sigma))|\,\mathbf{X}^\star_{t(\sigma)}\right]+J^\star\left(\sigma^\uparrow(\sigma,\mathbf{X}^\star_{t(\sigma)})\right)\\
&\leq \mathbb{E}\left[R(\sigma,\pi^\star(\sigma))|\,\mathbf{X}^\star_{t(\sigma)}\right]+J^\star(\sigma^\uparrow(\sigma,0))\\
&\leq \mathbb{E}\left[R(\sigma,\pi^\star(\sigma))|\,\mathbf{X}^\star_{t(\sigma)}\right]+g\left(\bar{C}(\sigma),\bar{C}(\sigma)+\mathbf{X}^g_{t(\sigma)}\right)+J^\star\left(\tilde{\sigma}(\sigma^\uparrow(\sigma,0),\mathbf{X}^g_{t(\sigma)})\right)\\
&=\mathbb{E}\left[R(\sigma,\pi^\star(\sigma))|\,\mathbf{X}^\star_{t(\sigma)}\right]+\mathbb{E}\left[R(\sigma,\pi^g(\sigma))|\mathbf{X}^g_{t(\sigma)}\right]+J^\star\left(\tilde{\sigma}\left(\sigma^\uparrow(\sigma,0),\mathbf{X}^g_{t(\sigma)}\right)\right)\\
&=\mathbb{E}\left[R(\sigma,\pi^\star(\sigma))|\,\mathbf{X}^\star_{t(\sigma)}\right]+\mathbb{E}\left[R(\sigma,\pi^g(\sigma))|\,\mathbf{X}^g_{t(\sigma)}\right]+J^\star\left(\sigma^\uparrow\left(\sigma,\mathbf{X}^g_{t(\sigma)}\right)\right)\\
&\leq \mathbb{E}\left[R(\sigma,\pi^\star(\sigma))|\,\mathbf{X}^\star_{t(\sigma)}\right]+\mathbb{E}\left[R(\sigma,\pi^g(\sigma))|\,\mathbf{X}^g_{t(\sigma)}\right]+2\cdot J^g\left(\sigma^\uparrow\left(\sigma,\mathbf{X}^g_{t(\sigma)}\right)\right)\ ,
\end{aligned}
$$

where the first inequality follows from Lemma EC.1, and the second inequality follows from Lemma EC.2 and the fact that the state is equal to $\sigma^\uparrow(\sigma,0)$ and $\alpha = \mathbf{X}^g_{t(\sigma)}$; the first equality proceeds by noting that, using the definitions of $\tilde{\sigma}$ and $\sigma^\uparrow$, we have $\tilde{\sigma}\left(\sigma^\uparrow(\sigma,0),\mathbf{X}^g_{t(\sigma)}\right) = \sigma^\uparrow\left(\sigma,\mathbf{X}^g_{t(\sigma)}\right)$. The last inequality immediately follows from the induction hypothesis.

To conclude, note that

$$\begin{aligned}
J^\star(\sigma) &= \mathbb{E}\left[\mathbb{E}\left[R\left(\sigma,\pi^\star(\sigma)\right)\middle|\,\mathbf{X}^\star_{t(\sigma)}\right] + J^\star\left(\sigma^\uparrow(\sigma,\mathbf{X}^\star_{t(\sigma)})\right)\right] \\
&\le \mathbb{E}\left[\mathbb{E}\left[R(\sigma,\pi^\star(\sigma))\middle|\,\mathbf{X}^\star_{t(\sigma)}\right]\right] + \mathbb{E}\left[\mathbb{E}\left[R(\sigma,\pi^g(\sigma))\middle|\,\mathbf{X}^g_{t(\sigma)}\right] + 2J^g\left(\sigma^\uparrow\left(\sigma,\mathbf{X}^g_{t(\sigma)}\right)\right)\right] \\
&\le \mathbb{E}\left[R(\sigma,\pi^\star(\sigma))\right] + \mathbb{E}\left[R(\sigma,\pi^g(\sigma))\right] + 2\cdot\mathbb{E}\left[J^g\left(\sigma^\uparrow\left(\sigma,\mathbf{X}^g_{t(\sigma)}\right)\right)\right] \\
&\le 2\cdot J^g(\sigma)\ ,
\end{aligned}$$

where the last inequality uses the fact that $\mathbb{E}\left[R(\sigma,\pi^\star(\sigma))\right] \le \mathbb{E}\left[R(\sigma,\pi^g(\sigma))\right]$ by the definition of the greedy policy, and the definition of $J^g(\sigma)$. $\square$

EC.2.2. Proof of Proposition 2

To establish the tightness of the 1/2 approximation, we exploit the hardness result in Kapralov et al. (2013) for the online welfare maximization problem.

In the online welfare maximization problem, there are n items arriving online from a set N; each item should be allocated upon arrival to one of m agents (suppliers). Each agent $i \in [m]$ is endowed with a valuation function $v_i : 2^n \to \mathbb{R}_+$. The goal is to maximize $\sum_{i=1}^m v_i(C_i)$, where $C_i \subseteq N$ is the subset of items allocated to agent i. We now introduce the class of *coverage functions*, which imposes further restrictions on the agents' valuation functions.

Definition EC.2 (Coverage valuations). A valuation function $v : 2^{|N|} \to \mathbb{R}_+$ is called a *coverage valuation* if there is a set system $(\mathcal{U}, \{U_j\}_{j\in N})$ given by a finite set of elements $\mathcal{U}$ and a family $\{U_j\}_{j\in N}$ of subsets of $\mathcal{U}$ such that $v(C) = |\cup_{j\in C} U_j|$ for all $C \subseteq N$.

As stated in the next claim, Kapralov et al. (2013) show that, under a plausible complexity assumption, there exists no algorithm better than 1/2-competitive for the online welfare maximization problem with coverage valuation functions.

Theorem EC.1 **(Kapralov et al., 2013)**. *Unless NP = RP, there is no $(1/2+\delta)$-competitive polynomial-time algorithm (even randomized, against an oblivious adversary) for the online welfare maximization problem with coverage valuations and constant $\delta \ge 0$.*

Proof of Proposition 2. We first establish the result without any assumptions on ϵ. Specifically, for any $\delta > 0$ we show that there is no $(1/2+\delta)$-competitive polynomial-time algorithm for the online two-sided assortment problem. The proof proceeds by contradiction. Suppose that there exists a $\delta > 0$ such that there is a $(1/2+\delta)$-competitive polynomial-time algorithm for the online two-sided assortment problem. Let π be such an algorithm. We will argue below that we can use π to compute a $(1/2+\delta)$ approximation for the online welfare maximization problem with coverage valuations, thereby contradicting Theorem EC.1.

Specifically, let $\mathcal{I} = (N_{\mathcal{I}}, M_{\mathcal{I}}, \{v_i\}_{i \in M_{\mathcal{I}}})$ be an instance of the online welfare maximization problem with coverage valuations, where $N_{\mathcal{I}}$ is the set of items, $M_{\mathcal{I}}$ denotes the set of agents, and $\{v_i\}_{i \in M_{\mathcal{I}}}$ are their corresponding coverage valuations defined over a set system $(\mathcal{U}, \{U_j\}_{j \in N_{\mathcal{I}}})$. As in the proof of Theorem EC.1 in Kapralov et al. (2013), each item $j \in N_{\mathcal{I}}$ is associated with a set in $\{U_j\}_{j \in N_{\mathcal{I}}}$ and each valuation v_i is defined by a permutation $\alpha_i : N_{\mathcal{I}} \to N_{\mathcal{I}}$ over the sets such that $v_i(C) = \left|\cup_{j \in C} U_{\alpha_i(j)}\right|$ for any $C \subseteq N_{\mathcal{I}}$. Next, we construct an instance $\tilde{\mathcal{I}}$ of the online two-sided assortment problem as follows:

- Define the time horizon T as $T = |N_{\mathcal{I}}|$.

- Let $\mathcal{Z} = N_{\mathcal{I}}$. That is, for each item $j \in N_{\mathcal{I}}$, add customer type j defined by (i) a choice model without an outside option; i.e., for each set of suppliers S we require that $\phi_0(S) = 0$ whenever $S \neq \emptyset$, and that (ii) customer type j be associated with the set U_j, which we use below when constructing the suppliers' preferences.

- Let $\mathcal{S} = M_{\mathcal{I}}$. That is, for every agent $i \in M_{\mathcal{I}}$, add a supplier (denoted also by i) to the set $\mathcal{S}$ and define $\mathcal{S}_t = \mathcal{S}$ for every $1 \le t \le T$. Each supplier $i \in \mathcal{S}$ has rank-based preferences, generated as follows. Construct the preference lists $L_1^i, L_2^i, \ldots, L_{|\mathcal{U}|}^i$ (one for each element in the ground set), where L_k^i is formed by the elements $\{j : k \in U_{\alpha_i(j)}\}$ ranked in an arbitrary order. Supplier i's preferences are governed by the uniform distribution over lists $L_1^i, L_2^i, \ldots, L_{|\mathcal{U}|}^i$. That is, the probability that supplier i matches given a set of customer requests C is equal to the hitting probability $w_i(C) = \frac{1}{|\mathcal{U}|} \cdot \sum_{k \in \mathcal{U}} \mathbb{I}[C \cap L_k^i \neq \emptyset]$.

There are two crucial observations. First, for every $i \in M_{\mathcal{I}}$ and $C \subseteq N_{\mathcal{I}}$, we have

$$v_i(C) = \left|\cup_{j \in C} U_{\alpha_i(j)}\right| = \sum_{k \in \mathcal{U}} \mathbb{I}\left[k \in \left(\cup_{j \in C} U_{\alpha_i(j)}\right)\right] = \sum_{k \in \mathcal{U}} \mathbb{I}\left[C \cap L_k^i \neq \emptyset\right] = |\mathcal{U}| \cdot w_i(C) \ . \quad \text{(EC.4)}$$

Second, let $OPT(\mathcal{I})$ denote the (offline) optimal welfare in instance $\mathcal{I}$, and let $\{C_i^\star\}_{i \in M_{\mathcal{I}}}$ denote an optimal solution. As in the online two-sided assortment instance $\tilde{\mathcal{I}}$, the set of suppliers $\mathcal{S}$ is constant over time and customers do not have outside options, which means that an algorithm that shows to each arriving customer j the assortment $U_j = \{i\}$ if and only if $j \in C_i^\star$ is guaranteed to end with the set of requests $\bar{C}^T = \{C_i^\star\}_{i \in M_{\mathcal{I}}}$. Thus, for any sequence of arrivals $\{z_t\}_{t=1}^T$ in the online two-sided assortment instance $\tilde{\mathcal{I}}$, we have

$$\mathcal{M}^{\pi^C}\left(\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T\right) \ge \sum_{i \in \mathcal{S}} w_i(C_i^\star) = \sum_{i \in N_{\mathcal{I}}} \frac{1}{|\mathcal{U}|} \cdot v_i(C_i^\star) = \frac{1}{|\mathcal{U}|} \cdot OPT(\mathcal{I}) \ . \quad \text{(EC.5)}$$

Now, we devise an algorithm $\bar{\pi}$ for instance $\mathcal{I}$ of the online welfare maximization problem that mimics algorithm π. Let s^π be π's random seed and n_1 denote the first arrival in instance $\mathcal{I}$. Let A_1^π be the assortment shown by algorithm π to customer $z_1 = n_1$ in instance $\tilde{\mathcal{I}}$ (i.e., when $\mathcal{H}_1^\pi = \sigma(s^\pi, z_1, \mathcal{S}_1)$). Then, algorithm $\bar{\pi}$ assigns item n_1 to agent i with probability $\phi_i^{z_1}(A_1^\pi)$. Let ξ_1 denote the resulting assignment, and let n_2 be the incoming item. We define $\mathcal{H}_2^\pi = \sigma(\mathcal{H}_1^\pi, A_1^\pi, \xi_1(A_1^\pi), z_2 = n_2, \mathcal{S})$ and use A_2^π to construct the next randomized assignment of $\bar{\pi}$ as above, and so on. We specify algorithm $\bar{\pi}$ by iteratively repeating this process. By construction, we have

$$\frac{\mathbb{E}[\sum_{i\in M_\mathcal{I}} v_i(C_i^{T+1,\bar{\pi}})]}{OPT(\mathcal{I})} = \frac{\mathbb{E}[\sum_{i\in\mathcal{S}} |\mathcal{U}| \cdot w_i(C_i^{T+1,\pi})]}{OPT(\mathcal{I})} \geq \inf_{\{z_t\}_{t=1}^T} \frac{\mathcal{M}^\pi\left(\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T\right)}{\mathcal{M}^{\pi^C}\left(\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T\right)} \geq \frac{1}{2} + \delta \ ,$$

where the first equality proceeds from (EC.4) and the first inequality holds by (EC.5). We have just shown that the algorithm $\bar{\pi}$ achieves a competitive ratio of $1/2 + \delta$, thereby contradicting Theorem EC.1.

We next extend our result by imposing the ϵ-small probability assumption. Specifically, fix any $\epsilon > 0$ and any $\delta > 0$. A close inspection of the proof of Kapralov et al. (2013) reveals that the negative result they obtain for the online welfare maximization problem stems from the NP-hardness of distinguishing between certain YES-NO Max k-cover instances. We refer the reader to Kapralov et al. (2013, Section 2) for a formal statement of this NP-hardness result, which was established in Feige (1998). The crucial observation is that this NP-hardness result holds even under additional restrictions on the set systems $(\mathcal{U}, \{U_{\alpha_i(j)}\}_{j\in N_\mathcal{I}})$. Specifically, for large enough parameters $s, N_0 \in \mathbb{N}$, we may require the following properties:

Property 1: The subsets $\{U_j\}_{j\in N_\mathcal{I}}$ are all of size s.

Property 2: The number of elements in the ground set $|\mathcal{U}|$ is greater than or equal to N_0.

Consequently, the instance of the online two-sided assortment problem $\tilde{\mathcal{I}}$ constructed by our reduction satisfies, for every customer type $j \in \mathcal{Z}$,

$$w_i(\boldsymbol{e}_j) = \frac{1}{|\mathcal{U}|} \cdot \sum_{k\in\mathcal{U}} \mathbb{I}\left[\boldsymbol{e}_j \cap L_k^i\right] = \frac{1}{|\mathcal{U}|} \cdot \sum_{k\in\mathcal{U}} \mathbb{I}\left[k \in U_{\alpha_i(j)}\right] = \frac{|U_{\alpha_i(j)}|}{|\mathcal{U}|} \leq \frac{s}{N_0} \ ,$$

where the inequality is a direct consequence of Properties 1 and 2. Hence, the ϵ-small probability condition is met when $\frac{s}{N_0} \leq \epsilon$. Consequently, for every $\delta, \epsilon > 0$, we infer that there exists no polynomial-time randomized algorithm with a competitive ratio better than $1/2 + \delta$ in the ϵ-small probability regime. $\square$

EC.2.3. Proof of Proposition 3

Let p_z denote the probability that an arriving customer is of type $z \in \mathcal{Z}$ under distribution $\mathcal{D}$. Recall that a customer set $C \in \mathbb{N}_0^\mathcal{Z}$ is a vector with coordinates $z \in \mathcal{Z}$, where C_z denotes the number of match requests from customers of type z. In the remainder of this section, we use the shorthand $\mathcal{C} = [T]^\mathcal{S}$ to denote the collection of all possible customer sets generated over a time horizon of T periods.

We establish Proposition 3 through a sequence of lemmas. To this end, we introduce the following linear program:

$$
\begin{aligned}
(UB) \qquad & \max_{\mathbf{x},\mathbf{y}} && \sum_{i\in\mathcal{S}}\sum_{C\in\mathcal{C}} x_i(C)\cdot w_i(C) && \\
& \text{s.t.} && \sum_{C\in\mathcal{C}} x_i(C)\cdot C_z \le \sum_{A\subseteq\mathcal{S}} \phi_i^z(A)\cdot y_z(A)\cdot p_z T && \forall i\in\mathcal{S},\ \forall z\in\mathcal{Z} \\
& && \sum_{A\subseteq\mathcal{S}} y_z(A) = 1 && \forall z\in\mathcal{Z} \\
& && \sum_{C\in\mathcal{C}} x_i(C) = 1 && \forall i\in\mathcal{S} \\
& && x_i(C)\ge 0 && \forall i\in\mathcal{S}, C\in\mathcal{C} \\
& && y_z(A)\ge 0 && \forall z\in\mathcal{Z}, A\subseteq\mathcal{S}\ .
\end{aligned}
$$

Hereinafter, we denote by UB the optimal value of the linear program (UB).

Lemma EC.3. *The optimal value of (UB) is an upper bound on the optimal expected number of matches in the stochastic i.i.d. model with time-invariant supplier sets, where the expectation is taken over the arrival process. That is, for all instances satisfying the conditions in the statement of Proposition 3, we have*

$$\text{UB} \ \ge\ \mathbb{E}_{\{z_t\}_{t=1}^T}\left[\mathcal{M}^{\pi^C}\left(\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T\right)\right]\ .$$

Proof. Fix an arbitrary arrival sequence $\{z_t\}_{t=1}^T$ and an algorithm. Let the random variables $A_1,\ldots,A_T$ denote the sequence of assortments offered by the algorithm, and let $\xi_1,\ldots,\xi_T$ be the random variables corresponding to the customers' choices. Note that A_t can depend on $A_1,\ldots,A_{t-1}$ and $\xi_1,\ldots,\xi_{t-1}$, and we allow $C_1^{T+1},\ldots,C_m^{T+1}$ to depend on $A_1,\ldots,A_T$ and $\xi_1,\ldots,\xi_T$ as well. Note that the expected number of matches is given by

$$\mathbb{E}\left[\sum_{i\in\mathcal{S}}\sum_{C\in\mathcal{C}} w_i(C)\cdot\mathbb{I}\left[C_i^{T+1}=C\right]\right] = \sum_{i\in\mathcal{S}}\sum_{C\in\mathcal{C}} w_i(C)\cdot\Pr\left[C_i^{T+1}=C\right]\ .$$

For every $C\in\mathcal{C}$, let $\bar{x}_i(C)=\Pr[C_i^{T+1}=C]$, and let $\bar{\mathbf{x}}=(\bar{x}_i(C))_{i\in\mathcal{S},C\in\mathcal{C}}$. Note that, by definition, we have $\bar{x}_i(C)\ge 0$ for all $i\in\mathcal{S}$ and $C\in\mathcal{C}$, and $\sum_{C\in\mathcal{C}}\bar{x}_i(C)=1$ for all $i\in\mathcal{S}$. Therefore, it suffices to show that there exists $\bar{\mathbf{y}}$ such that $(\bar{\mathbf{x}},\bar{\mathbf{y}})$ is a feasible solution to (UB).

To construct such a vector $\bar{\mathbf{y}}$, we remark that, with probability 1, supplier i cannot see more requests of customers of type z than those directed to him. That is, with probability 1, we have

$$\sum_{C\in\mathcal{C}} C_z\cdot\mathbb{I}\left[C_i^{T+1}=C\right] \le \sum_{t=1}^{T}\sum_{A\subseteq\mathcal{S}}\mathbb{I}\left[A_t=A,\xi_t=i,z_t=z\right].$$

By taking expectations, on both sides, we have

$$\begin{aligned}\sum_{C\in\mathcal{C}} C_z\cdot\bar{x}_i(C) &\le \mathbb{E}\left[\sum_{t=1}^{T}\sum_{A\subseteq\mathcal{S}}\mathbb{I}\left[A_t=A,\xi_t=i,z_t=z\right]\right]\\ &= \mathbb{E}\left[\phi_i^z(A)\cdot\sum_{t=1}^{T}\sum_{A\subseteq\mathcal{S}}\mathbb{I}\left[A_t=A,z_t=z\right]\right]\\ &= \left(\sum_{t=1}^{T}\sum_{A\subseteq\mathcal{S}}\phi_i^z(A)\cdot\Pr\left[A_t=A|z_t\right]\Pr\left[z_t=z\right]\right)\\ &= p_z\cdot\left(\sum_{t=1}^{T}\sum_{A\subseteq\mathcal{S}}\phi_i^z(A)\cdot\Pr\left[A_t=A|z_t\right]\right)\ ,\end{aligned}$$

where the first equality follows from the tower property of conditional expectation, by noting that $\mathbb{E}[\mathbb{I}\left[\xi_t=i\right]|A_t=A,z_t=z]=\phi_i^z(A)$. The second equality follows from the definition of $\mathbb{E}[\mathbb{I}\left[A_t=A,z_t=z\right]]$. Lastly, the third equality holds since the arrivals are i.i.d.

Now, let $\bar{y}_z(A)=\frac{1}{T}\cdot\sum_{t=1}^{T}\Pr\left[A_t=A|z_t\right]$ and $\bar{\mathbf{y}}=(\bar{y}_z(A))_{z\in\mathcal{Z},A\subseteq\mathcal{S}}$. It remains to show that $\bar{\mathbf{y}}$ satisfies the constraints in UB. First, note that, as $\Pr\left[A_t=A|z_t=z\right]\ge 0$ for all A, we have $\bar{y}_z(A)\ge 0$. Moreover, by definition, it is necessary that $\sum_{A\subseteq\mathcal{S}}\Pr\left[A_t=A|z_t=z\right]=1$. Therefore, both constraints are satisfied, which completes our proof. $\square$

LEMMA EC.4. *For every time period $t\in[T]$ and vector of customer sets $(C_i')_{i\in\mathcal{S}}\in[T]^{\mathcal{S}\times\mathcal{Z}}$, conditional on the event $\bigcap_{i\in\mathcal{S}}\{C_i^t=C_i'\}$, the expected increase in the number of matches from the random request made by the $t+1$-th arriving customer is at least $\frac{1}{T}\cdot(\mathrm{UB}-\sum_{i\in\mathcal{S}}w_i(C_i'))$.*

Proof. In the remainder of the proof, we condition on the event $\bigcap_{i\in\mathcal{S}}\{C_i^t=C_i'\}$, meaning that the current vector of customer sets is precisely $(C_i')_{i\in\mathcal{S}}$. Let $(\mathbf{x},\mathbf{y})$ be any feasible UB solution and let $n_{iz}=\sum_{C\in\mathcal{C}}x_i(C)\cdot C_z$. We use the notation $C_i'+C$ to denote the union of the multisets C_i' and C, adding up the multiplicities of each customer type. By the property of diminishing returns (Assumption 2), we have

$$w_i(C_i'+C)-w_i(C_i')\le\sum_{z\in\mathcal{Z}}C_z\cdot\left(w_i\left(C_i'+\boldsymbol{e_z}\right)-w_i(C_i')\right)\ .$$

By the inequality $x_i(C)\ge 0$ for all $i\in\mathcal{S}$ and all $C\in\mathcal{C}$, we have

$$\begin{aligned}\sum_{C\in\mathcal{C}}x_i(C)\cdot(w_i(C_i'+C)-w_i(C_i')) &\le \sum_{C\in\mathcal{C}}\sum_{z\in\mathcal{Z}}x_i(C)\cdot C_z\cdot(w_i(C_i'+\boldsymbol{e_z})-w_i(C_i'))\\ &= \sum_{z\in\mathcal{Z}}n_{iz}\cdot(w_i(C_i'+\boldsymbol{e_z})-w_i(C_i'))\ .\end{aligned}$$

Since $\sum_{C\in\mathcal{C}}x_i(C)=1$ by feasibility and $w_i(C_i'+C)\ge w_i(C)$ by monotonicity, we obtain

$$\sum_{C\in\mathcal{C}}x_i(C)\cdot w_i(C)-w_i(C_i)\le\sum_{z\in\mathcal{Z}}n_{iz}\left(w_i\left(C_i'+\boldsymbol{e_z}\right)-w_i(C_i')\right). \tag{EC.6}$$

Now, consider the following randomized assortment decision: if the next customer is of type z, show her assortment A with probability $y_z(A)$. (By the feasibility of $(\mathbf{x}, \mathbf{y})$, we have $y_z(A) \geq 0$ and $\sum_{A \subseteq \mathcal{S}} y_z(A) = 1$.) Then, the random gain (i.e., increase in number of matches) from using this assortment decision is equal in expectation to

$$
\begin{aligned}
\mathbb{E}[\text{random gain}] &= \sum_{z \in \mathcal{Z}} p_z \cdot \sum_{A \subseteq \mathcal{S}} y_z(A) \left(\sum_{i \in \mathcal{S}} \phi_i^z(A) \left(w_i(C_i' + \boldsymbol{e_z}) - w_i(C_i') \right) \right) \\
&= \sum_{i \in \mathcal{S}} \sum_{z \in \mathcal{Z}} \left(w_i(C_i' + \boldsymbol{e_z}) - w_i(C_i') \right) \cdot \left(\sum_{A \subseteq \mathcal{S}} p_z \cdot y_z(A) \cdot \phi_i^z(A) \right) \\
&\geq \sum_{i \in \mathcal{S}} \sum_{z \in \mathcal{Z}} \left(w_i(C_i' + \boldsymbol{e_z}) - w_i(C_i') \right) \cdot \frac{n_{iz}}{T} \\
&\geq \frac{1}{T} \cdot \left(\sum_{i \in \mathcal{S}} \sum_{z \in \mathcal{Z}} x_i(C) \cdot w_i(C) - w_i(C_i') \right) .
\end{aligned} \tag{EC.7}
$$

The first inequality follows from the feasibility of $(\mathbf{x}, \mathbf{y})$, where we have

$$
\sum_{C \in \mathcal{C}} x_i(C) \cdot C_z \leq \sum_{A \subseteq \mathcal{S}} \phi_i^z(A) \cdot y_z(A) \cdot p_z T \qquad \forall i \in \mathcal{S}, z \in \mathcal{Z}.
$$

The second inequality follows from (EC.6).

To complete the proof, note that a greedy assortment decision gains as much as the randomized assortment decision for all feasible solutions $(\mathbf{x}, \mathbf{y})$. Therefore,

$$
\mathbb{E}[\text{greedy gain}] \geq \mathbb{E}[\text{random gain}] \geq \frac{1}{T} \cdot \left(\text{UB} - w_i(C_i') \right),
$$

where the last inequality follows from replacing $\mathbf{x}$ with the optimal solution of (UB) in (EC.7). $\square$

We now conclude the proof of Proposition 3.

Proof of Proposition 3. Denote the set of customer requests obtained after t periods as $\bar{C}^t = (C_1^t, \ldots, C_m^t)$. By Lemma EC.4, conditioned on $\bar{C}^t$, the expected number of matches after the next customer arrives and is offered an assortment is

$$
\mathbb{E}\left[\sum_{i \in \mathcal{S}} w_i(C_i^{t+1}) \,\middle|\, \bar{C}^t \right] \geq \sum_{i \in \mathcal{S}} w_i(C_i^t) + \frac{1}{T} \cdot \left(\text{UB} - \sum_i w_i(C_i^t) \right) .
$$

Taking the expection over $\bar{C}^t$, we obtain

$$
\mathbb{E}\left[\sum_i w_i(C_i^{t+1}) \right] \geq \sum_{i \in \mathcal{S}} \mathbb{E}[w_i(C_i^t)] + \frac{1}{T} \cdot \mathbb{E}\left[\left(\text{UB} - \sum_{i \in \mathcal{S}} w_i(C_i^t) \right) \right].
$$

Let $W(t) = \mathbb{E}[w_i(C_i^t)]$. By the last inequality, we have

$$
W(t+1) \geq W(t) + \frac{1}{T} \cdot (\text{UB} - W(t))
$$

or, equivalently,

$$\mathrm{UB}-W(t+1)\le\left(1-\frac{1}{T}\right)\cdot(\mathrm{UB}-W(t)).$$

Consequently, by a straightforward induction on $t\in[T]$, we have

$$\mathrm{UB}-W(t)\le\left(1-\frac{1}{T}\right)^{t}\cdot(UB-W(0))\le e^{-\frac{t}{T}}\cdot\mathrm{UB}.$$

Since the expected value of the solution found by the greedy algorithm at time T is $W(T)=\mathbb{E}\left[\sum_{i\in\mathcal{S}}w_i(C_i^{T+1})\right]$, we conclude that

$$W(T)\ge\left(1-\frac{1}{e}\right)\cdot\mathrm{UB}\ge\left(1-\frac{1}{e}\right)\cdot\mathbb{E}_{\{z_t\}_{t=1}^T}\left[\mathcal{M}^{\pi^C}\left(\{z_t\}_{t=1}^T,\{\mathcal{S}_t\}_{t=1}^T\right)\right].$$

Finally, we observe that the performance of the greedy algorithm is best-possible as a by-product of the following result in Kapralov et al. (2013):

THEOREM EC.2 **(Kapralov et al., 2013)**. *Unless NP = RP, there is no $(1-1/e+\delta)$-competitive polynomial-time algorithm for welfare maximization with coverage valuations in the i.i.d. stochastic model, for fixed $\delta>0$.*

Using the reduction in the proof of Proposition 2, we can show that a $(1-1/e+\delta)$-competitive polynomial-time algorithm for the online two-sided assortment problem in the i.i.d. setting of Proposition 3 would imply the existence of a $(1-1/e+\delta)$-competitive polynomial-time algorithm for welfare maximization with coverage valuations in the i.i.d. stochastic model. □

EC.3. Proofs of Section 4

EC.3.1. Proof of Lemma 1

We proceed by stating a series of valid inequalities that will prove the claim. Define $\tilde{f}_\kappa(x)=f_\kappa(\min\{1,x+\epsilon\})$.

$$\begin{aligned}
\mathcal{M}^{\pi^{f_\kappa}}\left(\{z_t\}_{t=1}^T,\{\mathcal{S}_t\}_{t=1}^T\right) &= \sum_{i\in S}\mathbb{E}_\omega\left[w_i\left(C_i^{T+1,\pi^{f_\kappa}}\right)\right]\\
&=\mathbb{E}_\omega\left[\sum_{t=1}^{T}\sum_{i\in A_t^{\pi^{f_\kappa}}}\Delta_i^t\cdot\phi_i^t\left(A_t^{\pi^{f_\kappa}}\right)\right]\\
&\ge\mathbb{E}_\omega\left[\sum_{t=1}^{T}\sum_{i\in A_t^{\pi^{f_\kappa}}}\Delta_i^t\cdot\frac{1-f_\kappa(w_i(C^{t,\pi^{f_\kappa}}))+\tilde{f}_\kappa(w_i(C^{t,\pi^{f_\kappa}}))}{1+\eta_\kappa\cdot\epsilon}\phi_i^t\left(A_t^{\pi^{f_\kappa}}\right)\right]\\
&\ge(1-\eta_\kappa\cdot\epsilon)\cdot\mathbb{E}_\omega\left[\sum_{t=1}^{T}\sum_{i\in A_t^{\pi^{f_\kappa}}}\Delta_i^t\cdot\phi_i^t\left(A_t^{\pi^{f_\kappa}}\right)\cdot\left(1-f_\kappa\left(w_i\left(C_i^{t,\pi^{f_\kappa}}\right)\right)\right)\right.
\end{aligned}$$

$$
\begin{aligned}
&\qquad\qquad + \sum_{t=1}^{T} \sum_{i \in A_t^{\pi^{f_\kappa}}} \Delta_i^t \cdot \phi_i^t \left(A_t^{\pi^{f_\kappa}}\right) \cdot \tilde{f}_\kappa \left(w_i \left(C_i^{t,\pi^{f_\kappa}}\right)\right) \Bigg] \\
&\geq (1 - \eta_\kappa \cdot \epsilon) \cdot \mathbb{E}_\omega \left[\sum_{t=1}^{T} \sum_{i \in A_t^{\pi^C}} \Delta_i^t \cdot \phi_i^t \left(A_t^{\pi^C}\right) \cdot \left(1 - f_\kappa \left(w_i \left(C_i^{t,\pi^{f_\kappa}}\right)\right)\right) \right.\\
&\qquad\qquad \left. + \sum_{t=1}^{T} \sum_{i \in A_t^{\pi^{f_\kappa}}} \Delta_i^t \cdot \phi_i^t \left(A_t^{\pi^{f_\kappa}}\right) \cdot \tilde{f}_\kappa \left(w_i \left(C_i^{t,\pi^{f_\kappa}}\right)\right) \right] \\
&= (1 - \eta_\kappa \cdot \epsilon) \cdot \sum_{i \in \mathcal{S}} \mathbb{E}_\omega \left[\sum_{t=1}^{T} \Delta_i^t \cdot \mathbb{I}\left[\xi_t \left(A_t^{\pi^C}\right) = i\right] \cdot \left(1 - f_\kappa \left(w_i \left(C_i^{t,\pi^{f_\kappa}}\right)\right)\right) \right.\\
&\qquad\qquad \left. + \sum_{t=1}^{T} \Delta_i^t \cdot \mathbb{I}\left[\xi_t \left(A_t^{\pi^{f_\kappa}}\right) = i\right] \cdot \tilde{f}_\kappa \left(w_i \left(C_i^{t,\pi^{f_\kappa}}\right)\right) \right] \\
&\geq (1 - \eta_\kappa \cdot \epsilon) \cdot \sum_{i \in \mathcal{S}} \mathbb{E}_\omega \left[\left(1 - f_\kappa \left(w_i \left(C_i^{T+1,\pi^{f_\kappa}}\right)\right)\right) \cdot \left(\sum_{t=1}^{T} \Delta_i^t \cdot \mathbb{I}\left[\xi_t \left(A_t^{\pi^C}\right) = i\right] \right) \right.\\
&\qquad\qquad \left. + \sum_{t=1}^{T} \Delta_i^t \cdot \mathbb{I}\left[\xi_t \left(A_t^{\pi^{f_\kappa}}\right) = i\right] \cdot \tilde{f}_\kappa \left(w_i \left(C_i^{t,\pi^{f_\kappa}}\right)\right) \right] , \\
&= (1 - \eta_\kappa \cdot \epsilon) \cdot \left(\sum_{i \in S} \mathbb{E}_\omega \left[R_i\right] \right) ,
\end{aligned}
$$

where the first inequality holds since $0 \leq \tilde{f}_\kappa(x) - f_\kappa(x) \leq \eta_\kappa \cdot \epsilon$ by the definition of $\tilde{f}_\kappa(\cdot)$ and η_κ. The third inequality follows from the optimality criterion used by the balance algorithm (see Eq. (2)), together with the fact that we assume $\epsilon \in (0, 1/\eta_\kappa)$ so that $1 - \eta_\kappa \cdot \epsilon > 0$. The next equality holds by the tower property of conditional expectations. Finally, the last inequality holds since the function $f(\cdot)$ is non-decreasing.

EC.3.2. Proof of Lemma 2

For simplicity, we drop the reference to the coupled probabilistic space ω. We start by noting that

$$
\begin{aligned}
&\sum_{t=1}^{T} \Delta_i^t \cdot \mathbb{I}\left[\xi_t \left(A_t^{\pi^C}\right) = i\right] \\
&\quad = \sum_{t=1}^{T} \left(w_i \left(C_i^{t,\pi^{f_\kappa}} + e_t \right) - w_i \left(C_i^{t,\pi^{f_\kappa}} \right) \right) \cdot \mathbb{I}\left[\xi_t \left(A_t^{\pi^C}\right) = i\right] \\
&\quad = \sum_{t=1}^{T} \frac{q_{i,t}}{(1 + \sum_{j \in C_i^{t,\pi^{f_\kappa}}} q_{i,j}) \cdot (1 + q_{i,t} + \sum_{j \in C_i^{t,\pi^{f_\kappa}}} q_{i,j})} \cdot \mathbb{I}\left[\xi_t \left(A_t^{\pi^C}\right) = i\right] \\
&\quad \geq \sum_{t=1}^{T} \frac{q_{i,t}}{(1 + \sum_{j \in C_i^{T+1,\pi^{f_\kappa}}} q_{i,j}) \cdot (1 + q_{i,t} + \sum_{j \in C_i^{T+1,\pi^{f_\kappa}}} q_{i,j})} \cdot \mathbb{I}\left[\xi_t \left(A_t^{\pi^C}\right) = i\right]
\end{aligned}
$$

$$
\begin{aligned}
&\geq \sum_{t=1}^{T} \frac{(1-\epsilon)\cdot q_{i,t}}{(1+\sum_{j\in C_i^{T+1,\pi^{f_\kappa}}} q_{i,j})^2} \cdot \mathbb{I}\left[\xi_t\left(A_t^{\pi^C}\right)=i\right] \\
&= (1-\epsilon)\cdot\left(1-w_i\left(C_i^{T+1,\pi^{f_\kappa}}\right)\right)^2 \cdot \left(\sum_{t=1}^{T} q_{i,t}\cdot \mathbb{I}\left[\xi_t\left(A_t^{\pi^C}\right)=i\right]\right) \\
&= (1-\epsilon)\cdot\left(1-w_i\left(C_i^{T+1,\pi^{f_\kappa}}\right)\right)^2 \cdot \frac{w_i(C_i^{T+1,\pi^C})}{1-w_i(C_i^{T+1,\pi^C})} \ ,
\end{aligned}
$$

where the first inequality holds since $C_i^{t,\pi^{f_\kappa}} \subseteq C_i^{T+1,\pi^{f_\kappa}}$ for every $t\in[T]$. For the second inequality, we use the fact that the small probability assumption together with the MNL assumption imply that $q_{i,j}\leq\epsilon$, plus the fact that $1/(1+\epsilon)\geq(1-\epsilon)$. The second-to-last equality uses the probabilistic structure of the MNL model, which implies that

$$
\frac{1}{(1+\sum_{j\in C_i^{T+1,\pi^{f_\kappa}}} q_{i,j})^2} = \left(1-w_i\left(C_i^{T+1,\pi^{f_\kappa}}\right)\right)^2 \ .
$$

The final equality follows from the fact that

$$
w_i(C_i^{T+1,\pi^C}) = \frac{\sum_{t=1}^{T} q_{i,t}\cdot\mathbb{I}\left[\xi_t\left(A_t^{\pi^C}\right)=i\right]}{1+\sum_{t=1}^{T} q_{i,t}\cdot\mathbb{I}\left[\xi_t\left(A_t^{\pi^C}\right)=i\right]} \ .
$$

EC.3.3. Proof of Lemma 3

Define $\tilde{f}_\kappa(x)=f_\kappa(\min\{1,x+\epsilon\})$. For simplicity, we drop the reference to the coupled probabilistic space ω. In addition, we observe that

$$
\begin{aligned}
&\sum_{t=1}^{T} \Delta_i^t\cdot\mathbb{I}\left[\xi_t\left(A_t^{\pi^{f_\kappa}}\right)=i\right]\cdot\tilde{f}_\kappa\left(w_i\left(C_i^{t,\pi^{f_\kappa}}\right)\right) \\
&= \sum_{t=1}^{T}\left(w_i\left(C_i^{t,\pi^{f_\kappa}}+e_t\right)-w_i\left(C_i^{t,\pi^{f_\kappa}}\right)\right)\cdot\mathbb{I}\left[\xi_t\left(A_t^{\pi^{f_\kappa}}\right)=i\right]\cdot f_\kappa\left(\min\left\{1,w_i\left(C_i^{t,\pi^{f_\kappa}}\right)+\epsilon\right\}\right) \\
&\geq \sum_{t=1}^{T}\mathbb{I}\left[\xi_t\left(A_t^{\pi^{f_\kappa}}\right)=i\right]\cdot\left(\int_{w_i(C_i^{t,\pi^{f_\kappa}})}^{w_i(C_i^{t,\pi^{f_\kappa}}+e_t)} f_\kappa(u)\,du\right) \\
&= \int_0^{w_i(C_i^{T+1,\pi^{f_\kappa}})} f_\kappa(u)\,du \ ,
\end{aligned}
\tag{EC.8}
$$

where the inequality follows by noting that

$$
\begin{aligned}
\left(w_i\left(C_i^{t,\pi^{f_\kappa}}+e_t\right)-w_i\left(C_i^{t,\pi^{f_\kappa}}\right)\right)\cdot f_\kappa\left(w_i\left(C_i^{t,\pi^{f_\kappa}}\right)+\epsilon\right) &= \int_{w_i(C_i^{t,\pi^{f_\kappa}})}^{w_i(C_i^{t,\pi^{f_\kappa}}+e_t)} f_\kappa\left(w_i\left(C_i^{t,\pi^{f_\kappa}}\right)+\epsilon\right)du \\
&\geq \int_{w_i(C_i^{t,\pi^{f_\kappa}})}^{w_i(C_i^{t,\pi^{f_\kappa}}+e_t)} f_\kappa(u)\,du \ ,
\end{aligned}
$$

since $f_\kappa(\cdot)$ is non-decreasing and $w_i(C_i^{t,\pi^{f_\kappa}}+e_t)\leq w_i(e_t)+w_i(C_i^{t,\pi^{f_\kappa}})\leq\epsilon+w_i(C_i^{t,\pi^{f_\kappa}})$ by the ϵ-small probability assumption plus the fact that $w_i(\cdot)$ satisfies the diminishing returns property.

EC.3.4. Proof of Theorem 2

By Lemma 1, we have

$$
\begin{aligned}
&\mathcal{M}^{\pi^{f_{\gamma,\kappa}}}\left(\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T\right) \\
&\geq (1-\eta_\kappa\cdot\epsilon)\cdot\sum_{i\in\mathcal{S}}\mathbb{E}_\omega\left[\left(1-f_\kappa\left(w_i\left(C_i^{T+1,\pi^{f_{\gamma,\kappa}}}\right)\right)\right)\cdot\left(\sum_{t=1}^T\Delta_i^t\cdot\mathbb{I}\left[\xi_t\left(A_t^{\pi^C}\right)=i\right]\right)\right. \\
&\qquad\left.+\sum_{t=1}^T\Delta_i^t\cdot\mathbb{I}\left[\xi_t\left(A_t^{\pi^{f_{\gamma,\kappa}}}\right)=i\right]\cdot\tilde{f}_\kappa\left(w_i\left(C_i^{t,\pi^{f_{\gamma,\kappa}}}\right)\right)\right] \\
&=(1-\eta_\kappa\cdot\epsilon)\cdot\sum_{i\in\mathcal{S}}\mathbb{E}_\omega\left[R_i(\omega)\right]\ .
\end{aligned}
\tag{EC.9}
$$

With respect to the optimal clairvoyant algorithm, we have

$$
\mathcal{M}^{\pi^C}\left(\{z_t\}_{t=1}^T, \{\mathcal{S}_t\}_{t=1}^T\right)=\sum_{i\in S}\mathbb{E}_w\left[w_i\left(C_i^{T+1,\pi^C}(\omega)\right)\right]\ . \tag{EC.10}
$$

The remainder of the proof proceeds by fixing a sample realization of ω, and by comparing the random variables appearing in the expectations (EC.9) and (EC.10) for each given supplier $i\in\mathcal{S}$. Specifically, the competitive ratio of Theorem 2 immediately follows from the next claim.

Lemma EC.5. $\frac{R_i(\omega)}{w_i(C_i^{T+1,\pi^C}(\omega))}\geq(1-\epsilon)\cdot\kappa$.

Proof. We begin by formulating an optimization problem that is closely related to our analysis. Define $B(\epsilon)$ as the set of non-negative vectors $\boldsymbol{s}=\{s_{t,k}\}_{t\in[T],k\in K}$ such that $\sum_{k\in K}(s_{t,k})^\gamma\leq\epsilon/(1-\epsilon)$ for every $t\in[T]$. Given a non-negative vector $\boldsymbol{x}=(x_1,\ldots,x_K)$ and $\alpha\in(0,1]$, the $(\boldsymbol{x},\alpha)$-*marginal nest allocation problem* is defined as follows:

$$
\begin{aligned}
\min_{\boldsymbol{s}\in B(\epsilon)}\ &\sum_{t=1}^T\left(\frac{\sum_{k=1}^K(x_k+s_{t,k})^\gamma}{1+\sum_{k=1}^K(x_k+s_{t,k})^\gamma}-\frac{\sum_{k=1}^K x_k^\gamma}{1+\sum_{k=1}^K x_k^\gamma}\right) \\
\text{s.t.}\ &\frac{\sum_{k=1}^K(\sum_{t=1}^T s_{t,k})^\gamma}{1+\sum_{k=1}^K(\sum_{t=1}^T s_{t,k})^\gamma}\geq\alpha\ .
\end{aligned}
\tag{EC.11}
$$

In what follows, we denote by $h(\boldsymbol{x},\alpha)$ the optimal value of the above mathematical program. In the next claim, we provide a lower bound on $h(\boldsymbol{x},\alpha)$, which is a key ingredient of our analysis.

Lemma EC.6. *For every $\boldsymbol{x}\geq 0$ and $\alpha\in(0,1]$, we have*

$$
h(\boldsymbol{x},\alpha)\geq(1-\epsilon)\cdot\nu(\boldsymbol{x})(1-\nu(\boldsymbol{x}))\cdot\left(\left(1+\left(\frac{\alpha\cdot(1-\nu(\boldsymbol{x}))}{\nu(\boldsymbol{x})\cdot(1-\alpha)}\right)^{1/\gamma}\right)^\gamma-1\right)\ ,
$$

where $\nu(\boldsymbol{x})=\sum_{k=1}^K x_k^\gamma/(1+\sum_{k=1}^K x_k^\gamma)$.

To avoid deviating from our main argument, the proof is deferred to Appendix EC.3.5. In the remainder of the proof, we use the shorthand $\alpha=w_i(C_i^{T+1,\pi^C})$. Let $\psi_{i,t,k}=\mathbb{I}[\xi_t(A_t^{\pi^C})=i,\ t\in N_k]$

for every $t \in [T]$ and $k \in [K]$. In addition, we specify the vectors $\boldsymbol{x} = (\sum_{t\in N_k} C_{i,t}^{T+1,\pi^{f_{\gamma,\kappa}}} \cdot q_{i,t})_{k\in K}$ and $\boldsymbol{s} = (q_{i,t} \cdot \psi_{i,t,k})_{t\in[T],k\in[K]}$.

Note that

$$
\begin{aligned}
&\sum_{t=1}^{T} \Delta_i^t \cdot \mathbb{I}\left[\xi_t\left(A_t^{\pi^C}\right) = i\right] \\
&= \sum_{t=1}^{T} \left(w_i\left(C_i^{t,\pi^{f_{\gamma,\kappa}}} + e_t\right) - w_i\left(C_i^{t,\pi^{f_{\gamma,\kappa}}}\right)\right) \cdot \mathbb{I}\left[\xi_t\left(A_t^{\pi^C}\right) = i\right] \\
&\geq \sum_{t=1}^{T} \left(w_i\left(C_i^{T+1,\pi^{f_{\gamma,\kappa}}} + e_t\right) - w_i\left(C_i^{T+1,\pi^{f_{\gamma,\kappa}}}\right)\right) \cdot \mathbb{I}\left[\xi_t\left(A_t^{\pi^C}\right) = i\right] \\
&= \sum_{t=1}^{T} \frac{\sum_{k=1}^{K}(\sum_{u\in N_k} C_{i,u}^{T+1,\pi^{f_{\gamma,\kappa}}} \cdot q_{i,u} + q_{i,t}\cdot\psi_{i,t,k})^\gamma}{1+\sum_{k=1}^{K}(\sum_{u\in N_k} C_{i,u}^{T+1} \cdot q_{i,u} + q_{i,t}\cdot\psi_{i,t,k})^\gamma} - \frac{\sum_{k=1}^{K}(\sum_{u\in N_k} C_{i,u}^{T+1,\pi^{f_{\gamma,\kappa}}} \cdot q_{i,u})^\gamma}{1+\sum_{k=1}^{K}(\sum_{u\in N_k} C_{i,u}^{T+1,\pi^{f_{\gamma,\kappa}}} \cdot q_{i,u})^\gamma} \\
&= \sum_{t=1}^{T} \frac{\sum_{k=1}^{K}(x_k + s_{t,k})^\gamma}{1+\sum_{k=1}^{K}(x_k+s_{t,k})^\gamma} - \frac{\sum_{k=1}^{K} x_k^\gamma}{1+\sum_{k=1}^{K} x_k^\gamma} \\
&\geq f(\boldsymbol{x},\alpha) \\
&\geq (1-\epsilon)\cdot\nu(\boldsymbol{x})(1-\nu(\boldsymbol{x}))\cdot\left(\left(1+\left(\frac{\alpha\cdot(1-\nu(\boldsymbol{x}))}{\nu(\boldsymbol{x})\cdot(1-\alpha)}\right)^{1/\gamma}\right)^{\gamma} - 1\right) \\
&= (1-\epsilon)\cdot w_i\left(C_i^{T+1,\pi^{f_{\gamma,\kappa}}}\right)\cdot\left(1-w_i\left(C_i^{T+1,\pi^{f_{\gamma,\kappa}}}\right)\right)\cdot \\
&\qquad \left(\left(1+\left(\frac{\alpha\cdot(1-w_i(C_i^{T+1,\pi^{f_{\gamma,\kappa}}}))}{w_i(C_i^{T+1,\pi^{f_{\gamma,\kappa}}})\cdot(1-\alpha)}\right)^{1/\gamma}\right)^{\gamma} - 1\right) ,
\end{aligned}
\tag{EC.12}
$$

where the first inequality proceeds from the submodularity of the aggregate matching function $w_i(\cdot)$. The second inequality follows by noting that the vector $\boldsymbol{s}$ is feasible with respect to the $(\boldsymbol{x},\alpha)$-marginal nest allocation problem (EC.11). Indeed, we have

$$
\alpha = w_i(C_i^{T+1,\pi^C}) = \frac{\sum_{k\in[K]}(\sum_{t\in[t]} s_{t,k})^\gamma}{1+\sum_{k\in[K]}(\sum_{t\in[t]} s_{t,k})^\gamma} .
$$

In addition, note that $\boldsymbol{s} \in B(\epsilon)$ since $\sum_{k\in[K]}(s_{t,k})^\gamma \leq q_{i,t}^\gamma \leq \epsilon/(1-\epsilon)$ where the last inequality follows from the ϵ-small probability assumption. The third inequality above proceeds from Lemma EC.6. Lastly, the final equality follows by noting that $\nu(\boldsymbol{x}) = w_i(C_i^{T+1})$.

Now, as in the proof of Theorem 1, we have

$$
\sum_{t=1}^{T} \Delta_i^t \cdot \mathbb{I}\left[\xi_t\left(A_t^{\pi^{f_{\gamma,\kappa}}}\right) = i\right] \cdot \tilde{f}_\kappa\left(w_i\left(C_i^{t,\pi^{f_{\gamma,\kappa}}}\right)\right) \geq \int_0^{w_i(C_i^{T+1,\pi^{f_{\gamma,\kappa}}})} f_{\gamma,\kappa}(u)\,du . \tag{EC.13}
$$

The latter inequality proceeds from the same reasoning as inequality (EC.8) in Appendix EC.3.3. By combining (EC.12) and (EC.13), we obtain

$$
\frac{R_i(\omega)}{w_i(C_i^{T+1,\pi^C}(\omega))}
$$

$$
\begin{aligned}
&\geq \frac{1}{\alpha} \cdot \left(\int_0^{w_i(C_i^{T+1,\pi^{f_{\gamma,\kappa}}})} f_{\gamma,\kappa}(u)\,du + (1-\epsilon) \cdot \left(1 - f_{\gamma,\kappa}\left(C_i^{T+1,\pi^{f_{\gamma,\kappa}}}\right)\right) \cdot w_i\left(C_i^{T+1,\pi^{f_{\gamma,\kappa}}}\right) \right.\\
&\qquad \left. \cdot \left(1 - w_i\left(C_i^{T+1,\pi^{f_{\gamma,\kappa}}}\right)\right) \cdot \left(\left(1 + \left(\frac{\alpha \cdot (1 - w_i(C_i^{T+1,\pi^{f_{\gamma,\kappa}}}))}{w_i(C_i^{T+1,\pi^{f_{\gamma,\kappa}}}) \cdot (1-\alpha)} \right)^{1/\gamma} \right)^{\gamma} - 1 \right) \right)\\
&\geq (1-\epsilon) \cdot \kappa \ ,
\end{aligned}
$$

where the second inequality holds since $f_{\gamma,\kappa}(\cdot)$ is a solution of inequality (MNL-ODI) with parameter κ. $\square$

EC.3.5. Proof of Lemma EC.6

Let $\tilde{h}(\boldsymbol{x}, \alpha)$ be the optimal value of a variant of the $(\boldsymbol{x}, \alpha)$-marginal nest allocation problem, defined as

$$
\begin{aligned}
\min_{\boldsymbol{s} \geq 0} \quad & \sum_{t=1}^{T} \sum_{k=1}^{K} (x_k + s_{t,k})^{\gamma} - x_k^{\gamma} \\
\text{s.t.} \quad & \frac{\sum_{k=1}^{K} (\sum_{t=1}^{T} s_{t,k})^{\gamma}}{1 + \sum_{k=1}^{K} (\sum_{t=1}^{T} s_{t,k})^{\gamma}} \geq \alpha \ .
\end{aligned}
\tag{EC.14}
$$

Let $\tilde{\boldsymbol{s}}$ be an optimal vector for (EC.14). Observe that, without loss of optimality, we may focus on solutions were for every $k \in [K]$ there exists $t_k \in [T]$ such that $\tilde{s}_{t,k} = 0$ for all $t \neq t_k$. Indeed, suppose that $\tilde{s}_{t_1,k}, \tilde{s}_{t_2,k} > 0$ for some $t_1 \neq t_2$ and $k \in [K]$. Consider the modified vector $\bar{\boldsymbol{s}}$ where $\bar{s}_{t_1,k} = 0$, $\bar{s}_{t_2,k} = \tilde{s}_{t_1,k} + \tilde{s}_{t_2,k}$, and $\bar{s}_{t',k'} = \tilde{s}_{t',k'}$ for every $(t', k') \notin \{(t_1, k), (t_2, k)\}$. The key observation is that the vector $\bar{\boldsymbol{s}}$ is feasible and has an objective value less than or equal to that of $\tilde{\boldsymbol{s}}$ since

$$
\begin{aligned}
&\left(\sum_{t=1}^{T} \sum_{k=1}^{K} (x_k + \bar{s}_{t,k})^{\gamma} - x_k^{\gamma} \right) - \left(\sum_{t=1}^{T} \sum_{k=1}^{K} (x_k + \tilde{s}_{t,k})^{\gamma} - x_k^{\gamma} \right) \\
&\qquad = (x_k + \tilde{s}_{t_1,k} + \tilde{s}_{t_2,k})^{\gamma} + x_k^{\gamma} - (x_k + \tilde{s}_{t_1,k})^{\gamma} - (x_k + \tilde{s}_{t_2,k})^{\gamma} \\
&\qquad \leq 0 \ ,
\end{aligned}
$$

where the last inequality holds since $\gamma \leq 1$. By repeating this procedure a finite number of times, we ensure that there exists an optimal vector that satisfies the above property. In what follows, we overload notation by referring to this vector as $\tilde{\boldsymbol{s}}$.

Moreover, the necessary KKT conditions for stationary points imply that there exists $\theta > 0$ such that, for every nest $k \in [K]$,

$$
\gamma \cdot \left(x_k + \tilde{s}_{t_k,k}\right)^{\gamma-1} = \theta \cdot \gamma \frac{(\sum_{u=1}^{T} \tilde{s}_{u,k})^{\gamma-1}}{(1 + \sum_{k=1}^{K} (\sum_{t=1}^{T} \tilde{s}_{t,k})^{\gamma})^2} \ .
$$

By rearranging the above equation, we infer that the ratio $(x_k + \tilde{s}_{t_k,k})/\tilde{s}_{t_k,k}$ is uniform across all the nests $k \in [K]$. Hence, it immediately follows that there exists $\tilde{\theta} > 0$ such that $\tilde{s}_{t_k,k} = \tilde{\theta} \cdot x_k$ for every $k \in [K]$. Consequently, $\tilde{h}(\boldsymbol{x}, \alpha)$ can be expressed in closed form as follows:

$$\begin{aligned}
\tilde{h}(\boldsymbol{x}, \alpha) &= \sum_{t=1}^{T} \sum_{k=1}^{K} (x_k + \tilde{s}_{t,k})^\gamma - x_k^\gamma \\
&= \sum_{k=1}^{K} \left(x_k + \tilde{s}_{t_k,k}\right)^\gamma - x_k^\gamma \\
&= \left(\sum_{k=1}^{K} x_k^\gamma\right) \cdot \left(\left(1 + \tilde{\theta}\right)^\gamma - 1\right) \\
&= \frac{\nu(\boldsymbol{x})}{1 - \nu(\boldsymbol{x})} \cdot \left(\left(1 + \left(\frac{\alpha \cdot (1 - \nu(\boldsymbol{x}))}{\nu(\boldsymbol{x}) \cdot (1 - \alpha)}\right)^{1/\gamma}\right)^\gamma - 1\right) ,
\end{aligned} \tag{EC.15}$$

where the fourth equality holds since (i) from the definition of ν, we immediately infer that $\sum_{k=1}^{K} x_k^\gamma = \frac{\nu(\boldsymbol{x})}{1-\nu(\boldsymbol{x})}$, and (ii) the optimality of $\tilde{\boldsymbol{s}}$ with respect to the modified optimization problem (EC.14) implies that $\tilde{\theta}^\gamma \cdot (\sum_{k=1}^{K} x_k^\gamma) = \sum_{k=1}^{K} (\sum_{t=1}^{T} \tilde{s}_{t,k})^\gamma = \alpha/(1-\alpha)$, where the latter equality proceeds by remarking that the constraint in (EC.14) is necessarily binding at optimality.

Now, we are ready to lower bound $h(\boldsymbol{x}, \alpha)$. To this end, let $\boldsymbol{s}^*$ be an optimal vector with respect to the original $(\boldsymbol{x}, \alpha)$-marginal nest allocation problem (EC.11). We obtain

$$\begin{aligned}
h(\boldsymbol{x}, \alpha) &= \sum_{t=1}^{T} \left(\frac{\sum_{k=1}^{K} (x_k + s^*_{t,k})^\gamma}{1 + \sum_{k=1}^{K} (x_k + s^*_{t,k})^\gamma} - \frac{\sum_{k=1}^{K} x_k^\gamma}{1 + \sum_{k=1}^{K} x_k^\gamma} \right) \\
&= \sum_{t=1}^{T} \frac{\sum_{k=1}^{K} (x_k + s^*_{t,k})^\gamma - \sum_{k=1}^{K} x_k^\gamma}{(1 + \sum_{k=1}^{K} (x_k + s^*_{t,k})^\gamma) \cdot (1 + \sum_{k=1}^{K} x_k^\gamma)} \\
&\geq \frac{1}{(1 + \sum_{k=1}^{K} x_k^\gamma + \sum_{k=1}^{K} s^{*\gamma}_{t,k}) \cdot (1 + \sum_{k=1}^{K} x_k^\gamma)} \cdot \left(\sum_{t=1}^{T} \sum_{k=1}^{K} \left((x_k + s^*_{t,k})^\gamma - x_k^\gamma \right) \right) \\
&\geq \frac{1 - \epsilon}{(1 + (1-\epsilon) \cdot \sum_{k=1}^{K} x_k^\gamma) \cdot (1 + \sum_{k=1}^{K} x_k^\gamma)} \cdot \left(\sum_{t=1}^{T} \sum_{k=1}^{K} \left((x_k + s^*_{t,k})^\gamma - x_k^\gamma \right) \right) \\
&\geq (1 - \epsilon) \cdot (1 - \nu(\boldsymbol{x}))^2 \cdot \left(\sum_{t=1}^{T} \sum_{k=1}^{K} \left((x_k + s^*_{t,k})^\gamma - x_k^\gamma \right) \right) \\
&\geq (1 - \epsilon) \cdot (1 - \nu(\boldsymbol{x}))^2 \cdot \tilde{h}(\boldsymbol{x}, \alpha) \\
&\geq (1 - \epsilon) \cdot \nu(\boldsymbol{x}) (1 - \nu(\boldsymbol{x})) \cdot \left(\left(1 + \left(\frac{\alpha \cdot (1 - \nu(\boldsymbol{x}))}{\nu(\boldsymbol{x}) \cdot (1 - \alpha)}\right)^{1/\gamma}\right)^\gamma - 1\right) ,
\end{aligned}$$

where the first inequality holds since $\gamma \leq 1$. The second inequality is implied by $\boldsymbol{s}^* \in B(\epsilon)$. The fourth inequality proceeds from the definition of the modified optimization problem (EC.14) relative to (EC.11). The last inequality immediately follows from (EC.15).

EC.3.6. Efficient construction of preference-aware balancing functions

Through the next lemma, we develop a numerical method that generates certifiable preference-aware balancing functions. This approach is implemented in the computer-aided proof program of Section 4.2.

LEMMA EC.7. *Fix $\gamma \in (0,1)$ and $\kappa \in [1/2, \kappa^*)$ such that there exists a K-Lipschitz non-decreasing discount function that satisfies the nonlinear ordinary differential inequality* (NL-ODI)*. For every $\kappa' \in [1/2, \kappa)$ and $\epsilon \in (0, \epsilon_0)$, where $\epsilon_0 = \min\{(\kappa-\kappa')/(K\kappa), \kappa'/(3+2\kappa/(\kappa-\kappa'))\}$, there is an algorithm that runs in time $O(\frac{1}{\epsilon^3} \cdot (\log\frac{1}{\epsilon})^2)$ and computes a continuous piecewise linear, non-decreasing discount function $f_\epsilon(\cdot)$ that satisfies the following nonlinear ordinary differential inequality in the variable $x \in [0,1]$:*

$$\min_{\alpha\in(0,1)} \frac{1}{\alpha} \cdot \left(\int_0^x f_\epsilon(u)du + (1-f_\epsilon(x)) \cdot x\,(1-x) \cdot \left(\left(1 + \left(\frac{\alpha\cdot(1-x)}{x\cdot(1-\alpha)} \right)^{\frac{1}{\gamma}} \right)^{\gamma} - 1 \right) \right) \geq \kappa' - \left(3 + \frac{2\kappa}{\kappa-\kappa'} \right) \epsilon \ . \quad \text{(EC.16)}$$

Proof. As in the statement of Lemma EC.7, we fix $\gamma \in (0,1)$ and $\kappa \in [1/2, \kappa^*)$ such that there exists a K-Lipschitz continuous, non-decreasing discount function $f_{\gamma,\kappa}(\cdot)$ that satisfies the nonlinear ordinary differential inequality (NL-ODI). Let $\kappa' \in [1/2, \kappa)$ and let $\epsilon = \frac{1}{N}$ for some integer $N \geq \frac{1}{\epsilon_0}$. In what follows, we formulate an approximate dynamic program that returns a piecewise linear function $f_\epsilon(\cdot): [0,1] \to [0,1]$ with respect to the sequence of intervals $\{I_k\}_{k=1}^{N^2}$, where $I_k = [(k-1)\epsilon^2, k\cdot\epsilon^2)$ for each $k \in [N^2]$, that meets the properties stated in Lemma EC.7.

Dynamic program. Each state of the dynamic program is described by the integer state variables $(k,d) \in [N^2] \times [0, \lceil \log_{1+\epsilon} N \rceil]$. For every $d \in [0, \lceil \log_{1+\epsilon} N \rceil]$, we define $w_d = \epsilon \cdot (1+\epsilon)^d$. Next, we define the value function $F(\cdot)$ through the following dynamic programming equations, where for every integers $2 \leq k \leq N^2$ and $0 \leq d \leq [0, \lceil \log_{1+\epsilon} N \rceil]$, we have

$$F(k,d) = \max_{\substack{d_0\in[0,d]: \\ F(k-1,d_0)\geq 0}} \begin{cases} F(k-1,d_0) + \epsilon^2 \cdot w_d \ , & \text{if } w_{d-1} \leq 1 - \Psi\left((k-1)\cdot\epsilon^2, F(k-1,d_0)\right) \\ -1 \ , & \text{otherwise.} \end{cases} \quad \text{(EC.17)}$$

Here, for every $x, y \in (0,1)$, we denote by $\Psi(x,y)$ an $(1-\epsilon)$-close underestimate of the optimal value of the following optimization problem:

$$\max_{\alpha\in(0,1)} \mathbb{I}[y \leq C_1] \cdot \frac{\kappa'\cdot\alpha - y/(1-\epsilon)}{(1-x)\cdot x \cdot ((1+(\frac{\alpha\cdot(1-x)}{x\cdot(1-\alpha)})^{\frac{1}{\gamma}})^{\gamma} - 1)} \ , \quad \text{(EC.18)}$$

where $C_1 = \kappa' \cdot (1-\epsilon)^2$. We will explain shortly how problem (EC.18) is approximated by proposing an efficient subroutine. To fully specify our dynamic program $F(\cdot)$, we impose the boundary conditions $F(1,d) = \epsilon^2 \cdot w_d$ if $(1-\epsilon)\cdot(1-\frac{\kappa'}{\kappa}) \leq w_{d-1} \leq (1-\kappa')$ and $F(1,d) = -1$ otherwise.

As revealed by our subsequent analysis, the value of $F(k,d)$ can be thought of as an "approximation" for the maximum value of the integral $\int_0^{k\epsilon^2} g(u)du$ over all continuous piecewise linear, non-decreasing functions $g(\cdot):[0,k\epsilon^2]\mapsto[0,w_d]$ that satisfy the following nonlinear ordinary differential inequality in the variable $x\in[0,k\epsilon^2]$:

$$\min_{\alpha\in(0,1)} \frac{1}{\alpha}\cdot\left(\int_0^x g(u)du+(1-g(x))\cdot x\,(1-x)\cdot\left(\left(1+\left(\frac{\alpha\cdot(1-x)}{x\cdot(1-\alpha)}\right)^{\frac{1}{\gamma}}\right)^{\gamma}-1\right)\right)\geq \kappa'-O(\epsilon)\ .$$

When no such function $g(\cdot)$ exists, the dynamic program takes the value $F(k,d)=-1$.

Computational aspects. In terms of running time, the dynamic program $F(\cdot)$ can be solved in time $O(|C_\Psi|\cdot N^2\log_{1+\epsilon}N)$, where $|C_\Psi|$ is an upper bound on the running time for approximately solving the optimization problem (EC.18) that defines the value of $\Psi(x,y)$ for any given inputs $x,y\in(0,1)$ generated by the dynamic programming equation. For this purpose, we utilize a binary search method on the variable $\alpha\in(0,1)$ as a subroutine to approximate problem (EC.18). Next, we establish an upper bound on the running time of this method.

Claim EC.1. $|C_\Psi|=O(\log\frac{1}{\epsilon})$.

Based on the above claim, the dynamic program F has an overall running time of $O(\frac{1}{\epsilon^3}(\log\frac{1}{\epsilon})^2)$.

Analysis. We proceed by analyzing the dynamic program $F(\cdot)$ to establish the properties stated by Lemma EC.7. Our analysis is twofold. First, we relate the value function $F(\cdot)$ of our dynamic program to the function $f_{\gamma,\kappa}(\cdot)$. Second, we show how to convert the optimal sequence of dynamic programming decisions into a piecewise linear, non-decreasing discount function $f_\epsilon(\cdot)$ that satisfies the properties of Lemma EC.7. These properties are established via the next claims.

Claim EC.2 **(Existence)**. *For every $k\in[N^2]$, we have $\max_{d\in[0,\lceil\log_{1+\epsilon}N\rceil]}F(k,d)\geq(1-\epsilon)\cdot(\int_0^{k\epsilon^2}f_{\gamma,\kappa}(u)du)$.*

Claim EC.3 **(Certifiability)**. *Let $(1,d_1^*),(2,d_2^*),\ldots,(N^2,d_{N^2}^*)$ be an optimal sequence of dynamic programming decisions with respect to $F(\cdot)$ and define $f_\epsilon(u)=w_{d_{k-1}^*-1}+(u-(k-1)\epsilon^2)/\epsilon^2\cdot w_{d_k^*-1}$, where $d_0^*=d_1^*$ and $k\in[N^2]$ is the unique integer such that $u\in I_k$. Then, the function $f_\epsilon(\cdot)$ is continuous piecewise linear and non-decreasing, and it satisfies the non-linear ordinary differential inequality* (EC.16).

Proof of Claim EC.1. Fix any inputs $x,y\in(0,1)^2$ of the optimization problem (EC.18), as generated by the dynamic programming equation. To upper bound $|C_\Psi|$, we make the following observations:

- Observation 1: Without loss of optimality, we can restrict the variable α to the domain $(C_2,1)$ where $C_2=\min\{y,C_1\}/(1-\epsilon)/\kappa'$. The function optimized by problem (EC.18) is unimodal with respect to α in the domain $(C_2,1)$.

- Observation 2: Let α^* be the optimal solution of problem (EC.18). By noting that $C_2 \leq (1-\epsilon)$, it is not difficult to verify that $\alpha^* \geq (1+\epsilon^3) \cdot C_2$.
- Observation 3: In each call to the subroutine for problem (EC.18), the argument x in $\Psi(x,y)$ satisfies $x \leq 1-\epsilon^2$; this observation proceeds by noting that the argument is of the form $x = (k-1)\epsilon^2$ where $k \in [N^2]$. Based on this observation and the fact that $C_2 \leq 1-\epsilon$, it is not difficult to verify that $\alpha^* \leq 1-\epsilon^4$.

By combining these observations, we conclude that our binary search method returns a $(1-\epsilon)$-underestimate of $\Psi(x,y)$ after $O(\log \frac{1}{\epsilon y})$ steps. Due to the boundary conditions on $F(1,\cdot)$, we have $y \geq \epsilon^3$ in each call to our subroutine, meaning that $\Psi(x,y)$ is determined in time $O(\log \frac{1}{\epsilon})$.

□

Proof of Claim EC.2. Define $d_k = \lceil \log_{1+\epsilon}(\frac{1}{\epsilon} \cdot \max\{f_{\gamma,\kappa}((k-1)\epsilon^2), 1-\frac{\kappa'}{\kappa}\})\rceil$ for every $k \in [N^2]$. We show that the sequence of dynamic programming decisions $(1,d_1),\dots,(N^2,d_{N^2})$ is feasible with respect to the dynamic programming equation (EC.17). It follows that

$$
\begin{aligned}
F\left(k,d_k\right) &\geq \sum_{q=1}^{k} \epsilon^2 \cdot w_{d_q} \geq \sum_{1=1}^{k} \int_{(q-1)\epsilon^2}^{q\epsilon^2} \max\left\{ f_{\gamma,\kappa}\left((q-1)\cdot\epsilon^2\right), 1-\frac{\kappa'}{\kappa}\right\} du \\
&\geq \sum_{q=1}^{k} \int_{(q-1)\epsilon^2}^{q\epsilon^2} \max\left\{ f_{\gamma,\kappa}\left(u\right) - K\epsilon^2, 1-\frac{\kappa'}{\kappa}\right\} du \geq \left(1-\frac{K\kappa}{\kappa-\kappa'}\epsilon^2\right)\int_0^{k\epsilon^2} f_{\gamma,\kappa}(u)du \\
&\geq (1-\epsilon)\cdot \int_0^{k\epsilon^2} f_{\gamma,\kappa}(u)du \ ,
\end{aligned} \tag{EC.19}
$$

where the third inequality holds since $f_{\gamma,\kappa}(\cdot)$ is K-Lipschitz and the last inequality ensues from the fact that $\epsilon \leq \epsilon_0 \leq \frac{K\kappa}{\kappa-\kappa'}$.

Now, we fix $k \in [N^2]$ and we seek to show that $w_{d_k-1} \leq 1-\Psi((k-1)\epsilon^2, F(k-1,d_{k-1}))$, meaning that the corresponding dynamic programming decision at state (k,d_k) is feasible. Suppose that $f_{\gamma,\kappa}((k-1)\epsilon^2) \geq 1-\frac{\kappa'}{\kappa}$. Here, we note that

$$
\begin{aligned}
w_{d_k-1} &\leq f_{\kappa}((k-1)\epsilon^2) \\
&\leq 1-\max\left\{0\ ,\ \max_{\alpha\in(0,1)} \frac{\kappa\cdot\alpha - \int_0^{(k-1)\epsilon^2} f_\kappa(u)du}{(1-(k-1)\epsilon^2)\cdot(k-1)\epsilon^2\cdot((1+(\frac{\alpha(1-(k-1)\epsilon^2)}{(k-1)\epsilon^2\cdot(1-\alpha)})^{\frac{1}{\gamma}})^{\gamma}-1)}\right\} \\
&\leq 1-\max_{\alpha\in(0,1)}\max\left\{0\ ,\ \frac{\kappa'\cdot\alpha - F(k-1,d_{k-1})/(1-\epsilon)}{(1-(k-1)\epsilon^2)\cdot(k-1)\epsilon^2\cdot((1+(\frac{\alpha(1-(k-1)\epsilon^2)}{(k-1)\epsilon^2\cdot(1-\alpha)})^{\frac{1}{\gamma}})^{\gamma}-1)}\right\} \\
&\leq 1-\Psi\left((k-1)\epsilon^2, F(k-1,d_{k-1})\right)\ ,
\end{aligned} \tag{EC.20}
$$

where the second inequality follows from the differential inequality (NL-ODI). The third inequality proceeds from inequality (EC.19). The last inequality holds since $\Psi(\cdot)$ is an underestimate of the optimal value of problem (EC.18).

Conversely, suppose that $f_{\gamma,\kappa}((k-1)\epsilon^2) < 1 - \frac{\kappa'}{\kappa} = w_{d_k}$. By multiplying the differential inequality (NL-ODI) by a factor of $(\kappa'/\kappa)/(1 - f_{\gamma,\kappa}((k-1)\epsilon^2))$, we obtain in particular that, for $x_k = (k-1)\epsilon^2$,

$$\min_{\alpha\in(0,1)} \frac{1}{\alpha} \cdot \left(\int_0^{x_k} f_{\gamma,\kappa}(u)du + \left(1 - w_{d_k}\right) \cdot x_k\left(1 - x_k\right) \cdot \left(\left(1 + \left(\frac{\alpha(1-x_k)}{x_k \cdot (1-\alpha)} \right)^{\frac{1}{\gamma}} \right)^{\gamma} - 1 \right) \right) \geq \kappa' .$$

Consequently, using a line of reasoning identical to that in the sequence of inequalities (EC.20), we infer that

$$\begin{aligned}
w_{d_k - 1} &\leq 1 - \max_{\alpha\in(0,1)} \max \left\{ 0 \ , \ \frac{\kappa' \cdot \alpha - \int_0^{(k-1)\epsilon^2} f_\kappa(u)du}{(1-(k-1)\epsilon^2)\cdot(k-1)\epsilon^2 \cdot ((1 + (\frac{\alpha(1-(k-1)\epsilon^2)}{(k-1)\epsilon^2\cdot(1-\alpha)})^{\frac{1}{\gamma}})^{\gamma} - 1)} \right\} \\
&\leq 1 - \max_{\alpha\in(0,1)} \max \left\{ 0 \ , \ \frac{\kappa' \cdot \alpha - F(k-1, d_{k-1})/(1-\epsilon)}{(1-(k-1)\epsilon^2)\cdot(k-1)\epsilon^2 \cdot ((1 + (\frac{\alpha(1-(k-1)\epsilon^2)}{(k-1)\epsilon^2\cdot(1-\alpha)})^{\frac{1}{\gamma}})^{\gamma} - 1)} \right\} \\
&\leq 1 - \Psi\left((k-1)\epsilon^2, F(k-1, d_{k-1})\right) .
\end{aligned}$$

□

Proof of Claim EC.3. By construction, it is immediate that the function $f_\epsilon(\cdot)$ is continuous and piecewise linear. The constraints of the dynamic program (EC.17) guarantee that $d_0^* \leq d_1^* \leq \ldots \leq d_{N^2}^*$, which implies that $f_\epsilon(\cdot)$ is non-decreasing. Consequently, it suffices to show that, for every $x \in [0,1]$ and $\alpha \in (0,1)$, we have

$$\begin{aligned}
\frac{1}{\alpha} \cdot \left(\int_0^{x} f_\epsilon(u)du + (1 - f_\epsilon(x)) \cdot x\left(1 - x\right) \cdot \left(\left(1 + \left(\frac{\alpha(1-x)}{x \cdot (1-\alpha)} \right)^{\frac{1}{\gamma}} \right)^{\gamma} - 1 \right) \right)& \\
\geq \kappa' - \left(3 + \frac{2\kappa}{\kappa - \kappa'}\right)\epsilon .& \qquad \text{(EC.21)}
\end{aligned}$$

To this end, we first verify the above inequality for all $x \in \{(k-1)\cdot\epsilon^2 : k \in [N^2]\}$ by leveraging the dynamic programming equation (EC.17). With respect to $x = (k-1)\cdot\epsilon^2$ for some $k \in [N^2]$ and an arbitrary $\alpha \in (0,1)$, we have

$$\begin{aligned}
f_\epsilon(x) &= w_{d_{k-1}^* - 1} \\
&\leq w_{d_k^* - 1} \\
&\leq 1 - \Psi\left((k-1)\epsilon^2, F\left(k-1, d_{k-1}^*\right)\right) \\
&= 1 - \Psi\left((k-1)\epsilon^2, \sum_{q=1}^{k-1} \epsilon^2 \cdot w_{d_q^* - 1}\right) \\
&\leq 1 - \Psi\left((k-1)\epsilon^2, \int_0^{(k-1)\epsilon^2} f_\epsilon(u)du\right)
\end{aligned}$$

$$\leq 1-(1-\epsilon)\cdot \max_{\alpha\in(0,1)} \max\left\{0\ ,\ \frac{(\kappa'-\epsilon)\cdot\alpha-(\int_0^x f_\epsilon(u)du)/(1-\epsilon)}{(1-x)\cdot x\cdot((1+(\frac{\alpha(1-x)}{x\cdot(1-\alpha)})^{\frac{1}{\gamma}})^{\gamma}-1)}\right\}$$

$$\leq 1-(1-\epsilon)\cdot \max_{\alpha\in(0,1)} \max\left\{0\ ,\ \frac{(\kappa'-2\epsilon)\cdot\alpha-\int_0^x f_\epsilon(u)du}{(1-x)\cdot x\cdot((1+(\frac{\alpha(1-x)}{x\cdot(1-\alpha)})^{\frac{1}{\gamma}})^{\gamma}-1)}\right\}$$

$$\leq 1- \max_{\alpha\in(0,1)} \max\left\{0\ ,\ \frac{(\kappa'-3\epsilon)\cdot\alpha-\int_0^x f_\epsilon(u)du}{(1-x)\cdot x\cdot((1+(\frac{\alpha(1-x)}{x\cdot(1-\alpha)})^{\frac{1}{\gamma}})^{\gamma}-1)}\right\}\ , \tag{EC.22}$$

where the first inequality proceeds from the dynamic programming constraint in (EC.17). The second equality is formed by repeatedly applying the dynamic programming equation. The third inequality holds since $f_\epsilon(u)\geq w_{d_q^*-1}$ if $u\in I_q$ by construction. The fourth inequality follows from the definition of $\Psi(\cdot)$ in (EC.18). By rearranging the above inequality, it immediately follows that inequality (EC.21) is satisfied when $x=(k-1)\cdot\epsilon^2$.

Now, it remains to consider the case of an arbitrary $x\in I_k$, meaning that in particular $(k-1)\cdot\epsilon^2\leq x\leq k\cdot\epsilon^2$. Here, we distinguish between two cases.

Case 1: $\boldsymbol{k\geq\frac{1}{\epsilon}}$. Let $x_k=k\cdot\epsilon^2$. Using a sequence of inequalities identical to that leading to inequality (EC.22), we obtain

$$\begin{aligned}
f_\epsilon(x) &\leq w_{d_{k+1}^*-1}\\
&\leq 1-\max_{\alpha\in(0,1)}\max\left\{0\ ,\ \frac{(\kappa'-3\epsilon)\cdot\alpha-\int_0^{x_k} f_\epsilon(u)du}{(1-x_k)\cdot x_k\cdot((1+(\frac{\alpha(1-x_k)}{x_k\cdot(1-\alpha)})^{\frac{1}{\gamma}})^{\gamma}-1)}\right\}\\
&= 1-\max_{\alpha\in(0,1)}\max\left\{0\frac{(\kappa'-3\epsilon)\cdot\alpha-\int_0^{x_k} f_\epsilon(u)du}{(1-x_k)\cdot((x_k^{\frac{1}{\gamma}}+(\frac{\alpha(1-x_k)}{1-\alpha})^{\frac{1}{\gamma}})^{\gamma}-x_k)}\right\}\\
&\leq 1-\max_{\alpha\in(0,1)}\max\left\{0\frac{(\kappa'-3\epsilon)\cdot\alpha-\int_0^{x_k} f_\epsilon(u)du}{(1-x)\cdot((x^{\frac{1}{\gamma}}+(\frac{\alpha(1-x)}{1-\alpha})^{\frac{1}{\gamma}})^{\gamma}-x)}\right\}\\
&= 1-\max_{\alpha\in(0,1)}\max\left\{0\frac{(\kappa'-3\epsilon)\cdot\alpha-\int_0^{x} f_\epsilon(u)du-\int_x^{x_k} f_\epsilon(u)du}{(1-x)\cdot((x^{\frac{1}{\gamma}}+(\frac{\alpha(1-x)}{1-\alpha})^{\frac{1}{\gamma}})^{\gamma}-x)}\right\}\\
&\leq 1-\max_{\alpha\in(0,1)}\max\left\{0\frac{(\kappa'-3\epsilon)\cdot\alpha-(1+2\epsilon\kappa/(\kappa-\kappa'))\cdot\int_0^{x} f_\epsilon(u)du}{(1-x)\cdot((x^{\frac{1}{\gamma}}+(\frac{\alpha(1-x)}{1-\alpha})^{\frac{1}{\gamma}})^{\gamma}-x)}\right\}\\
&\leq 1-\max_{\alpha\in(0,1)}\max\left\{0\frac{(\kappa'-3\epsilon-2\epsilon\kappa/(\kappa-\kappa'))\cdot\alpha-\int_0^{x} f_\epsilon(u)du}{(1-x)\cdot((x^{\frac{1}{\gamma}}+(\frac{\alpha(1-x)}{1-\alpha})^{\frac{1}{\gamma}})^{\gamma}-x)}\right\}\ ,
\end{aligned}$$

where the first inequality holds since $f_\epsilon(x)\in[w_{d_{k-1}^*-1},w_{d_k^*-1}]$. The third inequality proceeds by noting that $x\leq x_k$ and, for every $\theta\geq 0$, the function $z\mapsto(1-z)\cdot(z^{\frac{1}{\gamma}}+\theta(1-z)^{\frac{1}{\gamma}})^{\gamma}-z)$ is non-increasing over $z\in[0,1]$. The next inequality holds since $\int_x^{x_k} f_\epsilon(u)du\leq\epsilon^2\leq\epsilon\cdot x/(1-\epsilon)\leq\frac{2\epsilon\kappa}{\kappa-\kappa'}\cdot\int_0^x f_\epsilon(u)du$, where the second inequality holds since $x\geq(k-1)\cdot\epsilon^2\geq(1-\epsilon)\cdot\epsilon$ and the last inequality holds since $f_\epsilon(u)\geq w_{d_0^*-1}\geq(1-\epsilon)\cdot(1-\frac{\kappa'}{\kappa})$ for every $u\in[0,x]$.

Case 2: $\boldsymbol{k < \frac{1}{\epsilon}}$. Let $x_k = (k-1)\cdot\epsilon^2$. For every $\alpha \geq (\int_0^x f_\epsilon(u)du)/(\kappa' - 3\epsilon)$, we have

$$\begin{aligned}
&(1-x)\cdot\left(\left(1+\left(\frac{\alpha(1-x)}{x(1-\alpha)}\right)^{\frac{1}{\gamma}}\right)^{\gamma}-1\right)\\
&\geq (1-\epsilon)^2\cdot(1-x_k)\cdot\left(\left(1+\left(\frac{\alpha(1-x_k)}{x(1-\alpha)}\right)^{\frac{1}{\gamma}}\right)^{\gamma}-1\right),
\end{aligned}\tag{EC.23}$$

where we use the fact that $(1-x)\geq(1-\epsilon)\cdot(1-x_k)$ and the super-homogeneity of the function $z\mapsto(1+\theta\cdot z^{\frac{1}{\gamma}})^{\gamma}-1$ for every $\theta\geq 0$. Consequently, we have

$$\begin{aligned}
f_\epsilon(x) &\leq w_{d_k^*-1}\\
&\leq 1-\max_{\alpha\in(0,1)}\max\left\{0\ ,\ \frac{(\kappa'-3\epsilon)\cdot\alpha-\int_0^{x_k} f_\epsilon(u)du}{(1-x_k)\cdot x_k\cdot((1+(\frac{\alpha(1-x_k)}{x_k\cdot(1-\alpha)})^{\frac{1}{\gamma}})^{\gamma}-1)}\right\}\\
&\leq 1-(1-\epsilon)^2\cdot\max_{\alpha\in(0,1)}\max\left\{0\ ,\ \frac{(\kappa'-3\epsilon)\cdot\alpha/x_k-(\int_0^{x_k} f_\epsilon(u)du)/x_k}{(1-x)\cdot((1+(\frac{\alpha(1-x)}{x_k\cdot(1-\alpha)})^{\frac{1}{\gamma}})^{\gamma}-1)}\right\}\\
&= 1-(1-\epsilon)^2\cdot\max_{u\in(0,1/x_k)}\max\left\{0\ ,\ \frac{(\kappa'-3\epsilon)\cdot u-(\int_0^{x_k} f_\epsilon(u)du)/x_k}{(1-x)\cdot((1+(\frac{u(1-x)}{(1-u\cdot x_k)})^{\frac{1}{\gamma}})^{\gamma}-1)}\right\}\\
&\leq 1-(1-\epsilon)^2\cdot\max_{u\in(0,1/x)}\max\left\{0\ ,\ \frac{(\kappa'-3\epsilon)\cdot u-(\int_0^{x} f_\epsilon(u)du)/x}{(1-x)\cdot((1+(\frac{u(1-x)}{(1-u\cdot x)})^{\frac{1}{\gamma}})^{\gamma}-1)}\right\}\\
&= 1-(1-\epsilon)^2\cdot\max_{\alpha\in(0,1)}\max\left\{0\ ,\ \frac{(\kappa'-3\epsilon)\cdot\alpha/x-(\int_0^{x} f_\epsilon(u)du)/x}{(1-x)\cdot((1+(\frac{\alpha(1-x)}{x(1-\alpha)})^{\frac{1}{\gamma}})^{\gamma}-1)}\right\}\\
&\leq 1-\max_{\alpha\in(0,1)}\max\left\{0\ ,\ \frac{(\kappa'-5\epsilon)\cdot\alpha/x-(\int_0^{x} f_\epsilon(u)du)/x}{(1-x)\cdot((1+(\frac{\alpha(1-x)}{x(1-\alpha)})^{\frac{1}{\gamma}})^{\gamma}-1)}\right\},
\end{aligned}$$

where the second inequality immediately follows from (EC.22). The next inequality is a direct consequence of inequality (EC.23). The fourth inequality holds since $x\geq x_k$ and the function $f_\epsilon(\cdot)$ is non-decreasing. □

□

EC.4. An upper bound on the competitiveness of online algorithms under MNL preferences

Proposition EC.3. *No algorithm, deterministic or randomized, that does not know the sequence of customer types in advance can obtain a competitive ratio better than* 0.8074 *for the online two-sided assortment problem when suppliers have MNL preferences, even under the* ϵ*-small probability assumption.*

To establish Proposition EC.3 we construct a family of instances such that no online algorithm can obtain a competitive ratio better than 0.8074. Our instances follow the same structure as the "bad" instances in online matching and online one-sided assortment. In what follows, and to ease exposition, we will follow the steps in the analogous proof in Golrezaei et al. (2014).

The family of instances. To establish Proposition EC.3, we construct the following family of instances. Consider a setting with n suppliers. Let T be the time horizon; for simplicity, we think of T as a very large number (that would tend to infinity) and a multiple of n. Suppliers' choice models are symmetric and follow the MNL model with $q_{it} = q = n/T$ for every $i \in \mathcal{S}$ and $t \in [T]$. Observe that, for every $\epsilon > 0$, we can choose $T_0 = T_0(\epsilon)$ such that for all $T \geq T_0$, the ϵ-small probability assumption is satisfied.

Customers are divided into $2^n - 1$ types. Each type corresponds to a set $\Theta \neq \emptyset$ that contains the suppliers that a customer of that type equally likes.

The *arrival process* is defined as follows: customers arrive in n phases of equal length, that is, the number of customers in each phase is T/n . All the customers in each phase have the same type. We denote the type of the customers in phase j by Θ_j . We have $\Theta_1 = \{1, 2, \dots, n\}$. For $j = 2, \dots, n$, we have $\Theta_j = \Theta_{j-1} \backslash \theta_{j-1}$, where θ_{j-1} is a supplier chosen from Θ_{j-1} uniformly at random. In other words, the set of suppliers that a customer in phase j likes is equal to the set of suppliers liked by customers in phase $j-1$ minus one of those suppliers; i.e., customers in phase j randomly lose interest in one of the suppliers of interest in phase $j-1$. Therefore there are $n!$ sequences of customer arrivals, each with equal probability.

Towards the desired upper bound. Having described the arrival process, we now construct an upper bound on the competitive ratio of online algorithms.

Claim EC.4. *The number of expected matches achieved by the optimal clairvoyant algorithm under the arrival process defined above is at least* $n\frac{T/n \cdot n/T}{1+T/n \cdot n/T} = \frac{n}{2}$.

Proof. Observe that an online clairvoyant algorithm that shows $\{\theta_1\}$ to all customers in phase 1, $\{\theta_2\}$ to all customers in phase 2, and so on, achieves, in expectation, $\sum_{i=1}^{n} \frac{T/n \cdot q}{1+T/n \cdot q} = n\frac{T/n \cdot n/T}{1+T/n \cdot n/T} = n\frac{1}{2}$ matches. Therefore, the optimal clairvoyant algorithm achieves, in expectation, at least that many matches.[7] $\square$

Next, we establish the following lemma.

Lemma EC.8. *Consider the arrival process specified above. Then, an algorithm that offers to each customer the assortment consisting of all suppliers that (i) are of interest to her and (ii) have*

[7] In fact, one can show that the optimal clairvoyant algorithm is the one just described. While the proof is not complicated, we decided to skip it in the interest of space and because the weaker claim that lower bounds the number of expected matches by the optimal clairvoyant algorithm suffices to establish our results.

the lowest probability of matching given the requests they have received so far, is optimal among all deterministic algorithms. Moreover, there exists a sequence of customer types such that the expected number of matches achieved by the algorithm is at most $\sum_{i=1}^{n} \frac{\sum_{j=1}^{i} \frac{1}{n-j+1}}{1+\sum_{j=1}^{i} \frac{1}{n-j+1}}$.

Proof. The first part of the claim, namely that the algorithm that offers to each customer the assortment consisting of all suppliers with the lowest match probability that are of interest to her is optimal among all *deterministic* algorithms, follows by applying standard (albeit lengthy) arguments. For the sake of brevity, we refer the reader to the proof of Lemma 6 in Golrezaei et al. (2014) that establishes the same argument for their setting. The only minor modification is that our marginal rewards exhibit diminishing returns; however, this is not a problem as the main arguments in the proof rely on monotonicity and symmetry, two properties that our rewards (match probabilities) satisfy.

Hence, under the algorithm described above, in each phase, each remaining supplier that is of interest to the customers in that phase (which are all of the same type) receives an equal number of the customer requests. In phase j, this is equivalent to $\frac{T}{n}\frac{1}{n-j+1}$ requests per supplier that remains in that phase, as there are $n-j+1$ suppliers of interest in that phase. Therefore, supplier θ_1, i.e., the supplier that won't be of interest to the customers arriving in phase two and onward, will obtain $\frac{T}{n}\frac{1}{n}$ requests. Similarly, supplier θ_2 will obtain $\frac{T}{n}\frac{1}{n}+\frac{T}{n}\frac{1}{n-1}$ requests, and so on. Therefore, the expected number of matches achieved by the algorithm is at most $\sum_{i=1}^{n} \frac{\sum_{j=1}^{i} \frac{1}{n-j+1}}{1+\sum_{j=1}^{i} \frac{1}{n-j+1}}$, as desired. □

Combining Claim EC.4 and Lemma EC.8 we obtain the following corollary.

Corollary EC.1. *Consider the arrival process specified above. No deterministic algorithm that knows that the sequence of arrivals is as specified above can achieve a competitive ratio greater than*

$$\frac{2 \cdot \sum_{i=1}^{n} \frac{\sum_{j=1}^{i} \frac{1}{n-j+1}}{1+\sum_{j=1}^{i} \frac{1}{n-j+1}}}{n}.$$

Moreover, we have that $\lim_{n\to\infty} \frac{2 \cdot \sum_{i=1}^{n} \frac{\sum_{j=1}^{i} \frac{1}{n-j+1}}{1+\sum_{j=1}^{i} \frac{1}{n-j+1}}}{n} < 0.8074$.

Finally, by applying Yao's lemma (Yao 1977) to Corollary EC.1, we obtain the result in Proposition EC.3.

EC.5. Model robustness

In this section, we examine the robustness of our model when we consider relevant extensions to online labor marketplaces. In online marketplaces, it is not uncommon that the same worker simultaneously desires to be matched with multiple jobs/tasks. As discussed in Section 2, our modeling approach and algorithmic results can be directly leveraged when customers choose to send multiple matching requests; our only requirement, stated by Assumption 1, is that the single-stage assortment optimization problem needs to be computationally tractable. Nonetheless, one implicit assumption of our model is that, if the firm agrees to match in return, the worker proceeds with the job in question. In reality, this assumption is reasonable when the worker has sufficient capacity to fulfil all the tasks he is matched with, or if the work is completed by the time the match request is sent. For example, the latter assumption holds in contest design marketplaces such as DesignCrowd, where the workers' match request often corresponds to a full design submission, and there is no work to be done after the firms reveal which submissions are rewarded. Thus, for these settings, our approach and results are directly applicable.

In what follows, we examine a variant of our online matching problem for practical settings in which the above assumption is violated. Our modeling approach remains identical with the sole exception that workers can fulfil a single match, even if they submit multiple match requests. This setting is mostly relevant to online platforms for full-time employment, rather than part-time freelancing. The crucial difference with our model is that the matching process often involves multiple rounds where each side of the market sequentially expresses its preferences, to resolve the conflicts between match requests, as defined below. This type of mechanism is rarely studied in the context of online matching markets and, more broadly, when designing online algorithms, due to the difficulty of modeling repeated interactions between participants. Nonetheless, our goal is to quantify the degree to which our model gives a good approximation of the optimal expected number of matches in these types of markets. Consequently, our main objective is to measure the expected number of *conflicts*; i.e., a conflict occurs whenever two distinct match requests from the same customer are simultaneously accepted by suppliers. The resolution of these conflicts is overlooked by our model. We empirically measure the expected number of conflicts in synthetic instances, which may reflect realistic market regimes. In Section EC.5.1, we introduce our random generative model for these instances. In Section EC.5.2, we report our numerical results. We find that, in realistic market regimes, the number of conflicts is relatively small compared to the total expected number of matches.

EC.5.1. Synthetic instances

We generate random instances of the online two-sided assortment problem in which customers send match requests to multiple suppliers. This aspect is captured by a multi-purchase choice model. We implement our preference-aware balancing algorithm for the MNL model, with a numerical approximation of the discount function $f_{1,\kappa^*(1)}$. In what follows, we describe the instance parameters we use to conduct our simulations in Section EC.5.2:

Arrival process. The set $\mathcal{S} = [1, 100]$ is comprised of 100 suppliers and there are T arriving customers. These parameters are set to reflect that the number of freelancers in the market is typically much larger than the number of open jobs listed at any point in time.[8] We distinguish between two customer types $\mathcal{Z} = \{z_g, z_b\}$, which are respectively said to be good and bad. At each time period $t \in [T]$, the t-th arriving customer is good with probability θ, $\Pr[z_t = z_g] = \theta$, independently from the history. We vary the proportion of good customers in the market by picking $\theta \in \{5\%, 10\%, 15\%, 20\%\}$.

Customers' choice models. Contrary to the standard version of our model, customers now send multiple match requests. Reflecting this phenomenon, we specify the Bundle-MNL choice model, which corresponds to a natural generalization of the MNL model in multi-purchase settings (Tulabandhula et al. 2020). More specifically, we assume that each customer requests matches with κ distinct suppliers picked uniformly at random over $\mathcal{S}$. That is, for every customer-type $z \in \mathcal{Z}$ and an assortment $A \subseteq \mathcal{S}$, we assume that each bundle $B \subseteq A$ comprised of exactly κ suppliers is picked with a uniform probability $\phi^z_B(A) = (1 + \binom{|A|}{\kappa})^{-1}$, and the empty bundle $B = \emptyset$ is picked with the residual probability $\phi^z_B(A) = (1 + \binom{|A|}{\kappa})^{-1}$. Here, the parameter κ is varied in the set $\kappa \in \{1, 3, 5, 10, 15, 20\}$.

Suppliers' choice models. Finally, it remains to specify the suppliers' aggregate match functions $w_i(\cdot)$. We adopt the following MNL choice model: $w_i(C) = \frac{0.5C_g + 0.02C_b}{1 + 0.5C_g + 0.02C_b}$ with respect to the set of requests $C \in \mathbb{N}^{\mathcal{Z}}$. We assume that all suppliers leave at the end of the time horizon T. Given a policy π, each supplier $i \in \mathcal{S}$ picks one random customer $\zeta_i \in C^{T+1,\pi}$ with probability $\Pr[\zeta_i = t] = \frac{0.5 \cdot \mathbb{I}[z_t = z_g] + 0.02 \cdot \mathbb{I}[z_t = z_b]}{1 + 0.5 \cdot C_g^{T+1,\pi} + 0.02 \cdot C_b^{T+1,\pi}}$ for all $t \in C^{T+1,\pi}$, and remains unmatched with the remaining probability. (By a slight abuse of notation, we treat $C^{T+1,\pi}$ both as a count vector of customer types and a subset of arriving customers.)

The important quantity we wish to measure is the expected number of conflicts, defined as

$$\mathbb{E}\left[\sum_{t \in [T]} \max\left\{0, \sum_{i \in \mathcal{S}} \mathbb{I}\left[\zeta_i = t\right] - 1\right\}\right] .$$

If the expected number of conflicts is small, our model remains relevant, even if the matching process is incomplete after a single round. Intuitively, this metric is affected by two key dimensions of the problem: (i) the number of match requests κ made by customers, and (ii) the proportion θ of good versus bad customers in the market. For example, no conflict arises if $\kappa = 1$ as we assume in the remainder of the paper. Good customers attract more match requests from suppliers; this vertical differentiation means that match requests from suppliers are more concentrated on top customers,

[8] For example, in September 2021, DesignCrowd reported 999,462 freelance designers for 410,376 completed projects *since the platform's inception.*

which exacerbates the market congestion. Hence, we choose to vary both parameters κ and θ while leaving all other parameters constant. Additionally, we vary the number of arriving customers $T \in \{200, 400, 600, 800\}$. We also ran simulations varying other model parameters including the MNL weights of good customers and bad customers, and the weight of the outside option in the customers' choice models. The results are qualitatively similar, and thus omitted for brevity.

EC.5.2. Results

Our numerical results are reported in Table EC.1. The first metric "# matches" gives the average number of matches, "# conflicts" counts the average number of conflicts, and "% conflicts" expresses this quantity as a percentage of "# matches." Overall, we find that the number of conflicts ranges from 0% to 14% in proportion to the number of matches. As one might expect, if customers send a larger number κ of match requests, or if the proportion θ of good customers increases, then there are more conflicts in the market. The first parameter has a greater influence on the number of conflicts, but the relative marginal increase tends to diminish for $\kappa \geq 10$. In Table EC.2, we report the same metrics but this time we vary the number of customers (i.e., the customer/supplier imbalance) and the number of applications, while fixing the value of θ to 5%.

Based on the numerical findings in Table EC.1, we get a validation that the number of conflicts might be a relatively small effect in realistic regimes of parameters for online labor marketplaces. Overall, the existence of conflicts is a negligible concern when $\kappa \leq 5$, but it perhaps becomes significant when $\kappa \geq 10$. Nevertheless, the latter regime means that each worker submits a match request to 10% of the total number of listings available in the market. To interpret Table EC.1 in more concrete terms, in a large unnamed online labor market,[9] there are on average 400 new jobs posted every day. Each job receives 20 applications on average. The job fill rate is between 40% and 45%. There are 2.5 million freelancers listed but, naturally, we do not expect all of them to be active or to be compatible with every job. Observe that the combination of parameters in Table EC.1 that best fits this description (with the market size scaled down to 100 jobs) corresponds to $(\theta, \kappa) = (5\%, 3)$, for which the proportion of conflicts is less than 4%.

[9] All the information in this paragraph was obtained from the working paper by Horton and Vasserman (2021).

Table EC.1 Expected number of conflicts and matches as a function of θ and κ, for a fixed $T = 400$.

Parameters		Metrics		
θ	κ	# matches	# conflicts	% conflicts
5%	1	17.79	0	0%
5%	3	40.3	1.6	3.97%
5%	5	51.78	3.05	5.89%
5%	10	69.41	6.39	9.2%
5%	15	77.41	8.18	10.56%
5%	20	82.42	9.37	11.37%
10%	1	19.78	0	0%
10%	3	43.37	2.21	5.09%
10%	5	55.64	4.2	7.55%
10%	10	71.5	7.33	10.26%
10%	15	78.88	9.27	11.75%
10%	20	79.06	10.67	13.49%
15%	1	22.18	0	0%
15%	3	47.36	2.77	5.86%
15%	5	58.86	4.81	8.17%
15%	10	75.03	8.81	11.74%
15%	15	81.46	10.34	12.7%
15%	20	85.67	12	13.63%
20%	1	23	0	0%
20%	3	46.76	2.66	5.7%
20%	5	61.52	5.61	9.12%
20%	10	75.92	9.15	12.05%
20%	15	82.41	10.63	12.89%
20%	20	86.14	11.93	13.85%

Table EC.2 Expected number of conflicts and matches as a function of number of customers T and number of applications κ, for a fixed $\theta = 5\%$.

Parameters		Metrics		
T	κ	# matches	# conflicts	% conflicts
200	1	0.00	9.32	0.00%
200	3	1.26	24.28	5.19%
200	5	3.11	54.42	5.72%
200	10	7.79	54.42	14.32%
200	15	9.34	61.72	15.14%
200	20	12.91	70.24	18.39%
400	1	0.00	17.80	0.00%
400	3	1.60	40.30	3.97%
400	5	3.05	51.79	5.89%
400	10	6.39	69.42	9.21%
400	15	8.18	77.41	10.57%
400	20	9.38	82.42	11.38%
600	1	0.00	25.99	0.00%
600	5	3.11	62.68	4.96%
600	10	5.41	76.13	7.11%
600	15	6.60	83.86	7.87%
600	20	7.73	87.71	8.81%
800	1	0.00	30.48	0.00%
800	5	3.19	69.97	4.56%
800	10	4.82	82.33	5.85%
800	15	5.42	87.16	6.22%
800	20	5.85	90.31	6.48%