Two-Sided Assortment Optimization with Simultaneous and Sequential Matches

400

We study the two-sided assortment problem recently introduced by [Rios et al., 2022]. A platform must choose an assortment of profiles for each user in each period. Users can either like/dislike as many profiles as they want, and a match occurs if two users see and like each other, potentially in different periods. The platform's goal is to maximize the expected number of matches generated. We show that the problem is NP-hard, even if we consider two periods and only one-directional sequential matches (i.e., users on one side of the market reach out first, and then the ones on the other side respond). Given the complexity of the problem, we focus on the case with two periods and provide algorithms and performance guarantees for different variants of the problem. First, we show that if matches can only be formed sequentially (i.e., users cannot see each other in the same period), there is an approximation guarantee of $1/3$, which can be improved to $1 - 1/e$ if we restrict to one-directional matches. Second, we show that when we allow simultaneous matches in the first period, there is an approximation guarantee of $1/4$, which can be improved to $1/2$ if sequential matches are one-directional. Finally, we show that the general problem with simultaneous and sequential matches is more challenging, as the objective function is no longer submodular. Nevertheless, we provide performance guarantees for some specific cases, and we use data from our industry partner (a dating app in the US) to numerically show that the loss for not considering simultaneous matches in the second period when making first-period decisions is negligible.

Manuscript submitted for review to the 24nd ACM Conference on Economics & Computation (EC'23).

1 INTRODUCTION

A common feature of many marketplaces is their two-sided design, in which (1) both sides of the market have preferences, (2) both sides of the market can initiate an interaction with the other side, and (3) users must mutually agree to generate a match. Among many others, examples include freelancing platforms such as TaskRabbit or UpWork, ride-sharing apps such as Blablacar, accommodation companies such as Airbnb, and dating platforms such as Hinge and Bumble. In many of these platforms, the path towards a match starts with the platform eliciting preferences on both sides of the market. For instance, Airbnb requires guests to report the location and dates of their trips and add filters regarding the price and type of place. Similarly, Airbnb allows hosts to set the price and the minimum number of nights, as well as other preferences. After collecting this information, most of these platforms display a limited set of alternatives that users can screen before interacting with the other side of the market. Depending on the setting, this interaction can be one-directional, with one side of the market sending an initial request/message/like and the other side of the market either accepting or rejecting it; or it can be bi-directional, with both sides of the market being able to screen alternatives and move first. Airbnb and Blablacar are examples of the former, while Hinge and Upwork are examples of the latter. Finally, in all these platforms, a transaction takes place if, and only if, both sides of the market mutually accept/like each other.

As the previous discussion illustrates, one of the primary roles of these platforms is to select the subset of alternatives—the *assortment*—to display, considering the preferences and characteristics of the users on both sides of the market. We refer to this problem as a *two-sided assortment problem*. This problem is similar to the classic (one-sided) *assortment optimization problem*, in which a retailer must decide the subset of products to display to maximize the expected revenue obtained from a series of customers. However, the two-sidedness of the problem imposes some challenges. For instance, both users must mutually see and agree to each other. Hence, it is uncertain whether a match will occur, even after the first user likes the other. Moreover, the platform must carefully balance the trade-off between relevance– i.e., showing options that are likely to generate a match– and congestion, as the most popular users may get more requests than they can handle.

The present paper investigates how platforms should make this decision when both sides of the market can initiate the path towards a match. More specifically, the goal of this paper is twofold: (1) to present a two-sided assortment optimization model with bidirectional interactions, and (2) to provide algorithms and performance guarantees for this problem.

1.1 Our Contribution

We first describe a stylized model of a two-sided market mediated by a platform. The platform must choose which subset of profiles (i.e., the assortment) to show to each user. Then, users decide whether to like or dislike each option in their assortments. We extend previous work by allowing both sides of the market to start an interaction. As a result, matches can happen either (1) simultaneously, where the two users see and like each other in the same period, or (2) sequentially, where users see and like each other during different periods. Although our model is general enough to capture different settings, we focus on a dating platform to fix ideas and facilitate the exposition.[1]

We show that the problem faced by the platform is NP-hard, even if we consider two periods and only allow for one-directional sequential matches. Thus, we focus on studying the two-period version of the problem to facilitate the theoretical analysis. This case is relevant in practice since one-lookahead policies perform well, so this assumption is without major practical loss (see [Rios et al., 2022]). Moreover, the two-period version of the problem is theoretically interesting and

[1]This work results from an ongoing collaboration with a major dating app in the US, which provided us with real data to test our proposed algorithms. We keep the name of the app undisclosed as part of our NDA.

Table 1. Summary of results

		Sequential Matches	
		One directional	Two directional
Simultaneous Matches	None	$1 - e$	$1/3$
	First period	$1/2$	$1/4$
	Both periods	$\approx 1/4e$	-

Table 2. The result in the last row assumes a regime with small probabilities on one side.

challenging, as many common approaches (such as greedy, perfect matching, and any non-adaptive deterministic approach) can perform arbitrarily poorly in this setting. Despite these negative results, we provide constant factor approximations for different variants of the problem, as summarized in Table 1.

The first variant of the problem is when we only allow for sequential matches, i.e., no two users see each other in the same period. This case is particularly interesting because many two-sided matching markets (e.g., Airbnb and Blablacar) operate in this way, motivating recent work Aouad and Saban [2022], Ashlagi et al. [2022], Torrico et al. [2021]. In this case, we provide an algorithm with a performance guarantee of $1 - 1/e$. Moreover, if we relax the assumption that matches are initiated by only one side and allow both sides to reach out (as in platforms such as UpWork), we provide a performance guarantee of 1/3. Both results rely on submodular maximization techniques.

The second variant of the problem we analyze is when we allow for simultaneous matches only in the first period. This case captures settings such as dating apps, where users can get matches sequentially (as previously discussed) or simultaneously in the first period, which happens if both users see and like each other in that period. In this case, we provide a general performance guarantee of $1/4$, and we can improve it to $1/2$ if we restrict sequential matches to be initiated by only one side. This last setting is especially relevant as some major dating apps (e.g., Bumble) only allow one side of the market to reach out first, as this may improve social welfare [Kanoria and Saban, 2017].

Finally, the last variant of the problem is when we allow sequential and simultaneous matches in both periods. This problem is significantly more challenging, as it can no longer be modeled as a submodular maximization problem. Nevertheless, we show an approximation guarantee for the case of a large market with simultaneous matches in both periods and one-directional sequential matches initiated by the more selective side (i.e., with small like probabilities). Our one-directional assumption aligns with the results in [Kanoria and Saban, 2017, Shi, 2022a, Torrico et al., 2021] that show that market congestion can be reduced when the more selective side (or the one whose preferences are harder to describe) initiates the matchmaking process. To prove this guarantee, we show that the gains from allowing simultaneous matches in the second period are relatively small compared to when we allow simultaneous matches in the first period only. We empirically confirm this finding using our industry partner's data. Specifically, we find that considering simultaneous matches in the second period (while deciding the assortment to offer in the first period) does not play a significant role. Hence, we conjecture that the performance guarantees obtained for the case with simultaneous matches in the first period should translate to the more general case.

Organization of the paper. The paper is organized as follows. In Section 2, we discuss the most related literature. In Section 3, we introduce our model. In Section 4, we show that some natural approaches can perform arbitrarily poorly. In Section 5, we theoretically analyze the problem and provide our performance guarantees. In Section 6, we numerically compare our proposed algorithms with other relevant benchmarks. Finally, in Section 7, we conclude.

2 RELATED LITERATURE

Our paper is related to several strands of the literature. First, we contribute to the literature on assortment optimization. Most of this literature focuses on one-sided settings, where a retailer must choose the assortment of products to show in order to maximize the expected revenue obtained from a sequence of customers. This model, whose general version was introduced in [Talluri and van Ryzin, 2004], has been extended to include capacity constraints [Rusmevichientong et al., 2010], different choice models [Blanchet et al., 2016, Davis et al., 2014, Farias et al., 2013, Rusmevichientong et al., 2014], search [Wang and Sahin, 2018], learning [Caro and Gallien, 2007, Rusmevichientong et al., 2010, Sauré and Zeevi, 2013], personalized assortments [Berbeglia and Joret, 2015, Golrezaei et al., 2014], and also to tackle other problems such as priority-based allocations [Shi, 2022b]. We refer to [Kök et al., 2015] for an extensive review of the current state of the assortment planning literature in one-sided settings.

Over the last couple of years, a new strand of assortment optimization literature on two-sided markets has emerged. [Ashlagi et al., 2022] introduce a model where each customer chooses, simultaneously and independently, to either contact a supplier from the assortment offered to them or to remain unmatched. Then, each supplier can either form a match with one of the customers who contacted them in the first place or remain unmatched. The platform's goal is to select the assortment of suppliers to show to each customer to maximize the expected number of matches. The authors show that the problem is NP-hard, and they provide an algorithm that achieves a constant factor approximation. [Torrico et al., 2021] study the same problem and significantly improve the approximation factor obtained by [Ashlagi et al., 2022]. Moreover, [Torrico et al., 2021] provide the first performance guarantee to the problem with cardinality constraints. [Aouad and Saban, 2022] analyze the online version of the problem. Specifically, the platform must choose the assortment of suppliers to show to each arriving customer, who decide whether to contact one of the suppliers or to remain unmatched. Then, after some time, each supplier can choose to match with at most one of the customers that chose them. The authors show that when suppliers do not accept/reject requests immediately, then a simple greedy policy achieves a 1/2-factor approximation. [Aouad and Saban, 2022] also propose balancing algorithms that perform relatively well under the Multinomial and Nested Logit models. Notice that all these papers analyze sequential two-sided matching markets, where only one side can start the path towards a match. Users on the other side only respond by deciding which user to match with among those who contacted them in the first place. Moreover, customers are limited to choosing only one supplier in their assortments.

Within the assortment optimization literature, the closest paper to ours is [Rios et al., 2022]. The authors introduce a finite horizon model where a platform chooses an assortment for each user (on both sides of the market). Users can like/dislike as many of the profiles in their assortment as they want, and a match is formed if both users like each other. The authors show that the problem is NP-hard, and they propose a family of algorithms that account for the negative effect that the matches obtained today may have on future periods. However, [Rios et al., 2022] provide no theoretical guarantees for the algorithms they propose.

The second stream of literature related to our paper is on matching platforms. Starting with the seminal work of [Rochet and Tirole, 2003], this literature has focused on participation, competition, and pricing, highlighting the role of cross-side externalities. In the dating context, [Kanoria and Saban, 2017] study how the search environment can impact users' welfare and the performance of the platform. They find that simple interventions, such as limiting what side of the market reaches out first or hiding quality information, can considerably improve the platform's outcomes.

[Halaburda et al., 2018] show that two platforms can successfully coexist charging different prices by limiting the set of options offered to their users. They show that, depending on their outside option, users must balance two effects when choosing a larger platform: (1) a choice effect, whereby users are more likely to find a partner that exceeds their outside option; and (2) a competition effect, whereby agents on the other side of the market are less likely to accept a request as they have more options available. All these models consider a stylized matching market, where users interact with the other side of the market and leave the platform upon getting a match.

3 MODEL

Consider a two-sided market mediated by a platform, with two sets of users I and J representing each *side* of the market in a bipartite graph. We use indices i for specific users in I, j for specific users in J, and ℓ for a generic users in $I \cup J$. We denote by $E = \{\{i, j\} : i \in I,\ j \in J\}$ the set of all possible undirected edges between I and J, and by $\vec{E} = \vec{E}_I \cup \vec{E}_J$ the set where $\vec{E}_I = \{(i, j) : i \in I,\ j \in J\}$ and $\vec{E}_J = \{(j, i) : i \in I,\ j \in J\}$, that is, the sets of all directed arcs between I and J. For a given arc $a \in \vec{E}$, let $\overline{a}$ be the inverted arc, that is, $\overline{a} = (\ell', \ell)$ when $a = (\ell, \ell')$.

Given users' reported preferences and their observable characteristics, each user $\ell \in I \cup J$ has initially a set of potential partners $\mathcal{P}_\ell^1$—or simply *potentials*—i.e., $\mathcal{P}_i^1 \subseteq J$ for each $i \in I$, and $\mathcal{P}_j^1 \subseteq I$ for each $j \in J$. In each period $t \in \{1, \ldots, T\}$, the platform selects a subset of potentials—an *assortment*—to show to each user. Formally, let $S_\ell^t \subseteq \mathcal{P}_\ell^t$ be the assortment offered to user $\ell \in I \cup J$ in period $t \in \{1, \ldots, T\}$, where $\mathcal{P}_\ell^t$ is the set of available potentials for user ℓ at the beginning of period t. To mimic our industry partner's practice, we assume that the maximum size of the assortments is fixed and equal to K_ℓ, i.e., $|S_\ell^t| \leq K_\ell$ for all $\ell \in I \cup J$ and $t \in \{1, \ldots, T\}$.

For each user ℓ and a profile in their assortment $\ell' \in S_\ell^t$, we denote by $\Phi_{\ell\ell'}^t$ the binary random variable that indicates whether ℓ likes ℓ' in period t, i.e., $\Phi_{\ell\ell'}^t = 1$ if ℓ likes ℓ' and 0 otherwise. In addition, let $\phi_{\ell\ell'}^t = \mathbb{P}(\Phi_{\ell\ell'}^t = 1)$ be the probability that ℓ likes ℓ' in period t. We assume that these probabilities are known, and that they are independent across users, periods and of the assortment that is shown. The former assumption is reasonable because users do not know whether other users liked them before making their evaluations. The latter assumption is for simplicity, since [Rios et al., 2022] show that the number of matches obtained in the recent past affects users' future like behavior. Note that $\phi_{\ell\ell'}^t$ is not necessarily symmetric, namely, $\phi_{\ell\ell'}^t$ can be different from $\phi_{\ell'\ell}^t$. Finally, let us denote $\beta_{\ell\ell'}^t = \phi_{\ell\ell'}^t \phi_{\ell'\ell}^t$.

Let $\mathcal{B}_\ell^t$ be the backlog of user $\ell \in I \cup J$ at the beginning of period t, i.e., the subset of users that have liked user ℓ before period t, but have not been shown to ℓ yet. Then, we can now formalize the set of potentials and the backlog of user ℓ in period t as:

$$\mathcal{P}_\ell^t = \mathcal{P}_\ell^{t-1} \setminus (S_\ell^{t-1} \cup R_\ell^{t-1}),$$
$$\mathcal{B}_\ell^t = \left(\mathcal{B}_\ell^{t-1} \cup A_\ell^{t-1}\right) \setminus S_\ell^{t-1},$$

where the sets $A_\ell^{t-1} = \left\{\ell' : \ell \in S_{\ell'}^{t-1} \text{ and } \Phi_{\ell'\ell}^{t-1} = 1\right\}$ and $R_\ell^{t-1} = \{\ell' : \ell \in S_{\ell'}^{t-1} \text{ and } \Phi_{\ell'\ell}^{t-1} = 0\}$ correspond to the sets of users that liked and disliked ℓ in period $t - 1$, respectively.

A match between users ℓ and ℓ' occurs if both users see and like each other. Let $\mu_{\ell\ell'}^t = 1$ if a match between users ℓ and ℓ' happens in period t, and let $\mu_{\ell\ell'}^t = 0$ otherwise. Then, we know that $\mu_{\ell\ell'}^t = 1$ if and only if one of next two events holds:

(1) Simultaneous match: users see and like each other in the same period, i.e., $\ell \in S_{\ell'}^t$, $\ell' \in S_\ell^t$, and $\Phi_{\ell\ell'}^t = \Phi_{\ell'\ell}^t = 1$ for some period $t \in \{1, \ldots, T\}$.

(2) Sequential match: users see and like each other in different periods, i.e., for some period $t \in \{1, \ldots, T\}$,

$$\left\{\Phi^t_{\ell\ell'} = 1,\ \ell' \in \mathcal{B}^t_\ell\right\} \quad \text{or} \quad \left\{\Phi^t_{\ell'\ell} = 1,\ \ell \in \mathcal{B}^t_{\ell'}\right\}.$$

Notice that these conditions cannot hold simultaneously, as users are shown to each other at most once, and thus we cannot simultaneously have that $\Phi^t_{\ell'\ell} = 1$ and $\ell' \in \mathcal{B}^t_\ell$. The distinction between simultaneous and sequential matches will play a relevant role in the remainder of this paper, as the complexity of the problem increases considerably if we allow for simultaneous matches.

The platform aims to find a dynamic policy to select a feasible assortment for each user in each period to maximize the expected number of matches throughout the entire horizon. A policy π for this problem prescribes a sequence of assortments $\vec{S}^{t,\pi} = \left\{S^{t,\pi}_\ell\right\}$ for $t = 1, \ldots, T$ that depends on the initial sets of potentials, the history of assortments shown, and the realized like/dislike decisions.[2]

Problem 1. *The two-sided assortment optimization problem is the following:*

$$\max_{\pi \in \Pi} \mathbb{E}\left[\sum_{t=1}^{T} \sum_{i \in I} \sum_{j \in J} \mu^{t,\pi}_{i,j}\right]$$

where Π is the set of all admissible policies and the expectation is over the probabilistic choices made by the users and the possibly random selections made by the policy (if π not deterministic).

Note that Problem 1 can be formulated as a dynamic programming (DP) where the state of the system is fully characterized by the sets of potentials and backlogs; see Appendix A.1 for a general version of this DP. Our first result, formalized in Proposition 3.1, states the complexity of Problem 1. We present its proof in Appendix A.2.

Proposition 3.1. *Problem 1 is NP-hard, even if we consider $T = 2$ and only allow one-directional sequential matches.*

Given this complexity result, we study the two-period version of the problem in the remainder of the paper. This assumption is without major loss, as Rios et al. [2022] show that one-lookahead policies perform well in practice.[3] Hence, we can simplify the notation and denote by $\mathcal{B}_\ell = \mathcal{B}^2_\ell$ the subset of users that liked ℓ in the first period, i.e.,

$$\mathcal{B}_\ell = \{\ell' : \ell \in S^1_{\ell'},\ \Phi^1_{\ell'\ell} = 1 \text{ and } \ell' \notin S^1_\ell\}.$$

Note that the backlog in the second period is a random subset, since it is defined over the random realizations of the like events obtained from the first period.

Given that we are studying a two-period model, the expected number of matches obtained in the second period is a result of the assortments shown and the realized likes/dislikes in the first period. Hence, if we have a function that returns the expected number of matches in the second period given the assortments and decisions in the first period, we only need to consider as decision variables the assortments offered in the first period. The decisions in the first period correspond to a family of *feasible* assortments $\mathrm{S}^1 = (S^1_\ell)_{\ell \in I \cup J}$ such that, for given budgets K_ℓ, $S^1_\ell \subseteq \mathcal{P}_\ell$ and $|S^1_\ell| \leq K_\ell$ for all $\ell \in I \cup J$.

[2]Our model can be easily adapted to the setting where there is a utility value per match; the properties and insights that we prove in this work carry over.
[3]A one-lookahead policy is a policy that, in every period t, optimizes over the current and the next period, i.e., it considers as horizon $\{t, t+1\}$.

Given a family of feasible assortments S, let $\mathcal{M}_{\mathbf{S}}$ be the random variable that indicates the total number of matches achieved when S is shown in the first period. Now, we formally define our main optimization problem.

Problem 2. *The two-period version of Problem 1 is the following:*

$$\max\left\{\mathbb{E}\left[\mathcal{M}_{\mathbf{S}^1}\right]:\ \mathbf{S}^1=\left\{S^1_\ell\right\}_{\ell\in I\cup J},\ S^1_\ell\subseteq\mathcal{P}_\ell,\ |S^1_\ell|\le K_\ell,\ \textit{for every } \ell\in I\cup J\right\},$$

where the expectation is over the probabilistic choices made by the users.

In the remainder of the paper, we focus on deriving performance guarantees for different variants of the problem. The benchmark we will use to compare our algorithms is the two-period DP formulation and its optimal solution, denoted by OPT. We say that an algorithm is an α-*approximation* if it implements a feasible solution and the expected number of matches is at least an α fraction of OPT.

Notation. Hereafter, we write vectors in bold and their components in italic, e.g., $\mathbf{S}=(S_i)_i$ and $\mathbf{x}=(x_a)_a$.

4 NO GUARANTEES WITH COMMON APPROACHES

In this section, we show that some common approaches used in the literature can perform arbitrarily bad in our setting, i.e., their performance guarantees in the worst-case are asymptotically close to zero. Without loss of generality, consider for this section the problem with $K_\ell = 1$ for all $\ell\in I\cup J$. Then, the problem consists of deciding which profile to show to each user in each period to maximize the number of matches, where a match is formed either simultaneously or sequentially (from the backlog). There are many natural algorithms that could solve the problem. For instance:

(1) Local Greedy policy: Assign each user to the profile that maximizes the probability of a match.
(2) Sequential perfect match policy: Sequentially find a maximum weight perfect matching, where the weight of each edge is the probability of having a match between the users.
(3) Non-Adaptive policies: Any policy in which the second-period decisions do not take into account the realization of the backlogs resulting from the first-period decisions.

All these alternatives policies have a worst-case approximation guarantee equal to zero, as we formalize in Propositions 4.1, 4.2 and 4.3. The proofs for these propositions can be found in Appendix B.

Proposition 4.1. *The worst case approximation guarantee for the local greedy policy is zero.*

Proposition 4.2. *The worst case approximation guarantee for the sequential perfect match policy is zero.*

Proposition 4.3. *The worst case approximation guarantee for the optimal deterministic non-adaptive policy is zero.*

5 MAIN ALGORITHMS

As shown in the previous section, common approaches might fail to provide good approximation guarantees. With the goal of designing stronger algorithms for Problem 2, we use techniques from submodular optimization. Depending on the possibility of allowing simultaneous matches, we split our analysis in two different cases that we detail in what follows.

Note that we can formulate Problem 2 as the platform choosing edges in $E\cup\vec{E}$ (recall that $\vec{E}=\vec{E}_I\cup\vec{E}_J$), since this uniquely determines the corresponding assortments. For instance, if

$A \subseteq E \cup \vec{E}_I \cup \vec{E}_J$, then $S_\ell = \{\ell' : (\ell, \ell') \in A \text{ or } \{\ell, \ell'\} \in A\}$. Depending on whether we consider simultaneous or sequential matches, we denote by $\mathcal{E}$ the space of available edges (for the platform) and arcs between I and J.

In the remainder of this section, we use the following equivalent definitions of monotonicity and submodularity for a set function (with vector representation).

Definition 5.1 (Monotonicity and Submodularity). Consider a ground set of elements $\mathcal{E}$. A non-negative set function $g : \{0,1\}^{\mathcal{E}} \to \mathbb{R}_+$ is *submodular* if for every $\mathbf{x} \in \{0,1\}^{\mathcal{E}}$ and for every $u, v \in \mathcal{E}$ such that $x_u = x_v = 0$, we have $g(\mathbf{x} + \mathbb{1}_{\{u\}} + \mathbb{1}_{\{v\}}) - g(\mathbf{x} + \mathbb{1}_{\{v\}}) \leq g(\mathbf{x} + \mathbb{1}_{\{u\}}) - g(\mathbf{x})$, where $\mathbb{1}_{\{u\}} \in \{0,1\}^{\mathcal{E}}$ is the indicator vector with value 1 in component u and zero elsewhere. The function g is *monotone* if for every $\mathbf{x} \in \{0,1\}^{\mathcal{E}}$ and $u \in \mathcal{E}$ such that $x_u = 0$, we have $g(\mathbf{x} + \mathbb{1}_{\{u\}}) \geq g(\mathbf{x})$.

5.1 Sequential Matches Only

We consider the case where only sequential matches are allowed as a first step. Specifically, we assume that the platform chooses arcs in $\vec{E}$, i.e., $\mathcal{E} = \vec{E}$. Recall that $\vec{E}$ is the union of the arcs in $\vec{E}_I$ and $\vec{E}_J$. Depending on the realizations of the users' choices, the second-period assortments are those that maximize the final expected number of matches. This particular case is interesting for multiple reasons. First, as shown in [Kanoria and Saban, 2017], platforms can improve social welfare by restricting one side of the market to reach out first. Second, this particular case directly relates to the emerging literature of two-sided assortment optimization with sequential matches discussed in Section 2. Third, focusing on sequential matches simplifies the analysis and serves as a building block to study the more complex case with simultaneous matches. Our next result, formalized in Theorem 5.2, provides a performance guarantee for this case.

Theorem 5.2. *There exists a* $1/3$*-approximation algorithm for Problem 2 with no simultaneous shows.*

We prove Theorem 5.2 in two steps: (i) we provide an explicit reformulation of Problem 2 when no simultaneous shows are allowed, and (ii) we use appropriate submodular optimization techniques. To ease the exposition, in the remainder of this section, we avoid the subset notation for any assortment family S and we use its corresponding indicator variable $\mathbf{x} \in \{0,1\}^{\mathcal{E}}$ with $\mathcal{E} = \vec{E}$, where for any $a = (\ell, \ell') \in \mathcal{E}$ with $\ell \in I \cup J$ and $\ell' \in \mathcal{P}_\ell$, we have $x_a = 1$ if $\ell' \in S_\ell$ and zero otherwise.

First, we define the feasible region for the first-period decisions. Since we consider cardinality constraints and we do not allow for simultaneous matches, the feasible region can be characterized as:

$$P^1 = \left\{ \mathbf{x}^1 \in \{0,1\}^{\mathcal{E}} : \sum_{\ell' \in \mathcal{P}_\ell} x^1_{(\ell,\ell')} \leq K_\ell, \text{ for every } \ell \in I \cup J \right.$$
$$\left. x^1_a + x^1_{\bar{a}} \leq 1, \text{ for every } a \in \mathcal{E} \right\}. \tag{1}$$

Recall that $\bar{a}$ is arc a with the opposite orientation. The first family of constraints ensure our cardinality requirements, while the second family guarantees that there are no simultaneous matches, as ℓ and ℓ' cannot see each other in the same period (e.g., if $x^1_{(\ell,\ell')} = 1$, then $x^1_{(\ell',\ell)} = 0$).

Lemma 5.3. *Feasible region* P^1 *defined in* (1) *corresponds to the intersection of 2 partition matroids.*

The proof of this lemma can be found in Appendix C.2. Since we allow sequential matches to start from either I or J, note that the random family of backlogs $\mathcal{B} = (\mathcal{B}_\ell)_{\ell \in I \cup J}$ is such that $\mathcal{B}_\ell \subseteq \vec{E}_I$ when $\ell \in J$ and $\mathcal{B}_\ell \subseteq \vec{E}_J$ when $\ell \in I$. Let $B = (B_\ell)_{\ell \in I \cup J}$ be a realization of $\mathcal{B}$ and, in a slight abuse of notation, let $a \in \mathcal{B}$ represent that there exists ℓ such that $a \in \mathcal{B}_\ell$ (and similarly for $a \in B$). Given a first-period decision $\mathbf{x}^1 \in \{0,1\}^{\mathcal{E}}$, since users behave independently, the distribution of $\mathcal{B}$ is

$$\mathbb{P}_{\mathbf{x}^1}(\mathcal{B} = B) = \prod_{a \in B} \phi_a^1 x_a^1 \prod_{a \notin B} (1 - \phi_a^1 x_a^1). \tag{2}$$

We now focus on the second-period objective, which will be the maximum expected number of matches given the realized backlogs that result from the first period decisions. Specifically, given a fixed realization of the backlogs $\mathcal{B} = B$, let $f(B)$ be the expected number of matches, i.e.,

$$f(B) = \max_{\mathbf{x}^2 \in P_B^2} \left\{ \sum_{a \in \mathcal{E}} \phi_a^2 x_a^2 \right\}, \tag{3}$$

where the feasible region is given by

$$P_B^2 = \left\{ x^2 \in \{0,1\}^{\mathcal{E}} : \sum_{\ell' \in \mathcal{P}_\ell} x^2_{(\ell,\ell')} \leq K_\ell, \text{ for every } \ell \in I \cup J, x_a^2 \leq 1_{\{\overline{a} \in B\}}, \text{ for every } a \in \mathcal{E} \right\},$$

and $1_{\{\overline{a} \in B\}}$ is the indicator function that takes value one if $\overline{a} \in B$, and zero otherwise. Note that P_B^2 corresponds to the set of feasible assortment in which: (i) each assortment is of size at most K_ℓ; and (ii) each user $\ell \in I \cup J$ is included in the second-period assortment of $\ell' \in \mathcal{P}_\ell^1$ only if ℓ is in the backlog of ℓ', i.e., $1_{\{(\ell,\ell') \in B\}} = 1$.

Lemma 5.4. *For any backlog realization B, the feasible region P_B^2 corresponds to a partition matroid.*

The proof of Lemma 5.4 follows directly from the proof of Lemma 5.3. Then, given that P^2 is a matroid, we have the following result:

Corollary 5.5 (Proposition 3.1 in [Fisher et al., 1978]). *Function f defined in* (3) *is non-negative, monotone and submodular.*

Moreover, as it only requires to optimize a linear function over a partition matroid to evaluate f, we know that this can be efficiently done using a greedy algorithm. Finally, the objective for Problem 2 in the setting of no simultaneous shows can be written as

$$M_f(\mathbf{x}^1) = \sum_B f(B) \cdot \mathrm{P}_{\mathbf{x}^1}(\mathcal{B} = B), \tag{4}$$

where the sum is over all possible backlog realizations. Therefore, Problem 2 with no simultaneous shows can be written as

$$\max \left\{ M_f(\mathbf{x}^1) : \mathbf{x}^1 \in P^1 \right\}. \tag{5}$$

Since we can obtain the exact value of $f(B)$ for any realization of the backlogs B (and consequently, the corresponding second-period assortments), then an approximation algorithm for Problem (5) can be used to design its corresponding adaptive policy. Moreover, the guarantee of the approximation algorithm would imply a guarantee for its adaptive policy with respect to the DP formulation of the problem. This is because in Problem (5) we are optimizing the expected value over all possible backlogs. We formalize our dynamic policy in Algorithm 1.

Note that the approximation algorithm that we used in Algorithm 1 is the standard greedy algorithm for submodular maximization. The associated adaptive policy uses the solution of the greedy algorithm, and after observing the backlogs it optimizes function f defined in (3) to obtain the second stage assortments. Now, we are ready to prove Theorem 5.2.

Algorithm 1 Adaptive Policy for Sequential Shows

Input: An instance for a two-sided assortment problem.
Output: Feasible assortments: $\mathbf{x}^1$, $\mathbf{x}^2$

1: Use the standard greedy algorithm for submodular maximization to obtain an approximate solution $\mathbf{x}^1$ of Problem (5).
2: Observe the backlogs B generated by assortments $\mathbf{x}^1$ according to (2)
3: Obtain $\mathbf{x}^2$ by solving to optimality Problem (3).

Proof of Theorem 5.2. Consider the Reformulation (5) of Problem 2. Since $\mathbf{x}^1$ is a 0-1 vector, our goal is to prove that the objective M_f is monotone submodular over elements in $\mathcal{E}$. Given this, we know that a vanilla greedy algorithm achieves a $1/(1+r)$-approximation for the problem of maximizing a monotone submodular function over the intersection of r matroids [Fisher et al., 1978]. This would finish the proof by noting that $r = 2$ thanks to Lemma 5.3.

Let us prove then that M_f is a non-negative monotone submodular function. Recall that f is a non-negative, monotone submodular function due to Corollary 5.5. Since f is non-negative, then M_f is also non-negative. For $\mathbf{x} \in \{0,1\}^{\mathcal{E}}$, denote by $\text{supp}(\mathbf{x}) = \{a \in \mathcal{E} : \ x_a = 1\}$ the support of $\mathbf{x}$. Then, we can write M_f as follows

$$M_f(\mathbf{x}) = \sum_{B \subseteq \text{supp}(\mathbf{x})} f(B) \prod_{a \in B} \phi_a \prod_{a \in \text{supp}(\mathbf{x}) \setminus B} (1 - \phi_a)$$

On the other hand, for any $e \notin \text{supp}(x)$ we have

$$\begin{aligned} M_f(\mathbf{x} + \mathbb{1}_{\{e\}}) = \phi_e \cdot & \sum_{B \subseteq \text{supp}(\mathbf{x})} f(B+e) \prod_{a \in B} \phi_a \prod_{a \in \text{supp}(\mathbf{x}) \setminus B} (1 - \phi_a) \\ & + (1-\phi_e) \cdot \sum_{B \subseteq \text{supp}(\mathbf{x})} f(B) \prod_{a \in B} \phi_a \prod_{a \in \text{supp}(\mathbf{x}) \setminus B} (1 - \phi_a) \end{aligned}$$

Therefore,

$$M_f(\mathbf{x} + \mathbb{1}_{\{e\}}) - M_f(\mathbf{x}) = \phi_e \sum_{B \subseteq \text{supp}(\mathbf{x})} [f(B+e) - f(B)] \prod_{a \in B} \phi_a \prod_{a \in \text{supp}(\mathbf{x}) \setminus B} (1 - \phi_a)$$

Since f is monotone, then $f(B+e) - f(B) \geq 0$ for all $B \subseteq \text{supp}(\mathbf{x})$ and $e \notin \text{supp}(\mathbf{x})$, which implies $M_f(\mathbf{x} + \mathbb{1}_{\{e\}}) - M_f(\mathbf{x}) \geq 0$.

Now, let us prove that M_g is submodular. Consider any $\mathbf{x} \in \{0,1\}^{\mathcal{E}}$ and $e, e' \notin \text{supp}(\mathbf{x})$. Our goal is to show that $M_f(\mathbf{x} + \mathbb{1}_{\{e\}}) - M_f(\mathbf{x}) \geq M_f(\mathbf{x} + \mathbb{1}_{\{e\}} + \mathbb{1}_{\{e'\}}) - M_f(\mathbf{x} + \mathbb{1}_{\{e'\}})$. Note that we have the following expression for $M_f(\mathbf{x} + \mathbb{1}_{\{e\}} + \mathbb{1}_{\{e'\}})$

$$\begin{aligned} M_f(\mathbf{x} + \mathbb{1}_{\{e\}} + \mathbb{1}_{\{e'\}}) = \phi_e \phi_{e'} \cdot & \sum_{B \subseteq \text{supp}(\mathbf{x})} f(B+e+e') \prod_{a \in B} \phi_a \prod_{a \in \text{supp}(\mathbf{x}) \setminus B} (1 - \phi_a) \\ & + \phi_{e'}(1-\phi_e) \cdot \sum_{B \subseteq \text{supp}(\mathbf{x})} f(B+e') \prod_{a \in B} \phi_a \prod_{a \in \text{supp}(\mathbf{x}) \setminus B} (1 - \phi_a) \\ & + \phi_e(1-\phi_{e'}) \cdot \sum_{B \subseteq \text{supp}(\mathbf{x})} f(B+e) \prod_{a \in B} \phi_a \prod_{a \in \text{supp}(\mathbf{x}) \setminus B} (1 - \phi_a) \\ & + (1-\phi_{e'})(1-\phi_e) \cdot \sum_{B \subseteq \text{supp}(\mathbf{x})} f(B) \prod_{a \in B} \phi_a \prod_{a \in \text{supp}(\mathbf{x}) \setminus B} (1 - \phi_a) \end{aligned}$$

Analogously, we can compute $M_f(\mathbf{x} + \mathbb{1}_{\{e\}})$ and $M_f(\mathbf{x} + \mathbb{1}_{\{e'\}})$. By deleting common terms, we can obtain the following

$$\begin{aligned}&M_f(\mathbf{x} + \mathbb{1}_{\{e\}}) - M_f(\mathbf{x}) - M_f(\mathbf{x} + \mathbb{1}_{\{e\}} + \mathbb{1}_{\{e'\}}) + M_f(\mathbf{x} + \mathbb{1}_{\{e'\}})\\ &= \phi_e \phi_{e'} \cdot \sum_{B \subseteq \mathrm{supp}(\mathbf{x})} [f(B+e) - f(B) - f(B+e+e') + f(B+e')] \prod_{a \in B} \phi_a \prod_{a \in \mathrm{supp}(\mathbf{x}) \setminus B} (1 - \phi_a),\end{aligned}$$

from which submodularity follows due to submodularity of f. □

A natural concern is whether we can evaluate the function $M_f(\cdot)$ efficiently in every iteration of the greedy algorithm. Indeed, by using standard sampling techniques [Calinescu et al., 2011], for any point $\mathbf{x}$ we can obtain a $(1 - \epsilon)$ approximation of $M_f(\mathbf{x})$.

5.1.1 Improved Factor for One Directional Sequential Matches. In this section, we show that the approximation factor obtained in Theorem 5.2 can be improved when we restrict the problem to the setting in which there are one directional sequential matches, i.e., there is only one side of the market initiating the sequential matches. One directional sequential matching models have been studied in the past [Ashlagi et al., 2022, Torrico et al., 2021]; for more details, please refer to the discussion in Related Literature 2. Our main result is the following:

THEOREM 5.6. *There exists a $(1 - 1/e)$-approximation algorithm for Problem 2 with one directional sequential shows and no simultaneous shows.*

Without loss of generality, we assume that sequential matches can only be initiated by agents in I. This means that the platform must choose in the first stage edges only from $\mathcal{E} = \vec{E}_I$. Recall that in the previous section we also allowed arcs in $\vec{E}_J$. Given this, we can reformulate Problem 2 as follows. First, the feasible region for the first-period assortment decisions is the following:

$$P^1 = \left\{ \mathbf{x}^1 \in \{0,1\}^{\mathcal{E}} : \sum_{j \in \mathcal{P}_i} x^1_{(i,j)} \leq K_i, \text{ for every } i \in I \right\}. \tag{6}$$

Note that the only family of constraints that we need is the one corresponding to the cardinality constraints for agents in I. From the proof of Lemma 5.3, we can easily conclude the following result.

LEMMA 5.7. *Feasible region P^1 defined in* (6) *corresponds to a single partition matroid.*

Second, since in the second stage only agents in J respond back, then for any realized backlog B, the function $f(B)$ is given by

$$\begin{aligned}f(B) = \max \Bigg\{ \sum_{a \in \vec{E}_J} \phi^2_a x^2_a : \sum_{i \in \mathcal{P}_j} x^2_{(j,i)} \leq K_j, \text{ for every } j \in J,\\ x^2_a \leq 1_{\{\bar{a} \in B\}}, \text{ for every } a \in \vec{E}_J, \quad \mathbf{x} \in \{0,1\}^{\mathcal{E}} \Bigg\},\end{aligned}$$

which is slightly different to the one in the previous section since we only allowed edges in $\vec{E}_J$. Therefore, Problem 2 for the setting with one-sided sequential matches corresponds to

$$\max \left\{ M_f(\mathbf{x}) : \mathbf{x} \in P^1 \right\} \tag{7}$$

Observe that the standard greedy algorithm guarantees a $1/2$-approximation for Problem (7), since the objective is submodular and constraints in P^1 form a partition matroid. In the following, our goal is to improve this factor to $1 - 1/e$.

To prove Theorem 5.6, we use appropriate submodular optimization and randomized rounding tools to theoretically analyze the reformulated problem. To formalize this analysis, we first introduce some additional notation.

Definition 5.8 (Multilinear Extension). Consider a non-negative set function $g : \{0,1\}^{\mathcal{E}} \to \mathbb{R}_+$. The *multilinear extension* of g, denoted by $M^{\text{ext}}_g : [0,1]^{\mathcal{E}} \to \mathbb{R}_+$, is defined for any $y \in [0,1]^{\mathcal{E}}$ as the expected value of $g(O_y)$, where O_y is the random set generated by drawing independently each element $s \in \mathcal{E}$ with probability y_s. That is,

$$M^{\text{ext}}_g(y) = \sum_{O \subseteq \mathcal{E}} g(O) \prod_{s \in O} y_s \prod_{s \in \mathcal{E} \setminus O} (1 - y_s).$$

Note that for any $\mathbf{x} \in \{0,1\}^{\mathcal{E}}$, we have $M^{\text{ext}}_f(\phi^1 \mathbf{x}) = M_f(\mathbf{x})$. The difference between M^{ext} and M_f is that M^{ext} can be evaluated in $[0,1]^{\mathcal{E}}$, rather than only in $\{0,1\}^{\mathcal{E}}$.

Consider the following optimization problem given by

$$\begin{aligned} \max \quad & M^{\text{ext}}_f(\mathbf{z}) && && \text{(8a)} \\ \text{s.t.} \quad & \sum_{j \in J : \phi^1_{ij} > 0} \frac{z_{ij}}{\phi^1_{ij}} \le K_i && \text{for every } i \in I, && \text{(8b)} \\ & 0 \le z_a \le \phi^1_a && \text{for every } a \in \mathcal{E}. && \text{(8c)} \end{aligned}$$

Lemma 5.9. *The optimal value of* (8a)-(8c) *is an upper bound on the optimal value of* (7).

Proof. Consider a feasible solution $\mathbf{x}^1 \in \{0,1\}^{\mathcal{E}}$ for the optimization problem (7), that is, $\sum_{j \in J} x^1_{(i,j)} \le K$ for every $i \in I$. Let $\mathbf{z} = \phi^1 \mathbf{x}^1$. Observe that

$$\sum_{j \in J : \phi^1_{ij} > 0} z_{ij} / \phi^1_{ij} = \sum_{j \in J : \phi^1_{ij} > 0} x^1_{(i,j)} \le K_i$$

for every $i \in I$, and $z_a = x^1_a \phi^1_a \le \phi^1_a$ for every $a \in \mathcal{E}$. Therefore $\mathbf{z}$ is a feasible solution for the problem (8a)-(8c). Since the objective value of $\mathbf{x}^1$ in (7) is equal to the objective of $\mathbf{z}$ in (8a)-(8c), we conclude the lemma. □

Since f is monotone submodular, we can obtain the following result for Problem (8a)-(8c):

Corollary 5.10 ([Vondrák, 2008]). *There exists an efficient algorithm that computes a point $\tilde{\mathbf{z}}$ that satisfies* (8b) *and* (8c) *such that* $M^{ext}_f(\tilde{\mathbf{z}}) \ge (1 - 1/e) \cdot \max\{M^{ext}_f(\mathbf{z}) : \ \mathbf{z} \text{ s.t. (8b) and (8c)}\}$.

In what follows, we discuss the algorithm used to prove Theorem 5.6, which we formalize in Algorithm 2. Our algorithm is based mainly on two different tools: Theorem 5.10 by [Vondrák, 2008], and the dependent randomized rounding algorithm by [Gandhi et al., 2006]. We include a brief description of this method in Appendix C.1.

Our algorithm utilizes the dependent randomized rounding algorithm of [Gandhi et al., 2006]. Given the fractional solution $\tilde{\mathbf{z}}$ satisfying the guarantee in Corollary 5.10, let $\tilde{\mathbf{x}}_i \in [0,1]^J$ be the vector such that the j-th entry is equal to $\tilde{z}_{ij}/\phi^1_{ij}$ when $\phi^1_{ij} > 0$ and zero otherwise. Observe that thanks to constraint (8c) we have $\tilde{\mathbf{x}}_i \in [0,1]^J$ for each $i \in I$. Then, independently for each user $i \in I$, by the algorithm in [Gandhi et al., 2006] it is possible to efficiently compute an integral random vector $\hat{\mathbf{x}}_i \in \{0,1\}^J$ satisfying the following conditions:

(1) $\sum_{j \in J} \hat{x}_{ij} \le \lceil \sum_{j \in J} \tilde{x}_{ij} \rceil$, and
(2) $\mathbb{E}[\hat{x}_{ij}] = \tilde{x}_{ij}$ for each $i \in I$ and $j \in J$.

Algorithm 2 Approximation Algorithm for One Directional Sequential Shows

Input: An instance for a two-sided assortment problem.
Output: A feasible assortment $\mathbf{x}^1$.
1: Compute a solution $\tilde{\mathbf{z}}$ for the problem (8a)-(8c) using the algorithm from Corollary 5.10.
2: For each $i \in I$ and $j \in J$, set $\tilde{x}_{ij} = \tilde{z}_{ij}/\phi^1_{ij}$ when $\phi^1_{ij} > 0$ and zero otherwise.
3: Independently for each user $i \in I$, run the dependent randomized rounding algorithm [Gandhi et al., 2006] on the fractional vector $\tilde{\mathbf{x}}_i$ to compute an integral random vector $\hat{\mathbf{x}}_i \in \{0,1\}^J$.
4: For each $i \in I$ return $S^1_i = \{j \in J : \hat{x}_{ij} = 1\}$.

Note that Algorithm 2 can be used to design an adaptive policy analogous to Algorithm 1. We are ready to prove our main result.

Proof of Theorem 5.6. Thanks to condition (1) of the randomized rounding algorithm, for each $i \in I$ we have $\sum_{j\in J} \hat{x}_{ij} \le \left\lceil \sum_{j\in J} x_{ij} \right\rceil = \lceil \sum_{j\in J:\phi_{ij}>0} \tilde{z}_{ij}/\phi_{ij} \rceil \le K_i$, where the last inequality holds since K_i is integral and $\tilde{\mathbf{z}}$ satisfies constraint (8b). Therefore, our algorithm gives a feasible solution for Problem (7). We now analyze the approximation guarantee.

$$\begin{aligned}
\mathbb{E}_{\hat{x}}[M^{\text{ext}}_f(\hat{\mathbf{x}}^1)] &= \sum_B f(B) \cdot \mathbb{E}_{\hat{\mathbf{x}}}\left[\prod_{a\in B} \phi^1_a \hat{x}^1_a \prod_{a\notin B}(1-\phi^1_a\hat{x}^1_a)\right] \\
&= \sum_B f(B) \cdot \prod_{a\in B} \phi^1_a \mathbb{E}_{\hat{\mathbf{x}}}[\hat{x}^1_a] \prod_{a\notin B}(1-\phi^1_a \mathbb{E}_{\hat{\mathbf{x}}}[\hat{x}^1_a]) \\
&= \sum_B f(B) \cdot \prod_{a\in B} \phi^1_a \cdot \frac{\tilde{z}_a}{\phi^1_a} \prod_{a\notin B}(1-\phi^1_a \cdot \frac{\tilde{z}_a}{\phi^1_a}) = M_f(\tilde{z}),
\end{aligned}$$

where the second equality comes from the fact that $\hat{\mathbf{x}}_i$ is independent from $\hat{\mathbf{x}}_{i'}$ for every $i, i' \in I$ with $i \ne i'$ and the third equality comes from condition 2 of the randomized rounding procedure. Finally, Lemma 5.9 states that $\text{OPT}' \ge \text{OPT}$, where OPT' is the optimal value of Problem (8a)-(8c), so we conclude the proof by using Corollary (5.10). □

Note that Algorithm 2 is similar to the approach taken in [Torrico et al., 2021]. However, we are able to obtain a better approximation guarantee since the probabilities do not depend on the assortments. Instead, in [Torrico et al., 2021] the authors consider multinomial logit choice probabilities, which worsens the approximation to a $(1-1/e)/2$-factor.

5.2 Adding Simultaneous Matches in the First Period

One of the limitations of the usual approaches in sequential two-sided assortment optimization is the inability of users to evaluate profiles in each period. To accommodate this feature, we now consider the possibility that users see and like each other in the same period. For simplicity, in this section, we will focus on allowing simultaneous shows in the first period. Recall that, in Section 5.1, we considered decision variables $x^1_{(i,j)}$ and $x^1_{(j,i)}$. Also, we imposed that $x^1_{(i,j)} + x^1_{(j,i)} \le 1$ since we did not allow simultaneous matches. If we remove that family of constraints, a simultaneous match between users ℓ and ℓ' happens in the first period with probability $\phi^1_a\phi^1_{\bar{a}}x^1_a x^1_{\bar{a}}$. Therefore, an extra term has to be added to the optimization problem (5). Formally, Problem 2 with simultaneous matches in the first period can be formulated as follows:

$$\max\left\{\sum_{a\in\vec{E}} \phi^1_a\phi^1_{\bar{a}}x^1_a x^1_{\bar{a}} + M_f(\phi^1 x^1) : \mathbf{x}^1 \in \tilde{P}^1\right\},$$

where $\tilde{P}^1$ corresponds to P^1 defined in (1) without the second family of constraints and M_f defined in (4). Note that the objective function in this formulation is not necessarily submodular, so we cannot follow the same approach described in the previous section. To overcome this, we reformulate the objective function to make it submodular by adding extra variables and constraints. Our main result is the following

Theorem 5.11. *There exists a* $1/4$*-approximation algorithm for Problem 2 with simultaneous shows only in the first period.*

Let us consider an extra binary variable w_e^1 for every $e \in E$ (undirected edges) that is equal to one if $i \in I$ and $j \in J$ with $e = \{i, j\}$ simultaneously see each other in the first period, and it is equal to zero otherwise. Notice that for a given edge $e = \{i, j\}$, and arc $a = (i, j)$, only one of the variables w_e^1, x_a^1 and $x_{\bar{a}}^1$ can be equal to one. Therefore, our feasible region is therefore the following:

$$Q^1 = \left\{ \mathbf{x}^1 \in \{0,1\}^{\vec{E}},\ \mathbf{w}^1 \in \{0,1\}^E : w^1_{\{i,j\}} + x^1_{(i,j)} + x^1_{(j,i)} \leq 1, \qquad \text{for every } i \in I, j \in J, \right.$$
$$\sum_{j \in J} \left(x^1_{(i,j)} + w^1_{\{i,j\}} \right) \leq K_i, \qquad \text{for every } i \in I,$$
$$\left. \sum_{i \in I} \left(x^1_{(j,i)} + w^1_{\{i,j\}} \right) \leq K_j, \qquad \text{for every } j \in J, \right\}. \tag{9}$$

Intuitively, each edge $e \in E$ has three possible states: without direction (simultaneous shows) or either of both directed versions (sequential shows). In the first period, we can only choose one of these states for each edge $e \in E$, which corresponds to the first family of constraints in P^1. Each user $\ell \in I \cup J$ can only be shown at most K_ℓ users in the opposite side. This can happen in two types of states: a directed arc or an edge. Therefore, for each user we have to choose at most K_ℓ of these states, which corresponds to the second and third family of constraints. Before proving our main result, consider the following definition of r-extendible systems (see e.g., [Mestre, 2006]).

Definition 5.12 (r-extendible system). A subset system $\mathcal{I}$ is r-extendible if the following holds: for any $X \subset Y$, $X, Y \in \mathcal{I}$ and $u \notin A$ such that $A \cup \{u\} \in \mathcal{I}$, then there is a set $Z \subseteq B \setminus A$ such that $|Z| \leq r$ and $B \setminus Z \cup \{u\} \in \mathcal{I}$.

For example, matroids are 1-extendible systems and matchings are 2-extendible systems.

Lemma 5.13. *Feasible region Q^1 defined in* (9) *corresponds to a* 3*-extendible system.*

The proof for this lemma can be found in the Appendix C.2. Now, we can formulate our problem in the following way:

$$\max \left\{ \sum_{e \in E} \phi^1_{a_e} \phi^1_{\overline{a_e}} w^1_e + M_f(\mathbf{x}^1) : (\mathbf{x}^1, \mathbf{w}^1) \in Q^1 \right\}. \tag{10}$$

Observe that the objective corresponds to the sum of a linear function plus a submodular function, which gives a submodular function.

Proof of Theorem 5.11. This result follows by using the standard greedy algorithm for submodular maximization over a 3-extendible system which gives a $1/4$ approximation guarantee [Fisher et al., 1978]. □

In [Fisher et al., 1978] the authors prove a $1/(1 + r)$ approximation factor for r-systems which is a more general class than r-extendible systems.

5.2.1 Improved Factor for One Directional Sequential Matches. In this section, we show that the approximation factor obtained in Theorem 5.11 can be improved when we restrict the problem to the setting in which there are one-sided sequential matches, i.e., there is only one side starting the sequential matches but we still allow simultaneous shows in the first stage. Our main result is the following

THEOREM 5.14. *There exists a $1/2$-approximation algorithm for Problem 2 with one-sided sequential shows and simultaneous shows only in the first period.*

Without loss of generality, we assume that sequential matches can only be initiated by agents in I. This means that the platform must choose in the first stage edges only from $\mathcal{E} = E \cup \vec{E}_I$. Given this, we can reformulate Problem 2 as follows. First, the feasible region for the assortment decisions of the first stage is the following:

$$Q^1 = \left\{ \mathbf{x}^1 \in \{0,1\}^{\vec{E}_I},\ \mathbf{w}^1 \in \{0,1\}^E : w^1_{\{i,j\}} + x^1_{(i,j)} \leq 1, \qquad \text{for every } i \in I, j \in J, \right.$$
$$\left. \sum_{j \in J} \left(x^1_{(i,j)} + w^1_{\{i,j\}} \right) \leq K_i, \qquad \text{for every } i \in I. \right\} \tag{11}$$

Note that the only family of constraints that we need are the ones corresponding to agents in I. As we show in Lemma 5.15, the region Q^1 is a laminar matroid (see the proof in Appendix C.2).

LEMMA 5.15. *Feasible region Q^1 defined in* (11) *corresponds to a laminar matroid.*

Since in the second stage only agents in J respond back, then we define function f as in Section 5.1.1. Finally, the proof of Theorem 5.14 follows by using the greedy algorithm for submodular maximization over a single matroid [Fisher et al., 1978].

5.3 Adding Simultaneous Matches in both Periods

Consider now the general two-period version of our problem, where matches can happen either sequentially or simultaneously in any of the two periods.

Given the assortments defined by $\mathbf{x}^1$ and $\mathbf{w}^1$ (i.e., sequential and simultaneous shows in the first period) and a realized backlog $\mathcal{B}^2 = B$ at the beginning of the second period, let $\tilde{f}(B, \mathbf{x}^1, \mathbf{w}^1)$ be the optimal expected number of matches that can be obtained given the constraints of the problem, i.e., that assortments have a fixed size and that users can see each other at most once.[4] Formally, given $\mathbf{x}^1$ and $\mathbf{w}^1$ feasible points in Q^1 defined in (9) we have the following expression for $\tilde{f}$

$$\tilde{f}(B, \mathbf{x}^1, \mathbf{w}^1) = \max_{\mathbf{x}^2 \in \tilde{P}^2_B} \left\{ \sum_{a \in \vec{E}} \phi^2_a x^2_a + \sum_{e \in E} \phi^1_{a_e} \phi^1_{\overline{a_e}} w^1_e \right\}, \tag{12}$$

where the feasible region is given by

$$\tilde{P}^2_B = \left\{ \mathbf{x}^2 \in \{0,1\}^{\vec{E}}, \mathbf{w}^2 \in \{0,1\}^E : x^2_a \leq 1_{\{\overline{a} \in B\}}, \qquad \text{for every } a \in \mathcal{E} \right.$$
$$w^2_e \leq 1 - x^1_{a_e} - x^1_{\overline{a_e}} - w^1_e, \qquad \text{for every } e \in E,$$
$$\sum_{j \in J} \left(x^2_{(i,j)} + w^2_{\{i,j\}} \right) \leq K_i, \qquad \text{for every } i \in I,$$

[4]When clear from the context, and in a slight abuse of notation, we sometimes omit the dependence on $\mathbf{x}^1$ and $\mathbf{w}^1$ and simply write $\tilde{f}(B, \mathbf{x}^1, \mathbf{w}^1)$ as $\tilde{f}(B)$.

$$\sum_{i \in I} \left(x^2_{(j,i)} + w^2_{\{i,j\}} \right) \leq K_j, \qquad \text{for every } j \in J, \quad \Big\},$$

Recall that, for any edge $e \in E$, a_e and $\overline{a_e}$ are the corresponding directed arcs. It is easy to see that no pair of agents $i \in I$ and $j \in J$ will see each other more than once across periods. Moreover, as we show in Proposition 5.16 (see the proof in Appendix C.2), the function $\tilde{f}(B, \mathbf{x}^1, \mathbf{w}^1)$ can be efficiently evaluated by solving a linear program.

Proposition 5.16. *The function $\tilde{f}(B)$ as defined in* (12) *can be evaluated solving a linear program.*

As a result, Problem 2 can be reformulated as

$$\max \left\{ \sum_{e \in E} \phi^1_{a_e} \phi^1_{\overline{a_e}} w^1_e + M_{\tilde{f}}(\mathbf{x}^1, \mathbf{w}^1) : (\mathbf{x}^1, \mathbf{w}^1) \in Q^1 \right\}, \tag{13}$$

where

$$M_{\tilde{f}}(\mathbf{x}^1, \mathbf{w}^1) = \sum_{B} \tilde{f}(B, \mathbf{x}^1, \mathbf{w}^1) \mathbb{P}_{\mathbf{x}^1}(\mathcal{B} = B).$$

Note that Proposition 5.16 implies that the problem in the second period can be easily solved if we know the realized backlog after the first period. However, as discussed above, the number of backlogs that can result from the assortments shown in the first period grows exponentially in the number of subjects, and thus it is not possible to efficiently evaluate the function $M_{\tilde{f}}$.

An alternative approach to solve this problem would be to rely on similar submodular optimization techniques as those used in previous sections. Unfortunately, as we show in Proposition 5.17, the function $\tilde{f}(B)$ is not submodular.

Proposition 5.17. *Function $\tilde{f}$ as defined in* (12) *is not submodular in B.*

Even though we cannot use submodular optimization techniques, we can show the following result for one directional sequential matches (see the proof in Appendix C.2).

Theorem 5.18. *Consider the case with one-directional sequential matches starting from side I and simultaneous matches in both periods. Suppose that probabilities are time-independent, i.e., $\phi^1_{\ell\ell'} = \phi^2_{\ell\ell'} = \phi_{\ell\ell'}$ and that $\phi_{ij} \leq 1/n$ for any $i \in I, j \in J$, where α is a constant and $n = |I|$. Denote by OPT_1 the optimal solution of Problem* (10) *(with simultaneous matches only in the first period) and by OPT_2 the optimal value of Problem* (13) *(simultaneous matches in both stages). Then, we have*

$$OPT_1 \geq \left(\frac{1}{2e} - o(1) \right) \cdot OPT_2.$$

More importantly, any γ-approximation algorithm for Problem (10) *guarantees a $\gamma \left(\frac{1}{2e} - o(1) \right)$ approximation for Problem* (13).

This result states that when the initiating side (of the sequential matches) is sufficiently picky and the market is sufficiently large, adding simultaneous matches in the second stage does not improve the optimal expected number of matches. In other words, the majority of matches come from sequential shows. As a result, we conjecture that the guarantees found for the case with simultaneous matches in the first period may extend to the more general case.

6 EXPERIMENTS

In this section, we evaluate the performance of the algorithms described above using real data and we assess the value of considering simultaneous matches.

6.1 Data

We use a dataset obtained from our industry partner to perform our experiments. This dataset includes all heterosexual users from Houston, TX that logged in between February 14 and August 14, 2020, and includes all the observable characteristics displayed in their profiles for each user in the sample, namely, their age, height, location, race, and religion. It also includes an attractiveness score—or simply score—that depends on the number of likes received and evaluations received in the past.[5] Finally, the dataset includes all the profiles that each user evaluated between February 14 and August 14, 2020, including the decisions made (like or dislike), which other profiles were part of the assortment, and relevant timestamps. As a result, we have a panel of observations, and we can fully characterize each profile evaluated by each user in the sample.

Using this dataset, for any pair of users $\ell \in I \cup J$ and $\ell' \in \mathcal{P}^1_\ell$, we compute the probability that ℓ likes ℓ' using the panel regression model described in detail in Appendix D.1. Since we assume that like probabilities are independent across periods (i.e., we assume that there is no effect of the history on the like behavior of users),[6] we estimate the like probabilities considering a logit model with fixed effects at the user level. In addition, we control for all the observable characteristics available in the data, namely, the characteristics of the profile evaluated and the interaction with those of the user evaluating. In Table 4 (see Appendix D.1), we report the estimation results. Then, by using the estimated coefficients and users' observable characteristics, we predict the probabilities $\phi_{\ell,\ell'}$ for all $\ell \in I \cup J$ and $\ell' \in \mathcal{P}^1_\ell$.

We perform our simulations considering a sample of the dataset described above to reduce the computational time. In Table 3, we report several summary statistics of the sample (standard deviations in parenthesis), including the number of users, their average score, the average number of potentials available, their average backlog size, and their average like probabilities.

Table 3. Descriptives of Instance

	N	Score	Potentials	Backlog	Like Prob.
Women	173	5.289	73.23	0.120	0.203
		(2.454)	(15.032)	(0.152)	(0.247)
Men	113	2.618	97.461	0.012	0.536
		(1.463)	(21.680)	(0.026)	(0.368)

6.2 Benchmarks

To assess the performance of the proposed algorithms, which we refer to as *Global Greedy*, we compare them with several benchmarks:[7]

[5] This score is measured on a scale from 0 to 10, where 10 represents the most attractive profiles and 0 the least attractive ones.

[6] Rios et al. [2022] find that the number of matches obtained in the recent past has a significant negative effect on users' future like/dislike decisions. However, the authors show that the primary source of improvement comes from taking into account the market's two-sidedness and choosing better assortments. Hence, we decided to focus on the latter and simplify the estimation of probabilities.

[7] We also tested other benchmarks such as Naive, Random, and also our partner's algorithm. However, since the results reported here significantly outperform these other benchmarks, we decided to omit them and focus on the results of the algorithms proposed above.

(1) Local Greedy: for each user, select the subset of profiles that maximizes their expected number of matches, i.e.,

$$S_\ell^t = \underset{S \subseteq \mathcal{P}_\ell \setminus \bigcup_{\tau=1}^{t-1} S_\ell^\tau : |S| \leq K_\ell}{\arg\max} \left\{ \sum_{\ell' \in S} \phi_{\ell\ell'} \cdot \left(\mathbb{1}_{\{\ell' \in \mathcal{B}_\ell^t\}} + \phi_{\ell'\ell} \cdot \mathbb{1}_{\{\ell' \notin \mathcal{B}_\ell^t\}} \right) \right\}$$

(2) Perfect Matching (PM): in each period $t \in [T]$, solve the perfect match problem (including possible initial backlogs)

$$\begin{aligned} \max \quad & \sum_{\ell \in I \cup J} \sum_{\ell' \in \mathcal{P}_\ell} y_{\ell\ell'}^t \phi_{\ell\ell'}^t + \frac{1}{2} w_{\ell\ell'}^t \phi_{\ell\ell'}^t \phi_{\ell'\ell}^t \\ st. \quad & y_{\ell\ell'}^t \leq \mathbb{1}_{\{\ell' \in \mathcal{B}_\ell^t\}}, && \text{for every } \ell' \in \mathcal{P}_\ell^t, \ell \in I \cup J \\ & x_{\ell\ell'}^t + y_{\ell\ell'}^t \leq 1, && \text{for every } \ell' \in \mathcal{P}_\ell^t, \ell \in I \cup J \\ & \sum_{\ell'} x_{\ell\ell'}^t + y_{\ell\ell'}^t \leq K_\ell, && \text{for every } \ell \in I \cup J \\ & w_{\ell\ell'}^t \leq x_{\ell\ell'}^t,\ w_{\ell\ell'}^t \leq x_{\ell'\ell}^t,\ w_{\ell\ell'}^t = w_{\ell'\ell}^t, && \text{for every } \ell' \in \mathcal{P}_\ell^t, \ell \in I \cup J \\ & x_{\ell\ell'}^t, w_{\ell\ell'}^t \in [0,1], && \text{for every } \ell' \in \mathcal{P}_\ell^t, \ell \in I \cup J, \end{aligned} \tag{14}$$

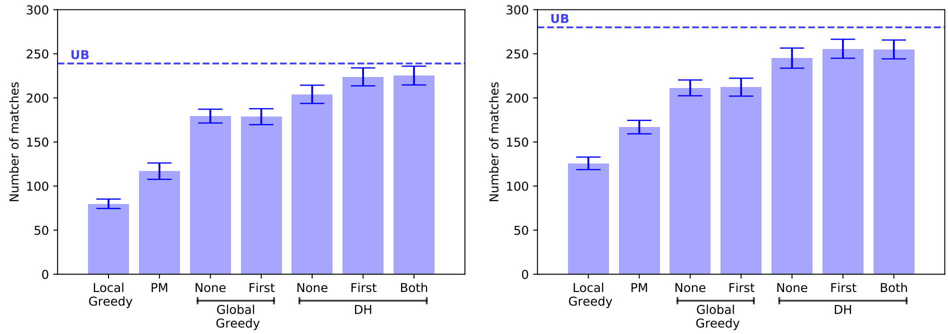
Then, set $S_\ell^t = \left\{ \ell' \in \mathcal{P}_\ell^t : x_{\ell\ell'}^t = 1 \text{ or } y_{\ell\ell'}^t = 1 \right\}$ for each $\ell \in I \cup J$. Notice that this is a particular case of (18) (see Appendix D.2), as it is the same formulation assuming a horizon of only one period.

(3) Dating Heuristics: we adapt the Dating Heuristic (DH) described in [Rios et al., 2022] to our setting. In Appendix D.2 we describe in detail how these heuristics work and how we adapt them to our setting. To assess the impact of considering simultaneous matches when making first period decisions, we consider a three special cases of Algorithm 3 (in Appendix D.2):
   (a) None: when solving Problem (18) to obtain the first period assortments, we avoid considering simultaneous matches in both periods, i.e., we force $w_{\ell,\ell'}^t = 0$ for all $t \in \{1, 2\}$ and all $\ell \in I \cup J, \ell' \in \mathcal{P}_\ell^1$.
   (b) First: when solving Problem (18) to obtain the first period assortments, we avoid considering simultaneous matches in the second period, i.e., we force $w_{\ell,\ell'}^2 = 0$ for all $\ell \in I \cup J, \ell' \in \mathcal{P}_\ell^1$.
   (c) Both: when solving Problem (18) to obtain the first period assortments, we consider simultaneous matches in both periods.

Moreover, we consider two variants of Global Greedy: (i) None and (ii) First. As for DH, in the former, we consider no simultaneous matches when choosing the first-period assortments (i.e., we use Algorithm 1 directly). In the latter, we extend this algorithm to allow simultaneous matches in the first period. Finally, we compare all these methods with the UB obtained from solving Problem (18).

6.3 Results

For each benchmark, we perform 100 simulations where, in each period, (i) we choose the assortment to show to each user considering $K_\ell = 3$ for all $\ell \in I \cup J$, (ii) we simulate the decisions of the users based on their like probabilities, and (iii) we update the state of the system before moving on to the next period. The results are summarized in Figure 1, where we report the average number of matches generated by each benchmark. In Figure 1a, we report the results considering no initial backlog for all users, i.e., $\mathcal{B}_\ell^1 = \emptyset$ for all $\ell \in I \cup J$, while in Figure 1b, we consider the actual backlogs as in August 16, 2020.



(a) Without Initial Backlog

(b) With Initial Backlog

Fig. 1. Matches by Benchmark

First, we observe that the performance of PM and Local Greedy are considerably better than their worst-case performance. Second, we observe that Global Greedy largely outperforms these benchmarks. Moreover, we find no significant differences in the number of matches obtained when considering simultaneous matches in the first period (i.e., comparing Global Greedy None vs. First). These results suggest that allowing simultaneous matches in the first period does not make a significant difference. Third, we find that DH and its variants outperform all the other benchmarks. Interestingly, if we only allow sequential matches (as in DH-None), the performance is relatively similar when both simultaneous and sequential matches are permitted. Indeed, when we assume no initial backlogs, DH-None achieves 90.56% of the matches generated by DH, while this number increases to 96.46% when considering initial backlogs. In addition, we observe that DH-First and DH-Both lead to almost identical results, suggesting that considering simultaneous matches in the second period while making first-period decisions plays no significant role. These results suggest that most matches are generated either sequentially and that considering simultaneous matches in the second period when choosing first-period assortments plays no significant role. Hence, these simulation results support the conjecture that the performance guarantees obtained for the case with simultaneous matches in the first period are similar to those for the general problem.

7 CONCLUSIONS

We theoretically study a two-period version of the two-sided assortment optimization problem introduced in [Rios et al., 2022]. This problem departs from previous work as it allows both sides of the market to start an interaction, as is the case in many matching markets, including dating, freelancing, car-sharing, etc. Using tools from submodular optimization, we provide the first performance guarantees for different variants of the problem, including the cases with sequential matches and also considering the general setting when both sequential and simultaneous matches are allowed. In addition, we show through simulations that the improvement obtained from considering simultaneous matches is limited. Hence, we conjecture that the performance guarantees obtained for the case with simultaneous matches in the first period are similar to those for the general case.

There are many exciting directions for future research. First, as mentioned above, we conjecture that the performance guarantee of $1/4$ also applies to the general case, so it would be worth exploring and confirming this is true. Second, extending our results to the case with multiple periods and when like probabilities depend on the assortment shown would be interesting. Finally, it is worth studying performance guarantees when users log in with some known probability.

REFERENCES

Ali Aouad and Daniela Saban. 2022. Online Assortment Optimization for Two-sided Matching Platforms. *Management Science* 0, 0 (2022).

Itai Ashlagi, Anilesh Krishnaswamy, Rahul Makhijani, Daniela Saban, and Kirankumar Shiragur. 2022. Technical Note - Assortment planning for two-sided sequential matching markets. *Operations Research* 70, 5 (2022), 2784–2803.

Gerardo Berbeglia and Gwenaël Joret. 2015. Assortment Optimisation Under a General Discrete Choice Model: A Tight Analysis of Revenue-Ordered Assortments. (2015), 36 pages.

Jose Blanchet, Guillermo Gallego, and Vineet Goyal. 2016. A Markov chain approximation to choice modeling. *Operations Research* 64, 4 (2016), 886–905.

Gruia Calinescu, Chandra Chekuri, Martin Pal, and Jan Vondrák. 2011. Maximizing a monotone submodular function subject to a matroid constraint. *SIAM J. Comput.* 40, 6 (2011), 1740–1766.

Felipe Caro and Jérémie Gallien. 2007. Dynamic assortment with demand learning for seasonal consumer goods. *Management Science* 53, 2 (2007), 276–292.

James M Davis, Guillermo Gallego, and Huseyin Topaloglu. 2014. Assortment Optimization Under Variants of the Nested Logit Model. *Operations Research* 62, 2 (2014), 250–273.

Vivek F Farias, Srikanth Jagabathula, and Devavrat Shah. 2013. A Nonparametric Approach to Modeling Choice with Limited Data. *Management Science* 59, 2 (2013), 305–322.

Marshall L Fisher, George L Nemhauser, and Laurence Wolsey. 1978. An analysis of approximations for maximizing submodular set functions–II. *Polyhedral combinatorics* (1978).

Rajiv Gandhi, Samir Khuller, Srinivasan Parthasarathy, and Aravind Srinivasan. 2006. Dependent rounding and its applications to approximation algorithms. *Journal of the ACM (JACM)* 53, 3 (2006), 324–360.

Negin Golrezaei, Hamid Nazerzadeh, and Paat Rusmevichientong. 2014. Real-Time Optimization of Personalized Assortments. *Management Science* 60, 6 (2014), 1532–1551.

Hanna Halaburda, Mikolaj Jan Piskorski, and Pinar Yildirim. 2018. Competing by Restricting Choice: The Case of Search Platforms. *Management Science* 64, 8 (2018), 3574–3594.

Yash Kanoria and Daniela Saban. 2017. Facilitating the Search for Partners on Matching Platforms: Restricting Agents' Actions. (2017).

A. Kök, M. Fisher, and R. Vaidyanathan. 2015. *Retail Supply Chain Management: Quantitative Models and Empirical Studies*. Springer US, 175–236.

Julián Mestre. 2006. Greedy in approximation algorithms. In *Algorithms–ESA 2006: 14th Annual European Symposium, Zurich, Switzerland, September 11-13, 2006. Proceedings 14*. Springer, 528–539.

Ignacio Rios, Daniela Saban, and Fanyin Zheng. 2022. Improving match rates in dating markets through assortment optimization. *Manufacturing & Service Operations Management* 0, 0 (2022).

Jean-Charles Rochet and Jean Tirole. 2003. Two-Sided Markets. *Journal of the European Economic Association* (2003), 990–1029.

Paat Rusmevichientong, Zuo-Jun Max Shen, and David B. Shmoys. 2010. Dynamic Assortment Optimization with a Multinomial Logit Choice Model and Capacity Constraint. *Operations Research* 58, 6 (2010), 1666–1680.

Paat Rusmevichientong, David Shmoys, Chaoxu Tong, and Huseyin Topaloglu. 2014. Assortment optimization under the multinomial logit model with random choice parameters. *Production and Operations Management* 23, 11 (2014), 2023–2039.

Denis Sauré and Assaf Zeevi. 2013. Optimal Dynamic Assortment Planning with Demand Learning. *Manufacturing & Service Operations Management* 15, 3 (2013), 387–404.

Peng Shi. 2022a. Optimal Matchmaking Strategy in Two-Sided Marketplaces. *Management Science* 0, 0 (2022), null.

Peng Shi. 2022b. Optimal Priority-Based Allocation Mechanisms. *Management Science* 68, 1 (January 2022), 171–188.

Kalyan Talluri and Garrett van Ryzin. 2004. Revenue Management Under a General Discrete Choice Model of Consumer Behavior. *Management Science* 50, 1 (2004), 15–33.

Alfredo Torrico, Margarida Carvalho, and Andrea Lodi. 2021. Multi-agent assortment optimization in sequential matching markets. (2021).

Jan Vondrák. 2008. Optimal approximation for the submodular welfare problem in the value oracle model. In *Proceedings of the 40th Annual ACM Symposium on the Theory of Computing (STOC)*. 67–74.

Ruxian Wang and Ozge Sahin. 2018. The Impact of Consumer Search Cost on Assortment Planning and Pricing. *Management Science* 64, 8 (2018), 3649–3666.

A APPENDIX TO SECTION 3

A.1 Dynamic programming formulation

Let $\mu^t_{\ell\ell'} = 1$ if a match between users ℓ and ℓ' happens in period t, and $\mu^t_{\ell\ell'} = 0$ otherwise. As discussed in Section 3, $\mu^t_{\ell,\ell'} = 1$ if and only if one of the following three (disjoint) events takes place: (i) $\left\{\Phi^t_{\ell\ell'} = 1, \ell' \in \mathcal{B}^t_\ell\right\}$, (ii) $\left\{\Phi^t_{\ell'\ell} = 1, \ell \in \mathcal{B}^t_{\ell'}\right\}$, or (iii) $\left\{\Phi^t_{\ell\ell'} = \Phi^t_{\ell'\ell} = 1\right\}$.

Given a set of potentials $\vec{\mathcal{P}} = \{\mathcal{P}_\ell\}_{\ell \in I \cup J}$ and a set of backlogs $\vec{\mathcal{B}} = \{\mathcal{B}_\ell\}_{\ell \in I \cup J}$, Problem 2 can be formulated as the following dynamic program:[8]

$$V^t(\vec{\mathcal{P}}, \vec{\mathcal{B}}) = \max_{\substack{S_\ell \in \mathcal{P}_\ell \\ |S_\ell| \le K_\ell \forall \ell \in I \cup J}} \left\{ \mathbb{E}\left[\sum_{\ell \in I \cup J} \sum_{\ell' \in S_\ell} \mu^t_{\ell\ell'} \cdot \mathbb{1}_{\{\ell' \in \mathcal{B}_\ell\}} + \frac{1}{2} \cdot \mu^t_{\ell,\ell'} \cdot \mathbb{1}_{\{\ell \notin \mathcal{B}_\ell\}} + V^{t+1}\left(\vec{\mathcal{P}} \setminus \left(\vec{S} \cup \vec{R}\right), \vec{\mathcal{B}} \cup \vec{A} \setminus \vec{S}\right) \right] \right\}$$

$$V^{T+1}(\vec{\mathcal{P}}, \vec{\mathcal{B}}) = 0$$

where $\vec{A} = \{A_\ell\}_{\ell \in I \cup J}$ and $\vec{R} = \{R_\ell\}_{\ell \in I \cup J}$ are such that A_ℓ and R_ℓ represent the sets of users that liked and disliked ℓ in that period, i.e.,

$$A_\ell = \left\{\ell' \in \mathcal{P}_\ell : \ell \in S'_\ell, \Phi_{\ell'\ell} = 1\right\}$$
$$R_\ell = \left\{\ell' \in \mathcal{P}_\ell : \ell \in S'_\ell, \Phi_{\ell'\ell} = 0\right\}$$

Note that the first term in the summation ($\mu^t_{\ell\ell'} \cdot \mathbb{1}_{\{\ell' \in \mathcal{B}_\ell\}}$) captures sequential matches, i.e., matches produced by cases (i) and (ii) mentioned above, while the second term ($\mu^t_{\ell,\ell'} \mathbb{1}_{\{\ell \notin \mathcal{B}_\ell\}}$) captures simultaneous matches generated produced by case (iii).[9] The latter term is multiplied by 1/2 to avoid double-counting simultaneous matches.

A.2 Complexity

Proof of Proposition 3.1. To show that the problem is NP-hard, we show that we can reduce the set covering problem (SCP) to ours. An instance of SCP consists of a set of elements $\mathcal{U}$, a collection $C = \{C_1, \ldots, C_m\}$ such that $C_i \subseteq \mathcal{U}$ for all $i \in [m]$ and $\bigcup_{i \in [m]} C_i = \mathcal{U}$, and an integer k. The decision problem is whether there exists a cover of size at most $k < m$, i.e., a subset $\mathrm{C} \subseteq C$ such that $|\mathrm{C}| \le k$ and $\mathcal{U} \subseteq \bigcup_{C \in \mathrm{C}} C$.

Given an instance of *SCP*, we construct the following instance of our problem. We consider two sides of the market, $I = \{1, \ldots, m\}$ and $J = \mathcal{U} \cup \bigcup_{i=1}^m D_i$, where D_i is a set of unique duplicates of the elements of C_i. Each user $i \in I$ has an initial set of potentials $\mathcal{P}_i = C_i \cup D_i$, while each user $j \in J$ is such that $\mathcal{P}_j = I$. For each $i \in I$, we assume that

$$\phi_{ij} = \begin{cases} 1 & \text{if } j \in C_i \\ \frac{1}{m \cdot \bar{c} \cdot |C_i|} & \text{if } j \in D_i \\ 0 & \text{otherwise} \end{cases}$$

where $\bar{c} = \max_{i \in [m]} |C_i|$. In contrast, we assume that $\phi_{j,i} = 1$ for all $j \in J$ and $i \in \mathcal{P}_j$. Finally, we assume that the assortment size of each user $i \in I$ is $K_i = |C_i|$, while the assortment size for each user $j \in J$ is $K_j = 1$, and we assume no initial backlogs, i.e., $\mathcal{B}_\ell = \emptyset$ for all $\ell \in I \cup J$.

We show that a set cover exists for the instance $(\mathcal{U}, C, k)$ if the number of matches obtained from the instance of the assortment problem described above is at least $|\mathcal{U}| + \frac{(m-k)}{m \cdot \bar{c}}$.

[8] In a sloght abuse of notation, we assume that the set operations in $V^{t+1}(\cdot)$ are component-wise, i.e., $\vec{\mathcal{P}} \setminus \left(\vec{S} \cup \vec{R}\right) = \{\mathcal{P}_\ell \setminus (S_\ell \cup R_\ell)\}_{\ell \in I \cup J}$ and $\vec{\mathcal{B}} \cup \vec{A} \setminus \vec{S} = \{\mathcal{B}_\ell \cup A_\ell \setminus S_\ell\}_{\ell \in I \cup J}$.

[9] This holds because it cannot happen that $\ell \in \mathcal{B}_{\ell'}$ and $\ell' \in \mathcal{B}_\ell$.

$\Rightarrow$ Suppose that there exists a set cover of size k, $\mathbf{C} = \left\{C_{i_1}, \ldots, C_{i_k}\right\}$. for convenience, let $I' = \{i_1, \ldots, i_k\}$ the set of indices that provide the cover $\mathbf{C}$. Then, we construct a solution for the assortment problem that achieves an expected number of matches of at least $|\mathcal{U}| + \frac{(m-k)}{m \cdot \bar{c}}$ as follows. Each user $i \in I'$ observes an assortment $S_i^1 = C_i$ in the first period, and an empty assortment in the next period. Each user $i \in I \setminus I'$ observes an assortment $S_i^1 = D_i$ in the first period, and an empty assortment in the second period. Finally, each user $j \in J$ observes an empty assortment in the first period and, in the second period, an assortment containing one of the users that saw them in the first period (if any), i.e., $S_j^2 \in \{i \in I : j \in S_i\}$. If multiple users in I see the same $j \in J$, then we assume that j sees each of them back with probability $1/\left|i \in I : j \in S_i^1\right|$.

First, note that the cardinality constraints are satisfied, since S_i^1 is either C_i or D_i, both of which have cardinality $|C_i| = K_i$, and users $j \in J$ see at most one profile in the second period. Moreover, note that all assortments $S_\ell^t \subseteq \mathcal{P}_\ell$ for all $\ell \in I \cup J$, and thus they are feasible. Second, since $\mathbf{C}$ is a cover, we know that $\mathcal{U} \subseteq \bigcup_{C \in \mathbf{C}} C$, and thus each element in $\mathcal{U}$ is seen by at least one user $i \in I$. Since $\phi_{i,j} = 1$ and $\phi_{j,i} = 1$ for $i \in I$ and $j \in C_i$, we know that each $j \in \mathcal{U}$ will be liked by someone in I', and that they will like back at least one of the users that saw them. Hence, all users in $\mathcal{U}$ will get one match with users in the set I'. Finally, note that all users $i \in I \setminus I'$ observe assortments $S_i^1 = D_i$, and like each of these profiles with probability $\frac{1}{m \cdot \bar{c} \cdot |C_i|}$. Then, these users are liked back with probability 1, since $\phi_{ji} = 1$ for all $j \in J,\ i \in \mathcal{P}_j$ and the fact that these duplicates are seen by at most one user. Hence, the users $i \in I \setminus I'$ obtain a expected number of matches equal to

$$\sum_{j \in S_i^t} \frac{1}{m \cdot \bar{c} \cdot |C_i|} \cdot 1 = |D_i| \cdot \frac{1}{m \cdot \bar{c} \cdot |C_i|} = \frac{1}{m \cdot \bar{c}},$$

where the first equality follows from $S_i^t = D_i$ and the second follows from D_i been a duplicate of the set C_i, and thus $|D_i| = |C_i|$. Hence, the expected number of matches generated by users in $I \setminus I'$ is $(m-k) \cdot \frac{1}{m \cdot \bar{c}}$, and thus the total expected number of matches is $|\mathcal{U}| + \frac{(m-k)}{m \cdot \bar{c}}$.

$\Leftarrow$ Suppose that there exists a solution to the assortment problem $\bar{S} = \left(\{S_i\}_{i \in I}, \left\{S_j\right\}_{j \in J}\right)$ that produces an expected number of matches of at least $|\mathcal{U}| + \frac{(m-k)}{m \cdot \bar{c}}$. First, note that the only way to obtain at least $|\mathcal{U}|$ matches is by covering all the elements of $\mathcal{U}$, since every user in J sees at most one profile (and thus belongs to at most one match), and every user in $J \setminus \mathcal{U}$ generates a match with probability at most $\frac{1}{m \cdot \bar{c} \cdot \underline{c}} \ll 1$, where $\underline{c} = \min_{i \in [m]} \{C_i\}$. Hence, we know that $\left\{S_i^1\right\}_{i \in I}$ is a cover of $\mathcal{U}$. It remains to show that we can find a k elements of $\left\{S_i^1\right\}_{i \in I}$ that are enough to cover $\mathcal{U}$. To see this, suppose that this is not the case, i.e., we need at least $k+1$ elements in $\left\{S_i^1\right\}_{i \in I}$ to cover $\mathcal{U}$. Then, there are less than $m-k$ users in I who see assortments containing exclusively duplicated elements, i.e., $\left|\left\{i \in I : S_i^1 \subseteq D_i\right\}\right| < m-k$. Then, the expected number of matches generated by these users is strictly less than $(m-k) \cdot \frac{1}{m \cdot \bar{c}}$. Finally, since every element in $\mathcal{U}$ generates at most one match in expectation, we conclude that the total number of matches is strictly less than $|\mathcal{U}| + (m-k) \cdot \frac{1}{m \cdot \bar{c}}$, which contradicts our hypothesis. □

B APPENDIX TO SECTION 4

Proof of Proposition 4.1. Suppose that there are n users on each side of the market, i.e., $I = \{i_1, \ldots, i_n\}$ and $J = \{j_1, \ldots, j_n\}$. In addition, suppose that $\mathcal{P}_i = J$ for every $i \in I$, $\mathcal{P}_j = I$ for every $j \in J$. Let us set the probabilities: $\beta_{ij}^1 = 1$ for $j = j_1$ and for all $i \in I$, $\beta_{ij}^1 = 1 - \varepsilon$ for all $i \in I$ and $j \neq j_1$, while $\beta_{ij}^2 = 0$ for all $i \in I \cup J$, $j \in \mathcal{P}_i$. In this setting, the greedy policy will chose $S_i^1 = \{j_1\}$ for every user i, and therefore only one match will take place in expectation. In contrast, an optimal solution is to assign $S_{i_k}^1 = \{j_k\}$, which leads to $1 + (n-1)(1-\varepsilon)$ matches in expectation. Then, the performance of the greedy policy is given by $1/(1+(n-1)(1-\varepsilon)) \to 0$, as $n \to \infty$ for ε sufficiently small. □

Proof of Proposition 4.2. Suppose that $|I| = 2n$, $|J| = 2$, that $\mathcal{P}_i^1 = J$ for every $i \in I$, $\mathcal{P}_j^1 = I$ for every $j \in J$, and that $\phi_{ij}^t = p$ while $\phi_{ji}^t = q$ for all $i \in I$, $j \in J$, and $t \in \{1,2\}$. Then, it is easy to see that the sequential perfect match policy leads to $4pq$ matches in expectation. On the other hand, consider the policy where: (i) In $t = 1$, $\{i_1, \ldots, i_n\}$ see j_1, $\{i_{n+1}, \ldots, i_{2n}\}$ see j_2, j_1 sees i_{2n} and j_2 sees i_1, and (ii) in $t = 2$, $\{i_1, \ldots, i_n\}$ see j_2, $\{i_{n+1}, \ldots, i_{2n}\}$ see j_1, j_1 sees any profile that liked her in $t = 1$, and same for $j = 2$. Given this policy, the matches (i_{2n}, j_1) and (i_1, j_2) happen with probability pq each. On the other hand, j_1 matches with someone in $\{i_1, \ldots, i_n\}$ with probability $(1-(1-p)^n)q$, and the same for j_2 matching with someone in $\{i_{n+1}, \ldots, i_{2n}\}$. Then, the total expected number of matches is $2pq + 2q(1-(1-p)^n)$, which is optimal for this instance. Then, the sequential perfect match policy achieves a performance of $4pq/(2q(p+1-(1-p)^n)) \to 2/(1+n)$ when $p \to 0$, and since $2/(1+n) \to 0$ when $n \to \infty$, we conclude the proof. □

Proof of Proposition 4.3. Consider the following instance: $I = \{1, \ldots, n\}$, $J = \{j\}$, $K_i = K_j = 1$, $\phi_{i,j} = 1/n$ and $\phi_{ji} = 1$ for all $i \in I$. If we only allow sequential shows, then any deterministic non-adaptive policy would have an expected value of at most $1/n$, since we are choosing pairs i, j in advance without looking at the backlogs. However, the optimal adaptive policy consists of: showing j to every $i \in I$ in the first stage, and if there is at least one agent $i \in I$ that likes j then show this agent to j in the second stage. The optimal value in this case is the probability that j is liked by at least one person, i.e., $1-(1-1/n)^n \to 1-1/e$. If we allow simultaneous shows, the ratio is still arbitrarily close to zero. □

C APPENDIX TO SECTION 5

C.1 Dependent randomized rounding

In the following, we describe the dependent rounding approach introduced in [Gandhi et al., 2006]. Let $x_i \in [0,1]^J$ be the fractional vector that we want to round and $\hat{x}_i \in \{0,1\}^J$ be the final output. Denote by $\tilde{x}_i \in [0,1]^J$ the current version of the rounded vector, where $\tilde{x}_i$ is initially x_i. The random choices in each step of the algorithm are independent of all choices made in the past. If $\tilde{x}_{ij} \in \{0,1\}$, we are done, so let us assume there exists at least one component of $\tilde{x}_i$ that is fractional, say $\tilde{x}_{i1} \in (0,1)$. Then, we have two cases:

(1) $\tilde{x}_{i1}$ is the only fractional component. In that case, we round $\tilde{x}_{i1}$ to 1 with probability $\tilde{x}_{i1}$ and to 0 with probability $1-\tilde{x}_{i1}$. Note that $\mathbb{E}[\hat{x}_{i1}] = \tilde{x}_{i1}$.
(2) There are at least two components with fractional values, say $\tilde{x}_{i1}$ and $\tilde{x}_{i2}$. Let ϵ and δ be such that: (i) $\tilde{x}_{i1} + \epsilon$ and $\tilde{x}_{i2} - \epsilon$ lie in $[0,1]$ with at least one of them lying in $\{0,1\}$; and (ii) $\tilde{x}_{i1} - \delta$ and $\tilde{x}_{i2} + \delta$ lie in $[0,1]$ with at least one of them lying in $\{0,1\}$. Observe that such $\epsilon > 0$ and $\delta > 0$ exist. The update $\tilde{x}_{i1}$ and $\tilde{x}_{i2}$ is as follows: with probability $\delta/(\epsilon+\delta)$ set $(\tilde{x}_{i1}, \tilde{x}_{i2}) \leftarrow (\tilde{x}_{i1} + \epsilon, \tilde{x}_{i2} - \epsilon)$ and with the complementary probability $\epsilon/(\delta+\epsilon)$ set $(\tilde{x}_{i1}, \tilde{x}_{i2}) \leftarrow (\tilde{x}_{i1} - \delta, \tilde{x}_{i2} + \delta)$.

The algorithm continues until all components lie in $\{0,1\}$. Since each step of the algorithm rounds at least one additional component, then we need at most $|J|$ iterations. Note that the order in which entries are chosen in case 2. above (whenever there are more than 2 fractional components) does not affect the properties of the rounding method: preservation of the marginal values, feasible output with probability 1 and negative correlation.

C.2 Missing proofs.

Proof of Lemma 5.3. Our ground set of elements is $\mathcal{E} = \vec{E}_I \cup \vec{E}_J$. The first partition consists of the following parts: $\mathcal{E}_\ell = \{a : a = (\ell, \ell') \text{ for every } \ell' \in \mathcal{P}_\ell\}$ for all $\ell \in I \cup J$. It is easy to check that $\mathcal{E} = \cup_{\ell \in I \cup J} \mathcal{E}_\ell$ and $\mathcal{E}_\ell \cap \mathcal{E}_{\ell'} = \emptyset$ for every ℓ, ℓ' such that $\ell \neq \ell'$. Finally, the budget for each part $\mathcal{E}_\ell$ is K_ℓ. Now, let us construct the second partition matroid. For every pair $i \in I$ and $j \in J$, we define a part $\mathcal{E}_{i,j}$ as the set $\{(i, j), (j, i)\}$. Indeed this forms a partition of $\mathcal{E}$. Finally, the budget for each part $\mathcal{E}_{i,j}$ is 1. □

Proof of Lemma 5.13. Consider edges in $\mathcal{E}$ as the ground set and the subset system $\mathcal{I}$ defined by the feasible solutions in Q^1, i.e., subset of edges and arcs satisfying the constraints in Q^1. Consider two subsets $A \subset B$, $A, B \in \mathcal{I}$ and $u \in \mathcal{E}$ such that $A \cup \{u\} \in \mathcal{I}$. Suppose u is of the form (i, j) for some pair $i \in I, j \in J$. Given the constraints in Q^1, it means that: (1) $(j, i), \{i, j\} \notin A$, because of the first constraint; (2) $|\{u' \in A : u' = (i, \ell)\}| + |\{u' \in A : u' = \{i, \ell\}\}| < K_i$, because of the second constraint. Therefore, in the worst case $(j, i) \in B$ and $|\{u' \in B : u' = (i, \ell)\}| + |\{u' \in A : u' = \{i, \ell\}\}| = K_i$, which implies that there exists $u' \in B \setminus A$ adjacent to i, so we have to remove at most two elements to include $u = (i, j)$. Suppose now that u is of the form $\{i, j\}$ for some pair $i \in I, j \in J$. Since $A \cup \{u\} \in \mathcal{I}$ and given the constraints in Q^1 we have the following: (1) $(j, i), (i, j) \notin A$, because of the first constraint; (2) $|\{u' \in A : u' = (i, \ell)\}| + |\{u' \in A : u' = \{i, \ell\}\}| < K_i$, because of the second constraint; (3) $|\{u' \in A : u' = (j, \ell)\}| + |\{u' \in A : u' = \{\ell, j\}\}| < K_j$, because of the third constraint. Therefore, in the worst case the second and third constraints are tight for B in nodes i and j, which implies that there exist $u', u'' \in B \setminus A$ adjacent to i and j respectively. If this is the case, we have to remove at most three elements, u', u'' and possibly (j, i) (if present in B). □

Before proving the next result, let us consider the following definition:

Definition C.1 (Laminar Matroid). A family $\mathcal{X} \subseteq 2^{\mathcal{E}}$ over a ground set of elements $\mathcal{E}$ is called laminar if for any $X, Y \in \mathcal{X}$ we either have $X \cap Y = \emptyset$, $X \subseteq Y$ or $X \subseteq Y$. Assume that for each element $u \in \mathcal{E}$ there exists some $A \in \mathcal{X}$ such that $A \ni u$. For each $A \in \mathcal{X}$ let $c(A)$ a positive integer associated with it. A laminar matroid $\mathcal{I}$ is defined as $\mathcal{I} = \{A \subseteq \mathcal{E} : |A \cap X| \leq c(X) \ \forall X \in \mathcal{X}\}$.

Proof of Lemma 5.15. First, let us defined our laminar family $\mathcal{X}$. For every pair $i \in I, j \in J$ consider $X_{i,j} = \{(i, j), \{i, j\}\}$, also for every $i \in I$ consider $Y_i = \{(i, \ell) : \ell \in \mathcal{P}_i\} \cup \{\{i, \ell\} : \ell \in \mathcal{P}_i\}$. Indeed, this is a laminar family, two sets of type X do not intersect and two sets of type Y also do not intersect. Sets of type X and Y intersect only if they correspond to the same $i \in I$ in which case $X_{i,j} \subseteq Y_i$. Finally, for every $X_{i,j}$ we have $c(X_{i,j}) = 1$ and for each Y_i we have $c(Y_i) = K_i$. Therefore, for this laminar family $\mathcal{S}$ and values $c(\cdot)$ we have that $\mathcal{I} = \{A \subseteq \mathcal{E} : |A \cap X| \leq c(X) \ \forall X \in \mathcal{S}\}$ coincides with feasible region Q^1. □

Proof of Proposition 5.16. Given a realized backlog B and a set of potentials P, define a bipartite graph with two sides U, V with $U = I \cup J$ and $V = B \cup \left\{(i, j) \in I \times J \ : \ j \in P_i, i \in P_j\right\}$, i.e., U contains the set of users and V the set of arcs that could be displayed in the second period. Let $E \subseteq U \times V$ be the set of edges. Then, a pair $(\ell, (\ell', \ell'')) \in U \times V$ belongs to E if and only if

$$(\ell', \ell'') \in B_\ell \quad \text{or} \quad [(\ell', \ell'') \notin B_\ell, \ell \in \{\ell', \ell''\}] .$$

In words, an edge between $\ell \in U$ and $(\ell', \ell'') \in V$ exists if and only if the edge (ℓ', ℓ'') is either in the backlog of ℓ or both users ℓ and ℓ' can see each other simultaneously. Now, for any pair of nodes $(u, v) \in E$ such that $v \in B$, we define a variable $y_{u,v}$ that is equal to 1 if $v = (u, u') \in B_u$ and u sees u', and zero otherwise. Similarly, for any pair $(u, v) \in E$ such that $v = (u, u') \in V \setminus B$, we define a variable $x_{u,v}$ that is equal to 1 if u sees u', and zero otherwise. Note that here we do a slight abuse of notation and assume that if $x_{i,(i,j)} = 1$, then $i \in I$ sees profile $j \in P_i \setminus B_i$ and, similarly, if

$x_{j,(i,j)} = 1$, then $j \in J$ sees $i \in P_j \setminus B_j$. For convenience, we will use in these cases that $i \in (i,j)$ and $j \in (i,j)$. Finally, for any pair $(u,v) \in E$ such that $u \in I$ and $v \in V \setminus B$, let $w_{u,v} = 1$ if both users involved in v see each other simultaneously, i.e., if $v = (i,j) \in V \setminus B$, then $x_{i,(i,j)} = x_{j,(i,j)} = 1$.

Using these variables, we can formulate the second period problem as follows:

$$
\begin{aligned}
f(B) := \max \quad & \sum_{u\in U}\sum_{v\in B} y_{u,v}\cdot\phi_v + \sum_{i\in I}\sum_{\delta(i)\cap V\setminus B} w_{i,v}\cdot\beta_v && \text{(15a)}\\
st. \quad & \sum_{v\in\delta(u)\cap B} y_{u,v} + \sum_{v\in\delta(u)\cap V\setminus B} x_{u,v} \le K_u, \quad \forall u \in U && \text{(15b)}\\
& w_{i,(i,j)} - x_{u,(i,j)} \le 0, \quad \forall (i,j) \in V\setminus B, i \in I, u \in (i,j) && \text{(15c)}\\
& w_{i,(i,j)} - x_{u,(i,j)} \le 0, \quad \forall (i,j) \in V\setminus B, i \in I, u \in (i,j) && \text{(15d)}\\
& x_{u,v} \in \{0,1\}, \quad v \in V\setminus B, u \in v && \text{(15e)}\\
& y_{u,v} \in \{0,1\}, \quad \forall u \in U, v \in \delta(u)\cap B && \text{(15f)}\\
& w_{i,v} \in \{0,1\}, \quad \forall v \in V\setminus B, i \in v && \text{(15g)}
\end{aligned}
$$

where $\beta_v = \phi_v \cdot \phi_{\overline{v}}$ is the match probability between the users in the pair v and $\delta(u)$ is the set of edges incident to node $u \in U$. Let Q^2 be the set of constraints in (15b), (15c) and (15d). Note that each variable appears at most twice in Q^2, and that every time they appear they are multiplied by either 1 or -1. Thus, to show that the matrix of constraints if totally unimodular, it remains to show that the constraints in Q^2 can be separated in two subsets such that (i) if a variable appears twice with different signs, then the constraints belong to the same subset, and (ii) if a variable appears twice with the same sign, then the constraints belong to different subsets. Let Q_I^2 and Q_J^2 be the subsets of constraints of Q^2 involving $u \in I$ and $u \in J$, respectively. Then, observe that

- each $w_{i,(i,j)}$ appears in two constraints with the same sign ($w_{i,(i,j)} - x_{i,(i,j)} \le 0$ and $w_{i,(i,j)} - x_{j,(i,j)} \le 0$), but these constraints belongs to Q_I^2 and Q_J^2, respectively.
- each $x_{i,(i,j)}$ appears in two constraints with different signs ($\sum_{v\in\delta(i)\cap B_i} y_{i,v} + \sum_{v\in\delta(i)\cap V\setminus B_i} x_{i,v} \le K_i$ and $w_{i,(i,j)} - x_{i,(i,j)} \le 0$), but these constraints belong both to the same subset Q_I^2. Similarly, $x_{j,(i,j)}$ appears in two constraints with different signs, but both constraints belong to Q_J^2.

Hence, using Hoffman's sufficient condition, we conclude that the constraints in Q^2 can be written as the product of a totally unimodular matrix and our vector of decisions variables. Finally, since the right-hand sides of the constraints are integral, we conclude that the feasible region of the problem is an integral polyhedron, and thus we can solve its linear relaxation. □

Proof of Proposition 5.17. Let $\mathcal{Y}(B)$ and $\mathcal{Z}(B)$ be the sets of backlog and non-backlog pairs shown in the optimal solution of the second period problem. Consider the following example. Let $I = \{i_1, i_2\}$, $J = \{j_1, j_2\}$, and the following probabilities:

$$\phi_{i_1,j_1} = 1, \phi_{j_1,i_1} = \epsilon, \phi_{i_2,j_2} = \epsilon, \phi_{j_2,i_2} = 1, \beta_{i_1,j_2} = 1/2, \beta_{i_2,j_1} = 1/2.$$

If $B = \emptyset$, then $\mathcal{Z}(B) = \{(i_1,j_2),(i_2,j_1)\}$. If we add $b = (i_2,j_2)$ to the backlog, then $\mathcal{Z}(B \cup \{(i_2,j_2)\}) = \{(i_2,j_1)\}$ and $\mathcal{Y}(B \cup \{(i_2,j_2)\}) = \{(j_2,i_2)\}$. Hence,

$$f(B \cup \{(i_2,j_2)\}) - f(B) = 1 + 1/2 - (1/2 + 1/2) = 1/2.$$

On the other hand, if $B' = \{(j_1,i_1)\}$, $\mathcal{Z}(B') = \{(i_2,j_1)\}$ and $\mathcal{Y}(B') = \{(i_1,j_1)\}$. If we add $b = (i_2,j_2)$ to B', then $\mathcal{Z}(B' \cup \{(i_2,j_2)\}) = \{(i_2,j_1)\}$ and $\mathcal{Y}(B' \cup \{(i_2,j_2)\}) = \{(i_1,j_1),(j_2,i_2)\}$. Then,

$$f(B' \cup \{(i_2,j_2)\}) - f(B') = 1 + 1 + 1/2 - (1 + 1/2) = 1.$$

Hence, we have that

$$B \subset B' \quad \text{and} \quad f(B \cup \{(i_2, j_2)\}) - f(B) < f(B' \cup \{(i_2, j_2)\}) - f(B'),$$

so we conclude that $f(B)$ is not submodular. □

C.3 Proof of Theorem 5.18

Recall function f defined in (3) as

$$f(B) = \max_{\mathbf{x}^2 \in P_B^2} \left\{ \sum_{a \in \mathcal{E}} \phi_a^2 x_a^2 \right\},$$

and M_f defined for $\mathbf{x}^1$ as

$$M_f(\mathbf{x}^1) = \sum_B f(B) \cdot \mathbb{P}_{\mathbf{x}^1}(\mathcal{B} = B).$$

Recall also that this sum is over all possible backlogs, i.e., all edges that belong to $\vec{E}$. Note that $\mathbb{P}_{\mathbf{x}^1}(\mathcal{B} = B)$ "restricts" the valid backlogs that come from the first-period assortments $\mathbf{x}^1$, i.e., this probability will be zero for any B that contains an element $e \in \vec{E}$ with $x_e^1 = 0$. This means that effectively we are summing over subsets of $\text{supp}(\mathbf{x}^1) = \{e \in \vec{E} : \ x_e = 1\}$.

Lemma C.2. *Function M_f can be reformulated as*

$$M_f(\mathbf{x}^1) = \sum_{B \subseteq \vec{E}} f(B \cap \text{supp}(\mathbf{x}^1)) \mathbb{P}(\mathcal{B} = B),$$

where $\mathbb{P}(\mathcal{B} = B) = \prod_{a \in B} \phi_a^1 \prod_{\vec{E} \setminus B} (1 - \phi_a^1)$.

Note that in the result above the distribution of the backlogs does not depend on the first-stage assortments.

Proof. First note that for any $A \subseteq \text{supp}(\mathbf{x}^1)$ we have that any $B \subseteq \vec{E}$ such that $A = B \cap \text{supp}(\mathbf{x}^1)$ we have $f(B) = f(A)$.

$$\begin{aligned}
M_f(\mathbf{x}^1) &= \sum_{B \subseteq \vec{E}} f(B \cap \text{supp}(\mathbf{x}^1)) \mathbb{P}(\mathcal{B} = B) \\
&= \sum_{A \subseteq \text{supp}(\mathbf{x}^1)} \sum_{B \subseteq \vec{E} : B \cap \text{supp}(\mathbf{x}^1) = A} f(B \cap \text{supp}(\mathbf{x}^1)) \mathbb{P}(\mathcal{B} = B) \\
&= \sum_{A \subseteq \text{supp}(\mathbf{x}^1)} \sum_{B \subseteq \vec{E} : B \cap \text{supp}(\mathbf{x}^1) = A} f(A) \mathbb{P}(\mathcal{B} = B) \\
&= \sum_{A \subseteq \text{supp}(\mathbf{x}^1)} f(A) \sum_{B \subseteq \vec{E} : B \cap \text{supp}(\mathbf{x}^1) = A} \mathbb{P}(\mathcal{B} = B) \\
&= \sum_{A \subseteq \text{supp}(\mathbf{x}^1)} f(A) \prod_{e \in A} \phi_a^1 x_a \prod_{e \in \text{supp}(\mathbf{x}^1) \setminus A} (1 - \phi_a^1 x_a),
\end{aligned}$$

which is the description of M_f that we studied in the proof of Theorem 5.2 □

In the same way, we can redefine $M_{\tilde{f}}$ for $\tilde{f}$ defined in (12). Now, we are ready to prove our main result.

Proof of Theorem 5.18. For simplicity, we prove the result for the case when $K_\ell = 1$ for all $\ell \in I \cup J$, as we can always duplicate nodes accordingly. Consider an optimal solution $(\mathbf{x}^{1,\star}, \mathbf{w}^{1,\star})$ of Problem (10) and an optimal solution $(\tilde{\mathbf{x}}^{1,\star}, \tilde{\mathbf{w}}^{1,\star})$ of Problem (13), leading to OPT_1 and OPT_2 expected matches, respectively. When clear from the context, we will drop superindices to ease the notation and exposition.

Our goal is to lower bound the following ratio

$$\frac{\text{OPT}_1}{\text{OPT}_2} = \frac{\sum_{e\in E} \phi_{a_e}\phi_{\overline{a_e}} w_e + M_f(\mathbf{x})}{\sum_{e\in E} \phi_{a_e}\phi_{\overline{a_e}} \tilde{w}_e + M_{\tilde{f}}(\tilde{\mathbf{x}}, \tilde{\mathbf{w}})}.$$

Since like probabilities are time-independent, we can use Lemma C.2 to redefine both objectives to consider backlog distributions that are independent of the decisions in the first period. Namely,

$$\frac{\sum_{e\in E} \phi_{a_e}\phi_{\overline{a_e}} w_e + M_f(\mathbf{x})}{\sum_{e\in E} \phi_{a_e}\phi_{\overline{a_e}} \tilde{w}_e + M_{\tilde{f}}(\tilde{\mathbf{x}}, \tilde{\mathbf{w}})} = \frac{\sum_B \left(\sum_{e\in E} \phi_{a_e}\phi_{\overline{a_e}} w_e + f(B\cap \text{supp}(\mathbf{x}))\right) \mathbb{P}(\mathcal{B} = B)}{\sum_B \left(\sum_{e\in E} \phi_{a_e}\phi_{\overline{a_e}} \tilde{w}_e + \tilde{f}(B\cap \text{supp}(\tilde{\mathbf{x}}), \tilde{\mathbf{x}}, \tilde{\mathbf{w}})\right) \mathbb{P}(\mathcal{B} = B)} \tag{16}$$

To lower bound this ratio, we now proceed to lower bound the expected contribution of each pair $(i, j) \in I \times J$. Given that we are assuming one directional sequential shows going from I to J, the expected contribution of the pair (i, j) depends on the first-period decisions:

- If $x_{(i,j)} = 1$, then the contribution of the pair (i, j) in the numerator is

$$\Delta_{ij} = \phi_{ij}\phi_{ji} \sum_{B: T^*_{B,\mathbf{x}} \ni (j,i)} \mathbb{P}(\mathcal{B} = B)$$

  where $T^*_{B,\mathbf{x}}$ is an optimal solution for objective f of the second stage for backlog B and assortments $\mathbf{x}$. Δ_{ij} is the product of: the probability that i liked j, the probability that i was shown to j in the second stage and the probability that j liked i. Observe that if no agent $i' \in I$ such that $x_{(i',j)} = 1$ liked j, then i would be part of the optimal solution of the second stage. In other words, the event in which no one (except i) likes j implies the event of i being part of the optimal solution. Therefore,

$$\Delta_{ij} = \phi_{ij}\phi_{ji} \sum_{B: T^*_{B,\mathbf{x}} \ni (j,i)} \mathbb{P}(\mathcal{B} = B) \geq \phi_{ij}\phi_{ji} \prod_{\ell \neq i: x_{(\ell,j)}=1} (1 - \phi_{\ell,j}) \geq \phi_{ij}\phi_{ji} \left(1 - \frac{1}{n}\right)^n$$

  where the last inequality is due to our assumption.
- If $w_{\{i,j\}} = 1$, then the contribution of the pair i, j in the numerator is $\phi_{ij}\phi_{ji}$.

Now, let us compare the contribution of (i, j) to OPT_1 relative to its contribution to OPT_2. As before, we have different cases depending on the solutions $\tilde{\mathbf{x}}$, $\tilde{\mathbf{w}}$, and their second-period responses:

- If $\tilde{x}_{(i,j')} = 1$ for some $j' \in \mathcal{P}_i$ with $j' \neq j$ (when $j' = j$ is analogous). In this case, the platform shows j' to i in the first stage as sequential show instead of j, hoping to get an extra simultaneous match in the second stage. Therefore, the contribution of $\tilde{x}_{(i,j')}$ in OPT' would be potentially composed by two terms. First, between i and j' we have

$$\tilde{\Delta}_{ij'} = \phi_{ij'}\phi_{j'i} \sum_{B: T^*_{B,\tilde{\mathbf{x}},\tilde{\mathbf{w}}} \ni (j',i)} \mathbb{P}(\mathcal{B} = B)$$

  where $T^*_{B,\tilde{\mathbf{x}},\tilde{\mathbf{w}}}$ is an optimal solution for objective $\tilde{f}$ of the second stage with backlog B, $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{w}}$. In the worst case, there is also a simultaneous match between i and j in the second period, which contributes (in expectation) with

$$\tilde{\Delta}_{ij} = \phi_{ij}\phi_{ji} \sum_{B:T^*_{B,\tilde{\mathbf{x}},\tilde{\mathbf{w}}} \ni \{i,j\}} \mathbb{P}(\mathcal{B} = B).$$

In the worst case we have the following contribution in the denominator

$$\tilde{\Delta}_{ij'} + \tilde{\Delta}_{ij} = \phi_{ij'}\phi_{j'i} \sum_{B:T^*_{B,\tilde{\mathbf{x}},\tilde{\mathbf{w}}} \ni (j',i)} \mathbb{P}(\mathcal{B} = B) + \phi_{ij}\phi_{ji} \sum_{B:T^*_{B,\tilde{\mathbf{x}},\tilde{\mathbf{w}}} \ni \{i,j\}} \mathbb{P}(\mathcal{B} = B) \leq 2\phi_{ij}\phi_{ji},$$

where the inequality is due to two facts: (i) $\tilde{\Delta}_{ij'} \leq \phi_{ij}\phi_{ji}$, otherwise in solution $\mathbf{x}$ with value OPT_1 we could show j' to i instead of j and obtain a better solution, which would contradict the optimality of $\mathbf{x}$; (ii) $\sum_{B:T^*_{B,\tilde{\mathbf{x}},\tilde{\mathbf{w}}} \ni \{i,j\}} \mathbb{P}(\mathcal{B} = B) \leq 1$.

- If $\tilde{w}_{\{i,j'\}} = 1$ for some $j' \in \mathcal{P}'$; j' potentially different than j. In this case, the platform decided to show j' to i simultaneously in the first stage. Then the contribution in the worst case is the following

$$\tilde{\Delta}_{ij'} + \tilde{\Delta}_{ij} = \phi_{ij'}\phi_{j'i} + \phi_{ij}\phi_{ji} \sum_{B:T^*_{B,\tilde{\mathbf{x}},\tilde{\mathbf{w}}} \ni \{i,j\}} \mathbb{P}(\mathcal{B} = B) \leq 2\phi_{ij}\phi_{ji}.$$

where the inequality can be justified analogous to the case above: we must have $\phi_{j'i} \leq \phi_{ji}$ (recall both $\phi_{ij'}$ and ϕ_{ij} are at most $1/n$), otherwise if $\phi_{j'i} > \phi_{ji}$ then we can change the solution in $\mathbf{x}$ and get a better objective since j' would be part of the second-stage optimal solution at least every time that j is.

- If $\tilde{x}_{(i,\ell)} = \tilde{w}_{\{i,\ell\}} = 0$ for all $\ell \in \mathcal{P}_i$. This case is similar to the previous one, but now the contribution would be composed only by a second-period term $\tilde{\Delta}_{ij}$, which is no worse than the other cases.

Therefore, for any pair (i,j) the ratio between each contribution is at least

$$\frac{\Delta_{ij}}{\tilde{\Delta}_{ij'} + \tilde{\Delta}_{ij}} \geq \frac{\phi_{ij}\phi_{ji}\left(1 - \frac{1}{n}\right)^n}{2\phi_{ij}\phi_{ji}} \geq \frac{1}{2e} - o(1)$$

Finally,

$$\frac{\sum_B \left(\sum_{e\in E} \phi_{a_e}\phi_{\overline{a_e}} w_e + f(B \cap \mathrm{supp}(\mathbf{x}))\right) \mathbb{P}(\mathcal{B} = B)}{\sum_B \left(\sum_{e\in E} \phi_{a_e}\phi_{\overline{a_e}} \tilde{w}_e + \tilde{f}(B \cap \mathrm{supp}(\tilde{\mathbf{x}}), \tilde{\mathbf{x}}, \tilde{\mathbf{w}})\right) \mathbb{P}(\mathcal{B} = B)} \geq \min_{i,j,j'} \frac{\Delta_{ij}}{\tilde{\Delta}_{ij'} + \tilde{\Delta}_{ij}} \geq \frac{1}{2e} - o(1).$$

Note that in our analysis we could be comparing more terms than actually needed to be, but we were looking for a lower bound. To conclude the second part of the theorem, we observe that we can always obtain more matches (in expectation) when we allow simultaneous matches in both periods rather than in the first period only. We formalize this in Lemma C.3.

Lemma C.3. *Consider a feasible solution* $(\mathbf{x}^1, \mathbf{w}^1)$ *in* Q^1 *as defined in* (9)*, then*

$$M_f(\mathbf{x}^1) \leq M_{\tilde{f}}(\mathbf{x}^1, \mathbf{w}^1),$$

where M_f *and* $M_{\tilde{f}}$ *are part of the objective function of* (10) *and* (13)*, respectively.*

Therefore, if we consider $\mathbf{x}^1, \mathbf{w}^1$ a γ-approximate solution for Problem (10). Then, we have

$$M_{\tilde{f}}(\mathbf{x}^1, \mathbf{w}^1) + \sum_{e\in E} \beta_e w_e^1 \geq M_f(\mathbf{x}^1) + \sum_{e\in E} \beta_e w_e^1 \geq \gamma \cdot \mathrm{OPT}_1 \geq \gamma \cdot \left(\frac{1}{2e} - o(1)\right) \cdot \mathrm{OPT}_2.$$

□

D APPENDIX TO SECTION 6

D.1 Estimation of the like probabilities

To estimate the probability that each user i likes a profile $j \in \mathcal{P}_i^t$, we use a logit model with user fixed effects:

$$y_{ijt} = \alpha_i + \lambda_t + X'_{ij}\beta + \epsilon_{ijt}. \tag{17}$$

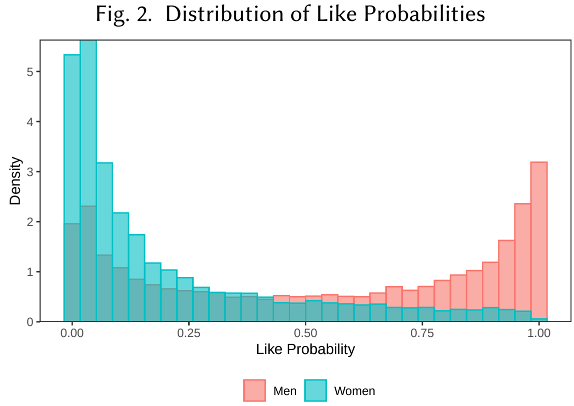
The dependent variable, y_{ijt}, is a latent variable that is related to the like decision Φ_{ij}^t according to

$$\Phi_{ij}^t = \begin{cases} 1 & \text{if } y_{ijt} > 0, \\ 0 & \text{otherwise.} \end{cases}$$

We control for users' unobserved heterogeneity by including user fixed-effects, α_i. We also control for period-dependent unobservables by including period fixed-effects, λ_t. The third term on the right-hand side, $X'_{ij}\beta$, controls for observable characteristics of profile j, and also for their interaction with user i's observable characteristics. Specifically, we control for the attractiveness score, age, height and education level (measured in a scale from 1 to 3) of the profile evaluated. In addition, for each of these variables we control for the square of the positive and negative difference between the value for the user evaluating and that of the profile evaluated. Finally, we also control for whether the users share the same race and religion. Finally, ϵ_{ijt} is an idiosyncratic shock that follows an extreme value distribution. In Table 4, we report the estimation results.

Using these coefficients, for each user i and profile $j \in \mathcal{P}_i^1$ we compute the probability ϕ_{ij}. In Figure 2 we plot the distribution of like probabilities separating by gender, estimated using the parameters from column (2) in Table 4. These are the probabilities we use on our simulation study.

Fig. 2. Distribution of Like Probabilities



D.2 Dating Heuristic (DH)

For each period $t \in [T]$, DH considers two steps:

Table 4. Estimation Results

	(1)	(2)
Batch size	-0.004***	-0.004***
	(0.0004)	(0.0004)
Score	0.832***	0.832***
	(0.014)	(0.014)
Score - Positive difference	0.012***	0.012***
	(0.003)	(0.003)
Score - Negative difference	-0.011***	-0.011***
	(0.003)	(0.003)
Age	-0.026***	-0.026***
	(0.004)	(0.004)
Age - Positive difference	-0.002***	-0.002***
	(0.0004)	(0.0004)
Age - Negative difference	-0.0009**	-0.0009**
	(0.0005)	(0.0005)
Height	0.055***	0.055***
	(0.008)	(0.008)
Height - Positive difference	0.002***	0.002***
	(0.0006)	(0.0006)
Height - Negative difference	-0.004***	-0.004***
	(0.0006)	(0.0006)
Education level	0.060**	0.061**
	(0.024)	(0.024)
Education level - Positive difference	-0.001	-0.0008
	(0.014)	(0.014)
Education level - Negative difference	-0.077***	-0.077***
	(0.016)	(0.016)
Share religion	0.078***	0.078***
	(0.013)	(0.013)
Share race	0.457***	0.458***
	(0.030)	(0.030)
User	✓	✓
Date		✓
Observations	396,226	396,226
Pseudo R^2	0.386	0.386

(1) Optimization: this step involves solving the following linear program:

$$
\begin{aligned}
\max \quad & \sum_{\tau=t}^{t+1} \sum_{\ell \in I \cup J} \sum_{\ell' \in \mathcal{P}_\ell^t} y_{\ell\ell'}^\tau \phi_{\ell\ell'}^\tau + \frac{1}{2} \cdot w_{\ell\ell'}^\tau \phi_{\ell\ell'}^\tau \phi_{\ell'\ell}^\tau \\
\text{s.t.} \quad & \sum_{\tau=t}^{t'} y_{\ell\ell'}^\tau \leq \mathbb{1}_{\{\ell' \in \mathcal{B}_\ell^t\}} + \sum_{\tau=t}^{t'-1} (x_{\ell'\ell}^\tau - w_{\ell\ell'}^\tau) \phi_{\ell'\ell}^\tau, && \text{for every } \ell' \in \mathcal{P}_\ell^t, \ell \in I \cup J, t' \in [t, t+1], \\
& \sum_{\tau=t}^{t+1} x_{\ell\ell'}^\tau + y_{\ell\ell'}^\tau \leq 1, && \text{for every } \ell' \in \mathcal{P}_\ell^t, \ell \in I \cup J, \\
& \sum_{\ell'} x_{\ell\ell'}^\tau + y_{\ell\ell'}^\tau \leq K_\ell, && \text{for every } \ell \in I \cup J, t \in [t, t+1], \\
& w_{\ell\ell'}^\tau \leq x_{\ell\ell'}^\tau, \ w_{\ell\ell'}^\tau \leq x_{\ell'\ell}^\tau, \ w_{\ell\ell'}^\tau = w_{\ell'\ell}^\tau, && \text{for every } \ell' \in \mathcal{P}_\ell^t, \ell \in I \cup J, t \in [t, t+1], \\
& x_{\ell\ell'}^\tau, \ y_{\ell\ell'}^\tau, w_{\ell\ell'}^\tau \in [0, 1], && \text{for every } \ell' \in \mathcal{P}_\ell^t, \ell \in I \cup J, t \in [t, t+1].
\end{aligned}
\tag{18}
$$

The decision variables $y^t_{\ell,\ell'}$ and $x^t_{\ell,\ell'}$ represent whether ℓ sees profile ℓ' in period t as part of a backlog and to initiate a sequential match, respectively. The objective is to maximize the expected number of matches obtained in periods $\{t, t+1\}$, including sequential matches (first term in the objective) and simultaneous matches (second term in the objective). The first constraint defines y and captures the evolution of the backlog. The second captures that a profile can be shown at most once, while the third constraint considers the assortment size. Finally, the last constraint captures the definition of $w^t_{\ell,\ell'}$, which accounts for simultaneous matches between ℓ and ℓ' in period t.

(2) Rounding: since the optimal decisions $x^{*,t}, y^{*,t}, w^{*,t}$ of (18) may be fractional, this step involves rounding them in order to decide the assortments to show in the current period. Specifically, the rounding process starts by adding to S^t_ℓ the profiles for which $y^t_{\ell,\ell'} > 0$ (in decreasing order). Then, if there is space left in the assortment, the rounding procedure add to S^t_ℓ the profiles for which $x^t_{\ell,\ell'} > 0$ (in decreasing order), making sure that the assortment size constraints are satisfied.

Notice that these two steps consider the current set of potentials $\mathcal{P}^t_\ell$ and backlog $\mathcal{B}^t_\ell$ for each user $\ell \in I \cup J$. Then, at the end of each period the sets of potentials and the backlogs must be updated considering the assortments shown and the like/dislike decisions. Specifically,

$$\mathcal{P}^{t+1}_\ell = \mathcal{P}^t_\ell \backslash \left(S^t_\ell \cup \left\{\ell' \in \mathcal{P}^t_\ell : \ell \in S^t_{\ell'},\ \Phi^t_{\ell'\ell} = 0\right\}\right), \quad \mathcal{B}^{t+1}_\ell = \left(\mathcal{B}^t_\ell \cup \left\{\ell' \in \mathcal{P}^t_\ell : \ell \in S^t_{\ell'},\ \Phi^t_{\ell'\ell} = 1\right\}\right) \backslash S^t_\ell.$$

This procedure in formally described in Algortihm 3.

Algorithm 3 Dating Heuristic (DH)

Input: An instance for a two-sided assortment problem.
Output: A feasible assortment.

1: **for** $t \in [T]$ **do**
2: Solve (18) and keep $x^{*,t}$, $y^{*,t}$.
3: For each user ℓ, sequentially add profiles ℓ' for which $y^{*,t}_{\ell,\ell'} > 0$ until the assortment size is reached. If there is space left in the assortment, add profiles ℓ' for which $x^{*,t}_{\ell,\ell'} > 0$ until the assortment size is reached.
4: Update potentials and backlogs.