ICIS 2019 Proceedings
Economics and IS

Your Preference or Mine? A Randomized Field Experiment on Recommender Systems in Two-sided Matching Markets

Lanfei Shi
University of Virginia, ls4tj@comm.virginia.edu

Kunpeng Zhang
University of Maryland, kzhang@rhsmith.umd.edu

Your Preference or Mine? A Randomized Field Experiment on Recommender Systems in Two-sided Matching Markets

Completed Research Paper

Lanfei Shi
University of Virginia
ls4tj@comm.virginia.edu

Kunpeng Zhang
University of Maryland College Park
kzhang@rhsmith.umd.edu

Abstract

The literature on recommender systems mainly focuses on product recommendation where buyer's preferences are considered. However, for user recommendation in two-sided matching markets, potential matches' preferences may also play a role in focal user's decision-making. Hence, we seek to understand the impact of providing potential candidates' preference in such settings. In collaboration with an online dating platform, we design and conduct a randomized field experiment and present users with recommendations based on i) their own preferences, ii) potential matches' preferences, or iii) mutual preferences. Interestingly, we find that users are sensitive to the provision of potential candidates' preferences, and they proactively reach out to those "who might prefer them" despite those candidates' relatively lower desirability. This leads to a greater improvement in matching. The findings provide valuable insights on how to design user recommendation systems beyond the current practice of recommendations based on focal user's preferences.

Keywords: Recommender systems, two-sided matching platforms, preference information, strategic behavior, randomized field experiments

Introduction

Peer-to-peer two-sided matching markets have become major players across many industries, e.g., labor markets, crowd-funding, and online dating. With the rapid growth of these markets, the choices for users expand exponentially exacerbating search frictions. Consequently, platforms resort to personalized recommender systems as one of the most effective approaches to improve the efficiency of search and matching. While researchers and practitioners generally focus on *product* recommendations in transactional markets, there is a dearth of research that studies *user* recommendations in two-sided matching markets.

User recommendation in two-sided matching markets differs from *product* recommendation in transactional markets due to some fundamental characteristics that distinguish the two types of markets. First, a match on a two-sided matching platform is a bilateral decision, as opposed to a purchasing decision in E-commerce, that eventually depends on the preferences of both sides of the markets - focal users on one side vs. potential matches on the other side (e.g., employers vs. employees in a labor market, or men vs. women in a dating market). Given this two-sided nature, focal users may make different choices when the recommendations are generated based on the other side's preference. Another distinction relates to the bandwidth issue of recommendations, especially those in high demand. In transactional markets, a popular item (say, a best-seller) can be recommended to multiple users. However, it is not ideal for matching markets to recommend the same popular candidate to many potential partners since only a few are likely to get a response. This congestion may lead to fewer matches for the platform as the whole.

Clearly, when designing *user* recommendations on two-sided matching platforms, more attention needs to be paid on the candidate pool regarding what preference information is used and how it impacts the platform as a whole. Yet, most online matching platforms provide recommendations similar to that of transactional platforms – their recommendations are largely based on the preferences of the focal user (Horton, 2017). It may not be the optimal practice, considering the differences between the two types of markets. Previous studies in Economics and Information Systems have provided some theoretical and empirical pointers to the potential benefits by including the other side's preferences. The findings suggest that providing the information about the other side's preferences can lead to strategic behaviors of the focal users and such provisioning is likely to improve matching outcomes (Avery & Levin, 2010; Bapna et al., 2016; Cole et al., 2013). However, there have been no field studies that empirically examine its implications for the design of recommender systems in two-sided matching markets.

We seek to bridge this gap and start by investigating how the usage of potential matches' preferences in recommender systems impacts focal users' decision-making. Specifically, we seek to understand what preference information should be used in recommender systems and how it impacts user decision-making and matching outcomes. From the perspective of choices and preferences, the research question can be viewed as in two-sided matching markets, whether and how people make different choices when the choice set includes or precludes the other side's preferences. From the perspective of information provisioning, we can think of the question as to whether and how users' decisions are affected when the other side' preferences are made available. The choice sets and the preference information presented to the focal users here are generated by recommender systems.

Three recommender systems are developed to examine how focal users respond differently. The first system uses focal users' preference ("Your Preferences"). The second system is based on potential matches' preferences ("Potential Matches' Preferences"). The third takes mutual preferences into consideration ("Mutual Preferences"). All other elements of the three recommender systems are held the same, including the feature set and the prediction model, to make comparisons meaningful. In other words, the three recommender systems only differ in the input data for candidate generation, by using the preference information of the different sides.

To examine the research question based on real user behavior, we collaborate with one of the leading online dating platforms in the U.S. and design and conduct a randomized field experiment. We choose online dating as the representative context not only due to its prevalence but also due to its flexibility to conduct randomized field experiments compared to other matching markets (Coles et al., 2010; Hitsch et al., 2010). Besides the three treatment groups that respectively implement the three recommendation systems, there is a control group with a baseline model. As opposed to the common baseline choice that shows a random list of users, we modify it to randomly show the top popular users to serve as a "higher" baseline. To avoid carry-over effects, we adopt a between-subject design that makes sure every subject is assigned only to one group of the experiment. Furthermore, the treatment contains two inseparable elements[1]: 1) recommendations generated by the assigned recommender system, which, compared to the recommender systems in other treatment groups, only differs in the preference information used as input, and 2) the associated title of the assigned recommender system to inform users what preference information is used to generate the recommendations. This study seeks to provide design implications for two-sided matching platforms on what preference information should be used, so the recommendation content needs to be truly based on different preferences rather than merely a manipulation of framings without changing the content. This is consistent with the literature on preference signals (Avery & Levin, 2010; Cole et al., 2013), which notes that the information has to 'transparent' to the focal users for it to be effective. Finally, users are also informed that the ordering of the candidates is based on the fitness of the designated preference. For instance, in "People who might prefer you" group, candidates on the top have a higher likelihood of preferring the focal users than the candidates on the bottom.

---

1 It would be interesting to separate the two elements and examine only one of them. However, these are beyond the interest and goal of this paper and can be pursued as future directions.

In examining the effects, we find that users are responsive to "the other side's preferences," which leads to both quantitative and qualitative impacts on the platform. Specifically, users in the "People who might prefer you" group and "Mutual Preferences" group are as proactive as users in the "Your Preferences" group; they all initiate more messages to the recommended candidates than in the control group. Interestingly, we further observe that focal users in the "Potential Matches' Preferences" group and "Mutual Preferences" groups tend to choose the candidates who are more likely to prefer them regardless of these candidates' desirability whereas people in "Your Preferences" group tend to seek highly desirable candidates. It is worth mentioning that in our study, users are not bounded by the limited recommendation choices; the platform provides a target search tool that ensures that every subject has the same opportunity to look for desirable partners. Interestingly, given the equal search access to all users on the platform, the qualitative difference of message receivers across groups only happens among those recommended candidates, not the candidates from search results. It further assures that those focal users in "Potential Matches' Preferences" and "Mutual Preferences" are not switching to lower desirable users in general but are responding to the recommended candidates who are more likely to "prefer" them. Our results indicate that users value the other side's preference and act upon it when such information is available.

The positive effect on message initiation by using "the other side's preference" is further amplified in the examination of responses and matches as the message receivers in "Potential Matches' Preferences" and "Mutual Preferences" groups respond more to the message proposals. We find that while "Your Preferences" group receives more responses than the control group, "Potential Matches' Preferences" and "Mutual Preferences" groups even outperform "Your Preferences" group. Therefore, providing the other side's preferences ("Potential Matches' Preferences" and "Mutual Preferences") does lead to more matches than only considering the focal user's preferences ("Your Preferences"). Such an increase may result from the novelty and diversity of choices generated by "Potential Matches' Preferences" and "Mutual Preferences" recommender systems, motivating focal users to explore and finally convert to matches. It may also be likely that users react strategically to the newly added information of the other side's preference; they tend to maximize the replies and matches. Further, there are heterogeneous effects wherein users who search broadly benefit more from "People who might prefer you" and "Mutual Preferences" recommendations, whereas "Your Preferences" benefits users who search narrowly.

This study contributes to the literature in several ways. First, to the best of our knowledge, it is among the first to examine the design and impact of *user* recommendation in a two-sided matching market that is fundamentally different from product recommendations. Second, it extends the literature on preference information disclosure and preference signaling to a new setting where the preference on the other side is provided by the platform's recommendation system. Further, we are among the first to design and conduct a randomized field experiment to investigate user recommendations in a two-sided matching market. It allows us to observe users' real-world choices and matching outcomes. These findings provide valuable implications to two-sided matching platforms and highlight the significance of including the other side's preference in the recommender systems. Such inclusion not only helps to improve user engagement and matching outcomes but also potentially reduces the disproportionate focus on the most popular users due to the diversity of users that are recommended as potential candidates for matching.

The paper is organized as follows. First, prior work is reviewed to outline our contributions. The research context is then described to provide details on the online dating platform as a representative of the two-sided matching markets. It is followed by a discussion of our experimental design as well as the details of the recommender systems we deploy. Variables and results are presented, and we conclude with managerial implications.

Prior Literature

Our paper closely relates to three streams of research. The first two streams of work examine recommender systems from different perspectives; one from the business perspective of recommender systems on how they impact users and platforms, and the other from the technical perspective of optimal design of recommender systems. The third stream of research draws upon studies on preference information provision and preference signaling to serve as the theoretical underpinning for how focal users may make choices differently when the recommendations are generated using different preference information.

There is emerging literature in the domains of information systems and economics that examines how recommender systems change users and online markets. Researchers have been focusing on either the quantitative or the qualitative side of the impact. On the one hand, researchers have found a positive effect of recommender systems on sales (Fleder and Hosanagar, 2009; De et al., 2010; Oestreicher-Singer and Sundararajan, 2012). On the other hand, some studies investigate how recommender systems shape consumers' choices – whether the introduction of recommender systems leads to more fragmented or unified choices collectively. However, mixed results are reported in different markets and contexts. For instance, Brynjolfsson et al. (2011) find that recommender systems lead to an increase in sales diversity while Hosanagar et al. (2014) find that it leads to an increase in commonality in music choices. Moreover, several studies have shown the co-existence of an increase in diversity and an increase in commonality, albeit on different levels of analysis (Fleder and Hosanagar, 2009; Lee and Hosanagar, 2014). The existing studies along this line focus mostly on product recommendations in transactional markets. Given the distinctive market characteristics of two-sided matching markets, our study seeks to be among the first to examine user recommendation in this setting. We complement this line of research by examining both the quantitative and qualitative impacts on user's decision-making in a two-sided matching market.

In contrast to studying the business impact, studies from computer science focus on the performance of recommender systems algorithms. Some recent papers have proposed recommendation algorithms for matching problems (Pizzato, et al., 2010; Xia, et al., 2015). As our focus is to investigate when using the same algorithm, how different sources of preferences would impact users' choices and matches differently, we adopt an existing algorithm in Pizzato et al.'s (2010) to obtain established recommendation performance. From the design aspect, we make additional effort to reduce biases of favoring popular users and confounding factors of inferring preference. From the evaluation aspect, existing studies evaluate new algorithms using secondary data while we design and conduct a randomized online field experiment to observe users' real choices and matches.

Further, we draw from the emerging literature on preference disclosure and preference signaling in Economics and Information Systems to provide supporting evidence that incorporating the other side's preferences in the recommender systems may be beneficial to the users and the platform. There is empirical evidence that presenting a focal user with information regarding the preferences of another user tends to increase the chance of a match between the two (Avery & Levin, 2010; Bapna et al. 2016). Such provision of the other side's preferences serves as a weak signal that prompts focal users to proactively connect with potential matches. Some theoretical work also suggests that the focal users would be more likely to accept one's proposal if it comes along with a credible signal of preferences (Cole et al., 2013). These theoretical and empirical pieces of evidence support the fact that a focal user's decision-making may be affected by the awareness of the other side's preference in a matching market. However, in the existing studies, the preference signal sent to the focal users is directly from another user, e.g., a proposal or a profile visit. It is not clear how focal users react when the preference of the other side is based on predictions, and when the preference signal is sent by the platform. Our study, therefore, extends this line of literature with a relaxed condition of predicted preference information.

Methods and Data

Research Context

We collaborate with one of the leading online dating platforms in the U.S, which has more than 1 million registered users. As with most online dating websites, it offers the following features to users. First, users can create their own online profiles to introduce themselves. User profiles typically also include photos. Moreover, the platform offers a decentralized search tool wherein users can filter profiles by age, location, and other demographic attributes to find potential matches. These targeted search results can be sorted based on location distance or user tenure from registration. Users are able to browse others' profiles without limitations and at no cost. There is no personalization or recommendation on this platform before our experiment.

Experimental Design

Based on preference information from the two sides on the platform (focal users on one side vs. potential matches on the other side), we compare three recommendation algorithms that are based on 1) focal users' preference ("Your Preferences"), 2) the other side's preferences ("Potential Matches' Preference"), and 3) mutual preferences ("Mutual Preferences") respectively. Since we are interested in how information about the different preferences impact user's decision-making, we use the same features and the same recommendation model but only alter the input information by leveraging preferences from different sides of the market. The control group outputs the top popular users in a random order to create a better baseline group than the commonly used benchmark - "generating a random list of users" as it makes sure the provided options are of high quality. The recommendation system is newly added to the platform, and the targeted search function remains in use without any changes. We also make careful design considerations to account for other factors contributing to users' choices. As suggested by the literature on decision-making, the size of the choice set plays a role. Therefore, we fix the number of recommendations for all the four recommendation algorithms. We also limit this number to be a reasonable size (i.e., 100 users) because too many choices may increase the complexity in decision-making due to bounded rationality.

To inform users how the recommendations are generated, a title is provided. "Your Preferences" group shows "People you might prefer" while "Potential Matches' Preferences" group uses "People who might prefer you." "Mutual Preferences" group displays "People who you might prefer and who might prefer you," and control group says "System Recommendation." It is important to disclose this information to make sure that users are aware of whether or not the choices are incorporated with the other side's preference. Otherwise, the strategic behavior documented in the previous literature would not be induced. In addition, the users are also informed regarding the sorting of the recommendations in each group, which is based on how compatible these candidates are with the designated preference. Specifically, the candidates shown at the top in "Your Preference" group have a higher chance of fitting focal users' own preferences than those candidates at the bottom. The recommendations displayed at the top in "Potential Matches' Preferences" are more likely to prefer the focal users than those ranked at the bottom.

Following a between-subject design, we randomly assign users to one of the four groups. We focus on the users who have interacted with others to be able to extract their revealed preference. Once a user is assigned to a group, we always generate recommendations using the assigned recommendation system to assure each subject is exposed to only one treatment. The recommender system refreshes every day, so each user will get updated recommendations on a daily basis. New users with no historical data will get their recommendations once they start engaging on the platform.

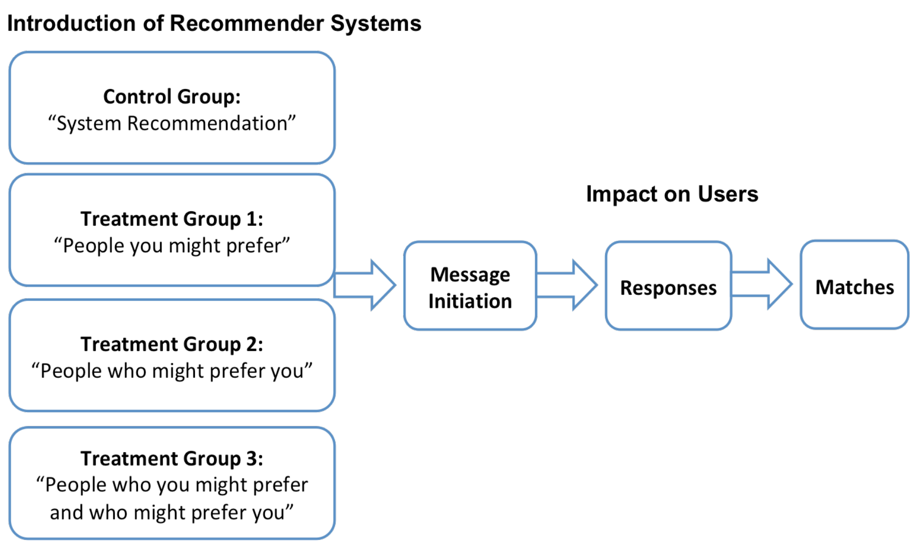

Figure 1. Experiment Design

Design of the Recommender Systems

Our interest in this study is not to design new algorithms but to leverage existing algorithms and investigate how different preference information would impact users' choices and matches differently. Researchers in computer science have used two types of models for matching problems; one is profile-based similarity ranking (Pizzato, et al., 2010) and the other is collaborative filtering (Xia, et al., 2015)[2]. With careful consideration, we decide to not use the collaborative filtering-based model as researchers have found that it tends to favor superstars (Lee and Hosanagar, 2014), which may potentially exacerbate the congestion among superstars in two-sided matching markets. Following the profile-based similarity ranking approach, we implement the recommendation algorithm based on Pizzato et al (2010)'s model, which looks for "similar" candidates who are compatible with the preference information based on user attributes on profile pages. The algorithm treats users with same profile attributes equivalent despite the fact that these users may differ in demand and may be considered differently using collaborative filtering. For details of the algorithm, please refer to the original paper (Pizzato, et al, 2010).

Specifically, the recommender system consists of three parts – input preference data, feature set, and model as shown in Figure 2. For the three recommender systems, we use the same feature set and recommendation model to make sure the only difference across "Your Preferences," "Potential Matches' Preferences" and "Mutual Preferences" recommender systems are the input preference data. Specifically, the input preference data in "Your Preferences" recommender system is extracted from those people who are visited or contacted by the focal users. The input preference data in "Potential Matches' Preferences" recommender system is extracted from those people who have initiated visits and messages to the focal users, and that in "Mutual Preferences" recommender system is extracted from the historical partners with back-and-forth visitation and conversations. We then extract features from these profiles to form preferences.

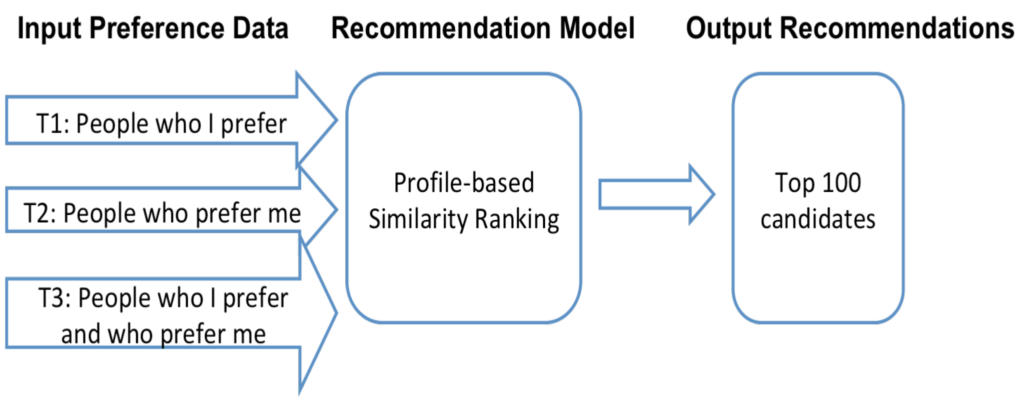

Figure 2. Components of Recommender Systems

The specific features we extract to represent one's preference are primarily based on the profile information since in general people rely on these profile attributes to make decisions. The features include the age difference between the focal user and potential candidates, location proximity, number of photos, income, education, length of self-introduction, and immigration status. We also include tenure-length as one feature as users, especially long-time users, are very familiar with all the other old users on the platform and thus they may pay more attention to new users.

---

[2] Other supervised machine learning techniques may also be applied to recommender systems in two-sided matching markets but, to the best of our knowledge, have not been studied in previous papers. At least in our context, there is data limitation that only positive cases are available (i.e., who likes whom) but no negative ones (i.e., who dislike whom) that make supervised models generally not feasible.

We create potential selection pool for each focal user using the active users over the last two weeks to make sure the potential candidates have been recently active, to maximize the response and engagement. For each user, we exclude the ones that they have visited over the last three months to create a customized selection pool for each user that aims to generate useful recommendations rather than redundant information. We calculate a compatibility score of each potential candidate within the selection pool and we output only the top one hundred compatible candidates for the focal user. The three recommender systems work in the same way and only differ in the source of the input preference data.

Although the focus of this paper is not to develop the best performing recommender systems for matching markets, we still seek to improve the existing Pizzato et al.'s model. Besides using the profile-based ranking to mitigate potential biases, the other improvements are listed as follows. First, while some prior work use stated preferences that are described by users in their profile, we mainly use the revealed preference based on historical behaviors of each user, which better reflect their true preferences. Along this line, we also carefully pre-process the historical information to pick only the initial visitation and messages between each pair as this indicates a strong preference compared to visiting back. Moreover, while existing studies evaluate new algorithms using secondary data, we design and conduct a randomized field experiment to observe the real choices of users. Finally, in order to ensure user engagement and observe how users use recommender systems, we update the recommendations on a daily basis.

Variables

As we focus on the impact on the focal users, we track their subsequent engagement behaviors upon receiving the experiment interventions. To obtain a comprehensive understanding on how different recommender systems may play a role in focal user's decision-making, we collect outcomes along the messaging funnel from message initiation to the other side's response and to the final match. We follow the previous literature in online dating to define matching as a three-round back-and-forth conversation since it indicates initial mutual interest of both sides (Bapna et al. 2016; Hitsch et al. 2010).

We are interested in not only the number of messages initiated by the focal users but also who the focal users send the messages to. The number of messages is a direct measure of user engagement to indicate the performance of the recommender systems while the qualitative aspect of these choices uncovers whether and how the focal users choose the candidates differently. These two dimensions working together provides us a better understanding of how the usage of different preference information in the recommender systems will impact the interaction and matching outcomes on the platform. We use *charm* to measure the overall desirability of each candidate, which is developed by the collaborating platform to track each user's popularity or demand.

We calculate these outcome variables within a certain time window denoted by *outcome_Xtime*. For instance, we focus on the outcomes within one week after treatment, so we calculate *msg_rec_1week*, *response_rec_1week*, and *match_rec_1week* to examine the messages initiated by the focal user, the messages responded by the message receivers, and the final matches formed between the focal users and the message receivers. We take a log transformation of these message counts. Further, as focal users can also use targeted search as an alternative way to identify potential candidate, we also look at the same series of outcomes initiated from the search. These outcomes in parallel serve as an additional check on how users are impacted by the introduction of recommender systems.

To further study the heterogeneous treatment effects, we link the experiment data to subjects' historical behaviors. We are particularly interested in categorizing the users based on their prior-experiment decision-making in searching. We segment the users based on search diversity – whether the users search broadly or narrowly. We speculate this may relate to their openness to explore "who are interested in me." For each user on the platform, we calculate the standard deviation of the charm scores of those who the user visits within two weeks prior to the experiment. We adopt a data-driven approach and choose the median of this distribution as the cutoff for "*broadness*." The subjects in the experiment are labeled as low or high in breadth depending on whether the value is below or above the threshold. The detailed coding of each variable is listed in Table 1.

Outcome Variables	**Description**
msg_rec_1week	the total messages sent to recommended candidates
response_rec_1week	the total responses of recommended candidates
match_rec_1week	the total matches of recommended candidates
msg_search_1week	the total messages sent to candidates from search
response_search_1week	the total responses of candidates from search
match_search_1week	the total matches of candidates from search
User Characteristics	**Description**
Charm	charm score based on popularity

Table 1. Individual-level Variable Description

Results and Discussion

Since the randomization is done at the focal user level, we use post-experiment individual-level data to run OLS regressions across experiment groups. We focus on the subsequent behaviors within one week after the treatment. Since gender difference has been noted in previous literature (Hitsch et al., 2010; Ravi et al., 2016), we block on gender in the randomization and run all the regression analyses for males and females respectively.

We first examine the quantity change across groups to see if the introduction of recommender systems leads to more message initiation from the focal users. Presumably, if the recommender system provides personalized choices that fit one's needs better, it should outperform the baseline algorithm even though we choose a relatively high baseline using top popular users with customization. As shown in Table 2, users in "Your Preferences" group on average initiate more messages than those in the control group, which further assures that the model and features in use work well in practice. More importantly, "Mutual Preference" and "Your Preference" groups also outperform the control group. There is no statistically significant difference in message initiation across the three treatment groups with different recommender systems, indicating that providing recommendations using the other side's preference has an equivalent scale of positive effect on the engagement of focal users.

VARIABLES	**Male** **msg_rec**	**Female** **msg_rec**
Your Preference	**0.0272***	**0.0238****
	(0.0155)	(0.0101)
Potential Matches' Preference	**0.0514*****	**0.0234****
	(0.0155)	(0.0101)
Mutual Preference	**0.0236***	**0.0336*****
	(0.0163)	(0.0106)
Constant	0.0725***	0.0231***
	(0.0109)	(0.00711)
Observations	5,559	5,196
p-value("Your"= "Potential")	0.119	0.968
p-value("Your"= "Mutual")	0.825	0.356
p-value("Potential"= "Mutual")	0.112	0.336

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2. The Number of Messages That Focal Users Initiate to Recommended Candidates

The result becomes even more interesting when we couple it with the qualitative analysis of the message receivers in each group, as shown in Table 3. *Chosen* is a dummy variable that documents whether a recommendation is chosen by the focal user for further contact. By comparing the chosen candidates with the unchosen candidates for each focal user across groups, we find that the average desirability or charm scores of the message receivers in "Your Preference" group are higher than the charm scores of the unchosen candidates. However, the charm scores of chosen candidates in "Potential Matches' Preference" group and "Mutual Preference" group are not always higher than the unchosen candidates. In other words, without the other side's information provided, focal users tend to pick the more desirable users from the list of recommended candidates. Yet when the matching side's preference is available, they value such information and are willing to choose those less desirable candidates who have a higher chance of preferring them. The increases in reaching out in "Potential Matches' Preferences" group and "Mutual Preferences" group are as significant as the increase in "Your Preference" group despite the fact that the increases are potentially driven by different mechanisms. Users in "Potential Matches' Preferences" and "Mutual Preferences" react on the access to candidates who are more likely to prefer themselves while the users in "Your Preferences" group become more proactive due to a good fit to their own preference. It is possible that users are curious about the novel and diverse choice sets generated by leveraging the other side's preferences, and browsing these profiles may lead to conversions to conversations. It is also possible that users tend to utilize the prediction of the other side's preference and act upon it to maximize the response rate.

VARIABLES	Male	Female
Your Preference	-0.0145**	-0.00628
	(0.00614)	(0.00416)
Potential Matches' Preference	0.00941	0.0117***
	(0.00610)	(0.00416)
Mutual Preference	0.00676	0.00313
	(0.00656)	(0.00442)
Chosen	0.0220***	0.0201***
	(0.00763)	(0.00576)
Your Preference & chosen	**0.0542*****	**0.0316*****
	(0.0114)	(0.00853)
Potential Matches' Preference & chosen	**-0.0492*****	**-0.0594*****
	(0.0111)	(0.00840)
Mutual Preference & chosen	**-0.0444*****	**-0.0375*****
	(0.0137)	(0.00954)
Constant	8.558***	7.854***
	(0.00431)	(0.00299)

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3. Desirability Comparison Between Chosen Candidates and Non-chosen Candidates across Groups

Furthermore, users are not bounded by the limited recommendation choices at all as they have a search tool to locate users they prefer. They have equal access to desirable partners with the same search cost using the generic target search. In other words, each user has full access to everyone on the platform using the search tool plus an additional subset of recommended candidates. The focal users contacting candidates with lower charm scores in "Potential Matches' Preferences" is not because users have no access to other more desirable candidates but because they intend to choose those candidates who may be "less popular" but are more likely to be "interested in themselves." As robustness checks, we further examine how the quantity and quality of message initiation using the search tool are affected at the meanwhile. As shown in Table 4, we find there is no statistically significant difference across treatment groups in both the numbers of initiated messages and the desirability of the message receivers. It means when the provided choices (e.g., from target search) do not contain the other side's preference, the focal users in "Potential Matches' Preferences" and "Mutual Preferences" groups make decisions similarly as their counterparts in "Your Preferences" group.

VARIABLES	**Male msg_search**	**Female msg_search**	**Male charm**	**Female charm**
My Preference	-0.00719	-0.000929	-0.0107	-0.0255
	(0.0208)	(0.0192)	(0.0176)	(0.0378)
Potential Matches' Preference	0.0167	0.00926	-0.00615	0.0188
	(0.0207)	(0.0192)	(0.0176)	(0.0377)
Mutual Preference	0.0251	0.00736	0.0114	0.0379
	(0.0218)	(0.0201)	(0.0185)	(0.0395)

Constant	0.261***	0.229***	8.026***	7.633***
	(0.0146)	(0.0135)	(0.0124)	(0.0266)
Observations	5,559	5,196	5,559	5,196
p-value (“Your”= “Potential”)	0.250	0.597	0.795	0.243
p-value (“Your”= “Mutual”)	0.140	0.681	0.234	0.110
p-value (“Potential”= “Mutual”)	0.700	0.925	0.344	0.629

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4. The Number of Messages That Focal Users Initiate to and the Desirability of Receivers Using Search

We further examine responses and matches along the messaging funnel. As shown in Table 5, we find that the positive effect of the recommender systems carries on, which leads to an increase in replies in all treatment groups. More importantly, “Potential Matches’ Preferences” and “Mutual Preferences” groups get even more responses than “Your Preferences” group. Similarly, we find the introduction of the recommender systems leads to an increase in final matches in all treatment groups, and “Potential Matches’ Preferences” and “Mutual Preferences” groups benefit from even more matches on average than “Your Preferences” group. Overall, providing the other side’s preferences (“Potential Matches’ Preferences” and “Mutual Preferences”) does lead to more matches than only using the focal user’s preferences (“Your Preferences”). This outcome gap in matching between “Your Preferences” and the other two groups is mainly driven by two aspects. Firstly, conditional on similar numbers of messages sent out, users in “Potential Matches’ Preferences” and “Mutual Preferences” groups are more likely to get a response, which plays an important role in the conversion of final matches. Secondly, the chosen candidates in “Your Preferences” group are more popular than those in “Potential Matches’ Preferences” and “Mutual Preferences” groups, and thus these candidates from “Your Preferences” group tend to have less bandwidth than those in the other two groups to deal with the extra incoming messages due to the introduction of recommender systems.

	Responses		**Matches**	
VARIABLES	**Male**	**Female**	**Male**	**Female**
Your Preference	**0.0120***	**0.00933****	0.00819	0.00566
	(0.00663)	(0.00396)	(0.00545)	(0.00349)
Potential Matches’ Preference	**0.0365*****	**0.0219*****	**0.0192*****	**0.0136*****
	(0.00661)	(0.00395)	(0.00543)	(0.00349)
Mutual Preference	**0.0259*****	**0.0188*****	**0.0199*****	**0.0134*****
	(0.00696)	(0.00414)	(0.00572)	(0.00365)
msg_rec_1week	0.0644***	0.118***	0.0575***	0.0869***
	(0.000861)	(0.00119)	(0.000708)	(0.00105)
Constant	0.0138***	-0.00255	0.00631	-0.00302
	(0.00467)	(0.00278)	(0.00384)	(0.00246)
Observations	5,559	5,196	5,559	5,196
p-value(“Your”= “Potential”)	<0.001	0.001	0.043	0.024
p-value(“Your”= “Mutual”)	0.046	0.023	0.042	0.036
p-value(“Potential”= “Mutual”)	0.131	0.446	0.906	0.956

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5. The Number of Responses and Matches That Focal Users Received from Recommendation

Finally, we investigate the heterogeneous effects of each recommender system on different user types to gain more insights on what user type would benefit the most from which recommender system. We do find that recommender systems have a differential impact on different users. We are particularly interested in segmenting users based on their search patterns. Specifically, the user type here is based on whether a user searches narrowly or broadly. Interestingly, as shown in Table 6, we find that users who search broadly have a significant increase in message initiation when offered with "Potential Matches' Preferences" and "Mutual Preferences" whereas "Your Preference" leads to a significant increase in message sending among users who search narrowly. This is consistent with the trend where users who search broadly are more likely to be more open-minded to candidates who are not typically their own "type" while people who are particular about choices and have a narrow search may tend to stick to their own preferences. In addition, given documented gender differences in online dating, it is worth mentioning that female and male users have a consistent pattern in response to each recommender system.

	Male		**Female**	
	Narrow	**Broad**	**Narrow**	**Broad**
VARIABLES	msg_rec	msg_rec	msg_rec	msg_rec
Your Preference	**0.0468****	-0.0139	**0.0546****	0.0230
	(0.0224)	(0.0414)	(0.0230)	(0.0154)
Potential Matches' Preference	0.0319	**0.251*****	0.0317	**0.0291***
	(0.0224)	(0.0418)	(0.0225)	(0.0154)
Mutual Preference	0.0193	0.0330	**0.0402***	**0.0432*****
	(0.0230)	(0.0436)	(0.0237)	(0.0160)
Constant	0.103***	0.0535*	0.0199	0.0326***
	(0.0158)	(0.0291)	(0.0159)	(0.0108)

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6. Messages That Focal Users Initiate to Candidates -Segmented by Search Broadness

Conclusion

User recommendation is often deemed as one of the keys to mitigate search friction and matching inefficiency in two-sided matching markets, but much less attention, both in industry and in academia, has been paid compared to product recommendations in transactional markets. With an emphasis on the fundamental characteristics of user recommendation in two-sided matching markets, our study seeks to fill this gap by starting at examining whether and how the provision of potential candidates' preference can positively impact users' decision-making and overall matching on the platform.

In collaboration with a leading online dating platform, we carefully design recommender systems with the same algorithm but only alter the preference information in use. We design and conduct a randomized field experiment to investigate how the recommender system using only the focal user's preference (i.e., "Users who you might prefer") plays a different role than the recommender systems using the other side's preference (i.e., "Users who might prefer you" and "Users who you might prefer and who might prefer you"). Very interestingly, we find that the focal users are willing to initiate messages to less desirable users than their counterparts when they are aware that these recommended candidates are likely to be interested in them. These focal users end up sending no fewer messages to these candidates "who may be interested in them" compared to their counterparts sending to those "who they may be interested in." Moreover, when it comes to responses and matches, the advantage of incorporating the other side's preference is further consolidated; the focal users in "Your Preference" group get a smaller increase in responses and matches than users in the other two groups. Clearly, users are sensitive and responsive to "the other side's preference" and value candidates who are likely to prefer them regardless of these candidates' desirability. It leads to a higher volume of matches since the message receivers in "Potential Matches' Preferences" and "Mutual Preferences" groups have a higher probability of responding. Further, these recommender systems display a differential impact on different users based on the diversity of their historical search. Users who search broadly are more responsive to "People who might prefer themselves" and "Mutual Preferences" while users who search narrowly are more interested in "People who we might prefer."

Our work contributes to the literature in several ways. First, we are among the first to acknowledge the fundamental characteristics of user recommender systems in two-sided matching markets and study the design and impact of user recommendations. The study extends the existing literature on the impact of recommender systems. Second, there is emerging literature that studies how the provision of the preference information from the sender will affect the decisions of the receiver, but there is no study examining the implications on recommender systems. Our findings, therefore, complement this line of work and add empirical evidence to a different setting where the preference information is prediction, and the preference signal is not directly sent by the sender. Finally, in terms of identification strategies, we are among the first to design and conduct a randomized field experiment to examine the impact of user recommendations in a two-sided matching market.

Our findings provide practical insights to the platform designers. The results suggest multiple benefits of incorporating the other side's preference into the recommender systems. Besides the greater volume of user engagement and final matches, more importantly, these recommender systems facilitate the discovery of "seemingly unusual" matches. Without any information on the other side, the focal user can only act on their own preferences and look for their preferred "types." However, with the other side's preference information available upfront, the focal users increase the efficiency of their search but also have access to a broader array of "types" they would not have reached out to otherwise. Driven either by curiosity or efficiency improvements, the focal users get a chance to learn about these candidates by browsing their profiles and talking to them, which in turn leads to more matching opportunities for the focal users. This also mitigates the overloading problem of superstars as more matches are discovered in this manner. Future work can examine how the different recommendation systems impact the user and effectiveness of targeted search mechanisms for different user segments.

References

Avery, C. and Levin, J.. 2010. "Early Admissions at Selective Colleges," *American Economic Review*, *100*(5), pp.2125-56.

Bapna, R., Ramaprasad, J., Shmueli, G. and Umyarov, A.. 2016. "One-way Mirrors in Online Dating: A Randomized Field Experiment," *Management Science*, *62*(11), pp.3100-3122.

Bruch, E.E. and Newman, M.E.J.. 2018. "Aspirational Pursuit of Mates in Online Dating Markets," *Science Advances*, *4*(8), p.eaap9815.

Brynjolfsson, E., Hu, Y. and Simester, D.. 2011. "Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales." *Management Science*, *57*(8), pp.1373-1386.

Coles, P., Kushnir, A. and Niederle, M.. 2013. "Preference Signaling in Matching Markets." *American Economic Journal: Microeconomics*, *5*(2), pp.99-134.

Fleder, D. and Hosanagar, K.. 2009. "Blockbuster Culture's Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity." *Management science*, *55*(5), pp.697-712.

De, P., Hu, Y. and Rahman, M.S.. 2010. "Technology Usage and Online Sales: An Empirical Study." *Management Science*, *56*(11), pp.1930-1945.

Hitsch, G.J., Hortaçsu, A. and Ariely, D.. 2010. "Matching and Sorting in Online Dating." *American Economic Review*, *100*(1), pp.130-63.

Horton, J.J.. 2017. "The Effects of Algorithmic Labor Market Recommendations: Evidence from a Field Experiment." *Journal of Labor Economics*, *35*(2), pp.345-385.

Lee, D. and Hosanagar, K.. 2014. "Impact of Recommender Systems on Sales Volume and Diversity."

Oestreicher-Singer, G. and Sundararajan, A.. 2012. "The Visible Hand? Demand Effects of Recommendation Networks in Electronic Markets." *Management science*, *58*(11), pp.1963-1981.

Pizzato, L., Rej, T., Chung, T., Koprinska, I. and Kay, J.. 2010. "RECON: A Reciprocal Recommender for Online Dating." In *Proceedings of the fourth ACM conference on Recommender systems* (pp. 207-214). ACM.

Xia, P., Liu, B., Sun, Y. and Chen, C.. 2015. "Reciprocal Recommendation System for Online Dating." In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*(pp. 234-241). ACM.