

Not All Matches Are Equally Valuable: An Online Experiment of Retention-Focused Recommendation in a Job-Matching Platform

Tatsuya Ute¹, Chiaki Ichimura¹ and Yuta Saito²

¹Wantedly, Inc., Tokyo, Japan.

²Hanjuku-kaso Co., Ltd., Tokyo, Japan.

Abstract

Recommender systems in two-sided matching platforms are commonly optimized for immediate engagement signals such as click-through rate, reply rate, or the total number of successful matches. However, in real-world marketplaces, maximizing matches alone may be misaligned with business goals such as user churn rate and platform revenue, especially when users with fewer matches are at substantially higher risk of churn. In this paper, we study a real job matching platform and show that users with very few recent matches are indeed much more likely to leave the platform, while additional matches for already successful users provide limited marginal value for retention. Motivated by this empirical finding, we formulate a retention-aware recommendation problem and implement a simple post-processing method that adjusts the baseline match-focused ranking to prevent user churn. Specifically, the implemented method gives a score boost to churn-risk users with the goal of increasing their likelihood of obtaining matches and improving retention. We evaluate this practical approach in an online experiment on a real job-matching platform. The treatment group showed directionally lower user churn than the control group, although the estimated effect was not statistically significant at conventional levels, while company-side churn showed no evidence of deterioration. To our knowledge, this is among the first online experimental studies to investigate retention-focused recommendation in a real reciprocal job-matching platform.

Keywords

Reciprocal Recommendation, Online Experimentation, User Retention, Objective-driven Recommendation

1. Introduction

Recommender systems in two-sided matching platforms are often optimized for immediate engagement signals such as click-through rate, reply rate, or the total number of successful matches [1, 2, 3]. This design is reasonable because these objectives are directly measurable and closely related to short-term platform activity. However, in real-world matching markets, optimizing only the total number of matches may not align with the platform’s business objectives. This is because additional matches generated for users who have already achieved sufficient matching success often contribute little to retention, whereas users with few matches remain at high risk of churn and may benefit much more from additional matches [4, 5, 6].

This paper is motivated by a case study of a job matching service, *Wantedly Visit*¹, where a recruiter views recommended candidates in the context of a specific job opening and may send scouting messages to selected candidates (Figure 1 illustrates the platform’s recommendation problem). A prior recommendation policy implemented on the platform ranked users for each company in descending order of predicted match probability estimated by a machine learning model. While this strategy was effective in maximizing the total number of matches, it also concentrated exposure on users who were already likely to match, leaving a substantial number of users with fewer matches underserved, a phenomenon that is broadly consistent with prior observations in marketplace recommendation and reciprocal recommendation [2, 7, 6]. As shown in Figure 2, our analysis indeed revealed that users with very few recent matches were substantially more likely to churn, whereas providing additional matches to users who have already obtained sufficient matches had only limited marginal value

for retention. This empirical observation suggests that the recommendation problem should not be framed solely as maximizing aggregate matches. Instead, the platform should explicitly account for *retention risk* and prioritize users in a high-risk churn region. In other words, the key objective is not simply to create one more match anywhere in the platform, but to allocate opportunities so as to reduce the number of users whose match count remains below a critical churn-risk threshold.

Based on this insight, we implement a simple retention-focused post-processing strategy on top of the baseline match-probability ranking. Specifically, for relevant users judged to be at risk of churn, the method gives a score boost, thereby allocating them greater exposure than they would otherwise receive under the baseline ranking. For other users, the baseline ranking is left unchanged. This lightweight intervention aims to improve user retention without sacrificing other key business metrics.

We conducted a large-scale online experiment on *Wantedly Visit*, a real job-matching platform. The treatment group exhibited directionally lower user churn than the match-maximizing baseline, although the estimated difference was statistically inconclusive at conventional levels, while we found no evidence of deterioration in company-side churn. To the best of our knowledge, this study provides one of the first online experimental investigations of retention-focused recommendation in a real reciprocal job-matching platform.

2. Related Work

Our work is related to several lines of research in recommender systems and marketplace design. First, our study is related to the literature on recommender systems for matching markets and reciprocal recommendation. In such environments, a successful interaction depends on the preferences and behaviors of both sides, and naive one-sided relevance optimization can produce inefficient or unbalanced outcomes. Prior work has shown that explicitly modeling both sides is essential for improving matching quality and

RecSys in HR '26: The 6th Workshop on Recommender Systems for Human Resources, in conjunction with the 20th ACM Conference on Recommender Systems, September 28–October 2, 2026, Minneapolis, MN, United States.

✉ tatsu@wantedly.com (T. Ute); ichimura@wantedly.com (C. Ichimura); saito@hanjuku-kaso.com (Y. Saito)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

¹<https://www.wantedly.com>

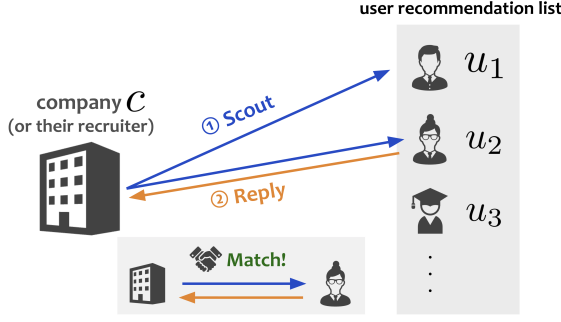


Figure 1: An example of the reciprocal recommendation setting considered in this study. The platform recommends a list of users for companies (or their recruiters), and companies then decide whether to send scouts to some of the users. Users who receive a scout then decide whether to respond. Users who receive a scout may in turn respond to it. Matches are recorded according to the platform’s operational definition. We omit platform-specific details of the reply interface and match definition, as they are not essential to the setting studied here.

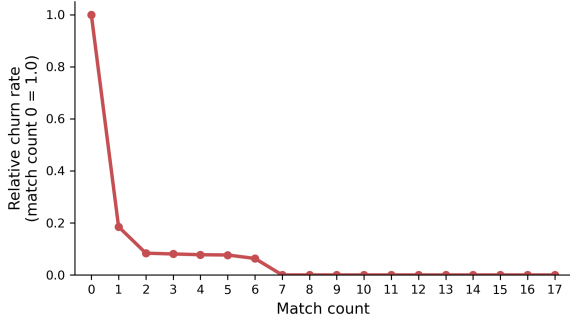


Figure 2: Relationship between users’ recent match counts and churn risk. Users with very few matches are much more likely to churn, whereas the marginal retention gain from additional matches is limited for users who have already obtained several matches.

marketplace health [2, 1, 3].

Second, our work is connected to long-term and retention-aware recommendation. While many industrial recommender systems optimize short-term engagement, prior studies have emphasized the importance of delayed outcomes such as satisfaction, continued participation, and long-term user value [4, 5, 8, 7, 6, 9]. Our approach belongs to this family of methods but differs in two respects. First, it emphasizes a simple operational intervention that prioritizes users whose recent match volume places them in a high-risk churn region, and second, it evaluates this intervention through an online experiment on a real two-sided matching platform.

Finally, our method is also related to exposure allocation and fairness in recommendation [10, 11, 12, 13, 14, 15, 16]. Methods such as FairRec [17] and FEIR [18] explicitly counteract concentration induced by relevance-only ranking by considering the distribution of exposure, utility, or competitive disadvantage across both sides of a marketplace. Our objective differs from these fairness-oriented approaches, i.e., rather than optimizing an exposure-parity constraint or another formal fairness criterion, we allocate a limited amount of additional exposure according to an empirically defined retention-risk signal.

3. Problem Formulation

We consider a two-sided job matching platform with a set of users $\mathcal{U}$ and a set of companies $\mathcal{C}$. For each user $u \in \mathcal{U}$ and company $c \in \mathcal{C}$, the system estimates a match probability

$$p_{\text{match}}(u, c) = p_{\text{scout}}(u, c) \cdot p_{\text{reply}}(u, c) \in [0, 1].$$

Here, $p_{\text{scout}}(u, c)$ denotes the probability that company c sends a scout to user u after the user is recommended, and $p_{\text{reply}}(u, c)$ denotes the conditional probability that user u replies to the scout from company c , given that a scout is sent. Their product $p_{\text{match}}(u, c)$ therefore represents the probability that recommending user u to company c leads to a successful match as operationally defined above.

In the standard formulation, including our previous recommender, the system ranks users for each company in descending order of $p_{\text{match}}(u, c)$ and optimizes the expected total number of matches,

$$\max_{\pi} \sum_{c \in \mathcal{C}} \sum_{u \in \pi(c)} p_{\text{match}}(u, c), \quad (1)$$

where $\pi(c)$ denotes the ordered list of recommended users for company (or recruiter) c .

Although this objective is reasonable from the perspective of short-term engagement, it does not capture user retention. Let $m_u(\pi)$ denote the number of matches obtained by user u under π in a predefined observation window. The empirical analysis for the target service, shown in Figure 2, suggested that users with very small values of $m_u(\pi)$ were substantially more likely to churn, whereas the marginal retention gain from increasing $m_u(\pi)$ for already successful users was limited. A similar trend was recently reported using real-world data from an online dating platform by Kishimoto et al. [6].³ Thus, in a real matching platform, there often exists a churn-risk threshold τ , defined as the minimum recent match count above which churn risk decreases substantially, such that users with $m_u(\pi) < \tau$ are considered to be at risk of churn.

This observation motivates a different optimization perspective compared to Eq. (1). Instead of treating every additional expected match equally, we seek to reduce the number of users whose match volume remains below the critical threshold τ . More specifically, let

$$\text{churn_risk}(u; \pi) = \mathbb{I}[m_u(\pi) < \tau]$$

be an indicator of whether user u is at risk of churn. Then, the platform-level objective can be formulated as minimizing the number of churn-risk users as follows.

$$\min_{\pi} \sum_{u \in \mathcal{U}} \text{churn_risk}(u; \pi) \quad (2)$$

The challenge is that the recommendation policy must strategically prioritize users who are at risk of churn to optimize this objective, even when they are not the users with the highest immediate match probability for companies. We examine this interesting trade-off between user-side churn and company-side satisfaction in the experiment section (Section 5).

²Depending on the platform implementation, c may represent either a recruiter evaluating candidates or a particular job opening. Whether c is defined at the recruiter or job-opening level does not affect our formulation, main contributions, or conclusions.

³Unlike our work, Kishimoto et al. [6] did not evaluate their retention-focused recommendation method through an online experiment on a real-world matching platform.

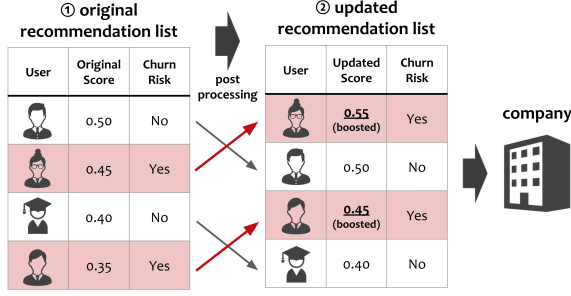


Figure 3: Illustration of the proposed retention-focused post-processing strategy. Starting from the original recommendation list produced by a baseline model, the method boosts the scores of users identified as being at risk of churn, while leaving other users unchanged. As a result, some churn-risk users move upward in the recommendation list after post-processing, increasing their chance of being exposed and obtaining a match.

4. Implemented Retention-Focused Strategy

This section describes our practical post-processing intervention for retention-focused recommendations. Rather than replacing the baseline recommender with an entirely new optimization procedure, we implement retention-focused recommendation as a lightweight post-processing step on top of the existing match-probability ranking model. This design choice keeps the method operationally simple and allows smooth deployment on the production system.

Let $s(u, c)$ denote the baseline recommendation score for user u and company c , which is often derived from the predicted match probability. Our method modifies this score only for users who are judged to be at risk of churn. Section 3 formulates retention risk in terms of the recent match count m_u . Since a match requires both a scout and a reply, the stage that the ranking most directly affects is scout receipt: it determines which candidates a recruiter sees. We therefore use s_u , the number of scouts a user received in the recent window, as the operational risk signal in the production system. A churn prediction model fit on historical log data with a range of pre-treatment behavioral features identified low recent scout receipt as a leading indicator of churn, and we accordingly treated users with fewer than three recent scouts as the high-risk region, $\mathcal{U}_{\text{risk}} = \{u \in \mathcal{U} \mid s_u < 3\}$. This cutoff was chosen from that analysis rather than through a formal breakpoint test, and the underlying relationship is predictive rather than causal; it should be interpreted as a platform-specific design choice rather than a universal churn threshold.

For users in $\mathcal{U}_{\text{risk}}$, we apply a score boost to give more exposure to those users. More specifically, let score_cutoff denote the minimum baseline score required to apply the boost, and let $n(u)$ denote the number of times user u has appeared in top-ranked positions (up to a predefined ranking cutoff) across rankings for all companies. Then, for each (u, c) -pair, we update the score as

$$s'(u, c) = \begin{cases} \text{clip}(s(u, c) + \beta \gamma^{n(u)}, -1, 1), & \text{if } u \in \mathcal{U}_{\text{risk}} \text{ and } s(u, c) > \text{score_cutoff}, \\ s(u, c), & \text{otherwise.} \end{cases}$$

where β is the base boost factor, γ is the diversity decay factor, and $\text{clip}(\cdot, -1, 1)$ truncates the final score to the valid

range. The values of β and γ can be selected through offline parameter tuning to optimize a predicted retention objective. After the post-processing step, users are ranked in descending order of the modified score $s'(u, c)$. Figure 3 illustrates the intuition of the proposed post-processing intervention.

This post-processing rule has three important properties. First, the boost is applied only to churn-risk users, reflecting our retention-oriented objective. Second, churn-risk users are not boosted indiscriminately; the boost is applied only when the predicted match probability for a (u, c) -pair is at least score_cutoff because boosting is of little value if the pair is unlikely to match. Third, the boost decays with the user’s overall exposure across companies, which prevents the method from excessively concentrating the boost on users who are already widely visible in the platform’s rankings. Thus, the method increases exposure for users who are at risk of churn but still have a reasonable likelihood of obtaining matches, while avoiding severe congestion in the recommendation space.

The main advantage of this approach is its practicality. Because the implemented method only adds a small adjustment to sufficiently promising users and leaves the rest of the ranking unchanged, it does not require retraining the underlying recommendation model or directly optimizing a complex retention objective. Instead, it introduces an interpretable and readily deployable intervention that strategically biases the ranking toward users with greater retention need.

5. Online Experiment

We conducted a large-scale online experiment on a recommender system implemented on *Wantedly Visit* to evaluate whether retention-focused recommendation can reduce user churn without sacrificing other key business indicators.

5.1. Experiment Design

The online experiment was conducted on the platform between November 10, 2025, and December 8, 2025. We randomized at the recruiter level (with each company potentially having multiple recruiters), resulting in approximately 3,000 recruiters per group. The analysis focused on users who had been active⁴ on at least one day during the 28-day pre-period (2025/10/13–2025/11/09) and received at least one impression from either a control or a treatment recruiter during the test period. Under this definition, the population consisted of approximately 20.9k user observations in the control-exposed group and 20.4k in the treatment-exposed group. Because randomization was performed at the recruiter level while the primary outcome was measured at the user level, a user could in principle be exposed to both groups. We therefore report a primary analysis allowing such overlap, as well as a robustness analysis restricted to users who were exposed to only one experiment group; in the latter sample, the analysis included approximately 6.6k control-only users and 6.0k treatment-only users.⁵

The baseline policy followed the conventional match-maximization strategy described in Eq. (1), whereas the

⁴A user is considered active on a given day if at least one login to the platform is recorded on that day.

⁵Note that the reported user and recruiter counts should be interpreted as approximate sample sizes for the AB-test analysis population, not as platform-wide totals.

Table 1

Effects of the implemented post-processing method on user churn. Bold values indicate estimates in the intended direction. None of the estimates is statistically significant at conventional levels.

Sample / subgroup	Effect (Odds Ratio)	<i>p</i> -value
Panel A: Main analysis		
Overlap allowed	0.957	0.128
Mutually exclusive only	0.948	0.223
Panel B: By prior number of received scouts		
0 scouts	0.956	0.200
1–2 scouts	0.948	0.359
3+ scouts	1.015	0.893
Panel C: By prior number of impressions		
0 impressions	0.845	0.139
1–5 impressions	0.903	0.095
6–20 impressions	0.973	0.564
21+ impressions	1.013	0.817

treatment policy applied the post-processing method presented in Section 4, which gives a score boost for churn-risk users while applying a smaller boost to users who already appear frequently in recruiters’ rankings. Our primary outcome was a binary churn flag, that is, among users who were active at least once during the pre-period, a user was considered to have churned if they did not become active at least once during the test period.

To estimate the treatment effect on user churn, we fit a user-level logistic regression with the treatment assignment indicator and 14 pre-treatment covariates. All continuous covariates were standardized before fitting the model. The covariates covered user attributes and preference settings, such as student status, scout acceptance preferences, and willingness for career change or side jobs; recent behavioral signals, such as the number of applications, matches, and received scouts; exposure volume, such as the number of ranking impressions; and activity-history variables, such as the number of days since the last active date. Because treatment was randomized, the coefficient on the treatment indicator can be interpreted as the average treatment effect up to the potential contamination induced by cross-group exposure, while the covariates improve estimation precision.

In addition to the churn outcome, we examined company-side guardrail metrics, including actual company churn and a model-based company churn score, to verify that improving user retention did not come at the expense of company-side outcomes. To the best of our knowledge, this is among the first online experiments that directly compare a retention-focused recommendation policy with a match-maximization policy in a real two-sided matching platform.

5.2. Results

This section reports and discusses the results of the online experiment.

Panel A of Table 1 summarizes the regression results for the user churn outcome. In both specifications, the estimated effect of the implemented post-processing method is in the desired direction: the odds ratio is below 1, indicating lower estimated odds of churn in the treatment group than in the control group. In the primary specification allowing overlap, the estimated odds ratio is 0.957, and in the mutually exclusive robustness sample, the odds ratio is 0.948. Although neither estimate is statistically significant

at conventional levels, the point estimates are directionally consistent across specifications. We therefore interpret these results as directionally positive but statistically inconclusive evidence regarding the effect of the intervention on user churn.⁶

We next conducted exploratory subgroup analyses to examine whether the estimated effects were directionally consistent with the intended target population (Panel B in Table 1). When users are grouped by the number of scouts received during the pre-period (Panel B in Table 1), the point estimates are more favorable among users with 0–2 prior scouts: the odds ratio is 0.956 for users with 0 prior scouts and 0.948 for users with 1–2 prior scouts, compared with 1.015 for users with 3 or more prior scouts. Panel C shows a similar descriptive pattern by prior impressions: the odds ratio is 0.845 for users with 0 prior impressions and 0.903 for those with 1–5 prior impressions, compared with 0.973 for users with 6–20 impressions and 1.013 for those with 21 or more impressions. These exploratory patterns are consistent with the intended targeting of underexposed users, but none of the subgroup estimates is statistically significant at conventional levels, and they should not be interpreted as evidence of treatment-effect heterogeneity.

We next examined overall matching outcomes. After covariate adjustment using Poisson regression with company-level controls, the point estimate for the total number of matches was 5.4% higher in the treatment group than in the control group. However, this difference was not statistically significant, and we therefore do not interpret it as evidence that the intervention increased overall matching performance. Importantly, the retention-focused intervention was not designed to increase the total number of matches, and the observed positive point estimate should be regarded as exploratory. One possible interpretation is that redistributing recruiter attention may alleviate congestion caused by repeated exposure of the same users and thereby affect conversion opportunities, consistent with prior work on congestion in job recommendation settings [21, 22].

We also assessed whether the proposed method harmed company-side retention. For this purpose, we analyzed companies whose contracts ended during the observation window among companies participating in the online experiment. Using weighted least squares with company-clustered robust standard errors, we found no evidence of deterioration in either actual company churn or a model-based company churn score: the estimated treatment–control coefficient difference was +0.006 for actual churn ($p = 0.77$) and -0.007 for the churn score ($p = 0.34$). Hence, within the precision of the current experiment, the user-side gains were not achieved at the expense of company-side outcomes.

6. Limitations

This study has several limitations. First, the experiment was conducted on a single job-matching platform over a four-week period, and the results may not generalize to other

⁶Our product decision criterion does not rely exclusively on the conventional $p < 0.05$ threshold. Recent work on experimentation programs argues that product decisions may also account for estimated effect sizes, opportunity costs, and other business considerations [19, 20]. Importantly, this decision criterion is distinct from statistical inference: the present estimates should be regarded as statistically inconclusive under conventional significance thresholds. We therefore separate our practical deployment decision from the strength of the statistical evidence reported in this paper.

marketplaces, recommendation interfaces, or longer-term retention outcomes.

Second, randomization occurred at the recruiter level, while outcomes were measured at the user level, and some users received exposure from recruiters assigned to both experimental arms. Although we report a mutually exclusive-user analysis as a robustness check, cross-arm exposure limits a clean user-level causal interpretation.

Third, the estimated effects on the primary churn outcome were not statistically distinguishable from the null with respect to $p < 0.05$, and the subgroup analyses were exploratory and insufficiently powered to establish treatment-effect heterogeneity. We also did not directly verify the mechanism through which the reranking intervention may affect retention. In particular, we did not separately establish whether boosted users received more effective exposure, scouts, or matches, nor whether changes in these intermediate outcomes mediated the estimated churn effect. Therefore, the observed churn estimates and the increase in aggregate matches should not be interpreted as causal evidence for a specific mediation pathway.

Fourth, the churn-risk threshold and score-boosting parameters were chosen from offline analyses rather than learned through a fully specified retention objective. Figure 2 is observational and does not strictly establish a causal effect of additional matches on churn.

Fifth, we compared the intervention only with the existing production baseline. We did not perform a head-to-head comparison with alternative congestion-aware or fairness-aware allocation methods that are relevant to our work; consequently, the experiment does not establish superiority over those approaches. Rather, the experiment is intended to illustrate the practical potential of the implemented retention-focused approach for a real matching platform.

Our intervention also raises fairness and transparency considerations. Because the method intentionally reallocates exposure toward users identified as being at risk of churn, some users with higher baseline match scores may receive less exposure than they would under a match-maximizing ranking, creating a potential trade-off among user retention, candidate-side exposure, and company-side utility. Although our company-side guardrail analysis found no evidence of deterioration in company churn, this does not by itself establish that the intervention is fair or that recruiters and users would perceive the policy as acceptable. We did not directly evaluate perceived fairness, transparency, or reactions to disclosure of the reranking policy; examining these aspects, as well as possible fairness-aware constraints, is an important direction for future work. Finally, privacy and commercial constraints limit the release of production logs and parts of the baseline model.

7. Conclusion and Future Directions

In this paper, we seek to design practical recommender systems that improve user retention rather than maximize total matches or satisfy fairness objectives. In the target platform, maximizing predicted matches alone was not well aligned with the business goal, because users with too few matches were more likely to churn. To address this issue, we formulated a retention-aware ranking problem and implemented a simple post-processing method that adjusts the baseline match-focused ranking for churn-risk users

through a small, controlled score boost. The proposed post-processing method yielded directionally lower user churn in the online experiment, although the estimated effect was statistically inconclusive at conventional levels. These results provide an initial industrial case study of retention-focused recommendation in a real reciprocal job-matching platform. More broadly, this case study suggests that marketplace recommenders should be carefully optimized for the true downstream objective rather than only for an immediate proxy like the number of matches.

As future work, we plan to extend the framework to optimize company retention as well as user retention, since company-side retention is more directly linked to platform revenue through mechanisms such as fees for job postings and user messaging. One possible approach is to apply a similar score-boosting intervention to systems that recommend companies or job postings to users. We also plan to improve the identification of users at risk of churn by incorporating a broader range of signals, as churn risk may depend on factors beyond the recent number of matches.

Declaration on Generative AI

During the preparation of this work, the authors used Claude Opus 4.8, Claude Opus 5, GPT-5.5 and GPT-5.6 Sol in order to: Grammar and spelling check, Paraphrase and reword, Drafting content. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

- [1] L. Pizzato, T. Rej, T. Chung, I. Koprinska, J. Kay, Reciprocal recommender system for online dating, in: *Proceedings of the 4th ACM Conference on Recommender Systems, RecSys '10*, 2010, pp. 207–214. doi:10.1145/1864708.1864787.
- [2] I. Palomares, C. Porcel, L. Pizzato, I. Guy, E. Herrera-Viedma, Reciprocal recommender systems: Analysis of state-of-art literature, challenges and opportunities towards social recommendation, *Information Fusion* 69 (2021) 103–127. doi:10.1016/j.inffus.2020.12.001.
- [3] C. Yang, S. Dai, Y. Hou, W. X. Zhao, J. Xu, Y. Song, H. Zhu, Revisiting reciprocal recommender systems: Metrics, formulation, and method, in: *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD '24*, 2024, pp. 3714–3723. doi:10.1145/3637528.3671734.
- [4] Q. Wu, H. Wang, L. Hong, Y. Shi, Returning is believing: Optimizing long-term user engagement in recommender systems, in: *Proceedings of the 26th ACM International Conference on Information and Knowledge Management, CIKM '17*, 2017, pp. 1927–1936. doi:10.1145/3132847.3133025.
- [5] Y. Wang, M. Sharma, C. Xu, S. Badam, Q. Sun, L. Richardson, L. Chung, E. H. Chi, M. Chen, Surrogate for long-term user experience in recommender systems, in: *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD '22*, 2022, pp. 4100–4109. doi:10.1145/3534678.3539073.
- [6] R. Kishimoto, R. Takehi, K. Tanaka, Y. Tomita, M. Nomura, R. Togashi, Y. Saito, Beyond match maxi-

- mization and fairness: Retention-optimized two-sided matching, in: *International Conference on Learning Representations*, volume 2026, 2026, pp. 80483–80501.
- [7] A. Acharya, S. Gao, B. Ocejio, K. Basu, A. Saha, K. Selvaraj, R. Mazumdar, P. Agrawal, A. Gupta, Promoting inactive members in edge-building marketplace, in: *Companion Proceedings of the ACM Web Conference 2023, WWW '23 Companion*, 2023, pp. 945–949. doi:10.1145/3543873.3587647.
 - [8] A. Acharya, S. Gao, A. Saha, B. Ocejio, K. Basu, K. Selvaraj, R. Mazumdar, Optimizing for member value in an edge building marketplace, in: *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management, CIKM '23*, 2023, pp. 5–14. doi:10.1145/3583780.3615000.
 - [9] D. Jannach, Multi-objective recommender systems: Survey and challenges, *CoRR* abs/2210.10309 (2022). URL: <https://arxiv.org/abs/2210.10309>.
 - [10] A. Singh, T. Joachims, Fairness of exposure in rankings, in: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '18*, 2018, pp. 2219–2228. doi:10.1145/3219819.3220088.
 - [11] B. Vassøy, A. Shekari, K. Verbert, Consumer-side fairness in recommender systems: A survey, *Artificial Intelligence Review* (2024). doi:10.1007/s10462-023-10663-5.
 - [12] Y. Tomita, T. Yokoyama, Fair reciprocal recommendation in matching markets, in: *Proceedings of the 18th ACM Conference on Recommender Systems, RecSys '24*, 2024. doi:10.1145/3640457.3688130.
 - [13] Y. Zhao, Y. Wang, Y. Liu, X. Cheng, C. C. Aggarwal, T. Derr, Fairness and diversity in recommender systems: A survey, *ACM Transactions on Intelligent Systems and Technology* (2024). URL: <https://arxiv.org/abs/2307.04644>, to appear / survey version available on arXiv:2307.04644.
 - [14] M. Zehlike, K. Yang, J. Stoyanovich, Fairness in ranking: A survey, *ACM Computing Surveys* (2022). doi:10.1145/3533379.
 - [15] A. Singh, T. Joachims, Policy learning for fairness in ranking, *Advances in Neural Information Processing Systems* 32 (2019).
 - [16] M. Morik, A. Singh, J. Hong, T. Joachims, Controlling fairness and bias in dynamic learning-to-rank, in: *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '20*, Association for Computing Machinery, 2020, pp. 429–438. doi:10.1145/3397271.3401100.
 - [17] G. K. Patro, A. Biswas, N. Ganguly, K. P. Gummadi, A. Chakraborty, FairRec: Two-sided fairness for personalized recommendations in two-sided platforms, in: *Proceedings of The Web Conference 2020*, Association for Computing Machinery, 2020, pp. 1194–1204. doi:10.1145/3366423.3380196.
 - [18] N. Li, B. Kang, J. Lijffijt, T. D. Bie, FEIR: Quantifying and reducing envy and inferiority for fair recommendation of limited resources, *ACM Transactions on Intelligent Systems and Technology* 15 (2024) 80:1–80:24. doi:10.1145/3643891.
 - [19] T. Sudijono, S. Ejdemyr, A. Lal, M. Tingley, Optimizing returns from experimentation programs, *arXiv preprint arXiv:2412.05508* (2024). URL: <https://arxiv.org/abs/2412.05508>. doi:10.48550/arXiv.2412.05508.
 - [20] W. Chou, C. Gray, N. Kallus, A. Bibaut, S. Ejdemyr, Evaluating decision rules across many weak experiments, in: *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2, KDD '25*, Association for Computing Machinery, 2025, pp. 4365–4374. doi:10.1145/3711896.3737217.
 - [21] Y. Mashayekhi, B. Kang, J. Lijffijt, T. D. Bie, Recon: Reducing congestion in job recommendation using optimal transport, in: *Proceedings of the 17th ACM Conference on Recommender Systems, RecSys '23*, Association for Computing Machinery, New York, NY, USA, 2023, pp. 696–701. doi:10.1145/3604915.3608817.
 - [22] Y. Su, M. Bayoumi, T. Joachims, Optimizing rankings for recommendation in matching markets, in: *Proceedings of the ACM Web Conference 2022, WWW '22*, Association for Computing Machinery, New York, NY, USA, 2022, pp. 328–338. doi:10.1145/3485447.3511961.