

Recommender systems and the value of user data^{*}

Gunhaeng Lee[†] Julian Wright[‡]

October 1, 2026

Abstract

Artificial intelligence (AI) increasingly shapes economic decisions through recommender systems, which personalize recommendations from user data. Yet, how the value from personalization scales with data remains unclear. We introduce a probability-based model designed to flexibly accommodate varying degrees of recommendation personalization. Using this model, we derive the value of recommender systems at different personalization levels, and decompose the value into customization, selection, and screening components. We also find a condition under which increased customization strictly benefits users. Our analysis reveals that deeper customization can yield non-diminishing returns and super-additivity. Empirical analyses using Bayesian predictions support our theoretical insights.

Keywords: recommender system, collaborative filtering, personalized recommendations, big data.

^{*}We thank Maximilian Schäfer, Juwon Seo and Tat-How Teh, as well as seminar participants at NUS, National Taiwan University, Academia Sinica, National Taipei University, SAET, APIOC, Baptist-Kyoto-NTU-Osaka-Sinica Workshop on Economic Theory. All errors are ours.

[†]Institute of Economics, Academia Sinica, glee@econ.sinica.edu.tw

[‡]Department of Economics, National University of Singapore, jwright@nus.edu.sg

1 Introduction

Artificial intelligence (AI) increasingly shapes economic decision making, with recommender systems emerging as a key example of how AI adds value to users. These AI-driven tools convert high dimensional user-item interaction data into personalized recommendations for products, media content, and services. Rather than peripheral features, recommender systems represent core engines of value creation for major platforms such as YouTube, TikTok, Instagram, Netflix, Spotify, and Amazon, directly influencing user choices and platform revenues. Despite their widespread adoption and economic significance, our understanding of how value is generated from and scales with the volume of data and customization remains limited, especially in settings with both within-user and across-user learning from large-scale data.

This paper addresses these gaps by proposing an economic model that captures how increases in data scale and algorithmic capacity translate into deeper customization, and how this personalization in turn generates value for users. The evolution of recommender systems illustrates a clear trajectory toward this goal. Early systems, limited by data processing capabilities, relied on simple metrics like popularity or aggregate ratings, which offered limited personalization. Over the last few decades, we have witnessed a steady, and at times radical, progression in recommender system technology. The move from naïve, popularity-based systems to more sophisticated methods like collaborative filtering marked a significant leap in personalization. More recently, the advances in generative AI models represent another step in this evolution, enabling even deeper levels of customization.¹ This expanded capability allows recommender systems to integrate diverse user behaviors and contextual features, creating significant value for users. For instance, music platforms can help a user discover novel songs matching their niche tastes, while online marketplaces can identify complementary products to complete a purchasing goal.²

In our stylized model, a platform aggregates historical user feedback to generate personalized probability estimates. While our framework can incorporate any observable user data, we focus specifically on user ratings for simplicity, as the model generalizes to other data forms through straightforward relabeling. The platform uses historical data to identify correlations across items and predicts the likelihood that a new user will like each item conditional on their past ratings on other items. User heterogeneity is represented in reduced form by the population joint dis-

¹For an extensive review comparing these paradigms of recommender systems, see Ayemowa et al. (2024) and Wu et al. (2024).

²For detailed insights, see Subbiah and Aggarwal (2024).

tribution of ratings: users are treated symmetrically *ex ante* as draws from a stable population, while persistent differences in tastes can be encoded in the correlation of a user’s ratings across items. A user’s observed history then provides information about where that user lies in the population preference distribution. Our recommendation framework closely mirrors the practical stages of modern systems: item retrieval, ranking based on predicted probabilities, and threshold-based filtering. Recommended items yield positive payoffs if tried and liked, negative payoffs if disliked, and zero otherwise.

Our study provides three main contributions. First, we introduce a tractable model that explicitly accommodates different degrees of customization. This flexible model allows for both theoretical analysis and empirical implications. Within this framework, we quantify the economic value created by personalization in recommendations, decomposing it into three distinct components: *customization*, *selection* and *screening*. Customization segments users based on their historical data; selection identifies the best item(s) to recommend from a pool of candidate items; screening determines if the selected item(s) are good enough to be recommended, filtering out suggestions that fall below a threshold level. We show that under the user-optimal threshold, additional customization weakly increases asymptotic expected user welfare. Conversely, for any interior threshold that differs from the user-optimal threshold, there always exist patterns in how items are correlated for which a nontrivial increase in customization strictly harms a user segment. We then derive a necessary and sufficient condition for an additional degree of customization to strictly increase user welfare.

Second, we characterize the learning curve as data grows along two distinct dimensions: breadth and depth. These dimensions play different roles. Greater breadth improves the precision of estimates for a fixed conditioning structure, whereas greater depth refines the information on which recommendations are conditioned and can raise the attainable value, while also spreading observations across more histories. We investigate how user welfare evolves as customization deepens. Unlike the breadth dimension, which yields diminishing returns in our fixed-customization benchmark, increasing data depth can produce non-diminishing and even increasing marginal returns. Moreover, the depth dimension can display super-additivity, where the gain from recommendations conditioned on multiple items surpasses the sum of gains from conditioning on each item individually, all measured relative to the same no-customization benchmark. We formalize the conditions under which this synergy occurs. Our simulations illustrate how these two dimensions can jointly create a learning curve composed of cascading, overlapping diminishing return curves. Increasing the breadth of data can raise user welfare at a diminishing rate while enabling deeper customiza-

tion, which can initiate a new, higher potential welfare curve. Consequently, as data expands along both dimensions, user welfare can progress upward through a series of these curves.

Our third contribution empirically illustrates and quantifies these mechanisms using a Bayesian probability estimation model applied to a dataset of over four million anonymous joke ratings from 73,421 users. Counterfactual analyses of the one-target-item screening comparison show that a rating system that does not offer customization increases simulated average utility by 30.4% (from 0.148 to 0.193) over one with no rating system, and customizing recommendations yields an additional 33.7% increase (to 0.258). These illustrative gains are conditional on our rating-to-utility calibration and evaluation sample. After examining how the marginal values of additional users and deeper customization vary with the available data, we assess the extent of complementarity or substitutability between these two dimensions of learning. We find the number of previous users exhibits strong complementarity with the degree of customization when both the number of users and the degree of customization are low, but this complementarity disappears quickly as the system accumulates data from more users. We also apply the framework to MovieLens 1M to compare value creation across data environments. Selection and screening remain important, while the complementarity–substitutability boundary shifts, indicating that the interaction between breadth and depth is data-dependent.

The remainder of the article proceeds as follows. Section 2 surveys the related literature. Section 3 formally defines the recommender system, develops the theoretical framework for analyzing its economic value, explores its implications, and presents key empirical results that illustrate our findings. Section 4 concludes. All proofs appear in Appendix A, and Appendix B summarizes the MovieLens results. Additional analyses are provided in the Online Appendix, with the full MovieLens analysis presented in Online Appendix F.

2 Related Literature

Our work contributes to a rapidly growing economic literature on artificial intelligence and recommender systems. It is arguably closest to Calvano et al. (2026), who study how collaborative filtering based recommendations on a platform affect various outcomes including the match between products and consumers, and the additional value generated for consumers. A key difference is that they incorporate how a platform’s recommendations can alter consumer search, and as a result, how competing firms price their products. This means market outcomes such as equilib-

rium prices and consumer surplus, which are their main focus, are endogenous in their setting. We abstract from endogenous market outcomes and focus just on the matching aspect of a recommender system. Methodologically, whereas Calvano et al. (2026) adopt a parametric taste-feature representation, we develop a non-parametric, probability-based model that works directly with the correlation structure between items and user histories. This lets us study theoretically how value scales with data depth as well as breadth, and it delivers conditions for non-diminishing returns and super-additivity from deeper customization.

Our distinction between data breadth and depth is related to the across-user and within-user learning studied by Hagiu and Wright (2023). They use reduced-form learning functions to analyze competitive advantage and data policies, focusing mainly on the two forms of learning separately, with a short extension combining them in special cases. In our setting, breadth captures learning from additional users, while depth captures the richer information about an individual user that supports customization. These inputs work together: data from previous users reveal how histories predict preferences, while the target user’s history allows the platform to apply these relationships to her. Our paper thus provides a statistical foundation for a joint across-user and within-user learning function in a setting of personalized recommendations. It explains how additional users improve estimation of population relationships, while richer histories permit finer customization and can raise attainable welfare. Hagiu and Wright’s joint formulation leaves unspecified whether the two learning inputs are complements or substitutes. Our numerical exercises illustrate how either pattern can arise depending on the available data. The recommendation technology helps explain this interaction: deeper customization offers greater scope for tailoring recommendations but also divides the available observations across more histories, making its value depend on the breadth of data available.

More broadly, our research connects to a considerable body of work examining how user data and firms’ algorithms influence market outcomes. Calvano et al. (2020) show in controlled simulations that Q-learning pricing agents can sustain supracompetitive prices via collusive strategies. Simulation studies of algorithmic pricing show that the design of learning protocols matters for market performance. Asker et al. (2023) demonstrate that different AI learning rules can lead to sharply different market prices. In contrast to these pricing-centric studies, our paper abstracts from prices and competition to isolate the intrinsic welfare value generated by deeper customization from user data. Beyond pricing, papers on learning consumer preferences and rankings study how algorithms infer heterogeneity from feedback signals; see, for example, Chen et al. (2018) and Feng

et al. (2022). On the other hand, contributions by Biglaiser et al. (2019), de Cornière and Taylor (2025), Fainmesser et al. (2023), and Aridor and Gonçalves (2022) consider equilibrium outcomes when firms use consumer data to enhance their offerings, with Aridor and Gonçalves (2022) explicitly addressing welfare impacts arising from platforms steering consumers toward proprietary products.

Finally, several recent empirical contributions explore the role and economic significance of recommender systems. Aridor et al. (2024) conduct a field experiment on a movie-recommendation platform, demonstrating that recommendations directly affect user beliefs about product quality and drive subsequent information acquisition. Relatedly, studies by Bajari et al. (2019), Peukert et al. (2023), and Yoganarasimhan (2020) quantify economic returns to data primarily through improvements in forecast accuracy. Particularly relevant to the distinction between data breadth and depth, Schäfer and Sapi (2023) investigate the impact of data size on search quality and find that both across-user and within-user learning significantly enhance search outcomes. Zielnicki et al. (2026) use Netflix data to decompose consumption differences into selection, exposure, and targeting, where selection refers to differences in baseline consumption propensities between targeted and average users rather than the recommender system’s choice among items.

3 Value of data

3.1 The model

A platform has $I \geq 1$ items to consider recommending to a target user and chooses an item to recommend to the user who can have either a positive or a negative experience with the item. Each of the I items is called a target item and we denote the set of target items by $\mathcal{I}$. The user has tried $C \geq 0$ items before and we call them conditioning items as the platform can customize its recommendation based on the user’s reported experiences with the C items. Although our analysis focuses on the scenario where conditioning items are past experiences, they can more generally represent any observable user features, such as demographics, geographic location, or browsing history. Thus, our model closely mirrors modern recommender systems that leverage diverse historical and contextual data for highly personalized recommendations.

In our framework, a user obtains a payoff of $v_1 > 0$ from a positive experience and $v_0 \leq 0$ from a negative experience with a target item. Similarly, the platform obtains w_1 and w_0 corresponding to the user’s positive or negative experiences. If the user does not try the item, both parties receive

a payoff of zero. The probability of either experience with a target item is initially unknown to the platform or to the target user. Therefore, the platform aims to estimate these probabilities based on accumulated user data before making a recommendation.

Data

The platform's data is represented by an $M \times N$ matrix, X , containing ratings from $N \geq 1$ users for M items. The first C rows represent ratings on conditioning items, and the remaining I rows correspond to target items. Thus, the set of conditioning items is $\{1, \dots, C\}$ and that of target items is $\mathcal{I} = \{C + 1, \dots, M\}$. Assume each item $i \in \{1, 2, \dots, C, C + 1, \dots, M\}$ has a finite number of possible ratings. For analytical convenience, we assume that the ratings for target items are binary (positive or negative). Some users may not have ratings on some of the M items. If a user has a missing rating of item i , we denote this by $\emptyset$. Thus, in the data X , a typical element x_{ij} records user j 's rating of item i , with $x_{ij} \in \{\emptyset, 0, 1\}$ for target items $i \in \mathcal{I}$. Without loss of generality, we reserve the first column of X to denote the target user's ratings. Naturally, we have $x_{i1} = \emptyset$, $\forall i \in \mathcal{I}$ given the target user's ratings of target items must be predicted.

Data X is said to be *complete* if we have $x_{ij} \neq \emptyset$, $\forall i, \forall j$ s.t. $i \leq C$ or $j \neq 1$. In other words, the platform has complete data if the only missing ratings are the target user's rating of the target items. On the other hand, if the condition for complete data is not satisfied, the data is said to be *partial*. For the welfare analysis below, we restrict attention to complete data.³

Statistical model

The platform estimates the likelihood of a target user having a positive user experience with each target item based on historical data. Formally, we define outcomes for each combination of conditioning items and a target item. An outcome is a vector of length $C + 1$ that records the user's ratings on C conditioning items and on a specific target item, i . The set of all possible outcomes corresponding to target item i is denoted by R^i .

Although the interpretation of an outcome vector depends on the target item i , the set of possible vectors is structurally identical across all target items. For notational simplicity, we can therefore refer to this common set of outcomes as simply R , letting the relevant target item be clear from the context. For example, consider a platform that conditions on the reported (binary)

³Online Appendix C derives the Bayesian prediction rule with partial data, whereas Online Appendix D analyzes finite-sample welfare with complete data.

rating of item 1 in making predictions about item 2 and item 3, i.e., $C = 1$ and $I = 2$. The main example we use throughout the article is represented in Table 1. The set of outcomes R in this case is given by $R = \{(1, 1), (1, 0), (0, 1), (0, 0)\}$.

	user 1	user 2	user 3	user 4	user 5
Rating on item 1	1	1	0	1	1
Rating on item 2	$\emptyset$	1	0	1	0
Rating on item 3	$\emptyset$	0	1	1	0

Table 1: Main example

The target user’s ratings on the C conditioning items can be represented by r' , which is a subvector of $r \in R$ whose length is C . We refer to r' as a history. The set of all such histories of length C is denoted by R' . When $C = 0$, we take $R' = \{\emptyset\}$. By slightly abusing notation, we denote the vector of dimension $C + 1$ created by adding 1 to the end of r' by $(r', 1)$. Similarly, $(r', 0)$ denotes the vector created by appending 0 to the end of r' . Each $r \in R$ corresponds uniquely to either $(r', 1)$ or $(r', 0)$ for some r' . For example, $r = (1, 0)$ is equivalent to $(r', 0)$ for $r' = (1)$.

We assume a stable population distribution and that users’ outcome vectors are independent draws from it. For each target item i , the outcome is governed by an unobserved probability vector $p^i = (p_r^i)_{r \in R}$, where p_r^i denotes the joint probability of a specific outcome r , representing the vector of ratings across the C conditioning items and target item i . We assume $p_r^i > 0$, for all i and r . This probability vector p^i is referred to as a *correlation structure* as it reveals how the user experience with the target item i and conditioning items are correlated with each other. Naturally, these probabilities satisfy $\sum_{r \in R} p_r^i = 1$. Additionally, the true probability associated with a given user history r' is denoted by $p_{r'}^i$, which equals the sum of probabilities for the corresponding positive and negative outcomes: $p_{r'}^i = p_{(r', 1)}^i + p_{(r', 0)}^i$. Because $p_{r'}^i$ is the marginal probability of the same conditioning-history event, it does not depend on the target item under consideration. Thus, by construction, $p_{r'}^i = p_{r'}^j$ for all $i, j \in \mathcal{I}$.

Interpretation of user heterogeneity

The vector p^i is a reduced-form population joint distribution of a randomly selected user’s history and experience with target item i . Users are symmetric ex ante, but may have different persistent tastes ex post. One possible microfoundation is a finite latent type θ , drawn once for each user,

such that

$$p_{(r',y)}^i = \sum_{\theta \in \Theta} \mu(\theta) \Pr(R' = r', Y_i = y \mid \theta).$$

The same θ may affect the user's experiences with items, so the type can be persistent; the conditional distribution can also accommodate idiosyncratic experience or reporting noise, provided the induced p^i satisfies the maintained full-support condition. We do not separately identify θ ; instead, we work directly with the induced population joint distribution. Additional users improve the estimation of this distribution when it is unknown, while greater depth can create value even if the distribution is known exactly, insofar as observed histories are informative about persistent differences in user tastes.

Recommender system

Given the correlation structure, the platform estimates the probability $z_i(r')$ that the target user will have a positive experience with target item i , conditional on their history r' .

$$z_i(r') = \frac{p_{(r',1)}^i}{p_{r'}^i}. \quad (1)$$

Assigning the same $z_i(r')$ to users with the same observed history reflects the information available to the system, rather than an assumption that those users have identical preferences. If histories are only weakly informative about latent tastes, $z_i(r')$ varies little across histories and customization creates correspondingly little value; richer histories allow finer differentiation.

Let $\hat{z}_i^N(r')$ be a pointwise estimator of $z_i(r')$ for a target user whose history is r' , based on data from $N - 1$ previous users. Additionally, the platform sets a threshold $\tau(r') \in [0, 1]$ that applies to users with history r' . If the estimated probability $\hat{z}_i^N(r')$ is below this threshold, the item i is deemed not suitable for the user with history r' , and therefore, the system does not recommend this item to the user. Formally, we define a recommender system as follows:

Definition 1 *A recommender system is a collection of functions*

$$\left\{ \left\{ \hat{z}_i^N(r') \right\}_{i \in \mathcal{I}}, \tau(r') \right\}_{r' \in R'}.$$

Given data X and the target user's history r' , one item is selected from $\arg \max_{i \in \mathcal{I}} \hat{z}_i^N(r')(X)$, with ties broken uniformly at random, and is recommended if $\max_{i \in \mathcal{I}} \hat{z}_i^N(r')(X) \geq \tau(r')$; otherwise, no item is recommended.

Note that the value derived from data is inherently linked to the specific recommender system employed. Different estimators within recommender systems can lead to varied inferences, user decisions, and subsequent learnings from identical data. Consequently, no single recommender system can capture the universal value that data might offer. Nevertheless, as the number of previous users N becomes very large, we can evaluate the value offered by a broad class of recommender systems termed *consistent recommender systems*. A recommender system is said to be consistent if, for each r' , the estimator $\hat{z}_i^N(r')$ is statistically consistent for $z_i(r')$. Formally, a consistent recommender system can be defined as follows.

Definition 2 *A recommender system is said to be consistent if $\hat{z}_i^N(r')$ converges in probability to $z_i(r') \in [0, 1]$ for each $i \in \mathcal{I}$ and $r' \in R'$.*

The existence of consistent recommender systems is guaranteed.⁴ Although consistency is an asymptotic property, it represents a fundamental requirement for reliable recommendation performance, as violating it would imply persistent prediction errors. In the following sections, we focus exclusively on consistent recommender systems, examining both the asymptotic value that such systems create for users and the incremental benefit arising from deeper customization. While user value from any consistent recommender system approaches its asymptotic limit with more data, finite-sample values are dependent on the specific system chosen.⁵ We detail the finite-sample implications for user value in Section 3.5, where we impose an additional unbiasedness condition on the estimator.

User behavior

We maintain that the target user follows the system’s recommendation: she tries the recommended item and does not try a target item when none is recommended. Suppose, hypothetically, the correlation structure is fully known to the platform. Because the user receives $v_1 > 0$ from a positive experience and $v_0 \leq 0$ from a negative experience, the user-optimal threshold is $\tau^u = \frac{-v_0}{v_1 - v_0}$. It is therefore optimal for the target user if the platform recommends an item with the highest probability of a positive experience whenever that probability is at least τ^u , and recommends nothing otherwise.

⁴We explicitly construct a Bayesian recommender system satisfying this consistency criterion in Online Appendix A.

⁵For fixed thresholds, convergence holds whenever the highest true success probability differs from the applicable threshold at each history. Under the user-optimal threshold, welfare convergence also holds at equality because the expected payoff from trying the item is zero.

On the other hand, the platform receives w_1 and w_0 from the user’s positive and negative experiences (with $w_1 > w_0$ and $w_1 > 0$), which yields the platform-optimal threshold $\tau^p = \max\{0, -w_0/(w_1 - w_0)\}$.

In the latent-type interpretation above, the user observes her history but does not directly observe θ or receive additional private signals about target-item payoffs. Our welfare analysis evaluates recommendation rules under the maintained-following assumption, rather than deriving optimal acceptance for arbitrary thresholds.

3.2 Value of data and its decomposition

We characterize and decompose the value created by personalized recommender systems. To make the decomposition explicit, we use the following terminology throughout. Customization (personalization) allows the predicted values and recommendation decision to depend on the target user’s history r' ; increasing C refines the partition of users on which the system conditions. Selection compares the target items and chooses the item with the highest predicted value for a given history. Screening withholds a recommendation when even the selected item fails to clear the threshold. These operations can be read directly from the first expression in Proposition 1: the summation over r' permits history-specific decisions, the maximization over $i \in \mathcal{I}$ captures selection, and the option 0 captures screening against not recommending any item. To isolate the gains from personalization, we contrast these systems with a simpler, non-personalized benchmark, which we refer to as a generic recommender system. This system relies solely on aggregate user data without personalizing recommendations based on individual user histories. Formally, the generic recommender system selects recommendations based on an estimated probability, $\hat{z}_i^G$ of the true probability $z_i^G = \sum_{r' \in R'} p_{(r', 1)}^i$, that an average user will positively experience an item, independent of any specific history.

Definition 3 *A generic recommender system is a recommender system $\{\{\hat{z}_i^G\}_{i \in \mathcal{I}}, \tau^G\}$ such that $\hat{z}_i^G$ is a consistent estimator for z_i^G .*

The generic system corresponds to early popularity or average-based rules that rely solely on across-user data and serve as a transparent benchmark for isolating the incremental value of conditioning. In contrast, the type of personalized recommender system we study represents a recommendation framework that leverages both across-user and within-user learning, providing a higher degree of customization. We analyze the ex-ante value created by these more sophisticated

personalized systems, averaging their expected benefits across all possible data realizations weighted by true probabilities.

Proposition 1 quantifies this value, separating it into two distinct components: (1) generic across-user learning and (2) additional gains from personalized recommendations. All the proofs are given in the Appendix.

Proposition 1 *1. For the user-optimal threshold, in any consistent recommender system, as $N \rightarrow \infty$, the expected utility to a user converges to*

$$\sum_{r' \in R'} \max_{i \in \mathcal{I}} \left\{ v_1 p_{(r',1)}^i + v_0 p_{(r',0)}^i, 0 \right\}.$$

2. For the user-optimal threshold, the expected utility created from a generic recommendation in the limit as $N \rightarrow \infty$ is

$$\max_{i \in \mathcal{I}} \left\{ \sum_{r' \in R'} \left(v_1 p_{(r',1)}^i + v_0 p_{(r',0)}^i \right), 0 \right\}.$$

Customization can be regarded as a finer segmentation of users according to observable features. To see why a finer segmentation leads to weakly higher asymptotic user welfare under the user-optimal threshold level, consider a history r' group of users whose average probability of liking a (unique) target item is given by z . Under the user-optimal threshold, the average utility of the history r' users is $\max\{v_1 z + v_0(1 - z), 0\}$. If a further segmentation of r' users into subgroups r'_a and r'_b is available to the system, the system makes two separate recommendations according to the two new average probabilities, z_a and z_b , of the two subgroups. Suppose a fraction p_a of r' users is assigned to r'_a subgroup. The expected utility of users under segmentation is then $p_a \max\{v_1 z_a + v_0(1 - z_a), 0\} + (1 - p_a) \max\{v_1 z_b + v_0(1 - z_b), 0\}$, where $z = p_a z_a + (1 - p_a) z_b$. The customization reveals the subgroup that is on-average better off trying (or not trying) the item, which in turn weakly increases the overall expected utility. In Figure 1, the expected utility without customization is zero while with customization it is $(1 - p_a)(v_1 z_b + v_0(1 - z_b))$.

The difference between the two expressions in the proposition is due to the different degree of customization each recommender system offers. A consistent recommender system makes history-dependent recommendations by selecting the best item for each group of history r' users. Instead, a generic recommender system only selects the item on average best for every user. Therefore, the value from generic recommendations in Proposition 1 can be attributed to pure across-user learning. On the other hand, the difference between the value from consistent recommendations

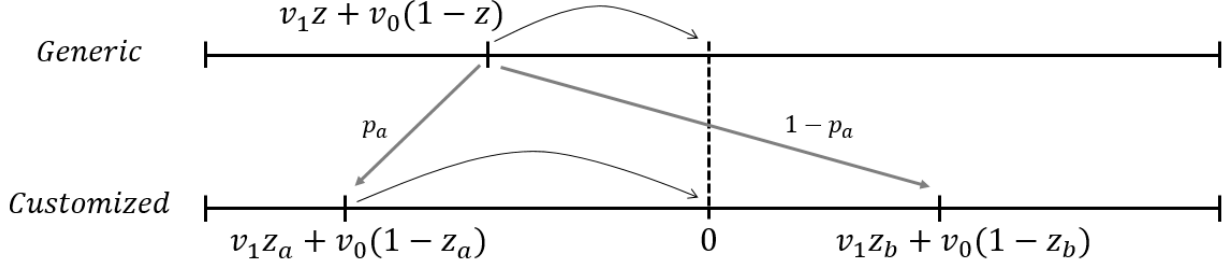


Figure 1: Customization under the user-optimal threshold

and the value from generic recommendations can be attributed to the customization benefits that the consistent recommendation provides. Under the user-optimal threshold, asymptotic expected utility is nonnegative under generic recommendations and weakly higher with customization. Thus, both pure across-user learning and customization contribute nonnegative value to users.

Both systems create value through selection and screening: they select the most suitable item and screen out items that do not meet the recommendation threshold. Under the user-optimal threshold, expanding the target-item pool weakly increases the asymptotic value created by either system. This threshold also maximizes the benefit from screening.

Example

To illustrate and better understand how a recommender system adds value through learning and customization, consider our main example in Table 1 and assume $(v_1, v_0) = (1, -1)$. In the example, the platform makes a recommendation between item 2 and item 3, or none based on the target user's rating on item 1. According to Proposition 1, when the threshold level is τ^u , the expected utility of a user under a consistent mechanism in the limit as $N \rightarrow \infty$ is given by

$$\underbrace{\max\{p_{(1,1)}^2 - p_{(1,0)}^2, p_{(1,1)}^3 - p_{(1,0)}^3, 0\}}_{\text{Recommendation for history 1 user}} + \underbrace{\max\{p_{(0,1)}^2 - p_{(0,0)}^2, p_{(0,1)}^3 - p_{(0,0)}^3, 0\}}_{\text{Recommendation for history 0 user}}. \quad (2)$$

Depending on the user's history, different items can be recommended. For example, if $p_{(1,1)}^2 - p_{(1,0)}^2 > \max\{p_{(1,1)}^3 - p_{(1,0)}^3, 0\}$ and $0 > \max\{p_{(0,1)}^2 - p_{(0,0)}^2, p_{(0,1)}^3 - p_{(0,0)}^3\}$, item 2 is recommended to users who had a positive experience with item 1 but no item is recommended to users with a negative experience with item 1. On the other hand, under a generic recommender system, which only makes use of across-user learning, users receive the same recommendation regardless of their history. The

value from a generic recommendation with the same threshold being applied is given by

$$\underbrace{\max \left\{ p_{(1,1)}^2 + p_{(0,1)}^2 - p_{(1,0)}^2 - p_{(0,0)}^2, p_{(1,1)}^3 + p_{(0,1)}^3 - p_{(1,0)}^3 - p_{(0,0)}^3, 0 \right\}}_{\text{No customization in recommendations}}. \quad (3)$$

It is clear that the value from customization is always weakly positive. On the other hand, it is strictly positive *if and only if* no single action—recommending a particular target item or recommending nothing—maximizes expected user utility for both history groups.

Asymptotic lower bound of prediction error

Our ex-ante analytical framework and explicitly introduced theoretical model of correlation structure enable us to derive a concise closed-form representation of the achievable asymptotic lower bound on prediction error for consistent recommender systems. It can also be shown that this lower bound of the error is weakly decreasing in the degree of customization. For a consistent recommender system $\{\{\hat{z}_i^N(r')\}_{i \in \mathcal{I}}, \tau(r')\}_{r' \in R'}$, let $\epsilon^m(X)$ be the number of wrong predictions across all target items for m target users. We treat recommending an item as a positive prediction and not recommending it as a negative prediction. False positives occur when a recommended item would be disliked, and false negatives occur when a non-recommended item would be liked, including when another item is recommended. The expected prediction error is defined to be $\mathbb{E}[\epsilon^m(X)/m]$, the expected average number of errors per user. The following corollary establishes the sharp lower bound of the prediction error, where the bound is attained asymptotically by any consistent recommender system with $\tau(r') = 1/2$ for every history.

Corollary 1 (Corollary to Proposition 1) *For $C \geq 0$ and $I \geq 1$,*

1. *The asymptotic lower bound of prediction error in a consistent recommender system is*

$$L_C = \sum_{r' \in R'} \min_{i \in \mathcal{I}} \min \left\{ p_{(r',1)}^i + \sum_{j \in \mathcal{I} \setminus \{i\}} p_{(r',1)}^j, p_{(r',0)}^i + \sum_{j \in \mathcal{I} \setminus \{i\}} p_{(r',1)}^j \right\}. \quad (4)$$

2. *L_C weakly decreases in C .*

The threshold attaining the bound coincides with the user-optimal threshold under symmetric payoffs; otherwise, prediction accuracy and user welfare are distinct objectives.

3.2.1 Empirical illustration: Decomposing the value of data

To provide an empirical illustration of our theoretical decomposition above, we conduct a simulation study using the Jester dataset (Goldberg et al., 2001), which contains over four million anonymous joke ratings from 73,421 users. We convert the continuous ratings into a binary format (positive/negative) to fit our framework. We estimate the correlation structure with a Bayesian Dirichlet-Multinomial model with a uniform prior.⁶ Using the estimates, we run counterfactual simulations to quantify the value created by the three key functions of a personalized recommender system: customization, selection, and screening.

To do so, we measure user utility under different system configurations. We compare a generic recommender system ($C = 0$) with a customized recommender system that conditions on user history ($C = 9$). We also vary the number of target items available for selection (I) and the screening threshold (τ). Table 2 summarizes the results. Full details of the dataset and our empirical methodology are available in Online Appendix B.

		<i>Average Utility</i>	<i>Standard Error</i>	<i>Min / Max</i>
$I = 1, \tau = 0$	<i>Generic RS ($C = 0$)</i>	0.148	(0.0077)	-0.567/0.616
	<i>Customized RS ($C = 9$)</i>	0.148	(0.0077)	-0.567/0.616
$I = 1, \tau = 1/2$ (<i>scr only</i>)	<i>Generic RS</i>	0.193	(0.0057)	-0.159/0.616
	<i>Customized RS</i>	0.258	(0.0049)	-0.051/0.622
$I = 3, \tau = 0$ (<i>sel only</i>)	<i>Generic RS</i>	0.342	(0.0048)	-0.284/0.617
	<i>Customized RS</i>	0.357	(0.0049)	-0.295/0.651
$I = 3, \tau = 1/2$ (<i>scr and sel</i>)	<i>Generic RS</i>	0.344	(0.0046)	-0.133/0.617
	<i>Customized RS</i>	0.371	(0.0042)	-0.013/0.634

Table 2: User surplus from recommender systems
scr: screening, *sel*: selection

Table 2 presents the estimated average utility that users receive from recommender systems with different parameter settings. Without selection and screening ($I = 1, \tau = 0$), the average utility is 0.148, which corresponds to the baseline where users try randomly selected items. On top of this, screening ($\tau = 1/2$) adds 0.045 to average utility with a generic recommender system, while customization ($C = 9$) adds a further 0.065, suggesting significant value from within-user learning. The value of customization is less pronounced when selection is wider (comparing $I = 3$

⁶Full details of this model and its properties are provided in Online Appendix A.

to $I = 1$ with $\tau = 1/2$), which is consistent with our theoretical predictions for items that are popular across many user types. When all three functions coexist ($C = 9, I = 3, \tau = 1/2$), the personalized recommender system adds 0.223 more value to users compared to the situation without any recommender system.

3.3 Harmful customization

While Proposition 1 shows that customization always weakly increases asymptotic expected user welfare under the user-optimal threshold, any fixed interior threshold that differs from it permits harmful customization under some correlation structures, provided the refinement is nontrivial.

Proposition 2 *For any $\tau \in (0, 1)$ with $\tau \neq \tau^u$, any $I \geq 1$, and any history r' over the first $C - 1$ conditioning items, there exists a collection of full-support correlation structures $q = \{q^i\}_{i \in \mathcal{I}}$ such that users with history r' have strictly lower asymptotic expected utility when recommendations additionally condition on item C than when they condition only on the first $C - 1$ items.*

Customization can benefit users even when the threshold is misaligned, but no misaligned interior threshold guarantees a weak welfare gain for every correlation structure. Thus, customization has the scope to hurt users if the platform’s interests cause its threshold level to diverge from the user-centric threshold. The proof is by construction: we identify q for each fixed $\tau \in (0, 1)$ with $\tau \neq \tau^u$. For a given correlation structure p , however, harmful thresholds need not exist. Online Appendix E characterizes such thresholds for a single target item and an additional binary conditioning item, allowing the set of harmful thresholds to be empty.

3.4 Marginal value of customization

In this section, we study how customization affects user value through the recommender system. Specifically, we are interested in the effect of changes in the degree of customization on the value generated by each target item under a recommender system. To do so, we isolate the target item by considering the case $I = 1$, and investigate the sources behind the value creation process of a recommender system.

Suppose that a consistent recommender system takes one more conditioning item into account in making a recommendation on the target item. One can equally think of a situation where the target user tries an item recommended by the platform, and the platform now can provide a customized recommendation that conditions also on the rating of the newly tried item. We denote

the new item and the target item respectively by item $C + 1$ and item $C + 2$. We take the new conditioning item's ratings to be binary. As the recommender system can condition its predictions on one more item, the best attainable asymptotic predictive accuracy of the target item weakly improves. However, this does not necessarily lead to higher user welfare because of the misalignment of interests between the user and the platform which is captured by the threshold level. Here, we study this marginal benefit or harm of customization. Because there is only one target item, we save notation by using p for p^{C+2} , the correlation structure associated with the $C + 1$ conditioning items and the target item.

Definition 4 For $r' \in R'$ over items $\{1, \dots, C\}$, item $C + 1$ and item $C + 2$ are positively correlated conditional on r' if $v_1 p_{(r',1,1)} + v_0 p_{(r',1,0)} \geq 0$ and $v_1 p_{(r',0,1)} + v_0 p_{(r',0,0)} \leq 0$, with at least one inequality holding strictly. If both inequalities are strict, we say that they are strictly positively correlated conditional on r' .

Here, correlation is defined in terms of the signs of conditional expected payoffs, rather than statistical dependence alone. When the target item and item $C + 1$ are positively correlated conditional on r' , users with history r' have nonnegative conditional expected utility from the target item if they liked item $C + 1$, and non-positive conditional expected utility if they did not. When $(v_1, v_0) = (1, -1)$, the condition is satisfied if we have $p_{(r',1,1)} \geq p_{(r',1,0)}$ and $p_{(r',0,1)} \leq p_{(r',0,0)}$ with at least one inequality holding strictly. In a parallel way we can define a (strict) negative correlation between item $C + 1$ and the target item. If item $C + 1$ and the target item are positively or negatively correlated conditional on $r' \in R'$, they are said to be correlated conditional on r' . For strict inequalities, they are said to be strictly correlated conditional on r' .

We focus on a universal threshold level $\tau = \tau(r')$, $\forall r'$, held fixed before and after customization, and compare asymptotic expected welfare.⁷ For each r' , there are two trivial cases in which an extra degree in customization yields only zero marginal customization effect: when the threshold level is too high or too low as stated below.

$$\tau < \min \left\{ \frac{p_{(r',1,1)}}{p_{(r',1)}}, \frac{p_{(r',0,1)}}{p_{(r',0)}} \right\} \text{ or } \tau > \max \left\{ \frac{p_{(r',1,1)}}{p_{(r',1)}}, \frac{p_{(r',0,1)}}{p_{(r',0)}} \right\}. \quad (5)$$

If the former is the case, the item is recommended to the user regardless of the history and the

⁷As in Lemma 1, when $\tau \neq \tau^u$ we exclude equality between τ and the true success probability conditional on r' or either refined history $(r', 0)$ or $(r', 1)$. Ties at the user-optimal threshold remain admissible because the expected payoff at its cutoff is zero.

introduction of the extra customization. This follows from the following equality.

$$\frac{p(r',1,1) + p(r',0,1)}{p_{r'}} = \frac{p(r',1)}{p_{r'}} \frac{p(r',1,1)}{p(r',1)} + \frac{p(r',0)}{p_{r'}} \frac{p(r',0,1)}{p(r',0)}.$$

Similarly, in the latter case, the item is not recommended regardless of the introduction of the extra degree in customization. As a result, the extra customization cannot affect user welfare if (5) holds. In what follows, we focus on histories with different refined success probabilities and thresholds strictly between them:

$$\min \left\{ \frac{p(r',1,1)}{p(r',1)}, \frac{p(r',0,1)}{p(r',0)} \right\} < \tau < \max \left\{ \frac{p(r',1,1)}{p(r',1)}, \frac{p(r',0,1)}{p(r',0)} \right\}.$$

In the next proposition, we find that a strict correlation between the newly added item and the target item is a necessary and sufficient condition under which an extra degree in customization is (on average) beneficial to users regardless of the misalignment of interests between the platform and users. The expected utility in the following proposition is averaged over the rating of item $C + 1$, conditional on the original history r' .

Proposition 3 *For a history r' , marginal customization strictly increases asymptotic expected utility for every admissible threshold in the interval specified above if and only if item $C + 1$ and the target item (item $C + 2$) are strictly correlated conditional on r' .*

When strict correlation exists, the system can use the item to segment r' users into subgroups based on their experiences with item $C + 1$. This segmentation enables the system to accurately identify and screen out subgroups unlikely to benefit from the target item, enhancing overall user welfare.

An important insight is that marginal value from additional customization does not always diminish; it can increase or decrease. Moreover, the marginal value of conditioning on multiple items simultaneously can even surpass the sum of individual values from conditioning on the items separately, demonstrating super-additivity in the value of customization.

To see this, let $(v_1, v_0) = (1, -1)$ and consider a platform that has three items under the user-optimal threshold level. The target item is fixed at item 3 and we compare user surplus under two cases depending on whether we condition on item 1 first or item 2 first.

The correlation structure p is given as follows:⁸

$$p = \left(\frac{2}{20}, \frac{1}{20}, \frac{1}{20}, \frac{7}{20}, \frac{1}{20}, \frac{2}{20}, \frac{5}{20}, \frac{1}{20} \right).$$

We can find the expected user utility under different degrees of customization. First, when the platform initially does not customize its prediction, the expected user utility is zero because the item is not recommended:

$$p_{(1,1,1)} + p_{(1,0,1)} + p_{(0,1,1)} + p_{(0,0,1)} = \frac{9}{20} < \frac{1}{2}.$$

On the other hand, if the platform conditions on both items in making a prediction about the target item, we can check that the expected utility to a user is $\frac{1}{4}$. In this case, users of history (1, 1) and (0, 0) try the item. Now, when the platform conditions only on item 1, the expected utility is

$$\max\{p_{(1,1,1)} + p_{(1,0,1)} - p_{(1,1,0)} - p_{(1,0,0)}, 0\} + \max\{p_{(0,1,1)} + p_{(0,0,1)} - p_{(0,1,0)} - p_{(0,0,0)}, 0\} = \frac{3}{20}.$$

Thus, in this case, the marginal value from customization is diminishing. It is $\frac{3}{20}$ from degree zero to degree one, and then $\frac{2}{20}$ from degree one to degree two.

On the other hand, if we customize the recommendation conditioning on item 2 first, the resulting expected utility is given as

$$\max\{p_{(1,1,1)} + p_{(0,1,1)} - p_{(1,1,0)} - p_{(0,1,0)}, 0\} + \max\{p_{(1,0,1)} + p_{(0,0,1)} - p_{(1,0,0)} - p_{(0,0,0)}, 0\} = 0.$$

Thus, in this case, the marginal value from customization is increasing: it is zero from degree zero to degree one, and then $\frac{1}{4}$ from degree one to degree two.

Another notable observation from this example is the super-additivity of values generated by customization. Specifically, the value created by simultaneously conditioning recommendations on multiple items can exceed the sum of the values generated by individually conditioning on each item separately. In the example, the value from conditioning on both items 1 and 2 simultaneously is $\frac{1}{4}$, while conditioning individually on either item 1 or item 2 yields values of $\frac{3}{20}$ or 0, respectively. The sum of the individual values, $\frac{3}{20}$, is thus strictly lower than the value from combined conditioning, which is $\frac{1}{4}$.

⁸Here $p = (p_{(1,1,1)}, p_{(1,1,0)}, p_{(1,0,1)}, p_{(1,0,0)}, p_{(0,1,1)}, p_{(0,1,0)}, p_{(0,0,1)}, p_{(0,0,0)})$.

At the user-optimal threshold, super-additivity occurs whenever a particular user history generates positive expected utility if considered alone, but this value is canceled out when aggregated with other histories under both one-item groupings. In the example, consider the history $r' = (1, 1)$. Individually, this segment's expected utility is positive and fully realized when customization involves both items. However, when this segment is combined with the $(1, 0)$ history (by conditioning on item 1 only), the aggregate expected utility becomes negative; when it is combined with the $(0, 1)$ history (by conditioning on item 2 only), the aggregate expected utility is zero. Both one-item groupings therefore mask its positive contribution. This illustrates that more precise segmentation can unlock substantial value that is otherwise hidden.

This intuition can be formalized in the following proposition, which provides a necessary and a sufficient condition for super-additivity. Consider two binary conditioning items, 1 and 2, and the user-optimal threshold level. Super-additivity compares gains relative to the same no-customization benchmark.

Proposition 4 *The value of customization exhibits super-additivity only if there exists a pair (i, j) such that*

$$\begin{aligned} & (v_1 p_{(i,1,1)} + v_0 p_{(i,1,0)}) (v_1 p_{(i,0,1)} + v_0 p_{(i,0,0)}) < 0 \text{ and} \\ & (v_1 p_{(1,j,1)} + v_0 p_{(1,j,0)}) (v_1 p_{(0,j,1)} + v_0 p_{(0,j,0)}) < 0. \end{aligned}$$

Furthermore, super-additivity is guaranteed if some pair (i, j) satisfies $v_1 p_{(ij1)} + v_0 p_{(ij0)} > 0$,

$$\begin{aligned} & v_1 (p_{(i,1,1)} + p_{(i,0,1)}) + v_0 (p_{(i,1,0)} + p_{(i,0,0)}) \leq 0 \text{ and} \\ & v_1 (p_{(1,j,1)} + p_{(0,j,1)}) + v_0 (p_{(1,j,0)} + p_{(0,j,0)}) \leq 0. \end{aligned}$$

This proposition formalizes the mechanism observed in the numerical example. The first part of the proposition establishes that a fundamental level of informational relevance—strict correlation—is required for super-additivity to be possible. The second part identifies a specific data structure—the existence of valuable information that is only achievable when multiple conditions are met—that is sufficient to guarantee a super-additive outcome. Recognizing that the value of data can scale in these synergistic ways, rather than only with simple diminishing returns, is crucial for understanding the full potential of modern AI-driven recommender systems.

3.5 Value of an additional user: the breadth and depth of data

This section studies how finite data affect recommendation value along two dimensions. Breadth, the number of previous users, improves the precision of estimates for a given level of customization. Depth, the number of conditioning items, can make recommendations more informative but also divides the data across more user histories. Deeper customization may therefore create greater potential value while requiring more data to realize it. Because finite-sample performance depends on the estimator, we impose a minimal restriction and study the marginal value of additional users. We first vary the number of previous users who share the target user’s history, and then compare customization levels holding the total training sample fixed.

To keep the analysis as simple as possible, we consider one target item and symmetric payoffs $(v_1, v_0) = (1, -1)$. The platform estimates the target user’s probability of a positive experience using data from N previous users who share her history r' .

In this case, the data X from N users can be summarized by $y = (y_{(r',1)}, y_{(r',0)})$, where y_r records the occurrences of outcome $r \in R = \{(r', 1), (r', 0)\}$ that appear in X . The true but unknown probability of a positive experience with the target item is $\frac{P_{(r',1)}}{P_{(r',1)} + P_{(r',0)}}$, which we denote by z . We focus on the recommender systems that satisfy the following condition.

Definition 5 *A recommender system is said to be unbiased if the target item is recommended to a target user with history r' if and only if $y_{(r',1)} \geq y_{(r',0)}$.*

An unbiased recommender system recommends the target item if and only if previous users with the target user’s history report at least as many positive as negative ratings.

We first examine the value of additional users while holding the target user’s history and the correlation structure fixed. For parts (1)–(2) of the following proposition, we avoid ties by considering odd matching sample sizes, $N = y_{(r',1)} + y_{(r',0)} = 2m - 1$, with $m \in \mathbb{N}$. Each increase in m therefore adds two matching users.

Let $v(m \mid r', p)$ denote the target user’s expected utility conditional on history r' and correlation structure p , averaging over samples of $N = 2m - 1$ matching users and the target user’s experience. Since trying the item yields expected payoff $2z - 1$, we have

$$v(m \mid r', p) = (2z - 1)P[y_{(r',1)} \geq y_{(r',0)}].$$

Parts (1)–(2) of the following proposition characterize how this expected value and its increment

change as matching observations accumulate, for every fixed correlation structure.

Part (3) of the following proposition instead compares customization levels using the same database of N total previous users and the same target item. Additional conditioning can reduce the number of matching observations; with no matches, the maintained rule recommends because both counts are zero. For this comparison, use p for the expanded correlation structure over the original C conditioning items, one additional binary conditioning item, and the target item.⁹

For part (3) only, draw p once from a distribution of full-support correlation structures, and draw all users independently conditional on it. Assume target-rating symmetry: the distribution is unchanged when $p_{(r',i,1)}$ and $p_{(r',i,0)}$ are interchanged simultaneously for every $r' \in R'$ and $i \in \{0, 1\}$. Let $V(C, N)$ denote expected utility at degree C with N total previous users, averaging over the correlation structure, training sample, and target user's history and experience. This averaging evaluates the same recommendation rule rather than changing it.

Proposition 5 *Let $\Delta v(m \mid r', p) = v(m + 1 \mid r', p) - v(m \mid r', p)$ denote the gain from adding two matching users.*

1. *Value weakly increases with matching-user data: for every p , $v(m \mid r', p)$ is weakly increasing in m .*
2. *The gain from additional matching-user data weakly diminishes: for every p , $\Delta v(m \mid r', p)$ is weakly decreasing in m .*
3. *Under target-rating symmetry, with one total previous user, $V(C + 1, 1) < V(C, 1)$ if and only if*

$$\sum_{r' \in R'} \text{Cov} \left(\mathbb{E}[y_{(r',0,1)} - y_{(r',0,0)} \mid p], \mathbb{E}[y_{(r',1,1)} - y_{(r',1,0)} \mid p] \right) > 0. \quad (6)$$

Part (1) shows that additional matching-user data improves screening. When $z > 1/2$, an additional pair of observations increases the probability of recommending a beneficial item; when $z < 1/2$, it reduces the probability of recommending a harmful item. These gains do not require expected utility itself to be positive. For every $0 < z < 1/2$ and finite m , expected utility is strictly negative because the system recommends with positive probability despite the negative expected payoff from trying the item. Additional matching observations reduce this expected loss. At $z = 1/2$, expected utility is zero at every sample size.

⁹Write $y_{(r',i,j)}$ for the count with original history r' , additional conditioning-item rating i , and target-item rating j . The coarse counts satisfy $y_{(r',j)} = y_{(r',0,j)} + y_{(r',1,j)}$, for $j \in \{0, 1\}$. In part (3), the inner expectations refer to the one-user training sample conditional on p , while the covariance is taken across correlation structures.

Part (2) establishes diminishing returns to this improvement in screening. Under the majority rule, two additional observations can change the recommendation only when the existing positive and negative counts differ by one. For a fixed z , these closely divided samples become less likely along the odd-sample sequence. Accounting for both directions in which the recommendation can change, the proof obtains

$$\Delta v(m \mid r', p) = (2z - 1)^2 \binom{2m - 1}{m} [z(1 - z)]^m.$$

This gain is nonnegative and weakly decreases in m . Thus, additional matching users improve recommendations at a diminishing rate without any restriction on the correlation structure beyond the maintained assumptions. The result follows from the majority rule, not from a tradeoff between prior information and new data.

Part (3) identifies when deeper customization lowers initial value by restricting which observations the system uses. A sufficient condition is nonnegative covariance between the two subgroups' conditional expected signed counts at every history, with strict covariance at some history. This captures common appeal across groups: observations from one subgroup can help predict whether the item benefits the other. With only one previous user, separating such groups can therefore discard useful information. The aggregate condition in (6), rather than nonnegative covariance at every history, is necessary and sufficient.

Taken together, these findings distinguish the returns to learning within a fixed history group from the gains available through finer customization. Breadth improves screening with diminishing increments in our matching-sample benchmark. Depth can raise attainable asymptotic utility under the user-optimal threshold, as established in Proposition 1, and the earlier examples show that its marginal benefits need not diminish and that customization gains can be super-additive. These differences motivate the possibility of cascading, overlapping learning curves as breadth and depth expand jointly. The simulations below illustrate such a path.

To illustrate joint expansion, we use the Jester dataset from Section 3.2. Training size increases one user at a time, and customization rises every 100 users, up to degree six. Figure 2 shows the resulting sequence of utility levels. Online Appendix B specifies the training-size grid, depth-switching rule, and nested user and item design.

Lastly, we assess whether the two dimensions of data, breadth and depth, are complements or substitutes in creating user value. We use our benchmark simulation setting, $(I, \tau) = (1, 1/2)$, and

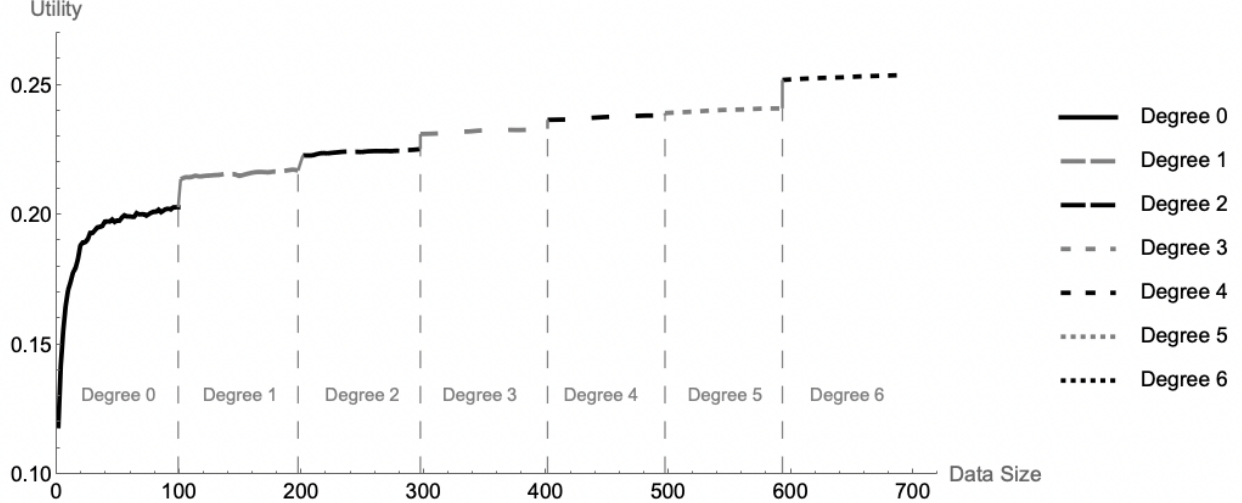


Figure 2: Cascading learning curves

run 1,000 simulations in which we vary the degree of customization C and the number of training users N . Let $V(C, N)$ denote the average utility generated by a recommender system that learns from data on C conditioning items and one target item for N previous users. We measure the interaction between breadth and depth using the following discrete cross-partial:

$$\Delta_{C,N} = [V(C + \delta_C, N + \delta_N) - V(C + \delta_C, N)] - [V(C, N + \delta_N) - V(C, N)].$$

Here, δ_C and δ_N denote increments in the degree of customization and the number of previous users, respectively. A positive value of $\Delta_{C,N}$ indicates local complementarity: increasing the degree of customization raises the marginal value of additional user data. A negative value indicates local substitutability. We set $\delta_C = 1$ and $\delta_N = 500$.

For visual interpretation, we treat values between -0.0005 and 0.0005 as a graphical neutral band. Values above this band indicate complementarity, while values below it indicate substitutability. This band is used only to organize the graphical discussion and should not be interpreted as a statistical confidence interval.

As shown in Figure 3, breadth and depth exhibit strong complementarity when both the degree of customization and the number of users are low. This pattern is consistent with the gains from joint expansion illustrated in Figure 2. The simulation underlying that figure, which covers data sizes up to 700 users and customization degrees up to six, lies largely within the region in which additional user data raises the marginal value of deeper customization.

The complementarity attenuates as the user base grows. Moreover, when customization is

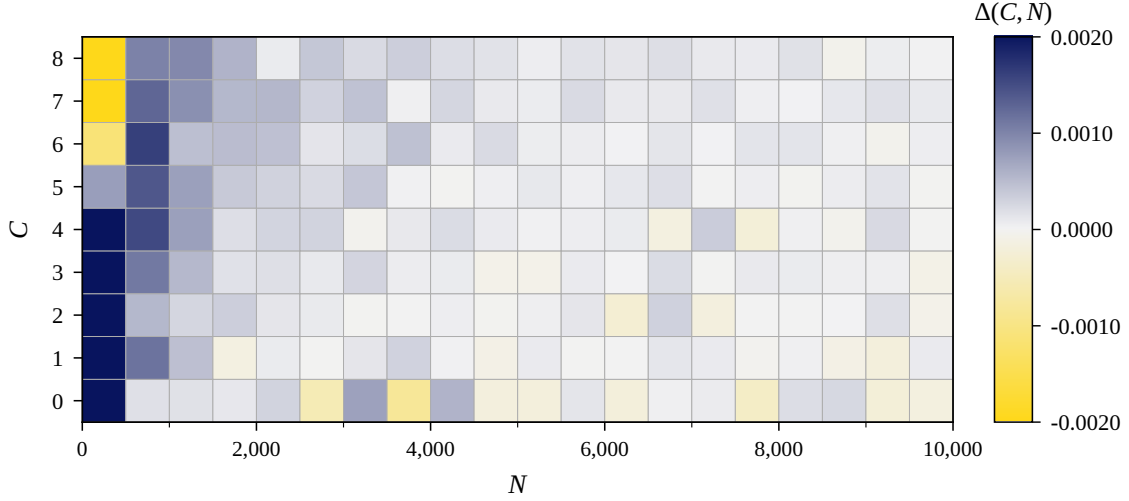


Figure 3: Interaction between data breadth and depth. The color scale is capped at -0.002 and 0.002 ; values outside this range are displayed using the respective endpoint colors.

high but the number of users is very small, the two dimensions become locally substitutable. For example, at $N = 0$, the estimated cross-partials are $\Delta_{6,0} = -0.0011$, $\Delta_{7,0} = -0.0026$, and $\Delta_{8,0} = -0.0035$.¹⁰ This pattern is consistent with the small effective sample sizes available for increasingly detailed histories when N is low and C is high.

A negative cross-partial does not, by itself, imply that increasing C reduces the level of expected utility. Rather, it means that breadth and depth are locally substitutable: additional user data has a smaller marginal effect at the higher degree of customization, or equivalently, the marginal value of customization falls as breadth increases over the relevant interval. Finally, the interaction becomes quantitatively small once the system has sufficiently many users from whom to learn. In our exercise, when $N \geq 5000$, the absolute value of every estimated cross-partial is below 0.0005 .

Appendix B summarizes the MovieLens 1M application, which complements the Jester analysis by providing a comparison across rating environments; the full analysis and results are presented in Online Appendix F. The same framework separates the contributions of selection, screening, and customization, while revealing a different pattern of breadth–depth interactions: at high customization, breadth and depth are substitutes at very low breadth, complements over an intermediate range, and have little interaction once sufficiently many observations are available. The comparison illustrates how these interactions can differ across applications rather than following a universal pattern.

¹⁰As in the definition of $\Delta_{C,N}$, these values compare the marginal value of increasing N from 0 to 500 at adjacent degrees of customization. A similar pattern arises when the training-sample comparison begins at $N = 1$.

4 Conclusion

In this paper, we propose a tractable economic model to understand and quantify the value created by recommender systems, highlighting the critical role of personalization. Our framework identifies three fundamental components underlying value creation: customization, which segments users based on their observable features (history); selection, which then identifies the best potential item for each group; and screening, which finally determines if that best item is good enough to be recommended. We show that additional customization weakly increases asymptotic expected user welfare under the user-optimal threshold, and provide conditions for strict gains. Empirically, we quantify the contributions of these three elements through counterfactual simulations using the Jester dataset, containing over four million anonymous joke ratings from 73,421 users, and the MovieLens 1M dataset, containing approximately one million ratings from 6,040 users.

Our findings reveal distinctive scaling properties along two dimensions: data breadth and data depth. Our benchmark exhibits diminishing returns to learning within user groups, whereas deeper customization can generate increasing returns and super-additive gains. With the total training sample instead held fixed, deeper customization can reduce expected user welfare. The empirical exercises illustrate how the marginal value of depth varies with available breadth across data and estimation settings. Jester exhibits strong complementarity at low breadth and low depth, whereas MovieLens displays substitutability for deep customization at very low breadth, complementarity over an intermediate range, and little interaction at high breadth. The MovieLens application thus complements the Jester analysis by illustrating how the same framework accommodates different sources of recommendation value and breadth–depth interactions.

The distinction between breadth and depth helps clarify the scope of data-access policies. Portability of an individual user’s history can help a rival platform customize its recommendations, while access to pooled data from other users can help it learn how that history predicts preferences. When these inputs are complementary, providing both can generate gains beyond the sum of providing each separately. Our framework provides a way to quantify the potential recommendation gains from such combined access. Non-diminishing returns and super-additivity in depth also mean that an incumbent’s advantage can stem from combining a richer set of user features, even when rivals have data on many users. Sharing only a narrow subset of features may therefore leave much of this advantage intact. The scope of a sharing obligation should consequently be assessed by which features it makes available for joint use, as well as by how many users it covers. Where the receiving

platform already has sufficient observations to estimate the relevant relationships accurately, access to richer histories may matter more than further expanding the user sample. Finally, the welfare effects of improved access depend on the receiving platform’s recommendation rule. More detailed information can harm users when the platform’s threshold leads it to recommend unsuitable items or withhold beneficial ones. Data-access policies should therefore be assessed by their effects on user welfare, taking account of how receiving platforms use the shared data to make recommendations.

Our flexible model, suitable for both theoretical and empirical analyses, can serve as a starting point for further exploration as we establish a theoretical upper bound on the surplus an ideal recommender system can confer to its users. Potential extensions include analyzing firms’ optimal pricing strategies in the presence of recommender systems, exploring the strategic manipulation of information to steer consumers toward more profitable items, and modeling platform competition employing different recommender technologies to analyze the resulting market outcomes and long-run market structure. Furthermore, our formalization of the conditions for super-additivity opens new avenues for research into optimal data acquisition strategies, as platforms could prioritize collecting data on features known to have strong, synergistic interaction effects.

A Appendix - Proofs of Propositions

A.1 Proof of Proposition 1

Proof. At the user-optimal threshold $\tau^u = -v_0/(v_1 - v_0)$, the true conditional expected payoff from trying item i is

$$u_i(r') = v_1 z_i(r') + v_0(1 - z_i(r')) = (v_1 - v_0)(z_i(r') - \tau^u).$$

Since $v_1 - v_0 > 0$, selecting an item with the highest success probability and recommending it if and only if that probability is at least τ^u maximizes expected utility over the target items and no recommendation. The history probability $p_{r'} = p_{(r',1)}^i + p_{(r',0)}^i$ is common to all items. Multiplying the optimal conditional utility by this probability gives the history's contribution to ex-ante expected utility:

$$p_{r'} \max_{i \in \mathcal{I}} \{u_i(r'), 0\} = \max_{i \in \mathcal{I}} \{v_1 p_{(r',1)}^i + v_0 p_{(r',0)}^i, 0\}.$$

At τ^u , a consistent recommender system maximizes estimated expected utility over the finite action set, including no recommendation. Consistency implies that the probability of choosing a strictly suboptimal action tends to zero. Bounded payoffs therefore imply convergence of expected utility to the optimum for each history, regardless of ties. Averaging over the finitely many histories proves part (1).

For part (2), treat the population as one group and apply the same argument using the unconditional success probabilities $z_i^G = \sum_{r' \in R'} p_{(r',1)}^i$. ■

A.2 Proof of Corollary 1

Proof. Conditional on history r' , the expected number of errors is $\sum_{j \in \mathcal{I}} z_j(r')$ when nothing is recommended. If item i is recommended, the expected number of errors is

$$1 - z_i(r') + \sum_{j \in \mathcal{I} \setminus \{i\}} z_j(r').$$

The first term counts the expected false positive on the recommended item, and the remaining terms count expected false negatives on non-recommended items. Recommending item i rather than nothing therefore changes expected error by $1 - 2z_i(r')$. Consequently, the error-minimizing

rule selects an item with the highest success probability and recommends it when that probability is at least $1/2$.

The history probability $p_{r'}$ is common to all target items, with $p_{r'} z_i(r') = p_{(r',1)}^i$ and $p_{r'}(1 - z_i(r')) = p_{(r',0)}^i$. Multiplying the minimum conditional error by $p_{r'}$ and summing over histories therefore gives L_C in (4).

The training data are independent of the target user's outcomes conditional on her history and the population distribution. Thus, any data-dependent or randomized recommendation rule has conditional expected error at least the minimum above. Hence L_C bounds expected errors per user from below, including with finite training data. Set $\tau(r') = 1/2$ for every history. The rule minimizes estimated conditional error, and consistency implies that these estimates converge to the true errors simultaneously over the finite action set. Thus, the true conditional expected error of the chosen action converges in probability to the minimum, including when minimizing actions are tied. Since errors per user are bounded by I , averaging over histories and taking expectations gives

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\frac{\epsilon^m(X)}{m} \right] = L_C.$$

This proves part (1). The error-minimizing threshold $1/2$ coincides with τ^u if and only if $v_1 = -v_0$.

For part (2), hold the target-item pool and population distribution fixed. A rule using an additional conditioning-item rating can ignore it and reproduce the original error-minimizing rule, with ex-ante expected error L_C . Minimizing error separately for each refined history can only weakly improve on this rule, so $L_{C+1} \leq L_C$. ■

A.3 Proof of Proposition 2

Proof. To begin, we introduce a lemma that characterizes the limiting expected utility from a consistent recommender system when there is only one target item and a general threshold is used.

Lemma 1 *Suppose there is only one item available to the target user, i.e., $i = C + 1$. For part (1), assume that $z_i(r') \neq \tau(r')$ at every history $r' \in R'$ where $\tau(r') \neq \tau^u$. For part (2), assume that $z_i^G \neq \tau^G$ whenever $\tau^G \neq \tau^u$.*

1. *For any consistent mechanism, the expected utility to the target user converges to*

$$\sum_{r' \in R'} \mathbf{1}\{p_{(r',1)}^i \geq \tau(r') p_{r'}^i\} (v_1 p_{(r',1)}^i + v_0 p_{(r',0)}^i).$$

2. For a generic recommender system, the expected utility to the target user converges to

$$\mathbf{1}\left\{\sum_{r' \in R'} p_{(r',1)}^i \geq \tau^G\right\} \sum_{r' \in R'} (v_1 p_{(r',1)}^i + v_0 p_{(r',0)}^i).$$

(Proof of Lemma 1.) With data from $N - 1$ previous users, the ex-ante expected utility is

$$\sum_{r' \in R'} \Pr\{\hat{z}_i^N(r') \geq \tau(r')\} (v_1 p_{(r',1)}^i + v_0 p_{(r',0)}^i),$$

where the probability is taken over the training data for the given target history. If $z_i(r') \neq \tau(r')$, consistency implies that the recommendation probability converges to $\mathbf{1}\{z_i(r') \geq \tau(r')\}$. If $z_i(r') = \tau(r')$, the assumption implies $z_i(r') = \tau^u$, so $v_1 p_{(r',1)}^i + v_0 p_{(r',0)}^i = 0$. Thus, this history contributes zero expected utility for every N , regardless of the recommendation probability. Taking limits in the finite sum proves part (1). For part (2), apply the same argument to the consistent estimator of z_i^G and threshold τ^G ; at $z_i^G = \tau^G = \tau^u$, the aggregate expected payoff from trying the item is zero. $\square$

To prove Proposition 2, fix $\tau \in (0, 1) \setminus \{\tau^u\}$ and $C \geq 1$. Item C has $n_C \geq 2$ possible ratings. First consider one target item, $M = C + 1$, and let R'' denote histories over the first $C - 1$ conditioning items. We compare limiting utilities conditional on each r'' , averaging over the additional rating of item C .

Choose positive probabilities $q_{r''}^M$ summing to one over R'' , and let $0 < z_L < z_H < 1$ and $w \in (0, 1)$ be as specified below. For each r'' and $k \in \{1, \dots, n_C\}$, set

$$\frac{q_{(r'',k)}^M}{q_{r''}^M} = \begin{cases} \frac{w}{n_C - 1}, & k < n_C, \\ 1 - w, & k = n_C, \end{cases} \quad \frac{q_{(r'',k,1)}^M}{q_{(r'',k)}^M} = \begin{cases} z_H, & k < n_C, \\ z_L, & k = n_C. \end{cases}$$

Set $q_{(r'',k,0)}^M = q_{(r'',k)}^M - q_{(r'',k,1)}^M$. This gives full support and, before conditioning on item C , success probability

$$\bar{z} = \frac{\sum_{k=1}^{n_C} q_{(r'',k,1)}^M}{q_{r''}^M} = wz_H + (1 - w)z_L.$$

Suppose first that $0 < \tau < \tau^u$. In this case, choose

$$0 < z_L < \tau < z_H < \tau^u, \quad 0 < w < \frac{\tau - z_L}{z_H - z_L}.$$

Then $\bar{z} < \tau$, so without the additional conditioning, the recommendation probability tends to zero. After refinement, it tends to one for $k < n_C$ and zero for $k = n_C$. Since $z_H < \tau^u$, the newly recommended item has negative expected payoff. By Lemma 1, the change in limiting utility conditional on r'' is

$$w(v_1 z_H + v_0(1 - z_H)) - 0 = w(v_1 - v_0)(z_H - \tau^u) < 0.$$

Next, if $\tau^u < \tau < 1$, choose

$$\tau^u < z_L < \tau < z_H < 1, \quad \frac{\tau - z_L}{z_H - z_L} < w < 1.$$

Then $\bar{z} > \tau$, so without the additional conditioning the recommendation probability tends to one. Refinement leaves this limit unchanged for $k < n_C$ but reduces it to zero for $k = n_C$. Since $z_L > \tau^u$, the latter users lose a recommendation with positive expected payoff. By Lemma 1, the change in limiting utility is

$$w(v_1 - v_0)(z_H - \tau^u) - (v_1 - v_0)(\bar{z} - \tau^u) = -(1 - w)(v_1 - v_0)(z_L - \tau^u) < 0.$$

In both cases the interval of admissible weights is nonempty because $z_L < \tau < z_H$, and no relevant success probability equals τ .

For $I > 1$, relabel the constructed item as $M = C + I$ and retain the history distribution. For each other target item j , set

$$q_{(r'', k, 1)}^j = \frac{\tau}{2} q_{(r'', k)}^M, \quad q_{(r'', k, 0)}^j = \left(1 - \frac{\tau}{2}\right) q_{(r'', k)}^M.$$

These structures have full support and common history marginals, and are jointly realizable by taking target-item outcomes independently conditional on history. Each added item's success probability is $\tau/2 < \tau$ before and after refinement. Consistency therefore makes their recommendation probabilities tend to zero, while item M is selected with probability tending to one whenever its success probability exceeds τ . The limiting utility comparisons are unchanged, proving the claim for every r'' and any $I \geq 1$. ■

A.4 Proof of Proposition 3

Proof. Fix a history r' with distinct refined success probabilities. For $b \in \{0, 1\}$, write

$$z_b = \frac{p(r', b, 1)}{p(r', b)}, \quad w_b = \frac{p(r', b)}{p_{r'}}, \quad u_b = v_1 z_b + v_0(1 - z_b).$$

Relabel the two ratings if necessary so that $z_0 < z_1$. Full support implies $w_0, w_1 > 0$, so the coarse success probability $\bar{z} = w_0 z_0 + w_1 z_1$ lies strictly between z_0 and z_1 .

All utilities below are asymptotic and conditional on r' . For thresholds satisfying Lemma 1, the utilities before and after refinement are, respectively,

$$\mathbf{1}\{\bar{z} \geq \tau\}(w_0 u_0 + w_1 u_1) \quad \text{and} \quad \sum_{b \in \{0, 1\}} w_b \mathbf{1}\{z_b \geq \tau\} u_b.$$

Subtracting gives the benefit of marginal customization:

$$\begin{cases} 0 & \text{if } \tau \leq z_0 \text{ or } \tau > z_1, \\ -w_0 u_0 & \text{if } z_0 < \tau \leq \bar{z}, \\ w_1 u_1 & \text{if } \bar{z} < \tau \leq z_1. \end{cases} \quad (\text{Diff})$$

At any retained threshold tie, $\tau = \tau^u$, the relevant coarse or refined group's expected payoff is zero, so these expressions remain valid even if its recommendation probability does not converge.

Strict correlation is equivalent to $u_0 < 0 < u_1$. Since both weights are positive, (Diff) is then strictly positive for every admissible threshold in (z_0, z_1) .

Conversely, without strict correlation, either $u_0 \geq 0$ or $u_1 \leq 0$. In the first case, choose an admissible threshold in $(z_0, \bar{z})$; the second line of (Diff) is nonpositive. In the second case, choose one in $(\bar{z}, z_1)$; the third line is nonpositive. Both intervals are nonempty, and only finitely many threshold ties are excluded, so these choices are possible. Thus, marginal customization is not strictly beneficial for every admissible threshold in (z_0, z_1) . ■

A.5 Proof of Proposition 4

Proof. All values below are asymptotic and evaluated at τ^u . For each history $(k, l) \in \{0, 1\}^2$, define its probability-weighted expected payoff from trying the target item by

$$x_{kl} = v_1 p_{(k,l,1)} + v_0 p_{(k,l,0)},$$

and write $a_+ = \max\{a, 0\}$. Under the user-optimal threshold, history (k, l) contributes $(x_{kl})_+$ to expected user utility. Measuring all customization gains relative to the same no-customization benchmark, the gain from conditioning on both items minus the sum of the gains from conditioning on each item separately is

$$\Delta = \sum_{k,l \in \{0,1\}} (x_{kl})_+ + \left(\sum_{k,l \in \{0,1\}} x_{kl} \right)_+ - \sum_{k \in \{0,1\}} (x_{k1} + x_{k0})_+ - \sum_{l \in \{0,1\}} (x_{1l} + x_{0l})_+.$$

Strict super-additivity is equivalent to $\Delta > 0$.

We first prove necessity by contraposition. If there is no pair (k, l) such that both $x_{k1}x_{k0} < 0$ and $x_{1l}x_{0l} < 0$, then either $x_{k1}x_{k0} \geq 0$ for every k , or $x_{1l}x_{0l} \geq 0$ for every l . In the first case, $(x_{k1})_+ + (x_{k0})_+ = (x_{k1} + x_{k0})_+$ for every k , so

$$\Delta = \left(\sum_{k,l \in \{0,1\}} x_{kl} \right)_+ - \sum_{l \in \{0,1\}} (x_{1l} + x_{0l})_+ \leq 0.$$

The inequality follows from $(u+v)_+ \leq u_+ + v_+$. In the second case, interchanging rows and columns gives the same cancellation and inequality. Hence, strict super-additivity requires the stated sign changes in both dimensions.

To prove sufficiency, suppose there exists (k, l) such that $x_{kl} > 0$, $x_{k1} + x_{k0} \leq 0$, and $x_{1l} + x_{0l} \leq 0$. For brevity, write

$$a = x_{kl} > 0, \quad b = -x_{k(1-l)} \geq a, \quad c = -x_{(1-k)l} \geq a, \quad d = x_{(1-k)(1-l)}.$$

The two adjacent cells are negative, and the row and column containing x_{kl} have nonpositive sums. Therefore,

$$\Delta = a + d_+ + (d - b - c + a)_+ - (d - b)_+ - (d - c)_+.$$

For $b \geq 0$, the function $t \mapsto t_+ - (t - b)_+$ is nondecreasing. Since $c \geq 0$, evaluating this function at d and $d - c$ gives

$$(d - b)_+ + (d - c)_+ \leq d_+ + (d - b - c)_+.$$

Substituting this inequality into the expression for Δ , we obtain

$$\Delta \geq a + (d - b - c + a)_+ - (d - b - c)_+ \geq a > 0,$$

where the second inequality follows from $a > 0$ and monotonicity of the positive-part function. This proves the sufficient condition. ■

A.6 Proof of Proposition 5

Proof. Let X_n count positive ratings among the first n independent matching observations. Then $X_n \sim \text{Binomial}(n, z)$ and $v(m \mid r', p) = (2z - 1) \Pr\{X_{2m-1} \geq m\}$. Adding two observations changes the decision from no recommendation to a recommendation only when $X_{2m-1} = m - 1$ and both new ratings are positive. The decision changes in the opposite direction only when $X_{2m-1} = m$ and both new ratings are negative. Independence therefore gives

$$\begin{aligned} & \Pr\{X_{2m+1} \geq m + 1\} - \Pr\{X_{2m-1} \geq m\} \\ &= z^2 \Pr\{X_{2m-1} = m - 1\} - (1 - z)^2 \Pr\{X_{2m-1} = m\} \\ &= \binom{2m-1}{m} [z^{m+1}(1 - z)^m - z^m(1 - z)^{m+1}] \\ &= (2z - 1) \binom{2m-1}{m} [z(1 - z)]^m, \end{aligned}$$

where the second equality uses $\binom{2m-1}{m-1} = \binom{2m-1}{m}$. Multiplying by the payoff $2z - 1$ yields

$$\Delta v(m \mid r', p) = v(m + 1 \mid r', p) - v(m \mid r', p) = (2z - 1)^2 \binom{2m-1}{m} [z(1 - z)]^m \geq 0,$$

proving part (1).

If $z = 1/2$, utility and all increments are zero. Otherwise, full support gives $0 < z < 1$, and the

increments are strictly positive. Their ratio is

$$\begin{aligned}\frac{\Delta v(m+1 \mid r', p)}{\Delta v(m \mid r', p)} &= \frac{\binom{2m+1}{m+1}}{\binom{2m-1}{m}} z(1-z) \\ &= \frac{2(2m+1)}{m+1} z(1-z) < 4z(1-z) \leq 1.\end{aligned}$$

Thus, increments strictly diminish for $z \neq 1/2$ and weakly diminish in all cases, proving part (2).

Lastly, to prove part (3), fix one total previous user and use p for the expanded correlation structure, as in Section 3.5. Conditional on p , the previous user and the target user are independent draws from the same population. All conditional expected counts below refer to this one-user sample, while outer expectations and covariances are taken across correlation structures. In particular,

$$\mathbb{E}[y_{(r', i, j)} \mid p] = p_{(r', i, j)}, \quad y_{(r', j)} = \sum_{i \in \{0, 1\}} y_{(r', i, j)}, \quad i, j \in \{0, 1\}.$$

For a target user with original history r' , the item is not recommended precisely when the previous user has the same history and a negative target-item rating. Thus, conditional on p , the recommendation probability is $1 - \mathbb{E}[y_{(r', 0)} \mid p]$. This accounts for both a matching positive observation and no matching observation, since the maintained rule recommends in the latter case. Averaging over the target user's history and experience, conditional independence therefore gives

$$V(C, 1) = \mathbb{E} \left[\sum_{r' \in R'} \mathbb{E}[y_{(r', 1)} - y_{(r', 0)} \mid p] (1 - \mathbb{E}[y_{(r', 0)} \mid p]) \right].$$

Under target-rating symmetry, simultaneously interchanging positive and negative target-item ratings leaves the distribution of p unchanged. This interchange reverses the sign of $\mathbb{E}[y_{(r', 1)} - y_{(r', 0)} \mid p]$ and exchanges the two conditional expected counts. Averaging the preceding expression with its rating-interchanged counterpart yields

$$V(C, 1) = \frac{1}{2} \mathbb{E} \left[\sum_{r' \in R'} (\mathbb{E}[y_{(r', 1)} - y_{(r', 0)} \mid p])^2 \right].$$

The same argument applied separately to the refined histories gives

$$V(C+1, 1) = \frac{1}{2} \mathbb{E} \left[\sum_{r' \in R'} \sum_{i \in \{0,1\}} (\mathbb{E}[y_{(r',i,1)} - y_{(r',i,0)} \mid p])^2 \right].$$

The coarse signed count is the sum of the two refined signed counts. Moreover, target-rating symmetry implies

$$\mathbb{E}[\mathbb{E}[y_{(r',i,1)} - y_{(r',i,0)} \mid p]] = 0 \quad \text{for every } r' \in R' \text{ and } i \in \{0,1\}.$$

Subtracting the two value expressions and expanding the square therefore gives

$$\begin{aligned} V(C, 1) - V(C+1, 1) &= \sum_{r' \in R'} \mathbb{E}[\mathbb{E}[y_{(r',0,1)} - y_{(r',0,0)} \mid p] \mathbb{E}[y_{(r',1,1)} - y_{(r',1,0)} \mid p]] \\ &= \sum_{r' \in R'} \text{Cov}(\mathbb{E}[y_{(r',0,1)} - y_{(r',0,0)} \mid p], \mathbb{E}[y_{(r',1,1)} - y_{(r',1,0)} \mid p]). \end{aligned}$$

Consequently, $V(C+1, 1) \leq V(C, 1)$ if and only if the covariance sum in (6) is nonnegative. The value comparison is strict if and only if that sum is strictly positive. This proves part (3).

Online Appendix D.1 extends this comparison to growing samples and gives conditions under which deeper customization eventually raises expected utility. ■

B MovieLens: summary of empirical results

MovieLens 1M contains 1,000,209 ratings from 6,040 users (Harper and Konstan, 2015). We code four- and five-star ratings as positive, one to three stars as negative, and treat missing ratings as unobserved. Table B1 shows substantial gains from selection and screening, with smaller additional gains from customization.

		$I = 1$			$I = 3$		
τ		Utility	Std. error	Min / Max	Utility	Std. error	Min / Max
0	Generic	-0.008	(0.0151)	-1.764/0.874	0.357	(0.0083)	-0.711/0.874
0	Customized	-0.008	(0.0151)	-1.764/0.874	0.360	(0.0082)	-0.711/0.874
τ^u	Generic	0.175	(0.0075)	-0.511/0.874	0.363	(0.0077)	-0.511/0.874
τ^u	Customized	0.204	(0.0069)	-0.406/0.817	0.377	(0.0062)	-0.225/0.825

Table B1: MovieLens counterpart of Table 2, using the 100 most-rated movies, $C = 7$ for customized RS.

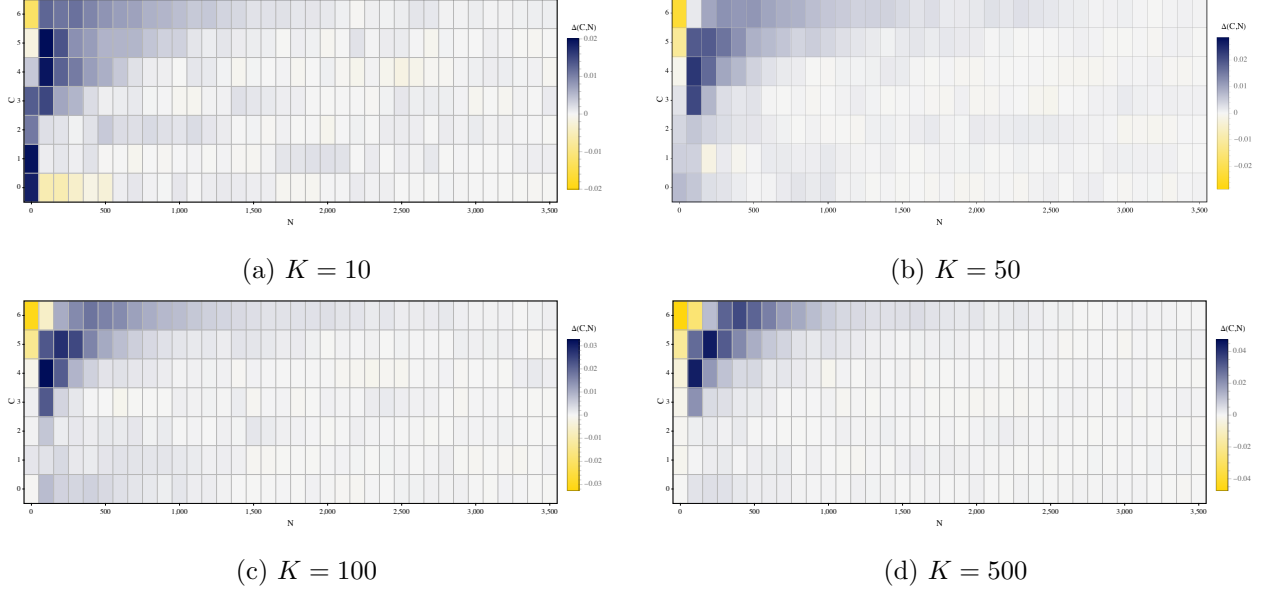


Figure B1: MovieLens counterparts of Figure 3, for pools of K most-rated movies. Here $\tau = \tau_K^u$, $\delta_C = 1$, and $\delta_N = 500$.

The MovieLens application complements the Jester illustration by showing that breadth–depth interactions differ across calibrated rating environments. All four pools exhibit high-depth substitutability at very low breadth, intermediate complementarity, and attenuation at the reported high-depth margin as breadth grows. The detailed MovieLens analysis, including the simulation design and additional results, is presented in Online Appendix F, with estimation details provided in Online Appendix C.

References

- G. Aridor and D. Gonçalves. Recommenders’ originals: The welfare effects of the dual role of platforms as producers and recommender systems. *International Journal of Industrial Organization*, 83:102845, 2022.
- G. Aridor, D. Gonçalves, D. Kluver, R. Kong, and J. Konstan. The informational role of online recommendations: Evidence from a field experiment. Working paper, 2024.
- J. Asker, C. Fershtman, and A. Pakes. The impact of artificial intelligence design on pricing. *Journal of Economics & Management Strategy*, 33(2), 2023.
- M. O. Ayemowa, R. Ibrahim, and M. M. Khan. Analysis of recommender system using generative artificial intelligence: A systematic literature review. volume 12, pages 87742–87766. IEEE Access, 2024.
- P. Bajari, V. Chernozhukov, A. Hortaçsu, and J. Suzuki. The impact of big data on firm performance: An empirical investigation. In *AEA Papers and Proceedings*, volume 109, pages 33–37, 2019.

- G. Biglaiser, E. Calvano, and J. Crémer. Incumbency advantage and its value. *Journal of Economics & Management Strategy*, 28(1):41–48, 2019.
- E. Calvano, G. Calzolari, V. Denicolò, and S. Pastorello. Artificial intelligence, algorithmic pricing, and collusion. *American Economic Review*, 110(10):3267–97, October 2020.
- E. Calvano, G. Calzolari, V. Denicolò, and S. Pastorello. Artificial intelligence, algorithmic recommendations and competition. Working paper, 2026.
- X. Chen, Y. Li, and J. Mao. A nearly instance optimal algorithm for “top-k” ranking under the multinomial logit model. pages 2504–2522. Proc. 29th Annual ACM-SIAM Symposium on Discrete Algorithms, Society for Industrial and Applied Mathematics, Philadelphia, 2018.
- A. de Cornière and G. Taylor. Data and competition: A simple framework. *RAND Journal of Economics*, 56(4):494–510, Winter 2025.
- I. P. Fainmesser, A. Galeotti, and R. Momot. Digital privacy. *Management Science*, 69(6):3157–3173, June 2023.
- Y. Feng, R. Caldentey, and C. T. Ryan. Robust learning of consumer preferences. *Operations Research*, 70(2):918–962, 2022.
- K. Goldberg, T. Roeder, D. Gupta, and C. Perkins. Eigentaste: A constant time collaborative filtering algorithm. *Information Retrieval*, 4(2):133–151, 2001.
- A. Hagiu and J. Wright. Data-enabled learning, network effects and competitive advantage. *RAND Journal of Economics*, 54(4):638–667, 2023.
- F. M. Harper and J. A. Konstan. The movielens datasets: History and context. *ACM Transactions on Interactive Intelligent Systems*, 5(4):1–19, 2015.
- C. Peukert, A. Sen, and J. Claussen. The editor and the algorithm: Recommendation technology in online news. *Management Science*, 70(9):5816–5831, 2023.
- M. Schäfer and G. Sapi. Complementarities in learning from data: Insights from general search. *Information Economics and Policy*, 65(101063), December 2023.
- A. Subbiah and V. Aggarwal. Transformers in music recommendation. Technical report, Google Research, <https://research.google/blog/transformers-in-music-recommendation/>, August 2024.
- L. Wu, Z. Zheng, Z. Qiu, T. Shen, C. Qin, C. Zhu, and E. Chen. A survey on large language models for recommendation. *World Wide Web*, 27(60), 2024.
- H. Yoganarasimhan. Search personalization using machine learning. *Management Science*, 66(3):1045–1070, 2020.
- K. Zielnicki, G. Aridor, A. Bibaut, A. Tran, W. Chou, and N. Kallus. The value of personalized recommendations: Evidence from netflix. *International Journal of Industrial Organization*, 108(103304), October 2026.