Online Appendix

Gunhaeng Lee and Julian Wright

We provide detailed proofs of material that is not included in the main text of our paper "Recommender systems and the value of user data".

A Making predictions from a partial data set

The Bayesian approach with complete data uses the conjugacy relationship of the multinomial distribution and the Dirichlet distribution. Since each user's review is an outcome of an i.i.d. multinomial distribution, the posterior distribution remains in the Dirichlet family. However, this conjugacy relationship no longer holds when some ratings are missing (i.e., *partial data*) reflecting that the outcomes are not drawn from an identical distribution. For example, suppose there are two items. The associated multinomial trial has the outcomes $R = \{(1,1),(1,0),(0,1),(0,0)\}$. If a user has an empty rating on item 2 while she has a positive rating on item 1, this outcome corresponds to an outcome of a binomial trial with success probability $(p_{(1,1)} + p_{(1,0)}, p_{(0,1)} + p_{(0,0)})$. It is theoretically possible to find a posterior distribution, which is a combination of 2^E Dirichlet distributions in this case, where E refers to the number of missing ratings. Thus, computationally, it becomes more and more intractable to use the conditional probability for a prediction as E grows. However, in what follows, we show that the Bayesian prediction can still be derived using a relatively simple formula.

Let γ_j denote the j^{th} user's ratings over the $C+1$ items. Note that with *partial* data, even a non-target user may not have ratings on all $C+1$ items. Thus, unlike $r \in R$, γ_j may contain $\emptyset$ in its elements.

To begin, we first define some notations that are used to express the prediction. As in the main model, suppose that we have N users and the number of conditioning items is C. Since we are interested in making a prediction about each target item, we let $I = 1$ for simplicity. The corresponding set of outcomes is again denoted by R. A typical element r of R is a length $C+1$ vector whose i^{th} element records a non-empty rating of item i, $i \in \{1, \cdots, C+1\}$. For notational simplicity, we denote $\{1, \cdots, N\}$ by $\bar{N}$. For $r' \in R$ and for a given collection of ratings $\{r_j\}_{j\in\bar{N}}$, define $N(r'|\{r_j\}_{j\in\bar{N}})$ to be the number of r' in $\{r_j\}_{j\in\bar{N}}$. For example, when the collection of ratings is given by $\{r_1, r_0, r_1\}$, we have $N(r_1|\{r_1, r_0, r_1\}) = 2$ and $N(r_0|\{r_1, r_0, r_1\}) = 1$.

Now, for partial data X, which is a collection $\{x_j\}_{j\in\bar{N}}$, we want to estimate the probability that the target user likes the target item, $C+1$. Let $\bar{X}$ be an augmented data that has a positive rating of the target user 1's item $C+1$ while all the other ratings including the empty ratings stay the same as in X. That is, the $(C+1) \times 1^{th}$ element in X is $\emptyset$ and it is 1 in $\bar{X}$. Using this notation, the prediction can be denoted by $P[\bar{X}|X]$. To describe the process of the prediction, let x_j be the collection of x_{ij} for $i \in \{1, \cdots, C+1\}$ which is the set of ratings by user j, and $R(x_j)$ be the set of outcomes consistent with the ratings left by the user. That is,

$$R(x_j) = \{r \in R | r_i = x_{ij},\ \forall i\ s.t.\ x_{ij} \neq \emptyset\}.$$

Recall that in the Bayesian model in Section 4, we define $R_i^+(x_j)$ as follows

$$R_i^+(x_j) = \{r \in R | r_{ij} = 1\ and\ r_{kj} = x_{kj},\ \forall k\ s.t.\ x_{kj} \neq \emptyset\}.$$

Proposition 1 *The prediction with partial data can be characterized as follows:*

$$P[\bar{X}|X] = \frac{\sum_{\gamma_1 \in R^+_{C+1}(x_1), \gamma_j \in R(x_j),\ j \neq 1} \prod_{r \in R} \Gamma(\alpha_r^0 + N(r|\{\gamma_j\}_{j \in \bar{N}}))}{\sum_{\gamma_j \in R(x_j),\ j \in \bar{N}} \prod_{r \in R} \Gamma(\alpha_r^0 + N(r|\{\gamma_j\}_{j \in \bar{N}}))}.$$

When the prior parameters α^0 is a vector of natural numbers, the expression for the prediction can be represented as follows.

Corollary 1

$$P[\bar{X}|X] = \frac{\sum_{\gamma_1 \in R^+_{C+1}(x_1), \gamma_j \in R(x_j),\ j \neq 1} \prod_{r \in R} (\alpha_r^0 - 1 + N(r|\{\gamma_j\}_{j \in \bar{N}}))!}{\sum_{\gamma_j \in R(x_j),\ j \in \bar{N}} \prod_{r \in R} (\alpha_r^0 - 1 + N(r|\{\gamma_j\}_{j \in \bar{N}}))!}.$$

We can illustrate this prediction mechanism using a simple example. The ratings are summarized by the following table when user 1 is the target user and the target item is item 2. The set of outcomes is $R = \{(1,0),(1,0),(0,1),(0,0)\}$ and the associated probability is $p =$

	user 1	user 2	user 3
item 1	1	$\emptyset$	1
item 2	$\emptyset$	1	0

Table 1: Example 2

$(p_{(1,1)}, p_{(1,0)}, p_{(0,1)}, p_{(0,0)})$. Suppose that we begin with a uniform prior, i.e., $p^0 \sim Dir(1,1,1,1)$ whose associated joint density is denoted by f^0.

Instead of calculating the prediction directly out of the expression in the proposition, we will compute the prediction step by step following the steps in the proof of the proposition. From the data, we first construct the sets of outcomes that are consistent with the data: $R(x_1) = \{(1,1),(1,0)\}$, $R(x_2) = \{(1,1),(0,1)\}$, $R(x_3) = \{(1,0)\}$. Based on the prior q^0, we know that $R(x_1)$ happens with probability $p^0_{(1,1)} + p^0_{(1,0)}$, $R(x_2)$ happens with probability $p^0_{(1,1)} + p^0_{(0,1)}$ and

$R(x_3)$ happens with probability $p^0_{(1,0)}$. Thus, the probability that this data is generated is

$$\begin{aligned} P[X] &= \int_{p\in[0,1]^3} \big(p_{(1,0)}(p_{(1,1)}+p_{(0,1)})(p_{(1,1)}+p_{(1,0)})\big) f^0(p)dp \\ &= \mathbb{E}_{f^0}[p_{(1,0)}p^2_{(1,1)} + p^2_{(1,0)}p_{(1,1)} + p_{(1,0)}p_{(0,1)}p_{(1,1)} + p^2_{(1,0)}p_{(0,1)}] \\ &= \mathbb{E}_{f^0}[p_{(1,0)}p^2_{(1,1)}] + \mathbb{E}_{f^0}[p^2_{(1,0)}p_{(1,1)}] + \mathbb{E}_{f^0}[p_{(1,0)}p_{(0,1)}p_{(1,1)}] + \mathbb{E}_{f^0}[p^2_{(1,0)}p_{(0,1)}] \\ &= \frac{\Gamma(4)}{\Gamma(7)\Gamma(1)}\big(\Gamma(2)\Gamma(3)+\Gamma(3)\Gamma(2)+\Gamma(2)\Gamma(2)\Gamma(2)+\Gamma(3)\Gamma(2)\big) \\ &= \frac{7}{120}. \end{aligned}$$

Similarly, letting $R_2(x_1) = \{(1,1)\}$,

$$\begin{aligned} P[\bar{X}] &= \int_{p\in[0,1]^3} \big(p_{(1,0)}(p_{(1,1)}+p_{(0,1)})p_{(1,1)}\big) f^0(p)dp \\ &= \mathbb{E}_{f^0}[p_{(1,0)}p^2_{(1,1)}] + \mathbb{E}_{f^0}[p_{(1,0)}p_{(0,1)}p_{(1,1)}] \\ &= \frac{\Gamma(4)}{\Gamma(7)\Gamma(1)}\big(\Gamma(2)\Gamma(3)+\Gamma(2)\Gamma(2)\Gamma(2)\big) \\ &= \frac{3}{120}. \end{aligned}$$

Thus, the predicted probability of user 3's liking item 2 is given by $\frac{3}{7}$.

Table 2 demonstrates predictions on x_{21} based on different historical ratings. The uniform prior is assumed to derive the predictions.

	user 1	user 2	user 3
item 1	1	$\emptyset$	1
item 2	3/11	0	0

	user 1	user 2	user 3
item 1	1	$\emptyset$	0
item 2	3/7	0	0

Table 2: Predictions with partial data

A.1 Proof of Proposition 1 and Corollary 1 (Online Appendix)

Proof. We first find the probability that the data X happens given the prior distribution. Cosider a review vector (possibly incomplete) x_j. For a given prior q^0, we have

$$P[x_j] = \sum_{r\in R(x_j)} q^0_r.$$

Since the review of j is independent of j' conditional on the true parameter, the probability of X can be represented as

$$P[X] = \int \prod_{j=1}^{N} (\sum_{r\in R(x_j)} q^0_r) f^0(q^0)dq^0,$$

where f^0 is a joint distribution of q^0 according to the Dirichlet prior distribution. Changing the order of the multiplication and the summation,

$$
\begin{aligned}
P[X] &= \int \prod_{j=1}^{N} \Big(\sum_{r \in R(x_j)} q_r^0 \Big) f^0(q^0) dq^0 \\
&= \sum_{\gamma_j \in R(x_j), j \in \bar{N}} \int \prod_{k=1}^{N} q_{\gamma_k}^0 f^0(q^0) dq^0 \\
&= \sum_{\gamma_j \in R(x_j), j \in \bar{N}} \mathbb{E}_{f^0}\Big[\prod_{k=1}^{N} q_{\gamma_k}^0\Big].
\end{aligned}
$$

Note that the product moments of a Dirichlet random variable has a representation using Gamma functions.[1] Using the representation, we have

$$
\begin{aligned}
P[X] &= \sum_{\gamma_j \in R(x_j), j \in \bar{N}} \frac{\Gamma(\sum_{r \in R} \alpha_r^0)}{\Gamma(\sum_{r \in R}(\alpha_r^0 + N(r|\{\gamma_j\}_{j \in \bar{N}})))} \prod_{r \in R} \frac{\Gamma(\alpha_r^0 + N(r|\{\gamma_j\}_{j \in \bar{N}}))}{\Gamma(\alpha_r^0)} \\
&= \sum_{\gamma_j \in R(x_j), j \in \bar{N}} \frac{\Gamma(\sum_{r \in R} \alpha_r^0)}{\Gamma(N + \sum_{r \in R} \alpha_r^0))} \prod_{r \in R} \frac{\Gamma(\alpha_r^0 + N(r|\{\gamma_j\}_{j \in \bar{N}}))}{\Gamma(\alpha_r^0)} \\
&= \frac{\Gamma(\sum_{r \in R} \alpha_r^0)}{\Gamma(N + \sum_{r \in R} \alpha_r^0)) \prod_{r \in R} \Gamma(\alpha_r^0)} \sum_{\gamma_j \in R(x_j), j \in \bar{N}} \prod_{r \in R} \Gamma(\alpha_r^0 + N(r|\{\gamma_j\}_{j \in \bar{N}})).
\end{aligned}
$$

Similarly, we can compute $P[\bar{X}]$ and it is given by

$$
P[\bar{X}] = \frac{\Gamma(\sum_{r \in R} \alpha_r^0)}{\Gamma(N + \sum_{r \in R} \alpha_r^0)) \prod_{r \in R} \Gamma(\alpha_r^0)} \sum_{\gamma_1 \in R_{C+1}^+(x_1), \gamma_j \in R(x_j),\ j \neq n} \prod_{r \in R} \Gamma(\alpha_r^0 + N(r|\{\gamma_j\}_{j \in \bar{N}})).
$$

Now, by the conditional probability formula, we have

$$
P[\bar{X}|X] = \frac{\sum_{\gamma_1 \in R_{C+1}^+(x_1), \gamma_j \in R(x_j),\ j \neq 1} \prod_{r \in R} \Gamma(\alpha_r^0 + N(r|\{\gamma_j\}_{j \in \bar{N}}))}{\sum_{\gamma_j \in R(x_j),\ j \in \bar{N}} \prod_{r \in R} \Gamma(\alpha_r^0 + N(r|\{\gamma_j\}_{j \in \bar{N}}))}.
$$

■

B Welfare analysis with a finite dataset

In the paper, we mainly focused on the asymptotic value that a recommender system creates. The value when there are a finite number of data points available depends on the recommender system that is in use. In this section, building upon the Bayesian recommender system we proposed in Section 4, we study the value a recommender system that learns from a finite dataset offers to

[1] If $Z \sim Dir(w_1, \cdots, w_m)$, $\mathbb{E}[\prod_{i=1}^{m} Z_i^{n_i}] = \frac{\Gamma(w_1 + \cdots + w_m)}{\Gamma(w_1 + n_1 + \cdots + w_m + n_m)} \prod_{i=1}^{m} \frac{\Gamma(w_i + n_i)}{\Gamma(w_i)}$.

users.

The following proposition presents the expected utility to users that the Bayesian recommender system offers under the user-optimal threshold. The user welfare under a different threshold can immediately be derived using the proposition. For simplicity, we focus on the case of $I = 1$. Consider complete data X. We can find y and r' such that y records the occurrences of r in X' as defined in the paper and r' corresponds to the target user's history. As the learning model is invariant to the order of the elements in X', y and r' characterize X.

Proposition 2 (Finite sample data) *When there are $N-1$ previous users, the value to the target user from item $C+1$ can be characterized as follows:*

1. *Ex-post value: Given y and r', the expected utility to a history r' user is*

$$\left(v_1\frac{p_{(r',1)}}{p_{r'}} + v_0\frac{p_{(r',0)}}{p_{r'}}\right)\mathbf{1}[v_1(y_{(r',1)} + \alpha^0_{(r',1)}) + v_0(y_{(r',0)} + \alpha^0_{(r',0)}) \geq 0].$$

2. *Interim value: The expected utility to a history r' user is*

$$\left(v_1\frac{p_{(r',1)}}{p_{r'}} + v_0\frac{p_{(r',0)}}{p_{r'}}\right)\sum_{k=0}^{N-1}\binom{N-1}{k}(1-p_{r'})^{N-1-k}p_{r'}^k\sum_{j=\lfloor\tau\rfloor+1}^{k}\binom{k}{j}\left(\frac{p_{(r',1)}}{p_{r'}}\right)^j\left(\frac{p_{(r',0)}}{p_{r'}}\right)^{k-j}.$$

3. *Ex-ante value: The expected utility to a user is*

$$\sum_{r'\in R'}(v_1p_{(r',1)} + v_0p_{(r',0)})\sum_{k=0}^{N-1}\binom{N-1}{k}(1-p_{r'})^{N-1-k}p_{r'}^k\sum_{j=\lfloor\tau\rfloor+1}^{k}\binom{k}{j}\left(\frac{p_{(r',1)}}{p_{r'}}\right)^j\left(\frac{p_{(r',0)}}{p_{r'}}\right)^{k-j},$$

where $\tau = \frac{-v_1\alpha_{(r',1)} - v_0(k+\alpha_{(r',0)})}{v_1 - v_0}$.

The derivation of the above expressions proceeds as follows. The target item is recommended to a user with history r' if and only if the consumption of the item is more likely to induce positive utility than not. That is, when the prediction is $\hat{x}$, the recommendation is made if and only if

$$v_1\hat{x} + v_0(1-\hat{x}) \geq 0.$$

Using the representation in Proposition 7, this condition is equivalent to requiring $v_1(y_{(r',1)} + \alpha^0_{(r',1)}) + v_0(y_{(r',0)} + \alpha^0_{(r',0)}) \geq 0$. For example, when $(v_1, v_0) = (1, -1)$ and $(\alpha^0_{(r',1)}, \alpha^0_{(r',0)}) = (1, 1)$, the item is recommended if and only if $y_{(r',1)} \geq y_{(r',0)}$ is observed in the data. Since the recommendation is unbiased, the user tries the item whenever she receives the recommendation. Once she tries the item, she receives $v_1\frac{p_{(r',1)}}{p_{r'}} + v_0\frac{p_{(r',0)}}{p_{r'}}$. The ex-ante value is derived from the interim value taking into account that the new user has a history r' with probability $p_{r'}$. From the ex-ante perspective, the probability that the item is recommended to the user with r' history is

$$\sum_{k=0}^{N-1} \binom{N-1}{k} (1-p_{r'})^{N-1-k} p_{r'}^k \sum_{j=\lfloor \tau \rfloor +1}^{k} \binom{k}{j} \left(\frac{p_{(r',1)}}{p_{r'}} \right)^j \left(\frac{p_{(r',0)}}{p_{r'}} \right)^{k-j}.$$

It is the sum of probabilities that k out of $N-1$ previous users have history r', and among the k users, j users have positive experiences with the item. Here, we sum only the cases in which j exceeds the threshold level for a recommendation, namely τ.

C Details on Figure 4

Figure C1 presents the average utility of users under the threshold level τ^p and the degree of customization from zero to nine, along with the corresponding 95% confidence intervals from running the simulation 1,000 times.

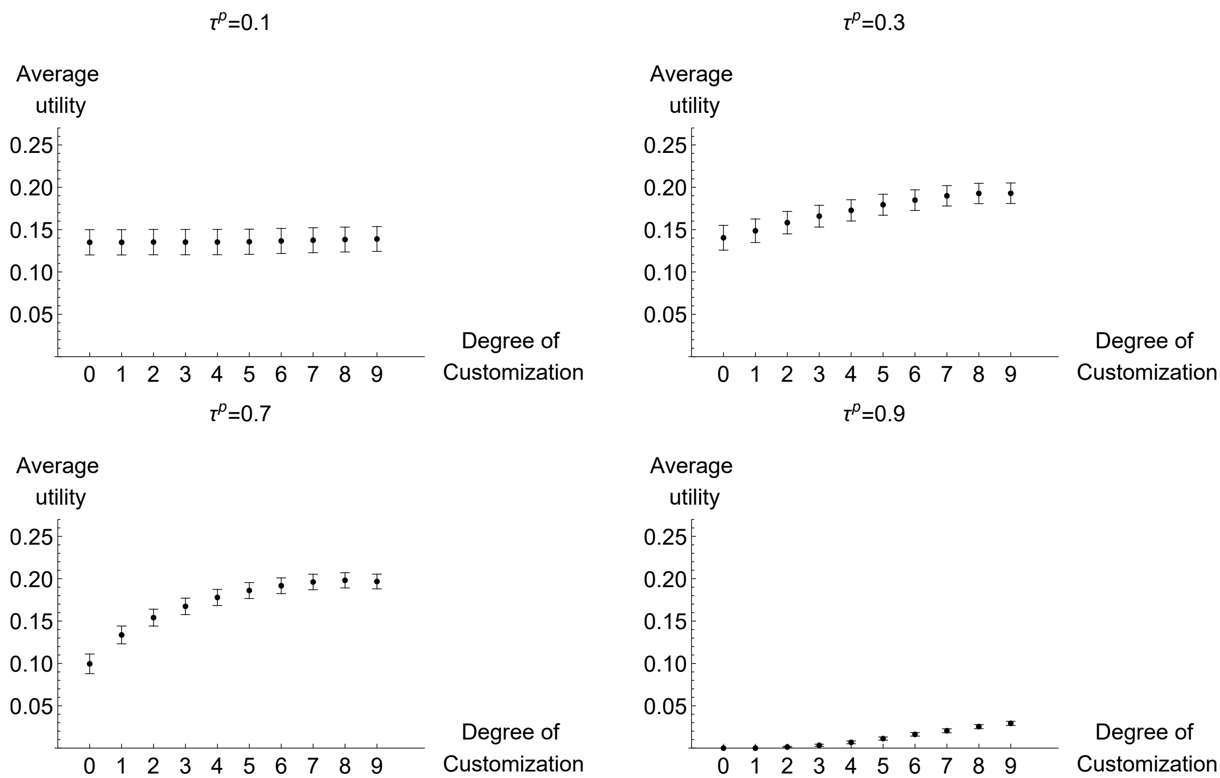

Figure C1: Different shapes of learning curves

D Within-user exploration

Consider the two-item platform in the Section 3.4 of the main text. The platform's data from $N-1$ users is summarized by $y = (y_{(1,1)}, y_{(1,0)}, y_{(0,1)}, y_{(0,0)})$. After Bayesian updating from the prior parameters α^0, the posterior Direchlet parameters accommodating y is denoted by $\alpha^1 = (\alpha^1_{(1,1)}, \alpha^1_{(1,0)}, \alpha^1_{(0,1)}\alpha^1_{(0,0)})$. Without loss of generality, we assume that $\alpha^1_{(1,0)} \geq \alpha^1_{(0,1)}$. So, item 1 is expected to generate a higher expected payoff for the user than item 2 does. We consider the expected utility that a recommender system generates for a new user, namely a target user, who lives for two periods and has a unit demand for each item. The user can consume at most one item in each period and we assume $(v_1, v_0) = (1, -1)$. We are interested in the dynamic recommendation policy over the two periods that maximizes the user's payoff. The user discounts the second period payoff with discount factor $\delta \in (0, 1]$.

Once the target user arrives, the platform makes its decisions over (1) which item to recommend first and (2) the threshold for a try-recommendation. We call the recommendation policy that recommends the highest myopic expected payoff first a *naïve order*. On the other hand, we call the recommendation threshold that induces the highest myopic expected payoff a *naïve threshold*. When the recommendation rule takes dynamic considerations optimally into account it will be called *dynamically optimal*. The next proposition predicts that, when there are two items, the *naïve order* is indeed dynamically optimal. However, the optimal threshold for a try-recommendation for item 1 is strictly below the level that maximizes the user's immediate surplus in many cases, i.e., the *naïve threshold* is not dynamically optimal.

Proposition 3 *For any two-item platforms and for any $\delta \in (0, 1]$,*

1. *The naïve order is optimal.*
2. *The dynamically optimal threshold for item 1 is strictly lower than $\frac{1}{2}$ if and only if*

$$\max\left\{\frac{\alpha^1_{(1,1)}}{\alpha^1_{(1,1)} + \alpha^1_{(1,0)}}, \frac{\alpha^1_{(0,1)}}{\alpha^1_{(0,1)} + \alpha^1_{(0,0)}}\right\} > \frac{1}{2}.$$

Proof. Assume $\alpha^1_{(1,0)} \geq \alpha^1_{(0,1)}$. Let $\bar{\alpha}^1 = \alpha^1_{(1,1)} + \alpha^1_{(1,0)} + \alpha^1_{(0,1)} + \alpha^1_{(0,0)}$ for notational simplicity. Regardless of the order of recommendation, for any user-optimal policy, the second item is recommended to the user if and only if its estimate is greater than 1/2. So, the naive order induces the following expected utility:

$$\begin{aligned}
&\frac{\alpha^1_{(1,1)} + \alpha^1_{(1,0)}}{\bar{\alpha}^1} - \frac{\alpha^1_{(0,1)} + \alpha^1_{(0,0)}}{\bar{\alpha}^1} \\
&+\delta\frac{\alpha^1_{(1,1)} + \alpha^1_{(1,0)}}{\bar{\alpha}^1}\left(\frac{\alpha^1_{(1,1)}}{\alpha^1_{(1,1)} + \alpha^1_{(1,0)}} - \frac{\alpha^1_{(1,0)}}{\alpha^1_{(1,1)} + \alpha^1_{(1,0)}}\right)\mathbf{1}\left\{\frac{\alpha^1_{(1,1)}}{\alpha^1_{(1,1)} + \alpha^1_{(1,0)}} \geq \frac{1}{2}\right\} \\
&+\delta\frac{\alpha^1_{(0,1)} + \alpha^1_{(0,0)}}{\bar{\alpha}^1}\left(\frac{\alpha^1_{(0,1)}}{\alpha^1_{(0,1)} + \alpha^1_{(0,0)}} - \frac{\alpha^1_{(0,0)}}{\alpha^1_{(0,1)} + \alpha^1_{(0,0)}}\right)\mathbf{1}\left\{\frac{\alpha^1_{(0,1)}}{\alpha^1_{(0,1)} + \alpha^1_{(0,0)}} \geq \frac{1}{2}\right\}.
\end{aligned}$$

On the other hand, the policy that recommends item 2 first yields

$$
\begin{aligned}
&\frac{\alpha^1_{(1,1)}+\alpha^1_{(0,1)}}{\bar{\alpha}^1}-\frac{\alpha^1_{(1,0)}+\alpha^1_{(0,0)}}{\bar{\alpha}^1}\\
&+\delta\frac{\alpha^1_{(1,1)}+\alpha^1_{(0,1)}}{\bar{\alpha}^1}\left(\frac{\alpha^1_{(1,1)}}{\alpha^1_{(1,1)}+\alpha^1_{(0,1)}}-\frac{\alpha^1_{(0,1)}}{\alpha^1_{(1,1)}+\alpha^1_{(0,1)}}\right)\mathbf{1}\left\{\frac{\alpha^1_{(1,1)}}{\alpha^1_{(1,1)}+\alpha^1_{(0,1)}}\geq\frac{1}{2}\right\}\\
&+\delta\frac{\alpha^1_{(1,0)}+\alpha^1_{(0,0)}}{\bar{\alpha}^1}\left(\frac{\alpha^1_{(1,0)}}{\alpha^1_{(1,0)}+\alpha^1_{(0,0)}}-\frac{\alpha^1_{(0,0)}}{\alpha^1_{(1,0)}+\alpha^1_{(0,0)}}\right)\mathbf{1}\left\{\frac{\alpha^1_{(1,0)}}{\alpha^1_{(1,0)}+\alpha^1_{(0,0)}}\geq\frac{1}{2}\right\}.
\end{aligned}
$$

Using these, the naive order is optimal *if and only if*

$$
\begin{aligned}
&\alpha^1_{(1,0)}-\alpha^1_{(0,1)}+\delta\mathbf{1}\{\alpha^1_{(1,1)}\geq\alpha^1_{(1,0)}\}(\alpha^1_{(1,1)}-\alpha^1_{(1,0)})-\delta\mathbf{1}\{\alpha^1_{(1,1)}\geq\alpha^1_{(0,1)}\}(\alpha^1_{(1,1)}-\alpha^1_{(0,1)})\\
&+\alpha^1_{(1,0)}-\alpha^1_{(0,1)}+\delta\mathbf{1}\{\alpha^1_{(0,1)}\geq\alpha^1_{(0,0)}\}(\alpha^1_{(0,1)}-\alpha^1_{(0,0)})-\delta\mathbf{1}\{\alpha^1_{(1,0)}\geq\alpha^1_{(0,0)}\}(\alpha^1_{(1,0)}-\alpha^1_{(0,0)})\geq 0.
\end{aligned}
$$

The value in the first line of the above inequality is positive. Firstly, if $\alpha^1_{(1,1)}\geq\alpha^1_{(1,0)}\geq\alpha^1_{(0,1)}$, the value is

$$\alpha^1_{(1,0)}-\alpha^1_{(0,1)}-\delta(\alpha^1_{(1,0)}-\alpha^1_{(0,1)})\geq 0.$$

Secondly, if $\alpha^1_{(1,0)}>\alpha^1_{(1,1)}\geq\alpha^1_{(0,1)}$, the value is

$$\alpha^1_{(1,0)}-\alpha^1_{(0,1)}-\delta(\alpha^1_{(1,1)}-\alpha^1_{(0,1)})>0.$$

The value is strictly positive because we have $\alpha^1_{(1,0)}>\alpha^1_{(1,1)}$. Lastly, if $\alpha^1_{(1,0)}\geq\alpha^1_{(0,1)}>\alpha^1_{(1,1)}$, the value is $\alpha^1_{(1,0)}-\alpha^1_{(0,1)}$, which is positive as we have $\alpha^1_{(1,0)}\geq\alpha^1_{(0,1)}$. Similarly, we can show the expression in the second line is positive. Thus, the naïve order is optimal.

Now, consider the threshold level of a try-recommendation for item 1. As there is no more item to consider after the consumption of item 2, the optimal threshold for item 2 is $\frac{1}{2}$. Suppose that the platform makes a recommendation of item 1 if and only if the expected value is above a threshold level τ. That is, a try-recommendation is made if and only if

$$\frac{\alpha^1_{(1,1)}+\alpha^1_{(1,0)}}{\bar{\alpha}^1}\geq\tau.$$

Since the left-hand side of the above inequality is discrete and the right-hand side is continuous, consider α^1 that satisfies

$$\frac{\alpha^1_{(1,1)}+\alpha^1_{(1,0)}}{\bar{\alpha}^1}\geq\tau>\frac{\alpha^1_{(1,1)}+\alpha^1_{(1,0)}-1}{\bar{\alpha}^1-1}.$$

For this α^1, the expected user surplus is

$$\begin{aligned}
&\frac{\alpha^1_{(1,1)}+\alpha^1_{(1,0)}}{\bar{\alpha}^1}-\frac{\alpha^1_{(0,1)}+\alpha^1_{(0,0)}}{\bar{\alpha}^1} \\
&+\delta\frac{\alpha^1_{(1,1)}+\alpha^1_{(1,0)}}{\bar{\alpha}^1}\left(\frac{\alpha^1_{(1,1)}}{\alpha^1_{(1,1)}+\alpha^1_{(1,0)}}-\frac{\alpha^1_{(1,0)}}{\alpha^1_{(1,1)}+\alpha^1_{(1,0)}}\right)\mathbf{1}\left\{\frac{\alpha^1_{(1,1)}}{\alpha^1_{(1,1)}+\alpha^1_{(1,0)}}\geq\frac{1}{2}\right\} \\
&+\delta\frac{\alpha^1_{(0,1)}+\alpha^1_{(0,0)}}{\bar{\alpha}^1}\left(\frac{\alpha^1_{(0,1)}}{\alpha^1_{(0,1)}+\alpha^1_{(0,0)}}-\frac{\alpha^1_{(0,0)}}{\alpha^1_{(0,1)}+\alpha^1_{(0,0)}}\right)\mathbf{1}\left\{\frac{\alpha^1_{(0,1)}}{\alpha^1_{(0,1)}+\alpha^1_{(0,0)}}\geq\frac{1}{2}\right\} \\
&=2\frac{\alpha^1_{(1,1)}+\alpha^1_{(1,0)}}{\bar{\alpha}^1}-1+\delta\frac{\alpha^1_{(1,1)}-\alpha^1_{(1,0)}}{\bar{\alpha}^1}\mathbf{1}\{\alpha^1_{(1,1)}\geq\alpha^1_{(1,1)}\}+\delta\frac{\alpha^1_{(0,1)}-\alpha^1_{(0,0)}}{\bar{\alpha}^1}\mathbf{1}\{\alpha^1_{(0,1)}\geq\alpha^1_{(0,0)}\}.
\end{aligned}$$

We want to find τ such that

$$2\tau-1+\delta\frac{\alpha^1_{(1,1)}-\alpha^1_{(1,0)}}{\bar{\alpha}^1}\mathbf{1}\{\alpha^1_{(1,1)}\geq\alpha^1_{(1,0)}\}+\delta\frac{\alpha^1_{(0,1)}-\alpha^1_{(0,0)}}{\bar{\alpha}^1}\mathbf{1}\{\alpha^1_{(0,1)}\geq\alpha^1_{(0,0)}\}=0$$

That is,

$$\tau=\frac{1}{2}-\delta\frac{\alpha^1_{(1,1)}-\alpha^1_{(1,0)}}{2\bar{\alpha}^1}\mathbf{1}\{\alpha^1_{(1,1)}\geq\alpha^1_{(1,0)}\}-\delta\frac{\alpha^1_{(0,1)}-\alpha^1_{(0,0)}}{2\bar{\alpha}^1}\mathbf{1}\{\alpha^1_{(0,1)}\geq\alpha^1_{(0,0)}\}$$

Clearly, $\tau<\frac{1}{2}$ unless we have both $\alpha^1_{(1,1)}<\alpha^1_{(1,0)}$ and $\alpha^1_{(0,1)}<\alpha^1_{(0,0)}$. ■

The second point in the proposition highlights that, unless item 2 is expected to generate a negative surplus regardless of the experience from item 1, the optimal threshold is always strictly lower than the threshold that maximizes the immediate surplus. Moreover, even though both items are expected to generate a negative surplus to a user, it can be optimal to recommend an item first, observe the user experience, and decide whether to recommend the other item. This latter result reflects the option value created from learning from the first item recommended.

Unlike the across-user exploration motive, the platform's within-user exploration motive does not disappear even after the platform eventually learns the true item correlation structure. Suppose the platform has learned the correlation structure $p=(p_{(1,1)},p_{(1,0)},p_{(0,1)},p_{(0,0)})$. We have the following corollary.

Corollary 2 *The optimal threshold for a try-recommendation for item 1 is strictly lower than* $\frac{1}{2}$ *if and only if*

$$\max\left\{\frac{p_{(1,1)}}{p_{(1,1)}+p_{(1,0)}},\frac{p_{(0,1)}}{p_{(0,1)}+p_{(0,0)}}\right\}>\frac{1}{2}.$$

Lastly, it should be noted that the *naïve order* does not have to be optimal when the platform offers more than two items. We address this issue by means of an example. Suppose the platform offers three items, and the user intends to consume all three items if it is expected to generate a positive surplus. Again, the user receives 1 from a positive experience and -1 from a negative experience. The posterior parameters updated using the previous users' data is given by

$$(\alpha^1_{(1,1,1)}, \alpha^1_{(1,1,0)}, \alpha^1_{(1,0,1)}, \alpha^1_{(1,0,0)}, \alpha^1_{(0,1,1)}, \alpha^1_{(0,1,0)}, \alpha^1_{(0,0,1)}, \alpha^1_{(0,0,0)}) = (4, 2, 1, 6, 4, 2, 1, 7).$$

From the posterior parameters, the platform can draw the expected utility from trying each item. It is $\frac{13}{27}$ from item 1, $\frac{12}{27}$ from item 2, and $\frac{10}{27}$ from item 3. The posterior parameters are constructed in a way that item 1 is most likely to induce the highest user surplus, but it is not closely related to the other items. By contrast, item 2 and item 3 are more correlated to each other.[2]

Suppose first that the platform uses the myopic policy, so that item 1 is recommended first. In this case, the immediate payoff to the user is $\frac{13}{27} - \left(1 - \frac{13}{27}\right)$. It can be checked that the optimal subsequent recommendation rule is to recommend item 2 regardless of the outcome from item 1 and then recommend item 3 if and only if the user has a positive experience with item 2. The value of this policy is as follows:

$$-\frac{1}{27} + \frac{13}{27}\left(-\frac{1}{13} + \frac{6}{13}\cdot\frac{1}{3} + \frac{7}{13}\cdot 0\right) + \frac{14}{27}\left(-\frac{2}{14} + \frac{6}{14}\cdot\frac{1}{3} + \frac{7}{14}\cdot 0\right) = 0.$$

Here, $-\frac{1}{27}$ is the immediate payoff to the user from trying item 1. The user has a positive experience with probability $\frac{13}{27}$. Once the user has a positive experience with item 1, her expected payoff from trying item 2 is $-\frac{1}{13}$. The user who has a positive experience with item 1 also has a positive experience with item 2 with probability $\frac{6}{13}$. Lastly, the expected payoff from trying item 3, given positive experiences with item 1 and item 2, is $\frac{1}{3}$. If the user has a negative experience with item 2, item 3 is not recommended to try, which induces 0 payoff. The events after a negative experience with item 1 can be interpreted similarly.

On the other hand, consider another policy that recommends item 2 first. In this policy, item 1 and item 3 are sequentially recommended to the user if she has a positive experience with item 2. If she has a negative experience with item 2, no item is further recommended. The user payoff from this non-myopic policy is as follows:

$$-\frac{3}{27} + \frac{12}{27}\left(\frac{4}{12} + \frac{8}{12}\cdot 0 + \frac{4}{12}\cdot 0\right) + \frac{14}{27}\cdot 0 = \frac{1}{27}.$$

The user surplus from the latter non-myopic policy is strictly larger than that from the myopic policy. As noted at the beginning of this example, this result arises because item 2 has a higher correlation with both item 1 and item 3 than item 1 does with the other items. So, in such a situation when an experience with an item reveals more information about experiences with other items, it can be better to let the user try the item first so as to tailor the subsequent recommendations for them.

[2] Specifically, $\alpha^1_{(1,1,1)}, \alpha^1_{(0,1,1)}, \alpha^1_{(1,0,0)}, \alpha^1_{(0,0,0)}$ are high while the rest of the parameters