E-commerce Recommendation with Personalized Promotion

Qi Zhao†, Yi Zhang†, Daniel Friedman‡, Fangfang Tan‡
† School of Engineering ‡ Economics Department
University of Calfornia, Santa Cruz, CA, 95060
qzhao2@ucsc.edu, yiz@soe.ucsc.edu,dan@ucsc.edu, tfangfan@ucsc.edu

ABSTRACT

Most existing e-commerce recommender systems aim to recommend the right products to a consumer, assuming the properties of each product are fixed. However, some properties, including price discount, can be personalized to respond to each consumer's preference. This paper studies how to automatically set the price discount when recommending a product, in light of the fact that the price will often alter a consumer's purchase decision. The key to optimizing the discount is to predict consumer's willingness-to-pay (WTP), namely, the highest price a consumer is willing to pay for a product. Purchase data used by traditional e-commerce recommender systems provide points below or above the decision boundary. In this paper we collected training data to better predict the decision boundary. We implement a new e-commerce mechanism adapted from laboratory lottery and auction experiments that elicit a rational customer's exact WTP for a small subset of products, and use a machine learning algorithm to predict the customer's WTP for other products. The mechanism is implemented on our own e-commerce website that leverages Amazon's data and subjects recruited via Mechanical Turk. The experimental results suggest that this approach can help predict WTP, and boost consumer satisfaction as well as seller profit.

Categories and Subject Descriptors

H.3.3 [**INFORMATION STORAGE AND RETRIEVAL**]: [Information filtering]

General Terms

Algorithms, Experimentation

Keywords

recommender system; e-commerce; consumer; promotion; crowdsourcing

1. INTRODUCTION

Recommender systems have achieved much commercial success and are becoming increasingly popular in a wide variety of practical applications. For example, online stores such as Amazon, iTunes and Walmart.com provide customized recommendations for additional products or services based on a consumer's history. Since it is widely believed that even minor improvements in recommender systems can boost the profitability of e-commerce companies, these systems have been studied intensively by researchers in industry and in academia.

An important limitation of existing studies is that they assume that the properties of items (i.e. products) are static. However, an e-commerce company could tailor some properties of a product for a particular customer, and that could dramatically improve the effectiveness of a recommendation. In particular, we argue that price is a controllable property that the recommender system should incorporate. A consumer might like a recommended product, but may reject it because the price is too high, and the purchasing decision could be changed by a personalized promotion. Of course, outside of the literature on recommender systems, the crucial role of pricing is widely recognized. Researchers in marketing, for example, have shown the importance of personalized promotion for increasing sales volume [17].

In this paper, we introduce personalized promotion into e-commerce recommender systems. The objective is to improve the effectiveness of product recommendations through customizing product price on an individual basis.

To achieve this goal, we create a novel auction/lottery mechanism to elicit consumer willingness to pay (WTP) for relevant products in an e-commerce setting. Using an online shopping website with products from Amazon, we recruit experimental subjects via Mechanical Turk and use responses elicited for some chosen products to predict a individual's WTP on other products. The predicted WTP and the given production cost enable us to find profitable personalized prices. The data indicate that the new approach achieves nearly 200% improvement in gross profit when compared with Amazon's default pricing strategy.

Our contribution is three-fold. First, to the best of our knowledge, our work is the first attempt to add personalized promotion to e-commerce recommender systems, and we measure recommendation performance with metrics of direct interest to industry. Second, we introduce a novel WTP elicitation procedure that is adapted to e-commerce systems. Most previous WTP elicitation studies are conducted in settings that are very different from an e-commerce

RecSys'15, September 16–20, 2015, Vienna, Austria.

DOI: http://dx.doi.org/10.1145/2792838.2800178.

website — typically only a handful of products in a narrow category are pre-selected by the experimenters — and so the applicability to e-commerce is unclear. Third, we have several methodological observations that could be very useful for further research on personalized pricing in e-commerce.

2. RELATED WORK

Despite the complexity of recommender systems, the most common approach is to reduce the problem to predicting the user's ratings or purchasing decision of items, and recommending items based on the prediction. In early work, the Grundy system built stereotypes for recommendation [28]. The Tapestry system let users identify similar users manually and made recommendations based on these similar users [16]. The advent of the internet resulted in large sets of user data. Consequently, manual constructions have been largely replaced by automatic recommendation algorithms, such as collaborative filtering algorithms, content-based filtering algorithms, and hybrid algorithms. **Collaborative filtering** is based on the assumption that users with similar tastes for previous items will have similar preferences for new items, so the algorithm recommends items ranked highly by users deemed similar to the current user. Such algorithms fall into two main categories [18] [10, 20, 21, 27]. **Content based filtering** is based on the assumption that the features (meta data, words in description, price, tags, etc.) used to describe the items that a user likes or dislikes tell much about the user's preferences. It usually recommends new items similar to previous items the user liked. The underlying research focuses on estimating a user's profile from explicit feedback on whether she liked previous items. State of the art text classification algorithms, such as Support Vector Machine (SVM), K nearest neighbors (K-NN), neural networks, logistic regression, and Winnow, have been used to solve this binary classification task [29, 36, 37, 30, 25, 35]. **Hybrid recommendation algorithms** combine collaborative filtering with content based filtering, and usually perform better than either filtering method alone [7].

Most existing research in recommender systems overlooks a factor that is extremely important in a consumer's decision making: the product price. One notable exception is [12], who show that price is a factor that is transferrable across categories to help recommendation system predictions. Another exception, [34], explores price's marginal net utility role in E-commerce recommendation. However, these works still assume price is given and are focusing on boosting recommendation prediction performance with price as additional information; they do not treat prices a as controllable factors to influence user decisions.

Two recent pilot studies examine pricing strategies in recommendation systems. Backhaus et al. [6] ask consumers questions over the phone, cluster consumers into different groups, use conjoint analysis to estimate customers' WTP, and then use the estimates to price a small set of intangible value-added services on B2B market when making recommendations. Massoud and Abo-Rizka[24] proposed a conceptual model of personalized pricing recommender systems. However, no experiments were done to implement or study the proposed conceptual model. Kamishma and Akaho [19] went a little further and proposed a method for adding price personalization to standard recommendation algorithms based on customer preference data and purchasing history. However, only artificially generated data (based on the MovieLens dataset) are used. It's not clear how the approaches proposed by these pilot studies can be applied in a real e-commerce system with a large and varied set of users and many products of different types, brands and price ranges.

There is an established literature in economics and marketing showing that personalized promotion is an important marketing tactic for increasing sales volume [17]. Especially important for our paper is prior research, also in economics and marketing, on estimating a consumer's WTP for a particular product. Roughly speaking, there are four empirical approaches: estimating WTP from transactions data, direct surveys, indirect surveys or laboratory auctions. Transaction data is incentive compatible in that it represents actual purchase decisions. However, the transaction price is only the lower bound on WTP, and equally relevant non-purchase decisions are missing from transactions data. Direct surveys (see, e.g., [31, 15]) estimate a consumer's WTP by directly asking indicate their maximum acceptable price for a given product [1]. However, as pointed out by Breidert et al. [9], direct surveys are not incentive compatible — the respondent is not motivated to reveal his or her true WTP, and often is motivated to understate it substantially. Indirect surveys offer respondents a list of alternative products (often hypothetical products with their properties varied independently) and ask them to either to rank them according to personal preference [23] as in conjoint analysis, or simply to choose the most favored alternative as in choice-based conjoint analysis (CBC)[22][2, 3, 4, 5, 13, 14]. Similar to direct surveys, the conjoint analysis and CBC are not incentive compatible. To overcome this problem, Ding et.al in [13] proposed an ICBC - incentive aligned choice-based conjoint method, which applies the BDM mechanism described below to the WTP inferred from conjoint analysis. Dong et al. [14] proposed ranking revealed products based on the WTP inferred from conjoint analysis data. The authors argue their proposed approach is incentive aligned as subjects will receive top ranked products.

The remaining empirical approach is WTP elicitation via laboratory auction. The famous Vickrey auction [33], requires each bidder to a submit sealed bid (not seen by other bidders); the bidder with highest bid wins the auction and pays a price equal to the second highest bid. This procedure is incentive compatible. The intuition is that a person's own bids only determine whether or not she wins the auction but never affects the price that she pays, and therefore she has no incentive to understate (or overstate) her WTP. Closely related to the Vickrey auction is the Becker-DeGroot-Marschak (BDM) mechanism to elicit WTP [8]. Under BDM, each bidder submits a bid to purchase a product. A sales price is randomly drawn from an interval which covers all plausible bids. If that sales price is lower than a participant's bid, then she receives the product and pays the sales price. The BDM is theoretically incentive compatible for the same reason as the Vickrey auction; indeed (although it was invented independently) it is equivalent to a Vickrey auction with two bidders, one human and the other an automaton who bids randomly. Our own approach exploits the fact that BDM is easier to operate in an e-commerce setting since, unlike the Vickrey auction, it does not require gathering all participants' bids in order to determine the winner.

Despite considerable research on WTP elicitation [26] and existing pilot research on pricing strategies when recommending, we still face several challenges when adapting existing approaches to real world e-commerce settings. Existing approaches often require an artificial laboratory setting with only a handful of products preselected by researchers, and/or with product data generated artificially. In e-commerce systems, however, consumers are exposed to a much larger number of products of different types and price ranges, and the system needs to estimate WTP for products that might be of interest to an individual user. The gaps between the settings makes it unclear whether the conclusions derived from laboratory experiments generalize to e-commerce settings.

3. METHODS

We envision a recommender system that not only predicts whether a consumer likes a product, but also predicts a consumer's *willingness-to-pay* (WTP) for it. WTP is the highest price the consumer is willing to pay for a product [32]. If the WTP is lower than the default product price, the system might give a personalized promotion to the consumer to increase the possibility of accepting the recommendation.

We propose that the e-commerce system run a lottery (related to the BDM mechanism just described) for customers to collect WTP for a small number of recommended products, and then use machine learning on each customer's lottery data to build a WTP prediction model. The WTP predictions then allow the recommender system to set personalized promotion prices automatically, potentially enhancing customer satisfaction as well as seller profit.

The rest of this section details this approach.

3.1 Bidding & Lottery Procedures

Our procedure elicits WTP for N products. To describe it clearly, we begin with the case of a single product ($N = 1$). The participant's true WTP, denoted y, is unknown to us. We endow her with a large amount of cash C that likely exceeds y, and tell her that the mechanism will draw a random price r uniformly from the range $[R_L, R_U]$, which is chosen to include all plausible values of her WTP. She is asked to state her actual WTP by entering a bid $y' \in [R_L, R_U]$, and told that the mechanism will sell her the product at price r only if that price is less than her bid.

Using the symbol P to denote probability, her payoff is

$$\mathcal{M} = \begin{cases} C & \text{if } r > y', P = \frac{R_U - y'}{R_U - R_L}, \\ C + y - r & \text{if } r \leq y', P = \frac{y' - R_L}{R_U - R_L}. \end{cases} \quad (1)$$

That is, she keeps the cash if the random price r exceeds her stated WTP y', and otherwise she purchases (and gains true benefit y) at the random price, which is lower (and hence a better deal) than her her stated WTP y'.

To see that this BDM procedure is incentive compatible, first note that the expected payoff is:

$$\begin{aligned} E\left[\mathcal{M}|y', y\right] &= P(r > y')C + P(r \leq y')E[C + y - r|r \leq y'] \\ &= C + P(r \leq y')(y - E[r|r \leq y']. \end{aligned} \quad (2)$$

Use the convenient property of the uniform distribution that the last conditional expectation is $E[r|r \leq y'] = \frac{R_L + y'}{2}$ to simplify the equation to:

$$E\left[\mathcal{M}|y', y\right] = C + \frac{2yy' - y'^2 - 2R_L y + R_L^2}{2(R_U - R_L)} \quad (3)$$

Equation 3 shows that the participant's expected payoff is a quadratic function of her decision variable y' whose unique global optimum is reached when $y' = y$. In other words, it is in the subject's best interest to state her WTP truthfully.

Note for later reference that when the participant indeed sets $y' = y$, her expected payoff can be written as

$$E\left[\mathcal{M}|y, y\right] = C + \frac{(y - R_L)^2}{2(R_U - R_L)} > C \quad (4)$$

In our experiment we set $R_L = 0$ and $R_U = p$, where p is the known outside (undiscounted) market price for the product. In this case,

$$E\left[\mathcal{M}|y, y\right] = C + \frac{1}{2}k^2 p \quad (5)$$

where $y = k * p$, and we refer to k as the normalized bidding price.

What if the subject wants to bid above R_U or below R_L? In principle, there is no problem; in the first case she obtains the object for sure and reveals that her true WTP exceeds R_U, and in the second case she keeps the cash for sure and reveals that her true WTP is below R_L.

A potential problem does arise when the number of products $N > 1$. Running a separate BDM elicitation for several products that are substitutes for each other will motivate participants to underbid, since the probability of winning at least one of several bids is higher than winning a single bid. We eliminate this problem by having each participant enter a bid for each product, but only drawing a random bid for one of the products chosen randomly. That is, we conduct a lottery over which of the products will actually be available to the participant. The expected payoff (or $1/N$ of the expected payoff) is still maximized by truthfully reporting WTP for each product.

An extension of this lottery approach enables us to economize payments in the experiment. We can also conduct a lottery over which participants actually get the cash C and perhaps one of the products. Again, that attenuates the expected payoff but doesn't change where it is maximized.

3.2 Personalized Promotion

Our lottery/auction procedure gives us training data for predicting a consumer's WTP on a range of products that might be recommended. Given the values of y elicited from consumer u for product i, we run the regression

$$y_{ui} = b_0 + b_u + f(\mathbf{x_{ui}}, \mathbf{w}), \quad (6)$$

where b_0 is the global bias and b_u is the user bias, $\mathbf{x}_{ui}$ is a feature vector representing information about consumer u, product i and their relationships. The functional form of the regression function f depends on the machine learning algorithm used, such as linear regression or gradient boosted trees. We used *linear regression* (LR) in our experiments. Thus our model parameters are:

$$(b_0, b_u, \mathbf{w}^*) \quad = \quad \underset{b_0, b_u, \mathbf{w}}{\arg\min} \sum_{u,i} L(y_{ui} - b_u - b_0 - f(\mathbf{x_{ui}}, \mathbf{w})) \quad (7)$$

where in our implementation the loss function L is quadratic.

Given the model parameters learned from the training dataset, we use Equation 6 to predict WTP for any user u and item i pair. To capture the uncertainty in our estimation, we assume that the true WTP value follows a Gaussian

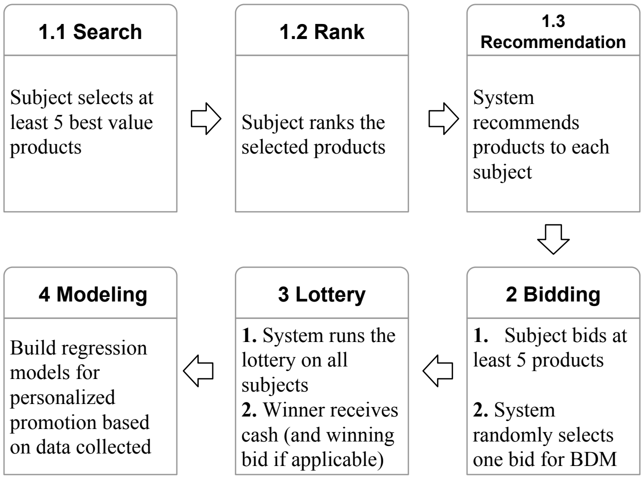

Figure 1: Overview of the experiment flow.

distribution centered on the predicted WTP:

$$Y_{ui} \sim \mathcal{N}(\mu_{ui}, \sigma^2) \tag{8}$$

where $\mu_{ui} = b_u + b_0 + f(\mathbf{x}_{ui}, \mathbf{w})$. The parameter σ^2 could be estimated, but for simplicity we assume that is the same for all user item pairs and approximate it using the variance of y on the training dataset in our experiments.

Based on the distribution of Y_{ui}, we can find the optimal price t^* that maximizes the expected seller profit as follows:

$$t^* = \arg\max_t (t - c)P(Y_{ui} \geq t) \tag{9}$$

where c is the production cost of product i and

$$\begin{aligned} P(Y_{ui} \geq t) &= \int_t^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\frac{(x-\mu_{ui})^2}{2\sigma^2}} dx \\ &= \frac{1}{2} - \frac{1}{2} erf\left(\frac{t - \mu_{ui}}{\sqrt{2}\sigma}\right) \end{aligned} \tag{10}$$

where $erf(x)$ is Gauss error function encountered in integrating the normal distribution. Equation 9 is solved numerically using standard optimization techniques.

4. EXPERIMENTAL PROCEDURES

Can the proposed lottery-auction mechanism and the prediction model actually work with normal e-commerce customers? Are those methods of practical value to sellers and customers? To begin to answer these questions we ran an experiment as follows.

4.1 Data collection

We first developed a website that links to amazon.com, a leading online shopping website. Our website implements a full fledged product search engine and displays the product information in a similar way to Amazon. The website hosts about 120k skin care products. We chose skin care products as they have short repurchase cycle, reducing problems that arise from unknown current holdings of the participants. The products are from amazon.com and the product information is synchronized with amazon.com at real time. For each product, Amazon provided both its list price and sales price.

Subjects recruited from Amazon Mechanical Turk were paid a participation fee of \$0.50 for around 10 minutes to complete the experiment. Figure 1 is an overview of the procedure. First, the subject sees a set of recommended products produced by a standard recommendation algorithm. This step would be straightforward for a typical e-commerce system, which can find relevant products based on subject's past purchasing or click data. Since our custom website don't have the recruited subject's purchasing history, we just let each subject search and identify at least 5 products which he or she thinks are worth buying; these are referred to as *best value* products. Then the subject proceeds to next step to rank their best value products in decreasing order of interest. Next, the system recommends a list of products using Amazon's "consumer who bought this also bought these" recommendations.

Having completed the steps labelled 1.1-1.3 in Figure 1, the recruited subject then participates in N BDM lotteries. That is, she enters a bid on each of $N \geq 5$ products that she chooses from those recommended in the previous step, knowing that at most one of the products will be randomly selected to go through the BDM procedure. Finally, one of the subjects recruited that day is randomly selected as the lottery winner, and one of her chosen N products is randomly selected. If her bid on that product exceeds the random buyout price r, then she receives the product and \$100 minus r. Otherwise she gets \$100 cash.

4.2 Subjects' training and selection

We showed earlier that truth telling (bidding one's actual WTP) is the unique optimal strategy in BDM. However, earlier empirical research has shown that people may have misconceptions about the game and may not use the optimal bidding strategy [11]. Guided by the earlier research, we provided detailed explanations about the game rule and encouraged our subjects to practice with a bidding simulator to see how the outcome changes with different bids. Then each subject was required to take a quiz to check whether the participant understands the optimal bidding strategy. Only the 79 subjects who passed the quiz (out of 130 who responded initially) were allowed to participate in the game.

4.3 Data collection

To clean the data, records with normalized bidding prices either below 10% or more than two times higher than the Amazon sale price are filtered out, as are the bids of subjects who took less than 5 minutes to complete the experiment. After cleaning, we ended up with 339 product bids from 54 subjects. The bid price is normalized using the product's Amazon sale price, as k in Equation 5.

To construct $\mathbf{x}_{ui}$, we used the 9 features described in Table 1. The features are based on information collected from each experiment step.

4.4 Evaluation Metrics

RMSE (Root Mean Square Error) is the most commonly used metric and is naturally adopted in our experiments. However, to understand the commercial value of personalized promotion based on the estimation of consumer's WTP, we go beyond RMSE and introduce *seller profit* to measure the effectiveness of WTP prediction. Seller profit of a single product purchase is simply the purchase price minus the cost of producing the product. For the proposed WTP pre-

Name	Description	Weight
list price	Amazon list price of the product	−0.0013
number of reviews	number of reviews on Amazon	−0.001
brand variety	the variety of the brands of consumer's best value products, measured by entropy. The larger value of this feature, the more variety.	0.018
discount	Amazon promotional discount	0.399
Amazon recommendation rank	rank in Amazon product recommendation list	0.02
average rating	Amazon average rating	0.044
user rank	rank in consumer's best value products	−0.0017
switch brand	whether product(s) of the same brand as the bidding product have been identified by consumer as best value products	−0.059
brand popularity	the number of times getting bidded	0.054
intercept	*intercept term of the LR model*	−0.058

Table 1: The features used to predict WTP modeling, and their weight coefficients learned by Equation 7.

diction method in Section 3.2, the profit is expected profit given pricing via Equation 9. Hence

$$\text{Profit}_{ui} = \begin{cases} t^*_{ui} - c_i & \text{if } y_{ui} \geq t^*_{ui} \\ 0 & \textit{otherwise} \end{cases} \tag{11}$$

where t^*_{ui} is the optimal personalized promotion price from Equation 9, y_{ui} is the consumer's true WTP and c_i is the unit production cost of product i.

5. EXPERIMENTAL RESULTS

We compared the proposed approach with Amazon sale price with 0, 10% and 20% discounts and ZeroR. The Amazon sale price methods serve as baselines. ZeroR algorithm is a simple learning based method that find the mean value of y on the training data and simply set the price at the learned mean value, which means a fixed discount rate. In our experiment, the normalized bidding price for product i by ZeroR is 0.72 (i.e. the learned discount rate is 0.28). Our proposed algorithm LR sets the price at the point where the expected profit is maximized (Equation 9). All algorithms are evaluated by 10 fold cross validation.

In our experiments, we don't know the production cost of each product. Because retail businesses commonly set price using a fixed markup on cost, we assume that production cost is a certain percentage of the sale price. We varied the percentage from 0 to 50% with an incremental step of 10%. Table 2 reports the corresponding profit and conversion rate for each algorithm at different production costs level.

Table 2 shows that our proposed method (i.e. LR) significantly outperforms Amazon's pricing strategy and simple learning method (i.e. ZeroR). All improvements are statistically significant ($p < 5\%$). The LR algorithm maximizes expected profit as in Equation 9, using $\sigma = 0.22$ based on the training dataset. Is it important to model the *uncertainty* of WTP and set the price by maximizing the expected profit? The answer is yes. We compared the proposed method with simply setting the price at the most likely WTP (i.e. $b_0 + b_i + f(x_{ui}, w)$), and we found the simple method is statistically significantly worse.

5.1 Further Analysis

The feature weights of normalized bidding price k in the LR model may be of interest. Table 1 indicates that:

List price: normalized WTP decreases slightly as product's list price increases.

Number of reviews: normalized WTP decreases as the number of reviews increases. This is counter intuitive, as consumers presumably would tend to perceive goods with more reviews as more valuable. Of course, the expression is reduced form and for normalized (not direct) WTP, so future work may be able to find an explanation.

Discount: Products with larger Amazon discounts have higher normalized WTP. Like the list price impact, this effect probably also works mainly through the normalization (again recall Equation 5) but it is much stronger.

Average rating: Consumers are willing to pay more for products rated higher.

User rank: Not surprisingly, a consumer is willing to pay more for products she ranks more highly.

Switch brand: A consumer switching to a new brand tends to have lower WTP. This seems closely related to the

Algorithm	RMSE	Profit ($) w.r.t production cost (percent of product price)					
		None	10%	20%	30%	40%	50%
Amazon price	0.4549	784.69	675.73	566.76	457.79	348.82	239.85
90% Amazon price	0.4012	1267.18	1091.21	915.24	739.27	563.30	387.33
80% Amazon price	0.3674	1884.11	1622.47	1360.83	1099.18	837.54	575.90
ZeroR	0.3586	2358.13	2030.66	1703.19	1375.72	1048.25	720.78
LR	**0.2200**	**2734.14**	**2260.44**	**1831.22**	**1442.21**	**1104.83**	**818.60**

Table 2: Performance of different WTP prediction algorithms in 10-fold cross validation setting. Profit is evaluated w.r.t production cost as percentage of sale price. Note that the flat discount rate is applied to Amazon market price (after promotion discount off the list price).

brand variety effect, as variety seekers are more likely to switch brand than loyal customers.

Brand popularity: Not surprisingly, the more popular the brand, the more consumer is willing to pay.

6. CAVEATS

The analysis above noted many technical simplifications (e.g., normally distributed estimation errors for normalized WTP with constant variance), most of which can be relaxed in future work. This may be a good juncture to mention more general complications that we have put to one side.

- Most products have substitutes (and complements); and we mostly sidestepped the resulting complications. As this line of research matures it may be useful to consider bundles of goods and sets of alternatives explicitly.
- Personalized discounts, if not properly framed, can provoke customer backlash. Large scale application of our suggestions must tread carefully to avoid causing resentment.
- Our profit metric only took into account unit production costs. These could easily include handling and shipping, and researchers with access to cost data should have no problem extending our approach. However, a different treatment would be required to account for the costs of running a recommender system, or the costs of eliciting WTP. Presumably these are an order of magnitude smaller than production costs but they still could be important.
- Our Mechanical Turk subjects were convenient and turned out to quite useful. There is no reason to believe that naturally occurring e-commerce customers will behave in qualitatively different ways, but that remains an empirical question to be tested by researchers with access to large numbers of such customers. True A/B testing on that subject pool would be the gold standard.

7. CONCLUSION

This paper proposes including personalized promotion in e-commerce recommender systems. We developed a lottery-auction mechanism to elicit consumers' willingness to pay on a small subset of products, and a machine learning model to predict each consumer's WTP on a wide range of other products. We demonstrated the feasibility of the proposed approach in an experiment with real world products from Amazon and subjects recruited from Amazon Mechanical Turks. The results suggest that personalized promotion leads to significantly higher profits for sellers compared to the baseline pricing.

Our work also shows the viability of doing e-commerce experiments via crowdsourcing; our Mechanical Turk subjects turned out to be quite useful. Besides recommender systems, the proposed approach also has practical implication for managerial marketing.

The work just presented is only first step; much remains to be done. First and foremost, personalized promotion and recommendation should be considered jointly within a unified framework, and not remain as separate problems. Second, we focused on seller's profit as the evaluation metric. However, we believe it is crucial to include consumer surplus into the objective function and we see no conceptual obstacles in doing so. Finally, as alluded to in the caveats, it will be crucial to study the longer term effects on consumer satisfaction with personalized promotion. We hope that researchers in industry, with better access to normal e-commerce consumers, will be inspired to do so.

8. ACKNOWLEDGMENTS

Part of this work is sponsored by the National Science Foundation under grant CCF-1101741 and IIS-0953908. Any opinions, findings, conclusions or recommendations expressed in this paper are the authors, and do not necessarily reflect those of the sponsors.

9. REFERENCES

[1] J. Abrams. A new method for testing pricing decisions. *The Journal of Marketing*, pages 6–9, 1964.

[2] G. M. Allenby, N. Arora, and J. L. Ginter. Incorporating prior knowledge into the analysis of conjoint studies. *Journal of Marketing Research*, pages 152–162, 1995.

[3] G. M. Allenby, N. Arora, and J. L. Ginter. On the heterogeneity of demand. *Journal of Marketing Research*, pages 384–389, 1998.

[4] N. Arora, G. M. Allenby, and J. L. Ginter. A hierarchical bayes model of primary and secondary demand. *Marketing Science*, 17(1):29–44, 1998.

[5] N. Arora and J. Huber. Improving parameter estimates and model prediction by aggregate customization in choice experiments. *Journal of Consumer Research*, 28(2):273–283, 2001.

[6] K. Backhaus, J. Becker, D. Beverungen, M. Frohs, O. Müller, M. Weddeling, R. Knackstedt, and M. Steiner. Enabling individualized recommendations and dynamic pricing of value-added services through willingness-to-pay data. In *Electronic Markets*, 2010.

[7] J. Basilico and T. Hofmann. A joint framework for collaborative and content filtering. In *27th Annual International ACM SIGIR Conference*, 2004.

[8] G. M. Becker, M. H. DeGroot, and J. Marschak. Measuring utility by a single-response sequential method. *Behavioral science*, 9(3):226–232, 1964.

[9] C. Breidert, M. Hahsler, and T. Reutterer. A review of methods for measuring willingness-to-pay. *Innovative Marketing*, 2(4):8–32, 2006.

[10] J. Canny. Collaborative filtering with privacy via factor analysis. In *SIGIR '02: Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 238–245, New York, NY, USA, 2002. ACM.

[11] T. N. Cason and C. R. Plott. Misconceptions and game form recognition of the bdm method: challenges to theories of revealed preference and framing. 2014.

[12] J. Chen, Q. Jin, S. Zhao, S. Bao, L. Zhang, Z. Su, and Y. Yu. Does product recommendation meet its waterloo in unexplored categories?: No, price comes to help. In *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*, SIGIR '14, pages 667–676, New York, NY, USA, 2014. ACM.

[13] M. Ding. An incentive-aligned mechanism for conjoint analysis. *Journal of Marketing Research*, pages 214–223, 2007.

[14] S. Dong, M. Ding, and J. Huber. A simple mechanism to incentive-align conjoint experiments. *International Journal of Research in Marketing*, 27(1):25–32, 2010.

[15] A. Gabor, C. W. Granger, and A. P. Sowter. Real and hypothetical shop situations in market research. *Journal of Marketing Research*, pages 355–359, 1970.

[16] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry. Using collaborative filtering to weave an information tapestry. *Commun. ACM*, 35(12):61–70, 1992.

[17] S. Gupta. Impact of sales promotions on when, what, and how much to buy. *Journal of Marketing research*, pages 342–355, 1988.

[18] T. Hofmann and J. Puzicha. Latent class models for collaborative filtering. In *Proceedings of the International Joint Conference in Artificial Intelligence*, 1999.

[19] T. Kamishima and S. Akaho. Personalized pricing recommender system: Multi-stage epsilon-greedy approach. In *Proceedings of the 2Nd International Workshop on Information Heterogeneity and Fusion in Recommender Systems*, HetRec '11, pages 57–64, New York, NY, USA, 2011. ACM.

[20] Y. Koren, R. Bell, and C. Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009.

[21] Y. J. Lim and Y. W. Teh. Variational Bayesian approach to movie rating prediction. In *Proceedings of KDD Cup and Workshop*, 2007.

[22] J. J. Louviere and G. Woodworth. Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data. *Journal of marketing research*, pages 350–367, 1983.

[23] Y. Marbeau. What value pricing research today. *Journal of the Market Research Society*, 29(2):153–182, 1987.

[24] M. Massoud and M. Abo-Rizka. A conceptual model of personalized pricing recommender system based on customer online behavior. 2012.

[25] P. McNamee, C. Piatko, and J. Mayfield. JHU/APL at TREC 2002: Experiments in filtering and arabic retrieval. In *Proceeding of the Eleventh Text REtrieval Conference (TREC-11)*, 2002.

[26] K. M. Miller, R. Hofstetter, H. Krohmer, and Z. J. Zhang. How should consumers' willingness to pay be measured? an empirical comparison of state-of-the-art approaches. *Journal of Marketing Research*, 48(1):172–184, 2011.

[27] A. Paterek. Improving regularized singular value decomposition for collaborative filtering. In *Proceedings of KDD Cup and Workshop*, 2007.

[28] E. Rich. User modeling via stereotypes. *Cognitive Science*, 3(4):329 – 354, 1979.

[29] S. Robertson and I. Soboroff. The TREC-10 filtering track final report. In *Proceeding of the Tenth Text REtrieval Conference (TREC-10)*, pages 26–37. National Institute of Standards and Technology, special publication 500-250, 2002.

[30] M. Srikanth, X. Wu, and R. Srihari. UB at TREC 11: Batch and adaptive filtering. In *Proceeding of the Eleventh Text REtrieval Conference (TREC-11)*, 2002.

[31] R. G. Stout. Developing data to estimate price-quantity relationships. *The Journal of Marketing*, pages 34–36, 1969.

[32] H. R. Varian and W. Norton. *Microeconomic analysis*, volume 2. Norton New York, 1992.

[33] W. Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance*, 16(1):8–37, 1961.

[34] J. Wang and Y. Zhang. Utilizing marginal net utility for recommendation in e-commerce. In *SIGIR 2011*. ACM.

[35] L. Wu, X. Huang, J. Niu, Y. Xia, Z. Feng, and Y. Zhou. FDU at TREC 2002: Filtering, Q&A, web and video tasks. In *Proceeding of the Eleventh Text REtrieval Conference (TREC-11)*, 2002.

[36] Y. Yang, S. Yoo, J. Zhang, and B. Kisiel. Robustness of adaptive filtering methods in a cross-benchmark evaluation. In *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2005.

[37] Y. Zhang. Using bayesian priors to combine classifiers for adaptive filtering. In *SIGIR '04: Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 345–352, New York, NY, USA, 2004. ACM Press.