learning consumer opinions from online reviews [23, 26]. In our work, through user modeling, we identify how users *behave* as a response to online reviews [2], and demonstrate how to extract demographic-specific preferences. Other studies proposed to combine popularity with user feedback or social annotations to refine search results [4, 14].

8. DISCUSSION AND FUTURE WORK

We presented a ranking algorithm that uses a behavioral model of consumers, based on utility maximization. The model generates an estimate of how much each product characteristic contributes to the product's overall utility, and estimates the sensitivity of consumers to changes in various product characteristics. The estimation models are privacy-friendly as they do not require individual consumer data but rather rely on aggregate data. Based on the generated models, we can estimate the surplus that each product generates for each consumer, and build rankings that capture the user preferences. We demonstrated, through extensive user studies, that our ranking schemes are better than any of the existing baselines. We also showed that personalized surplus-based rankings are even better than the non-personalized surplus-based rankings. By doing so, we are able to target at each individual customer, and offer products with the "*best value for money*" in response to consumer queries.

We should also note that our ranking scheme is "causal," in the sense that the model can predict what "should" happen when we observe changes in the market. For example, when we see a new product in the marketplace, we can rank it by simply observing its characteristics, without waiting to see the consumers' demand for the product. Also, we can dynamically change the rankings as a reaction to changes in the products. For example, if we observe a price change, or if we observe that a hotel closes its pool for renovations, we can adjust immediately the surplus values and re-estimate the rankings.

Also, in order to better understand the antecedents of consumer's decisions, future work can look not only at transaction data but also into their browsing history and learning behavior. For example, our current model assumes that consumers are engaging into optimal utility maximizing behavior. However, this is not always true, as some consumers are more thorough than others in their search. By leveraging browsing histories, we can build models that explicitly take into consideration the fact that some users are "utility optimizers" and some others simply engage into "satisficing." It would be also interested in examining the difference in the conversion rate of users, when presented with surplus-based rankings.

By examining product search through the "economic lens" of consumer behavior, we can leverage micro-economic theory and many theoretical models that have been developed over the years, which try to capture the decision-making process of humans. Economic theory provides a very solid basis upon which we can build further computer science research, which has a different focus than economic research. Our example is illustrating: while economists have been building utility models for years, their goal was to estimate demand for products and the notion of surplus was just "a means to an end" and never had of value by itself. By focusing on product ranking, we showed how surplus can improve product search. Our experimental results demonstrated a significant improvement in user satisfaction. Other economic models (e.g., measuring the utility of product bundles) can also be directly used in consumer-facing applications on the Web (e.g., search for "product bundles" instead of simple products). We are very optimistic that this interdisciplinary research direction can generate very interesting results in the future.

References

[1] Adomavicius, G., and Tuzhilin, A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE TKDE 17* (2005), 734–749.

[2] Archak, N., Ghose, A., and Ipeirotis, P. G. Show me the money!: deriving the pricing power of product features by mining consumer reviews. In *KDD* (2007), pp. 56–65.

[3] Balke, W.-T., and Güntzer, U. Multi-objective query processing for database systems. In *Proceedings of 28th International Conference on Very Large Data Bases (VLDB)* (2004), pp. 936–947.

[4] Bao, S., Wu, X., Fei, B., Xue, G., Su, Z., and Yu, Y. Optimizing web search using social annotations. In *WWW* (2007).

[5] Berry, S. Estimating discrete choice models of product differentiation. *RAND Journal of Economics 25* (1994), 242–262.

[6] Berry, S., Levinsohn, J., and Pakes, A. Automobile prices in market equilibrium. *Econometrica 63* (1995), 841–890.

[7] Berry, S., and Pakes, A. The pure characteristics demand model. *International Economic Review 48* (2007), 1193–1225.

[8] Chevalier, J. A., and Goolsbee, A. Measuring prices and price competition online: Amazon.com and BarnesandNoble.com. *Quantitative Marketing and Economics 1*, 2 (2003), 203–222.

[9] Forman, C., Ghose, A., and Wiesenfeld, B. Examining the relationship between reviews and sales: the role of reviewer identity disclosure in electronic markets. *ISR 19*, 3 (2008), 291–313.

[10] Ghose, A., , and Ipeirotis, P. G. Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. *IEEE TKDE* (2010).

[11] Ghose, A., Ipeirotis, P., and Sundararajan, A. Opinion mining using econometrics: A case study on reputation systems. In *ACL* (2007).

[12] Hansen, L. Large sample properties of generalized method of moments estimators. *Econometrica 50*, 4 (1982), 1029–1054.

[13] Heckman, J. Instrumental variables: A study of implicit behavioral assumptions used in making program evaluations. *Journal of Human Resources 32*, 3 (1997), 441–462.

[14] Jin, R., Valizadegan, H., and Li, H. Ranking refinement and its application to information retrieval. In *WWW* (2008).

[15] Lancaster, K. *Consumer Demand: A New Approach.* Columbia University Press, New York, 1971.

[16] Li, B., Ghose, A., and Ipeirotis, P. G. Stay elsewhere? improving local search for hotels using econometric modeling and image classification. In *WebDB* (2008).

[17] Marshall, A. *Principles of Economics*, Eighth ed. Macmillan and Co., London, 1926.

[18] McFadden, D. *Conditional Logit Analysis of Qualitative Choice Behavior.* Academic Press, New York, 1974.

[19] McFadden, D., and Train, K. Mixed MNL models of discrete response. *Journal of Applied Econometrics 15*, 5 (2000), 447–470.

[20] Mooney, R., and Roy, L. Content-based book recommending using learning for text categorization. In *ACM SIGIR Workshop Recommender Systems: Algorithms and Evaluation* (1999).

[21] Nelder, J. A., and Mead, R. A simplex method for function minimization. *The Computer Journal 7*, 4 (1965).

[22] Nie, Z., Wen, J.-R., and Ma, W.-Y. Webpage understanding: beyond page-level search. *SIGMOD Record 37*, 4 (2008), 48–54.

[23] Pang, B., and Lee, L. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval 2*, 1-2 (2008).

[24] Rosen, S. Hedonic prices and implicit markets: Product differentiation in pure competition. *J. of Political Econ. 82*, 1 (1974), 34–55.

[25] Song, M. A hybrid discrete choice model of differentiated product demand with an application to personal computers. FR 08-09, 2008.

[26] Ye, Q., Law, R., and Gu, B. The impact of online user reviews on hotel room sales. *Int. J. of Hosp. Mgmnt. 28*, 1 (2009), 180–182.

[27] Yee, K.-P., Swearingen, K., Li, K., and Hearst, M. Faceted metadata for image search and browsing. In *CHI* (2003), pp. 401–408.