Risky Purchases and Increased Choice: A Dynamic Model of Information Sharing and Purchase Decisions for Online Retail

Online retail, with its hundreds of billions of dollars in revenue (Statista 2018) and innovative business models, continues to present challenging but promising operational problems. In comparing the online consumer purchase process to the brick-and-mortar process, there are some apparent differences, notably (1) a significantly higher level of uncertainty about product quality along different dimensions such as durability, style, and color, and (2) a vast increase in choice alternatives, where subcategories that contain more than 1,000 products are becoming more the norm than the exception. Because of these two aspects, platforms rely on customers to write reviews of the items they purchase, and these reviews play a key role in the purchase decision.

As a sign of the importance of these reviews, platforms such as Amazon, Wayfair, and Overstock incentivize customers to leave reviews by sending them free items or giving them rewards points.

These operational facts raise interesting research questions: How should one conceptualize the influence of reviews on demand for products in a given category, especially when demand for different products is interdependent across the products? How do the information and demand dynamics work in this setting? Given a finite budget for incentivizing reviews, which users should be targeted with review invitations, and which products should they be invited to review?

How should the online platform quantify its objective, which includes trading off short-term gains, long-term gains via information gathering, and in some cases fairness considerations for its demand-side clients?

These questions relate to and build off of several streams of literature, including choice modeling and assortment optimization, models of customer search, economics observational models, and
models of sales dynamics such as in the case of network goods. Informed by sales data from an online platform, we model the demand and information dynamics of online sales and reviews for a product subcategory of durable goods. Customers have uncertain information about the true quality of products before their purchase decision. This uncertainty can be partially mitigated by reviews from earlier customers who have experienced the product already. Products are represented as elements in a multidimensional feature space, which are random from the point-of-view of the customers. Customers have heterogeneous preferences for different features. They perform Bayesian updating on their prior beliefs about the location of the products in the feature space, and then maximize their expected posterior utilities to make purchase decisions. Should they decide to purchase, they leave a review with some probability, and in the review, they mention each feature of the product with a probability that is increasing in the customer’s valuation for that feature.

The setup results in a dynamic choice model, where choice probabilities for each customer type depend on the number of reviews which mention their preferred features, and thus implicitly on the past demand for products. These dynamics result in similar effects to the dynamics in the cases of network goods (Wang and Wang 2016) or social influence (Hu et al. 2016), but we provide an alternative explanation for these effects as the result of information gain rather than network externalities or social utility, which we believe is helpful to explain demand dynamics for a wider class of products which may not exhibit network or social influence effects.

We show that under our model, given enough time periods customers will always eventually learn the true quality of underlying products, but the learning rate can be very different depending on the true quality and the make-up of the customer base (in terms of their preferences for different features). We quantify this rate of learning as a function of true product quality and
customer preference distribution, and use it to derive qualitative insights about the dynamics of the process. For example, we show that the more homogeneous the customer base, the more uneven the learning rate across different product features. We also show that products which are of high quality with respect to features that matter to most customers benefit from two positive effects: a learning effect, where true information is gathered faster, and a “Matthew effect” (Hu et al. 2016) where new customers choose the product more frequently because it has been purchased more frequently in the past. The low-quality products experience slower rates of learning about the relevant qualities, but the Matthew effect decreases their purchase rate quickly all the same.

Finally, we use this model to quantify for a platform the value of a review from a specific customer type of a specific product at a specific time, and use this to inform policies for soliciting reviews from customers.

