

Towards Automating the Pricing Power of Product Attributes: An Analysis of Online Product Reviews

Nikolay Archak, Anindya Ghose, Panagiotis G. Ipeirotis
Stern School of Business, New York University

The Internet has had a profound impact on at least two areas of life – the way people shop, and the way they exchange information. Both are relevant to consumer product reviews posted in IT-enabled electronic markets. Online consumer product reviews provide information that can facilitate economic exchange, which is the main purpose of electronic marketplaces. In particular, many users like to learn about the experiences of other customers with a product before purchasing the product. Online product reviews have been shown to influence product sales such as books and movies (Chevalier & Mayzlin 2006). Similarly, the volume of discussion about a product in blogs has recently been shown to correlate with the product's financial performance (Gruhl, Guha, Kumar, Novak & Tomkins 2005). We aim to extend these studies in different ways.

At a broader level, we plan to empirically estimate how the textual content of online reviews affects economic variables such as product prices and demand. We have been collecting price and sales rank data from Amazon for over 2 years now. Since the sales ranks can be converted into demand using a Pareto mapping parameter (Brynjolfsson, Hu and Smith 2003), we are able to econometrically impute the extent to which the textual content of online product reviews affects the sales of goods in a given product category. In our ongoing research, we use Amazon Web services to retrieve the reviews for all the products that exist in our dataset.

Specifically, our research involves a two step analysis. We use techniques similar to the ones described in (Hu & Liu 2004, Popescu & Etzioni 2005, Ghani, Probst, Liu, Crema & Fano 2006, Lee 2004) to identify product features and their related evaluations. These are the features that are most likely to be highly valued by consumers and hence, contribute to the aggregate price of that product. However, to date there exist no economic metrics that can isolate the economic impact associated with the various features of a multi-attribute product. For instance, while we can observe from the online word of mouth discussions whether customers like or dislike the zoom capabilities of a digital camera, we do not know whether these reviews actually affect a consumer's decision to buy the camera. Similarly, while we can observe that consumers value the differences in pixel resolutions across different brands of digital cameras, we do not know the extent to which the zoom capability of a camera is valued more than its pixel resolution by consumers. In other words, current information extraction systems that are used for opinion mining do not capture these subtle nuances embedded in the economics of consumer behavior.

In order to compute the economic value associated with each product feature and measure the extent to which they contribute to a manufacturer's pricing power for a given product, we use hedonic regressions (Rosen 1974, Feenstra 1995, Pakes 2002) to decompose a product into its constituent characteristics, and measure the economic

contribution of each of the product characteristics to the overall price of a product. By incorporating textual feedback in hedonic regressions, we can account for the effect of “qualitative” features, such as “ease of use” or “zoom quality,” while that have been traditionally ignored in prior economics and marketing research including the work on conjoint analysis.

Our research will allow us to not only quantify the economic value of reviews, but also identify which characteristics of the products are the most crucial and influential in determining the overall revenue from a product. From a managerial standpoint, our research can yield novel query interfaces on top of these econometrically quantified review data. This can be useful for users who are interested in determining the important characteristics of a category of products (e.g., digital cameras), and for firms who are interested in evaluating the optimality of the pricing strategies of their products.

At this point, we are in the middle of the experimental evaluation of our technique. By the end of February we expect to have a completed paper, which can be presented at the conference.