Deriving the Pricing Power of Product Features by Mining consumer Reviews

by Nikolay Archak, Anindya Ghose and Panagiotis Ipeirotis

Discussant: Ricardo Ribeiro (LSE)

NET Institute Conference

18th April 2008
Research Objective

Idea:

- Consumer product reviews impacts the demand of some product categories (e.g. Audio & Video and Camera & Photo).

The paper aims to:

- Identify the set of product features addressed on consumer’s reviews (e.g. lens, size, battery, ...).

- Identify the consumer’s opinion about each of the identified product features (e.g. bad, good, nice, amazing, ...).
• Estimate the weight that consumers place on each individual product feature (e.g. consumers may review the lens as a feature but not place any weight on it when buying the product).

• Estimate the implicit evaluation score of the consumer’s opinion about each feature (e.g. consumers may review the lens as amazing but that does not value much when buying the product).
Research Objective (Cont.)

How? Can we say that a review like "good battery" is better than "good lens" or than "nice battery"?

By deriving both the weight and evaluation score estimates in relation to changes in demand - a very clever idea!!
Econometric Strategy

Step 1: Identify the set of product features - using the part-of-speech tagger developed by the Stanford NLP Group.

Step 2: Identify the consumer’s opinion about each of the identified product features - using the syntactic dependency parser from the Stanford NLP toolkit.

Step 3: Estimate the non linear demand model by an iterative algorithm,

\[ \ln(D_{kt}) = \alpha + \beta p_{kt} + \gamma^T W(phrase, review_t, product_t) \delta + \varepsilon_{kt}. \]
A very well written paper with an extremely interesting research question!!

**Suggestion 1:** Identification discussion. Intuitively, demand for product $k$ may depend not only on its own-price, but also on the price of competing products.

For the above specification, those prices would be included in the error term. As prices of the different products tend to be correlated, there may be a correlation between a regressor and the error term. As a result the parameter estimates may exhibit a bias, which could justify some of the counterintuitive results on evaluations like "best camera", "perfect camera", ...

IV techniques may test this argument.
**Suggestion 2**: Extension towards a structural framework where demand is derived from a consumer utility function (e.g. a multinomial logit demand function: McFadden, 1978).

- Characteristic approach is still valid: see e.g. Berry, 1994 and Verboven, 1996.

- Can be applied even when market level quantity data is not publicly available, but only market share ranks: see e.g. Bajari, Fox & Ryan, 2007.

- Allows welfare analysis.