The Hong Kong University of Science and Technology
Department of Information Systems, Business Statistics and Operations Management

Seminar Announcement

Identifying and Overcoming Self-Selection Biases in Online Product Reviews

by

Professor Paul Pavlou
Fox School of Business
Temple University

Date: 9 July 2009 (Thursday)
Time: 11:00 am - 12:30 pm
Venue: Room 4379, ISOM Conference Room (Lift 17/18)

All interested are welcome

Abstract

Online product reviews help consumers infer product quality, and their mean (average) rating is often used as a proxy for product quality. However, we identify two self-selection biases that may render the mean rating a biased estimator. First, acquisition bias - mostly consumers with a favorable predisposition acquire the product and report a review; and second, under-reporting bias - consumers with extreme (either positive or negative) ratings are more likely to report their reviews. Econometric results from several Internet retailers show that virtually all online product reviews have an asymmetric, positively-skewed, bimodal (or 'J-shaped') distribution. In contrast, an experimental study where all respondents posted their reviews shows that the corresponding distributions are approximately normal.

Our analytical model explains the sources of the two self-selection biases and shows that they decrease consumer surplus by preventing the mean from reflecting true product quality. The model proposes an unbiased estimator of product quality that draws upon the entire distribution of product reviews, and not only the mean. The model analytically derives an unbiased product quality estimator by identifying the parameters needed to overcome: (1) acquisition bias - price and the standard deviation of the pre-acquisition quality perception, and (2) under-reporting bias - the two modes of the review distribution and the standard deviation of the utility shock that consumers experience after acquisition. These parameters render an unbiased estimator of product quality by recreating the underlying distribution from the J-shaped distribution. Numerical results show that the mean is not an unbiased estimator of product quality and it does not have a monotonic relationship with product quality. Thus, the mean rating cannot be used to either infer a product's absolute quality or to compare relative quality across products, and the proposed parameters are needed to infer true product quality.