MARKETING SCIENCE
Vol. 23, No. 3, Summer 2004, pp. 280–303
ISSN 0732-2399 | EISSN 1526-548X | 04 | 2303 | 0280

informs

DOI 10.1287/mksc.1040.0050

Modeling Browsing Behavior at Multiple Websites

Young-Hoon Park
Johnson Graduate School of Management, Cornell University, 330 Sage Hall, Ithaca, New York 14853-6201, yp34@cornell.edu

Peter S. Fader
Wharton School, University of Pennsylvania, 749 Huntsman Hall, 3730 Walnut Street, Philadelphia, Pennsylvania 19104-6340, faderp@wharton.upenn.edu

While there is a growing literature on investigating the Internet clickstream data collected for a single site, such datasets are inherently incomplete because they generally do not capture shopping behavior across multiple websites. A customer's visit patterns at one or more other sites may provide relevant information about the timing and frequency of his or her future visit patterns at the site of interest.

We develop a stochastic timing model of cross-site visit behavior to understand how to leverage information from one site to help explain customer behavior at another. To this end, we incorporate two sources of association in browsing patterns: one for the observable outcomes (i.e., arrival times) of two timing processes and the other for the latent visit propensities across a set of competing sites. This proposed multivariate timing mixture model can be viewed as a generalization of the univariate exponential-gamma model.

In our empirical analysis, we show that a failure to account for both sources of association not only leads to poor fit and forecasts, but also generates systematically biased parameter estimates. We highlight the model's ability to make accurate statements about the future behavior of the "zero class" (i.e., previous nonvisitors to a given site) using summary information (i.e., recency and frequency) from past visit patterns at a competing site.

Key words: Internet browsing behavior; data integration; multivariate duration models; customer acquisition

History: This paper was received May 27, 2002, and was with the authors 3 months for 3 revisions; processed by Scott A. Neslin.