The Impact of Location on Consumer Purchases in Electronic Markets

2nd Statistical Challenges in E-Commerce Research Symposium

Chris Forman
Tepper School of Business
Carnegie Mellon University

Anindya Ghose
Stern School of Business
New York University

Avi Goldfarb
Rotman School of Management
University of Toronto

Overview: It is well documented that the Internet retailing revolution has established a new distribution channel that represents a fundamental paradigm shift in consumer buying patterns. The rapid growth of alternative retail channels has transformed not only the competitive structure of several industries, but also the way in which consumers shop for products. Despite a wealth of research on electronic commerce, no prior research has measured how geographical location shapes consumer buying behavior in electronic markets. Do consumers in different locations derive different benefits from using the Internet in terms of selection, convenience, and price? This paper proposes a novel framework and data to address these issues. Our empirical strategy addresses an important problem that has been inadequately addressed in statistical e-commerce research: how consumers substitute between online and offline channels. We develop a methodology for understanding online/offline substitution by combining online purchase data with offline data on demographics and local channels.

Theory: Traditionally, retail markets for consumer products such as books, music, and DVDs have been predominately local. As local markets increase in size from small towns to large cities, consumer welfare is impacted in two ways. First, holding product offerings among retailers fixed, increases in location size encourage entry. Such entry leads to lower prices as well as improved service levels such as greater convenience. Moreover, larger markets also allow retailers to provide product offerings targeted to market segments that would be infeasible in smaller market settings. The emergence of new online retailing channels such as Amazon.com may act as a substitute for the benefits of urban concentration by both offering lower prices and by providing increased product differentiation for rural consumers. Like offline retailing, the key components of online retailing are selection, convenience, and price. We examine how the benefits of the online channel (due to increased convenience, wider selection, and lower prices) vary across locations.

This paper is related to multiple streams of research. First, it adds to the literature on the potential for Internet technology to reduce the costs associated with rural isolation (Forman, Goldfarb, and Greenstein 2005). Second, it adds to recent research on the consumer welfare benefits that online channels provide by lowering search costs (Ghose, Smith, and Telang 2006). Third, by looking at how consumers use online channels to substitute for offline retail supply deficiencies we contribute to the literature that uses spatial data to capture variations in supply-side factors—such as local retail competition—and demand-side factors—such as local consumer preferences (Bronnenberg and Mahajan 2001, Jank and Kannan 2006), each of which can influence consumer buying behavior across channels. Finally, we view this research as part of a wide literature that examines the impact of the Internet on consumer welfare more generally (e.g. Brynjolfsson and Smith 2000, Scott Morton, Zettlemeyer, and Silva-Risso 2001).

Data: The data that we use for this study come from the web pages on “Purchase Circles” on the Amazon.com web site. We used a JAVA spider to extract and parse data from Amazon’s website. Purchase Circles are specialized best-seller lists. The pages denote the top-selling books, music, and DVDs across large and small towns throughout the US. We collect weekly data on purchase circles over a period of ten months from April 2005 to January 2006. As Amazon.com is by far the largest online retailer (for instance, it has 70% of the market share in the book market-Forrester Research 2002), this paper provides an excellent way of measuring the use of online channels across locations within the US.
The Purchase Circles are organized in multiple layers - first, by state and then by town within a state. For each town, Amazon provides a list of top 20 sellers in the town for each product category. Amazon also provides on its website the national sales rank within that product category (books, music, etc.) for each product that it sells. Though previous studies have used data from Amazon, we believe we are among the first to use the Purchase Circles data.

To obtain data on location size and demographic characteristics, we match our data set to US Census data. We use Census County Business Patterns data on the number and size of many types of retail stores in each location to identify local offline retail supply. We also have data on openings and closings of stores including Wal-Mart, Barnes and Noble, and other major retailers.

**Broad Econometric Strategy:** We use the information on the top selling products in each location to determine how the characteristics of the location influence the online/offline purchase decision. In particular, we view the decision to buy a particular item online as the consequence of two separate decisions: (1) the decision to buy the item and (2) the decision to buy online rather than offline. Conditional on buying an item, the online/offline decision will depend on the local supply of the item. Our analysis controls for location and product fixed effects. The identification is off the interaction between location and product characteristics. Using this framework, we study how geographic variations in offline selection, convenience, and price influence online product choice.

**Contributions:** Utilizing a unique panel data set of online purchases of books, music, and DVDs by consumers across urban and rural locations in the US, we propose an econometric study to examine how location shapes consumer use of online channels. We develop a framework that controls for local product-specific tastes to determine how the benefits of purchasing online vary with local retail supply conditions. In particular, controlling for consumer preferences, we examine whether consumers with few local retail options purchase systematically more popular or less popular and more or less expensive products than urban customers. This in turn demonstrates how the selection, convenience, and price benefits of the Internet are different for customers in different types of locations.

The statistical and econometric challenges in examining these questions lead to two further methodological contributions. First we develop a model that allows us to study how online channels substitute for offline options without using offline purchase data. Second, we develop an identification strategy to separate the selection, convenience, and price benefits of the Internet.

**References**