DRAFT
COMMENTS WELCOME

The Great Equalizer?

Consumer Choice Behavior at Internet Shopbots

Erik Brynjolfsson* and Michael D. Smith**

May 2000
(Revised: April 13, 2001)
Available at http://www.heinz.cmu.edu/~mds/papers/tge

Acknowledgements: We thank Il-Horn Hann, John D.C. Little, Thomas Malone, Allen Montgomery, Nancy Rose, Robert Zeithammer, and seminar participants at Boston University, Carnegie-Mellon University, Indiana University, the University of Maryland, MIT, the University of Michigan, the University of Texas at Austin, Stanford University, The Wharton School, the Brookings conference on "Measuring E-Commerce" (Brookings Institution, September 24, 1999), and the Workshop on Information Systems and Economics (December 11, 1999) for valuable comments on this research. We thank Christoph Janz and Christopher Muenchhoff of EvenBetter.com for generously collecting the data necessary to conduct this research and for providing valuable insight into the shopbot market. Thomas Cheng and Frederic Mahoue provided excellent research assistance. The Center for eBusiness@MIT provided generous financial support under a grant from Fleet Bank.

---

* Sloan School of Management, MIT; email: erikb@mit.edu; web: http://ebusiness.mit.edu/erik
** H. John Heinz III School of Public Policy and Management, Carnegie Mellon University; email: mds@cmu.edu; web: http://www.heinz.cmu.edu/~mds

The Great Equalizer?
Consumer Choice Behavior at Internet Shopbots

ABSTRACT

Our research empirically analyzes consumer behavior at Internet shopbots — sites that allow consumers to make “one-click” price comparisons for product offerings from multiple retailers. By allowing researchers to observe exactly what information the consumer is shown and their search behavior in response to this information, shopbot data has unique strengths for analyzing consumer behavior. Furthermore, the method in which the data is displayed to consumers lends itself to a utility-based evaluation process, consistent with econometric analysis techniques.

We find that among these shopbot customers, while price is an important determinant of customer choice, branded retailers and retailers a consumer visited previously hold significant price advantages in head-to-head price comparisons. Our models accurately predict consumer behavior out of sample, suggesting that our analyses effectively capture relevant aspects of consumer choice processes and can form a useful basis for understanding consumer behavior at Internet shopbots.

1. Introduction

> "The Internet is a great equalizer, allowing the smallest of businesses to access markets and have a presence that allows them to compete against the giants of their industry."
>
> *Jim Borland, Knight Ridder (1998)*[1]

> "The cost of switching from Amazon to another retailer is zero on the Internet. It's just one click away."
>
> *Thomas Friedman, New York Times (1999)*[2]

> "Shopbots deliver on one of the great promises of electronic commerce and the Internet: a radical reduction in the cost of obtaining and distributing information."
>
> *Greenwald and Kephart (1999)*

Two decades ago information technology and bar code scanners radically reduced the cost of tracking and recording consumer purchases. A pioneering paper by Guadagni and Little (1983) used these data to estimate a multinomial logit model to analyze attribute-based consumer decision making in a retail environment. The results and extensions of their research (e.g., Kamakura and Russell 1989; Fader and Hardie 1996) have since been widely applied by academic researchers and by industry analysts for market forecasting, new product development, and pricing analysis.

Today continued reductions in computing cost and the rise of commercial uses of the Internet augur a similar revolution in retailing and consumer analysis. Our research seeks to apply multinomial logit models as a first step in understanding consumer behavior in Internet markets.

A better understanding of Internet markets could be particularly important in markets served by Internet shopbots. The Internet has been called "The Great Equalizer" because the technological capabilities of the medium reduce buyer search and switching costs and eliminate spatial competitive advantages that retailers would enjoy in a physical marketplace. Internet shopbots are emblematic of this capability.

---

[1] Borland, Jim. 1998. "Move Over Megamalls, Cyberspace Is the Great Retailing Equalizer." *Knight Ridder/Tribune Business News*, April 13.
[2] Friedman, Thomas L. 1999. "Amazon.you" *New York Times*, February 26, p. A21.

Shopbots are Internet-based services that provide one-click access to price and product information from numerous competing retailers. In so doing, they substantially reduce buyer search costs for product and price information.[3] They also strip away many of the accoutrements of a retailer's brand name by listing only summary information from both well- and lesser-known retailers.[4] Further, every retailer at a shopbot is "one click away," reducing switching costs accordingly. In each instance these factors should serve to increase competition and reduce retailer margins in markets served by shopbots — an effect that should be felt most strongly for homogeneous physical goods (e.g., Bakos 1997).

One wonders, then, what will happen to a retailer's brand equity and consumer loyalty in the presence of shopbots. Amazon.com has invested hundreds of millions of dollars in developing its online brand position. Likewise, brick-and-mortar retailers such as Barnes & Noble and Borders are attempting to transfer the value of their existing brand names to online markets.

Our research addresses these questions by analyzing consumer behavior through panel data gathered from an Internet shopbot. We use these data to study two major aspects of Internet shopbot markets. First, we analyze how consumers respond to the presence of retailer brand names. Second, we use Internet cookie data to analyze consumer loyalty to retailers they had visited previously. In addition, we analyze the correspondence between predicted and actual consumer behavior to assess the reliability of our models and the potential for retailers to use shopbot data to facilitate dynamic or personalized pricing strategies.

We find that branded retailers and retailers a customer had dealt with previously are able to charge $1.13 and more than their rivals, *ceteris paribus*. Our results also show that consumers are willing to pay an average of $2.49 more to buy from a retailer they have visited previously. Potential sources for the importance of brand and loyalty include service quality differentiation, asymmetric quality information, and cognitive lock-in. Further, we find a high correspondence between predicted and actual consumer behavior in our data suggesting that our models capture

[3] To illustrate this, we had a group of students compare the time needed to gather price quotes through various means. They found that gathering 30 price quotes took 3 minutes using a Internet shopbot, 30 minutes by visiting Internet retailers directly, and 90 minutes by making phone calls to physical stores. In practice, shopbots also introduce buyers to numerous retailers who would otherwise remain unknown to them.
[4] This characteristic of shopbots was the subject of recent litigation between eBay and BiddersEdge.com.

relevant aspects of consumer decision-making. We also note that retailers may be able to use the predictability of consumer behavior demonstrated in these models to facilitate personalized pricing strategies.

Our approach to analyzing electronic markets differs from recent empirical studies in that it examines the responses of actual consumers to prices set by retailers, not just the retailers' pricing behavior. Research analyzing retailer pricing strategies has been used to characterize the relative efficiency of electronic and physical markets (Bailey 1998; Brynjolfsson and Smith 2000), retailer differentiation strategies (Clay, Krishnan, Wolff, Fernandes 1999), and price discrimination strategies (Clemons, Hann, and Hitt 1998). However, retailer pricing strategies provide only second-order evidence of consumer behavior in electronic markets.

In this regard, shopbots provide Internet researchers with a unique opportunity to analyze actual consumer behavior in Internet markets. At Internet shopbots, thousands of consumers a day search for product information on different books. Their searches return comparison tables with a great deal of variation across retailers in relative price levels, delivery times, and product availability. Consumers then evaluate the product information and make an observable choice by clicking on a particular product offer. The result is a powerful laboratory where Internet researchers can observe snapshots of consumer behavior and, by tracking cookie numbers, consumer behavior over time.

The data available at Internet shopbots have several natural parallels to grocery store scanner data. First, shopbot data present consumer decisions made in response to a choice between several alternatives. Second, salient product attributes are observable by both consumers and researchers. Third, consumer behavior can be tracked over time.

The remainder of this paper is organized in four parts. Section 2 addresses the data we collect how it was collected and its strengths and limitations. Section 3 discusses the empirical models we use to analyze our data. Section 4 presents our results. Section 5 concludes, discusses implications of our results, and areas for future research.

2. Data

2.1. Data Source

We use panel data collected from EvenBetter.com to analyze consumer behavior at Internet shopbots. We selected EvenBetter for four reasons. First, EvenBetter sells books — well-defined homogeneous physical goods in a relatively mature Internet market. By analyzing shopping behavior in markets for homogeneous goods, we are able to control for systematic differences in the physical products through our methodological design. Additionally, homogeneous physical goods provide a useful reference point for the importance of brand and retailer loyalty because they should experience strong price competition in the presence of markets with low search costs (Bakos 1997). Examining relatively mature Internet markets ensures a sufficient number of consumers and retailers to draw meaningful conclusions.

A second reason for choosing EvenBetter is that their service offers consumers a more detailed list of product attributes than most other shopbots for books. This information includes separate fields for the total price, item price, shipping cost, sales tax, delivery time, shipping time, and shipping service. Third, EvenBetter does not offer priority listings to retailers who pay an extra fee (as do some other shopbots; e.g., MySimon.com). An unbiased listing of retailers provides a clearer interpretation of the factors driving consumers' choices. Fourth, EvenBetter.com has a revenue sharing arrangement with many of its retailers allowing us to compare descriptive statistics for the relative sales conversion ratios of the different retailers.

A disadvantage of using data gathered from Internet shopbots is that our analysis is restricted to consumers who choose to use a shopbot. Consumers who choose to use a shopbot are likely to be systematically different than consumers who visit Internet retailers directly. Thus, our logit model predictions must be understood as being conditioned on a consumer choosing to use a shopbot. Conditioning on prior consumer choice in this way does not bias multinomial logit results (Ben-Akiva and Lerman 1985). Furthermore, in analyzing the effect of this self-selection bias on our results, it seems reasonable to assume that shopbot consumers are more price sensitive than typical Internet consumers are. Thus, our estimates of brand and loyalty effects are likely to be lower bounds on the importance of brand and loyalty among the broader population of Internet consumers.

2.2. Data Characteristics

EvenBetter's shopbot operates similarly to many other Internet shopbots. A consumer who wants to purchase a book visits EvenBetter and searches on the book's title or author, ultimately identifying a unique ISBN as the basis for their search.[5] EvenBetter then queries 47 distinct book retailers checking to see if they have the book in stock and their price and delivery times. The prices and delivery times are queried in real-time and thus represent the most up-to-date data from the retailer. Because the prices are gathered directly from the retailers, they are the same prices that are charged to consumers who visit the retailer site directly.[6]

Prices are displayed in offer comparison tables (e.g., Figure 1). These tables list the total price for the book and the elements of price (item price, shipping cost, and applicable sales taxes) along with the retailer's name and the book's delivery information. If a retailer provides multiple shipping options at multiple prices (e.g., express, priority, book rate) the table lists separate offers for each shipping option.[7]

By default, the table is sorted by total price; however, the consumer can sort based on any of the 9 columns in the comparison table. After the consumer has evaluated the information, they can click-through on a particular offer and are taken directly to the retailer in question to finalize their purchase.

We collect four categories of data from EvenBetter.com: offer data, session data, consumer data, and choice data (Table 1). We define an offer as an individual price quote from a retailer — or equivalently an individual entry in an offer comparison table. Our offer data include separate variables for each of the nine columns in the offer comparison table: total price, item price, sales

---

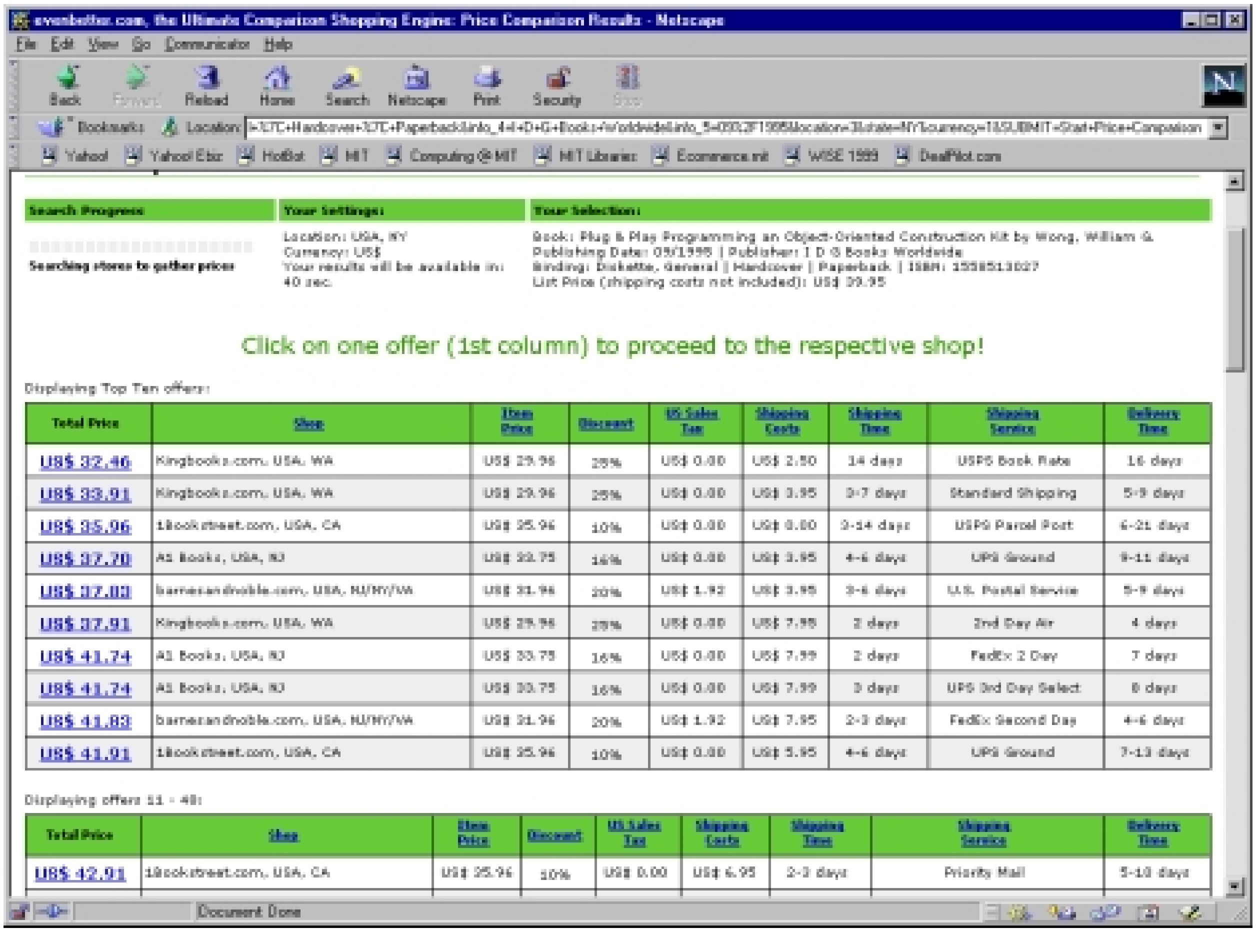
[5] International Standard Book Numbers (ISBNs) uniquely identify the individual version of the book (e.g., binding type, printing, and language). Because EvenBetter's search results are based on a single ISBN, all of the products returned in response to a search are physically identical.

[6] This fact is surprising as one might expect retailers to use shopbots as a price discrimination tool — charging lower prices to consumers who reveal a higher price sensitivity by virtue of using a shopbot.

[7] For example, in the offer comparison table in Figure 1, note that Kingbooks.com has separate listings for their book rate, standard, and 2-day shipping services.

tax (if applicable),[8] retailer name, shipping cost, shipping time, shipping service, and total delivery time.[9] Rank is the numerical position of the offer in the table.

Figure 1: Sample Screen from EvenBetter.com



We also track a variable we call "delivery 'N/A.'" In some instances, retailers are unable to determine how long it will take them to acquire the book from their distributor. When this occurs, EvenBetter lists "N/A" in the delivery time field (but still lists a numerical value in the shipping time field). We capture this situation with a dummy variable that takes on the value of 1

[8] The tax law during our study period stated that retailers had to charge sales tax only to consumers who lived in states where the retailer had a physical location (a.k.a. nexus). Furthermore, several companies have argued that their Internet operations are legally separate from the physical operations of the parent company. Thus, barnesandnoble.com must only charge tax in New York (where its headquarters is located) and New Jersey (where it has a distribution warehouse) even though its parent company, Barnes & Noble, has operations in all 50 states.

[9] Total delivery time is the sum of shipping time and acquisition time (the amount of time it takes for the retailer to obtain the book from their distributor).

whenever "N/A" is listed in the delivery time column. We model this by assuming that the consumer infers the total delivery time as the quoted shipping time plus an unknown constant (captured by the dummy variable).

Table 1: Shopbot Data Collected

Offer Data	
Total Price	Total price for the offer (item price plus sales tax plus shipping cost)
Item Price	The price for the item
Shipping Cost	The price for shipping
State Sales Tax	Sales tax (if applicable)
No Tax	=1 if there is no sales tax on the offer
Retailer	Retailer Name (used to create dummy variables for each retailer)
Shipping Time	Time to ship product from retailer to consumer (Min, Max, Average)
Acquisition Time	Time for retailer to acquire product (Min, Max, Average)
Delivery Time	Shipping time plus acquisition time (Min, Max, Average)
Shipping Method	Priority (1-day or 2-day), Standard (3-7 day), Book Rate (>7 day)
Delivery NA	=1 if retailer can't quote acquisition time on book
Rank	The position of the offer in the comparison table
Session Data	
Date/Time	Date and time search occurred
ISBN	ISBN number of book searched for (used to calculate book type)
Sort Column	Identifies which column the consumer sorted on (default is total price)
Consumer Data	
Cookie Number	Unique identifier for consumers who leave their cookies on
Cookies On	=1 if the consumer has their cookies on
Country	Which country the consumer says they are from
U.S. State	Which state the consumer says they are from (U.S. consumers only)
Choice Data	
Last Click-Through	=1 if the consumer's last click-through was on this offer
Click-Through	=1 if the consumer clicked on this offer
Loyalty Data	
Prior Last Click-Through	=1 if the consumer last clicked through on this retailer on most recent visit
Prior Click-Through	=1 if the consumer clicked through on this retailer on their most recent visit

From our offer data we impute two additional sets of dummy variables relating to the type of shipping associated with the offer and the position of the offer in the comparison table. To construct dummy variables associated with shipping service we use the fact that the shipping services offered by retailers generally fall into three categories: express shipping (typically a 1-2 day shipping time), priority shipping (3-6 day shipping time), and book rate (greater than 7 day shipping time). We generate dummy variables for each category of shipping service. We also generate dummy variables for the first offer in the comparison table and the first screen of offers displayed (i.e., the first 10 offers) in the comparison table.

Our second type of data is session data. We define a session as an individual search occasion for a book, or equivalently data that is common to an individual offer comparison table. Our session data include the date and time the book search occurred, the ISBN the consumer searched for, and whether the consumer chose to sort the offer comparison table based on a column other than total price (the default).

Our consumer data include fields for the consumer's unique cookie number,[10] whether the consumer had turned their cookies off (which occurred for 2.9% of the sessions), and the consumer's state and country location. The state and country data are self-reported and to allow the shopbot to accurately calculate local currency, taxes, and delivery times.

Our choice data are made up of two fields. A "click-through" field captures whether a consumer "examines" an offer from a particular retailer. Since 16% of the consumers in our sample look at multiple retailers, we use a separate field to record the last click-through made by each consumer during a session. We use this as a proxy for the offer selected by the consumer. As noted in section 2.4, the click-through variable does not appear to be biased with regard to sales in a way that would affect our conclusions.

Using our consumer and click-through data we construct two additional variables to help us control for consumer heterogeneity (Guadagni and Little 1983) and to track consumer loyalty over time: Prior Click, and Prior Last Click. Prior Click is a dummy variable taking on the value 1 for retailers the consumer clicked on in the most recent visit but did not "last click." Similarly, Prior Last Click is a dummy variable taking on the value 1 for retailers the consumer "last clicked" on in the most recent visit.

2.3. Descriptive Data

Our data set was gathered over 69 days from August 25 to November 1, 1999.[11] To simplify interpretation, we limit our analysis to prices for U.S.-based consumers (75.4% of sessions),

[10] The cookie number is a unique identifier that is stored on the computer's hard drive by the retailer or shopbot. The retailer can query this number on subsequent visits to the retailer's site and thereby uniquely identify the consumer's computer.

[11] We limited our sample to this time period to avoid potential bias resulting from the Christmas season. Nearing the Christmas holiday, consumers may become more sensitive to brand as a proxy for reliability in delivery time.

sessions that lead to at least one click-through (26.3% of remaining sessions), and sessions that return more than one retailer (99.9% of remaining sessions). The resulting data set contains 1,513,439 book offerings from 39,654 searches conducted by 20,227 distinct consumers. Included in this data set are 7,478 repeat visitors, allowing us to track consumer behavior over time.[12]

These data show a significant dispersion in prices, even for entirely homogeneous physical goods. The average difference in total price between the lowest priced offer and the tenth lowest priced offer is $10.77 in our data. In percentage terms, the tenth lowest priced offer is typically 32.3% more expensive than the lowest priced offer. These results are very similar to Brynjolfsson and Smith (2000, p. 575) who report an average range of 33% between the highest and lowest book prices obtained from 8 different Internet retailers in 1998-1999.

Table 2: Comparison of Retailers at a Shopbot

Retailer	*Internet Market Share (Est.)* *	*Shopbot Last Click Share*	*Proportion of Lowest Prices*	*Click-Sales Conversion Ratio*
Amazon.com	75%	8.6%	2.0%	.484
BN.com	8%	7.4%	3.1%	.461
Borders.com	5%	10.9%	9.8%	.456
A1Books	<1%	10.0%	12.5%	N/A
Kingbooks	<1%	9.8%	15.1%	.486
1Bookstreet.com	<1%	5.9%	8.3%	.509

* Internet market share is compiled from press reports and an analysis of click-through data from prior research (Brynjolfsson and Smith 2000).

Table 2 lists selected descriptive data statistics for our data from the 6 most popular retailers at EvenBetter. Column 1 lists estimates of market share in the broader Internet market and column 2 lists the share of last click-throughs for EvenBetter's consumers. Comparing these two columns yields two insights into the Internet shopbot market. First, shares of last click-throughs are significantly less concentrated than estimates of market share in the broader Internet market for books. Second, click-through shares strongly favor low priced retailers when compared to share estimates in the broader Internet market. For example, Amazon.com, a relatively high priced retailer, has approximately 75% of the total Internet book market yet holds only an 8.6% click-

[12] Limiting our data in this way allows us to focus our attention on a homogeneous consumer segment (U.S.-based consumers who reveal an intention to purchase). However, future research could analyze the differences between U.S. and foreign retailers, or the decision to click-through as a function of product price or product availability.

through share for EvenBetter's consumers. At the same time the share positions for three low priced, and relatively unknown, retailers are dramatically enhanced at EvenBetter.com.

One explanation for this difference is that the lower search costs offered by shopbots make it easier for consumers to locate and evaluate unbranded retailers and this changes their choice behavior from what it would have been if no shopbots were available. To the extent that this explanation holds, it supports the hypothesis that shopbots are a "great equalizer" in Internet markets, putting small retailers on a more equal footing with their larger and more well known competitors. It is also possible that because EvenBetter's consumers are highly price sensitive they are more inclined to shop at low priced retailers than consumers in the broader market.

However, while shopbot consumers appear to be price sensitive, 51% of them choose an offer that is not the lowest price returned in a search. Although the books offered are completely homogeneous, factors other than price influence consumer choice in this setting. Our descriptive data suggest that retailer brand identity is at least one of the factors influencing consumer behavior. This can be seen by comparing columns 2 and 3 in Table 2. These columns show that while branded retailers[13] have the lowest price for only 15% of the book searches they make up 27% of consumer choices. Likewise, the top three unbranded retailers, who have the lowest price 36% of the time, make up only 26% of consumer choices. The advantage held by branded retailers can also be seen by examining the offer price premium, the difference between the lowest priced offer and the price of the offer actually selected. For branded retailers this difference averages $3.99 while for unbranded retailers it averages $2.58, a difference of $1.41.

Our descriptive statistics also give insight into consumer purchase behavior. Because our choice data only track click-throughs, our empirical results only predict factors that drive traffic to a site — not necessarily factors that drive sales. However, the descriptive statistics in column 4 of Table 2 suggest that traffic is a relatively unbiased indicator of actual sales. These ratios are constructed by comparing the number of sales at a particular retailer during September and

[13] We refer to Amazon.com, Barnesandnoble.com, and Borders.com as "branded retailers." Using almost any reference point, these are the most heavily advertised and well-known retailers in the Internet book market. For example, based on a search of AltaVista.com, these 3 retailers make up 97% of the total number of Internet links to

October 1999 to the number of last click-throughs recorded for that retailer during the same time period.[14] These statistics do not vary significantly across branded and unbranded retailers — supporting the interpretation of our results with regard to the behavior that influences sales.

Descriptive statistics provide a useful first step in analyzing consumer choice data. However, definitive conclusions are only possible through systematic empirical models that control for the effect of other aspects of the product bundle. In the next section we discuss two systematic empirical models that can be used to analyze our research questions.

3. Methodology

As noted above, our research goal is to analyze how consumers respond to different aspects of a product bundle including brand name, retailer loyalty, partitioned prices, and contractible and non-contractible product characteristics. There are a variety of choice models available to analyze these questions in a multidimensional choice setting. We discuss the two most prominent models below — the multinomial logit and nested logit models — as an introduction to our analysis. We also provide brief descriptions of multinomial probit as an alternate empirical model and Hierarchical Bayesian Estimation as an alternate estimation technique.

As discussed below, the availability of a nested logit model to control for concerns about the independence of irrelevant alternatives, the applicability of aggregate response in the shopbot market, and the limited availability of longitudinal individual-level choice data leads us to conclude that logit-based models and maximum likelihood estimation techniques are the most appropriate analysis techniques for our research questions.

EvenBetter's retailers. Similarly, based on a search of Lexis-Nexis, these retailers make up 93% of the references in the press to EvenBetter's retailers.

[14] EvenBetter has associate program relationships with many retailers listed at their service. These programs provide EvenBetter with commissions on the sales driven through EvenBetter's site. As a reporting function, the retailers provide summaries of the sales that occurred through EvenBetter's service for a particular month, allowing us to create sales to click ratios statistics. A1Books does not have an associate program relationship based on sales and therefore we are unable to construct sales to click ratios for this retailer.

3.1. Multinomial Logit Model

Given the parallels between our data and scanner data, the multinomial logit model — the workhorse of the scanner data literature (e.g., Guadagni and Little 1983, Kamakura and Russell 1989, Fader and Hardie 1996) — provides a natural empirical starting point for our analysis. We describe the nature of this model briefly below and refer the interested reader to Ben-Akiva and Lerman (1985) or McFadden (1974) for more detailed treatments of the model.

In a choice setting, the multinomial logit model can be motivated by assuming consumers make choices by first constructing a latent index of utility (U_{it}) for each offer (t) in each session (i) based on the offer's characteristics and the consumer's preferences. We model the consumer's utility for each offer as the sum of a systematic component (V_{it}) and a stochastic component (ε_{it}):

$$U_{it} = V_{it} + \varepsilon_{it} \tag{1}$$

The stochastic disturbance can be motivated from a variety of perspectives (Manski 1973); for our purposes the two most natural motivations are (1) unobserved taste variation across consumers and (2) measurement error in evaluating offers.

We further express (V_{it}) as a linear combination of the product's attributes ($\mathbf{x}'_{it}$) and the consumer's preferences for those attributes (β). Equation (1) then becomes

$$U_{it} = \mathbf{x}'_{it}\beta + \varepsilon_{it} \tag{2}$$

To justify this starting point we note that, while modeling consumer choices in terms of latent utility indexes is accepted practice in the marketing and economics literature, its use may be particularly applicable in our setting. By listing offers in a comparison matrix with separate values for a variety of product attributes EvenBetter's comparison matrix lends itself to a rational, attribute-based evaluation by consumers.

The coefficients in (2) could be readily estimated using standard least squares techniques if the researcher could observe U_{it} directly. Unfortunately, this is not generally the case in practice. Instead we typically observe only the resulting choice in session i: $y_i = t$. However, under the assumption of utility maximization, we can infer that $y_i = t$ if and only if $U_{it} = \arg\max(U_{i1}, U_{i2,} ... U_{iT_i})$. Thus, we can write the probability that offer t is chosen in session i as:

$$P_t(\mathbf{x}_{it}, \beta) = \Pr\{U_{it} = \arg\max(U_{i1}, U_{i2,} ... U_{iT_i})\} \tag{3}$$

Using (2) this can be rewritten as:

$$\begin{aligned} &P_t(\mathbf{x}_{it}, \beta) = \\ &\Pr\{\varepsilon_{it} - \varepsilon_{i1} \geq -(\mathbf{x}_{it} - \mathbf{x}_{i1})'\beta, \varepsilon_{it} - \varepsilon_{i2} \geq -(\mathbf{x}_{it} - \mathbf{x}_{i2})'\beta, ... \varepsilon_{it} - \varepsilon_{iT_i} \geq -(\mathbf{x}_{it} - \mathbf{x}_{iT_i})'\beta\} \end{aligned} \tag{4}$$

The multinomial logit model assumes that the disturbance terms are independent random variables with a type I extreme value distribution

$$\Pr\{\varepsilon_j \leq \tau\} = e^{-e^{-\mu\tau}} \tag{5}$$

where μ is an arbitrary scale parameter. This distribution is motivated, in part, because it is an approximation to a normal distribution. However, the assumption has the even more desirable property that it dramatically simplifies (4) to the following form (McFadden 1974):

$$P_t(\mathbf{x}_i, \beta) = \frac{e^{\mu\beta'\mathbf{x}_{it}}}{\sum_{\tau=1}^{\tau_i} e^{\mu\beta'\mathbf{x}_{i\tau}}} \tag{6}$$

This formula has all the desirable properties of a purchase probability: it is always positive, it sums to 1 over all the τ_i offers in session i, and it is invariant to scaling.

Our goal is to determine the β vector — the weights on the consumers' evaluation of offers. Unfortunately, we estimate μβ. Since μ is present in each of the β terms it is not identifiable. However, since its purpose it to place a scale on the utility of the model, we can arbitrarily set it to any real number (Ben-Akiva and Lerman 1985, p. 107) to identify the β coefficients. While this is a benign assumption in the multinomial logit model, it has implications for our ability to compare coefficients in the nested logit model, which we now discuss.

3.2. Nested Logit Model

The parsimony of the multinomial logit formula comes at a cost. The assumption that errors are independent across offers gives rise to the Independence of Irrelevant Alternatives (IIA) characteristic in the multinomial logit model. Simply put the IIA problem is that the probability ratio of choosing between two offers depends only on the attributes of those two offers and not on the attributes of any other offers in the choice set. Using equation (6) this can be expressed as:

$$\frac{P_t(\mathbf{x}_i,\beta)}{P_s(\mathbf{x}_i,\beta)} = \frac{e^{\mathbf{x}'_{ti}\beta}}{e^{\mathbf{x}'_{si}\beta}} \qquad (7)$$

This restriction is violated if the error independence assumption does not hold. The error independence assumption might be violated if subsets of alternatives in the consumer's choice set are similar to one another. This problem may impact our data if consumers perceive different branded (or unbranded) retailers as offering similar service levels. For example, a consumer who placed a high value on offers from Amazon.com may also place a high value on offers from BarnesandNoble.com or Borders.com. In this case, the cross-elasticity between offers is not equal but rather is much higher among branded retailers than it is between branded and unbranded retailers (and potentially vice-versa).

The solution to this problem is to place similar offers in common groups — or nests — such that the IIA assumption is maintained within nests while the variance is allowed to differ between nests. Thus, the consumer can be modeled as facing an initial choice S (e.g., *S={branded retailers, unbranded retailers}*) followed by a restricted choice R (e.g., *R={{amazon, barnesandnoble, borders}, {a1books, kingbooks, 1bookstreet,…}}*).[15]

Given this decision model we represent the choice set for consumer n as the Cartesian product of the sets S and R minus the set of all alternatives that are infeasible for individual n, or $C_n = S \times R - C_n^*$. We further define the marginal brand choice set, S_n, to be the set of all brand options corresponding to at least one element of C_n and the conditional retailer choice set, R_{ns}, as

[15] A two-level nested model is chosen here for expositional simplicity and its applicability to our setting. Nested models containing 3 or more nests are simple extensions of the two-level nested logit model (see Goldberg 1995 for an empirical example of a five-level model).

the subset of all retailers available to consumer *n* conditional on the consumer making brand choice *s*.

We then model the utility associated with a choice of shipping type and retailer as

$$U_{sr} = V_s + V_r + V_{sr} + e_s + e_r + e_{sr} \tag{8}$$

where V_s and V_r are the systematic utilities associated with the choice of brand and retailer respectively and V_{sr} is the systematic utility associated with the joint choice of brand and retailer. The error terms are defined similarly as the random components of utility associated with the choice of brand, retailer, and the joint choice of brand and retailer.

We additionally assume that

1. $var(e_r)=0$, which is equivalent to assuming independence of choice alternatives in the bottom level nest (Guadagni and Little 1998);

2. e_s and e_{sr} are independent for brand and retailer selections in the consumer's choice set;

3. the e_{sr} terms are independent and identically Gumbel distributed with a scale parameter μ_r, and

4. the e_s terms are distributed such that $\max_{r \in R_{ns}} U_{rs}$ is Gumbel distributed with a scale parameter of μ_s.

Given these assumptions, the choice of retailer conditional on the choice of brand at the lower level nest becomes

$$p(r \mid s) = \frac{e^{\mu_r (V_{sr} + V_r)}}{\sum_{j \in R_{ns}} e^{\mu_r (V_{sj} + V_j)}} \tag{9}$$

which is simply the standard logit model.

Similarly, the choice of branded or non-branded retailers becomes

$$p(s) = \frac{e^{\mu_s (V_s + V_s')}}{\sum_{i \in S_n} e^{\mu_s (V_i + V_i')}} \tag{10}$$

where

$$V_s' = \frac{1}{\mu_r} \ln \sum_{r \in R_{ns}} e^{(V_r + V_{sr})} \tag{11}$$

As in the multinomial logit model, the coefficients we estimate are convoluted with the scale parameter (μ_r). Because the μ_r is constant within nests, it is possible to analyze the β parameters within nests. However, the scale parameter will not be constant across nests in general, making it impossible to directly compare coefficients across nests (Swait and Louviere 1993). However, it is possible to compare shared coefficients by normalizing to a common reference point. We discuss this in more detail in the analysis section.

3.3. Alternate Models and Estimation Techniques

The multinomial probit model (Hausman and Wise 1978) is the most recognized alternative to the logit-based models of choice described above. This model assumes that the discrete choice errors are normally distributed. The advantage of this assumption is two-fold. First it allows for more realistic correlation structures for the error components, eliminating the IIA problem. Second, and similarly, it allows for flexible modeling of taste variation across consumers (or other subsets of choice actors).

However, the normality assumption comes as a high cost. It is computationally intensive to evaluate the higher-order multivariate normal integrals used in the multinomial probit model. Several advances have been made in the evaluation these integrals. Hausman and Wise (1978) use a transformation of variables to reduce the dimensionality of the variance-covariance matrix by one. McFadden (1989) employs a method of simulated moments using Monte Carlo simulation to eliminate the need for direct estimation of the likelihood function. However, in spite of these advances, standard multinomial probit estimation using these techniques remains computationally infeasible for large samples or models with more than a handful of choice alternatives making it impractical in our setting. In its place, our use of the nested logit model should control for IIA concerns across branded and unbranded retailers.

Hierarchical Bayesian Estimation (McCulloch and Rossi 1994) provides an individual-level estimation alternative for both logit- and probit-based models. Hierarchical Bayesian Estimation uses Bayesian techniques to estimate individual-level responses for each consumer in a sample (along with aggregate level responses). Moreover, the model makes probit estimation feasible by using the Gibbs sampler to generate an exact posterior distribution of the multinomial probit model. This avoids the computational problems associated with estimation of the multinomial probit likelihood function while still allowing for a correlated error structure.

However, hierarchical Bayesian techniques are typically used to analyze individual level consumer response (e.g., Rossi, McCulloch, Allenby 1996; Montgomery 1997). Given the separation between shopbots and retailers, individualized pricing strategies are not currently used in shopbot markets making Hierarchical Bayesian techniques less appropriate for our analysis. Additionally, most of the customers in our data set make only a single purchase or have relatively short purchase histories, making individual level estimation less reliable. However, with longer purchase histories Hierarchical Bayesian Estimation may make a potentially useful area for future analysis, especially if shopbots develop individualized pricing regimes in the future.

4. Empirical Results

Our analysis addresses two empirical questions: consumer response to the presence of brand, and consumer loyalty to retailers they have visited previously. We also use the predictive characteristics of our models to assess the reliability of our results and to explore the potential for retailer-based personalized pricing strategies. We address each of these questions in turn below using multinomial logit and nested logit models.

4.1. Consumer Response to Brand

Retailer brand might matter to consumers of homogeneous physical goods if branded retailers provide objectively better service quality or if consumers are asymmetrically informed regarding individual retailer's service quality and are using brand as a proxy for quality. To analyze consumer response to brand, we capture brand name in two ways: first with a dummy variable that takes on a value of 1 for branded retailers, and second with separate dummy variables for each of these three retailers (Amazon.com, BarnesandNoble.com, Borders.com). Results for these models are presented in Table 3 along with other variables that may impact consumer choice: total price, average delivery time, and delivery "N/A." [16]

As noted above, the coefficients listed in Table 3 should be interpreted as preference weights in a latent utility function. Thus, the negative coefficient on price indicates that higher prices, *ceteris paribus*, lead to lower latent utilities and, as a result, to fewer consumer click-throughs. Likewise, longer delivery times and not being able to quote a specific delivery time (Delivery "N/A") lead to lower latent utility in the consumer's evaluation.

Table 3: Basic Models of Brand Choice

	1	*2*
Total Price	-.252 (.001)	-.253 (.001)
Average Delivery Time	-.011 (.001)	-.011 (.001)
Delivery "N/A"	-.417 (.015)	-.420 (.015)
Branded Retailers	.284 (.014)	
Amazon		.467 (.020)
BarnesandNoble		.179 (.023)
Borders		.186 (.020)
Log Likelihood	-100,706	-100,630
Adjusted U^2	.2693	.2698

* Standard Errors listed in parenthesis. All results are significant at $p<.05$. Adjusted U^2 = 1-(LL(*)-# of variables)/ LL(0) (Ben-Akiva Lerman 1985, p. 167). N=39,654 sessions.

At the same time, consistent with the descriptive data presented in section 2, we find that even after controlling for price and delivery time brand still has a significant positive effect on latent utility. Each of the coefficients on brand in specifications 1 and 2 are positive and highly significant suggesting that consumers are willing to pay more for offers coming from branded retailers.

Following Guadagni and Little (1983), we can use the absolute value of the ratio of the coefficient to the standard error (the t-statistic) to interpret the relative importance of each

[16] The range of quoted delivery times should also impact consumer choice (e.g., 3-7 days versus 1-9 days). However, measures of delivery range are collinear with delivery time. Because of this, we only analyze average times in this and subsequent results. Using minimum or maximum delivery times (as opposed to average time) does not substantively alter our results.

variable in the consumer's evaluation of an offer. This comparison is motivated by observing that larger coefficients indicate factors that are more important in the consumer's evaluation of the offer and more accurately estimated coefficients indicate factors where there is a high degree of uniformity in response to the variable. Using this comparison we note that the total price variable has a t-statistic of 176, which is nearly 10 times larger than the next closest t-statistic. This indicates that an offer's total price is by far the most important factor consumer's use to evaluate offers — supporting the inference that consumers are highly price sensitive in the shopbot setting.

We can use the relative sizes of the coefficients to gain an idea of the importance of brand name in dollar terms. This comparison exploits the fact that coefficients in the multinomial logit are product attribute weights in the consumer's latent utility function. Thus, we can construct counter-factual comparisons of varying offer characteristics to evaluate the importance of characteristics in dollar terms. For example, we can ask: Given two offers that are exactly the same with respect to all product attributes, if we added brand to one offer, how much would we need to decrease the price of the other offer to keep the latent utility constant? The answer, derived from equation (2) above is:

$$\Delta p = \frac{-\beta_{BRAND}}{\beta_{PRICE}} \tag{12}$$

Using this equation we can use the results from Table 3 column 1 to calculate that offers coming from one of the three branded retailers have a \$1.13 price advantage over unbranded offers. From column 2, we further infer that offers from Amazon.com have a \$1.85 advantage over unbranded retailers, *ceteris paribus*, and offers from Barnes and Noble and Borders have an advantage of approximately \$0.72 over unbranded retailers. Considering that the average total price of the books chosen by customers in our sample is \$36.80, these figures translate into 3.1% margin advantage for branded retailers (and a 5.0% margin advantage for Amazon.com) in head-to-head comparisons with unbranded retailers.

There are several possible explanations for the price advantage among branded retailers in Internet markets for homogeneous physical goods. First, branded retailers may provide objectively better service quality with regard to product delivery, web site ease-of-use, privacy policies, product return policies, or other service attributes. Retailer differentiation in these service characteristics is consistent with their strategic goal to mitigate direct price competition (de Figueiredo 2000).

Delivery service is likely to be one of the most important aspects of a retailer's service quality. While our empirical methodology will control for the quoted delivery time by each retailer, it is possible that branded retailers are more reliable in meeting their quoted delivery times. To investigate this possibility, we ordered 5 books, using various shipping services, from the 6 most popular retailers listed at EvenBetter.com and compared their actual and promised delivery times. Our results are displayed in Table 4 below. The first column displays the number of books (out of 5) that were delivered before the first day in the retailer's quoted delivery range. The second column displays the number of books that were delivered within the quoted delivery time (out of 5) including those that were delivered early. The third column displays the BizRate.com delivery rating (out of 5) for each retailer.[17] While each of the these ratings is an imperfect measure of the actual service quality delivered by these retailers, they do not indicate a dramatic differences in service quality between branded and unbranded retailers, suggesting that heterogeneity in this aspect of service quality may not explain the majority of brand response observed in our data.

Table 4: Retailer Delivery Accuracy

Retailer	*Early Delivery*	*On-Time Delivery (including early)*	*BizRate.com Delivery Rating*
Amazon	3	5	4.5
BarnesandNoble	1	5	4
Borders	5	5	4
A1Books	5	5	3.5
Kingbooks	1	5	4.5
1Bookstreet	1	5	4

A second possible explanation for the importance of brand concerns the information available to consumers in electronic markets. It is possible that service quality should be modeled as an experience good where consumers are asymmetrically informed, *ex ante*, regarding the quality they will receive for a particular order.

[17] Note that in the BizRate.com ratings, A1Books is rated by self-reported experiences from Internet shoppers whereas the ratings for the other 5 retailers are based on the experiences of BizRate.com staff members.

Erdem and Swait (1998) use an information economics framework to demonstrate that in markets with asymmetric information about quality, consumers use brand names as a signal of product quality. These signals reduce consumers' information acquisition costs, lower the risk they must incur when making purchases, and ultimately increase their expected. Brand signals can be communicated to consumers through advertising (Milgrom and Roberts 1986) and through prior personal evaluation (Erdem and Keane 1996).

Extending the information economics model of brand value to the Internet, Erdem et al (forthcoming), argue that the Internet may have a differential effect on brand value depending on the nature of the product: "We expect that for search goods the Internet reduces the importance of brand in its role of reducing perceived risk. For experience goods…we expect that the Internet will not reduce (and may well increase) the importance of a brand in its role of reducing perceived risk" (p. 269).

However, as noted above, the importance of service quality for physical products ordered over the Internet may cause these products to behave more like experience goods than search goods. This aspect of Internet markets may differ conceptually from physical world markets to the extent that the spatial and temporal separation between consumers, retailers, and products in Internet markets increases the importance of service quality and reduces consumers' ability to evaluate quality prior to making a purchase (Smith, Brynjolfsson and Bailey 2000). Under this explanation, retailer branding may remain an important source of competitive advantage for Internet retailers — even in markets served by shopbots.

It is also possible that our brand name results derive from unobserved loyalty. Because we do not observe consumer behavior for visits directly to the retailer or for visits to the shopbot outside of our sample window, consumers have prior unobserved relationships (and therefore loyalty) that disproportionately reside with branded retailers. In this case the loyalty effects discussed in section 4.2 will also apply to our brand coefficients.

4.2. Retailer Loyalty

Our data can also be analyzed to determine the effect of retailer loyalty. Consumers may be loyal to retailers for a variety of reasons. As noted above, in a setting with asymmetric information regarding retailer service quality, consumers may use prior experience with a retailer as a signal of service quality in subsequent purchase occasions. Consumers may also factor in the cost of time to learn how to use a new retailer site or to enter in the information necessary to establish an account with a new retailer. Johnson, Bellman, and Lohse (2000) refer to this effect as cognitive lock-in and find that it is a significant source of web site "stickiness."

We use the two variables Prior Click and Prior Last Click to analyze the effect of retailer loyalty in our setting. To simplify interpretation of the coefficients, we limit our analysis to repeat visitors. Our results adding these two variables to the previous models are shown in Columns 1 and 2 of Table 5. Here we find that consumers are much more likely to choose a retailer they have selected on a prior search (Prior Last Click). In dollar terms, retailers that a consumer had selected previously hold a $2.49 advantage over other retailers. We also find that consumers who had evaluated, but not selected, a brand (Prior Click) are statistically no more likely to select that brand on a subsequent visit. This suggests that, what they learned about the brand by visiting the retailer's site has, if anything, a negative effect on subsequent offer evaluations (consistent with their observed behavior on the initial visit).

Table 5: Basic Models of Brand Choice with Loyalty for Repeat Visitors

	1	*2*
Total Price	-.232 (.002)	-.233 (.002)
Average Delivery Time	-.011 (.001)	-.010 (.001)
Delivery "N/A"	-.368 (.018)	-.373 (.018)
Branded Retailers	.296 (.017)	
Amazon		.499 (.024)
BarnesandNoble		.252 (.028)
Borders		.130 (.025)
Prior Last Click	.577 (.028)	.579 (.028)
Prior Click	*-.096 (.064)*	*-.082 (.064)*
Log Likelihood	-67,356	-67,287
Adjusted U^2	.2612	.2620

* Standard Errors listed in parenthesis. Italicized results are insignificant at p<.05. N=26,390 sessions.

These findings are consistent with the importance of cognitive lock-in, web site convenience, and asymmetric information as sources of competitive advantage in electronic markets. They also help to quantify the importance of first mover advantage among Internet retailers. Moreover, these results are obtained from consumers who are likely to be among the least loyal consumers in Internet markets. According to shopbot managers, many customers use shopbots to locate retailers they are happy with and, after a period of good service, begin to visit the retailers, directly, bypassing the shopbot (and regrettably our data set. Thus, our loyalty results constitute a lower bound on loyalty among typical Internet customers.

The importance of loyalty in this setting also suggests that shopbots may provide an effective and low cost avenue for retailers to acquire new consumers and gain competitive advantage against their rivals. This factor may be particularly important for lesser-known retailers as reflected in the market and click-through share statistics presented in Table 2.

The selected variables include the differential response of consumers who sort on shipping columns to the product's item price, shipping price, average delivery time, and a dummy variable

4.3. Model Predictions

An additional aspect of understanding shopbot markets relates to how well the predictions of our models fit actual consumer behavior both within and outside the time sample. Accurate predictions of consumer behavior both confirm the validity of our findings and have implications for retailers considering differential pricing strategies for shopbot markets.

To avoid overfitting, it is important to analyze model predictions using a different data sample than the one used to estimate the model. To account for this, we divide our data into calibration and holdout samples. Our calibration sample is made up of 15,503 sessions conducted by consumers with odd numbered cookies between August 25, 1999 and October 18, 1999. We have two types of holdout samples. An intra-temporal holdout sample is made up of 15,503 sessions conducted by consumers with even numbered cookies between August 25, 1999 and October 18, 1999. The inter-temporal holdout sample is made up of 8,648 sessions conducted during the last two weeks of the data set: October 19, 1999 through November 1, 1999.

Table 6 presents the results from applying our calibration sample to an extended model specification. Column 1 presents a minimal model specification using only attribute specific dummy variables (Fader and Hardie 1996) to model different offers (alternatives). Our attribute specific dummy variables include the position of the offer in the comparison table and the retailer brand name for all retailers with greater than 3% last click-through share (12 retailers).

Column 2 adds coefficients for total price, average delivery time and delivery "N/A." Column 3 adds coefficients for prior last click and prior click behavior. Column 4 replaces the coefficient on total price with total price as a percentage of the lowest price available in the search. Allowing price to enter as a percentage of the lowest price in a search controls for prospect theoretic effects (Kahneman and Tversky 1979) — in this case the possibility that consumers may respond differently to a $1 price increase on a $5 book than on a $50 book.

Results from these more complete models are ostensibly the same as the results from the basic models in section 3.3.1. Consumers respond strongly to branded retailers and exhibit loyalty to retailers they have visited before. In evaluating the reliability of these models we note the standard errors are generally stable across specifications suggesting that collinearity is not a significant problem in our model specifications. This inference is confirmed in other standard tests of data collinearity. In the next section we discuss how to choose among these different specifications to determine the model that best combines explanatory power and parsimony.

4.3.1. Model Selection and Model Fit

Table 6 presents four different model specifications containing different independent variables. Various alternatives have been offered to choose among model specifications to best combine fit and parsimony. The most common model selection criteria fall into two categories. The first, Log likelihood-based criteria such as U^2 measures of fit (McFadden 1974; Ben-Akiva and Lerman 1985, p. 167) select the model that minimizes the log-likelihood value in maximum likelihood estimation, either ignoring issues of parsimony or accounting for parsimony by subtracting the number of parameters in the model. The second category, information theoretic criteria, selects models based on the amount of information in the data that is explained by the model. By using information theory, these models better account for both the fit and parsimony of the different candidate models. Notable information theoretic measures include the Akaike Information Criterion or AIC (Akaike 1973), Bayesian Information Criterion or BIC (Schwartz 1987, Raferty 1997), and information theoretic measure of complexity or ICOMP (Bozdogan 1990; Bearse, Bozdogan, Schlottmann 1997, a more recent test, which uses the Fisher information matrix. These criteria are discussed in more detail in Appendix C.

Table 6: Extensive Model of Consumer Behavior

Variables	*1*	*2*	*3*	*4*
Price				
Total Price		-.062 (.002)	-.061 (.002)	
Total Price/Min				-2.254 (.059)
Position in Table				
First Price Listed	2.507 (.019)	2.256 (.022)	2.257 (.022)	2.054 (.024)
In First 10 Prices	2.923 (.032)	2.358 (.035)	2.359 (.036)	2.117 (.036)
Delivery Time				
Delivery Avg.		-.029 (.001)	-.029 (.001)	-.028 (.001)
Delivery "N/A"		-.344 (.035)	-.362 (.036)	-.474 (.037)
Retailer Brand				
Amazon.com	1.079 (.039)	1.018 (.045)	.988 (.045)	.980 (.046)
BarnesandNoble	.787 (.042)	.591 (.049)	.560 (.050)	.565 (.050)
Borders	.212 (.039)	.194 (.047)	.166 (.047)	.186 (.048)
A1Books	.126 (.039)	.115 (.047)	*.090 (.047)*	.164 (.047)
Kingbooks	-.491 (.039)	-.335 (.044)	-.354 (.044)	-.339 (.045)
1Bookstreet	-.143 (.046)	*-.081 (.050)*	-.117 (.050)	-.370 (.053)
Alphacraze	*-.036 (.048)*	*.012 (.051)*	*.018 (.051)*	.129 (.052)
Alphabetstreet	-.847 (.049)	-1.087 (.057)	-1.095 (.058)	-.666 (.056)
Shopping.com	-.203 (.051)	-.356 (.055)	-.367 (.055)	-.301 (.056)
Fat Brain	*-.021 (.052)*	-.261 (.061)	-.274 (.061)	-.296 (.062)
Classbook.com	.587 (.056)	.368 (.069)	.344 (.070)	.348 (.067)
Books.com	-.739 (.056)	-.550 (.059)	-.548 (.059)	-.490 (.060)
Other Retailers	0	0	0	0
Prior Choices				
Prior Last Click			.729 (.049)	.644 (.051)
Prior Click			*-.112 (.113)*	*-.154 (.114)*
Log Likelihood	-31,255	-30,270	-30,158	-29,749
Adjusted U^2	.420	.439	.441	.448
AIC	4.034	3.907	3.893	3.840
BIC	-86,941	-88,882	-89,086	-89,903
ICOMP	62,513	60,558	60,337	59,515

* Standard Errors are listed in parenthesis. Italicized results are insignificant at p<.05. (N=15,503 sessions)

For each model in Table 6, we present the resulting log-likelihood values; Ben-Akiva and Lerman's adjusted U^2; and the AIC, BIC, and ICOMP information based measures of model selection. In spite of the very different nature of these selection criteria, they are unanimous in choosing specification 4 as the "best" specification.

Once a model has been selected as providing the best combination of explanatory power and parsimony, we can evaluate how well the predictions made by that model match observed behavior. To conduct this evaluation, we first calculate the hit rate — the proportion of times the prediction made by the model is the same as a choice made by the consumer (for the holdout sample) as

$$HitRate = \frac{\hat{y}'y}{N} \tag{19}$$

where $\hat{y}'$ is a vector which takes on the value of 1 for the offer that has the single highest predicted choice probability in each session and 0 otherwise, and *y* is a vector that takes on the value of 1 or 0 for the actual choices made by consumers.

Using this definition, we find a hit rate of .4873 intra-temporally and .4694 inter-temporally for specification 4 above. These hit rates compare very favorably to hit rates reported in the scanner data literature. While there is a slight drop in the hit rate for the inter-temporal holdout sample during the 2-week period following out estimation the hit rate during this 2-week period is still quite high.

Furthermore, this drop in hit rate can be explained by analysis of week-by-week predicted and actual choice share for EvenBetter.com's consumers. To analyze choice share in this way we use the holdout sample to calculate predicted share for each brand *j* in each week *k* as:

$$s_{jk} = \frac{1}{n_k} \sum_{i=1}^{n_k} p_i \tag{20}$$

(Guadagni and Little 1983, p. 224) where p_i is the predicted probability that the brand is chosen in each session and in each week and n_k is the number of sessions in each week. We also use the fact that the predicted offer selection is a binomially distributed random variable to calculate a standard error for the predicted share as

$$SE(s_{jk}) = \frac{1}{n_k} \left[\sum_{i=1}^{n_k} p_i (1 - p_i) \right]^{1/2} \tag{21}$$

We then graph the predicted and actual choice behavior along with a 90% confidence interval band ($\pm 1.64 \times SE(s_{jk})$) for each of the brands with more than 3% share. The graphs are presented in Appendix A. The vertical line in the graphs between weeks 8 and 9 represents the difference between the intra- and inter-temporal holdout samples.

As with the hit rate calculations above, these graphs show a strong consistency between predicted and actual share across retailers. Within the time period covered by the calibration sample, our predicted share is within a 10% error bound of the actual share 98% of the time. During the subsequent two weeks, the predicted share accuracy declines to 79% accuracy.

There are two aspects of the graphs that deserve further explanation. First, there is a strong decline in the actual (and predicted) share of BarnesandNoble.com during weeks 5 and 6. This drop in share is due to the fact that EvenBetter.com did not query BarnesandNoble during a significant portion of these two weeks because of concerns about the accuracy of BarnesandNoble.com's tracking of sales through their site. After talking with BarnesandNoble managers, EvenBetter realized that the discrepancy was due to an upgrade at BarnesandNoble's site and that all the data had been recorded correctly and they reinstated the retailer.

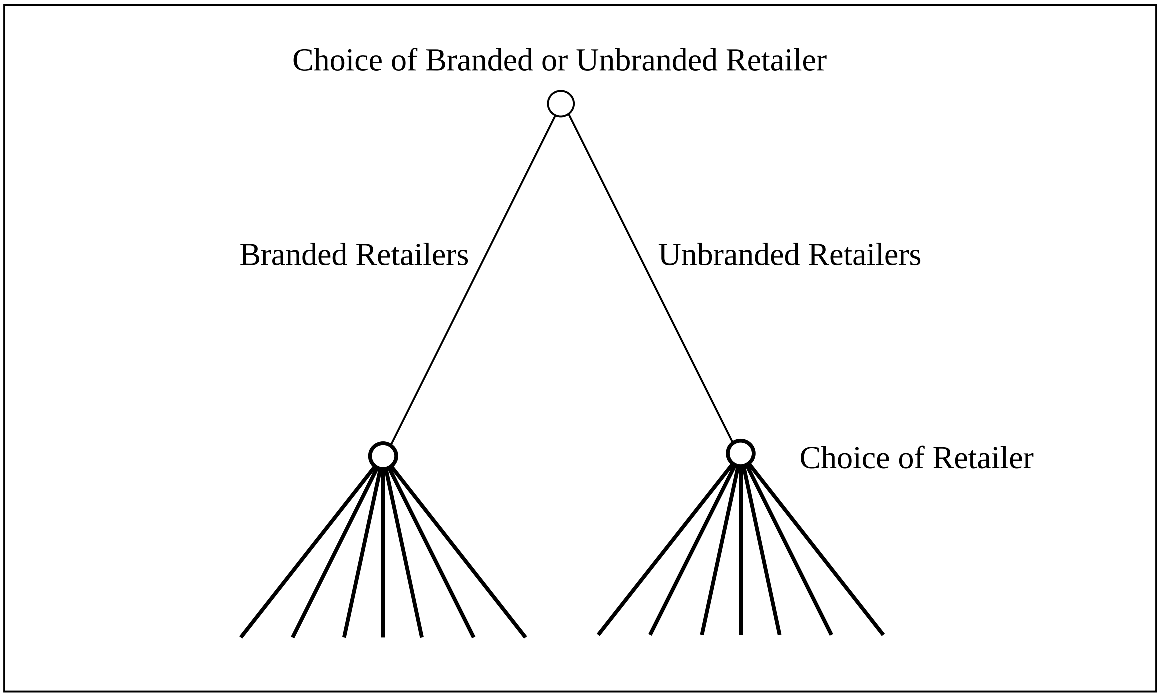
Second, there is a dramatic increase in Borders' actual share during week 10. Further analysis shows that on the last three days of the month of October, Borders' averages 21% of last click-throughs (see Figure A.13). During the first 65 days, Borders' share had averaged 10% (with a daily high of 13% and a low of 6%). This is displayed in Figure A.13, which shows the consistency of Borders' share until the end of the month and the return to a "normal" share value on November 1, the last date in our data sample. (Investigation of the data from November 2 to November 13 shows that Borders' share remained between 6-8%.)

These statistics, combined with the fact that there is no significant difference in Borders participation in sessions, pricing strategies, or shipping policies during this week, suggests that the source of the share jump is possibly a special temporary promotion on the part of Borders.com that we do not observe in our data. Unfortunately, efforts to verify this have been unsuccessful. Searches of press articles in Lexis-Nexis and USENET newsgroup messages during this time period have not revealed any mention of a special Borders promotion.

However, this change does highlight an interesting fact about this shopbot market. The increase in Borders' share appears to come at the expense of only Amazon.com and

BarnesandNoble.com's shares.[18] This suggests that there is a high cross-elasticity among the three branded retailers indicating that the IIA assumption, mentioned above, may be too restrictive for our market environment. In the next section, we attempt to address this concern by modeling the branded and unbranded retailers in separate nests of the nested logit model.

Figure 2: Nested Logit Decision Model



4.3.2. Nested Logit Models

As noted in section 3, the nested logit model offers an alternative modeling technique to control for correlation between the errors of different offers. Our results in section 4.2 suggest that there exist different error correlation structures for branded and unbranded retailer groups. Thus, a consumer who places a high value of offers from Amazon.com may also place a high value on offers from BarnesandNoble.com and Borders. To explore this possibility, we construct a nested logit model by supposing that consumers first choose whether to purchase from a branded or unbranded retailer and then choose which offers to select from the subset of offers in their choice set (Figure 2).

[18] The drop in BarnesandNoble.com share during weeks 5 and 6 did not result in a similar change in Amazon and Borders' shares because in the Borders case (we are arguing) that customers had different preferences for borders

At the top level, we model the choice between branded and unbranded retailers as arising from four variables. First, the difference between the lowest priced branded offer and the lowest priced unbranded offer when branded retailers have the lowest price and the analogous value when unbranded retailers have the lowest price. Second, whether the consumer last clicked (or clicked without last clicking) on a branded or unbranded retailer on their most recent visit. Third, a dummy variable for the lowest priced category (branded or unbranded). And fourth, a dummy variable for branded retailers. The variables in the bottom level nests are the same as those in column 4 of Table 6, except that we add a dummy variable for the offer with the best price in each nest ("Best Price In Nest").

We estimate our nested logit model sequentially as described in Ben-Akiva and Lerman (1985, pp. 297-298) and Guadagni and Little (1998). Sequential estimation produces consistent but asymptotically inefficient estimates, causing the standard errors to be too small (Amemiya 1978). However, it has been shown that in many applications the resulting standard errors are not significantly different from those resulting from Full-Information Maximum Likelihood estimation (Bucklin and Gupta 1992, p. 205). Given the strong significance of nearly all our coefficient estimates it is highly unlikely that Full Information Maximum Likelihood estimation would change our results.

Table 7: Nested Logit Model: Top Nests

Variable	*Coefficient*
Price Difference if Brand Lowest Price	.033 (.009)
Price Difference if Unbranded Lowest Price	.060 (.004)
Prior Last Click Brand	.358 (.056)
Prior Click Brand	-.323 (.097)
Lowest Priced Category	1.012 (.037)
Branded Retailer	.358 (.056)
Unbranded Retailer	0

* Standard Errors are listed in parenthesis. Italicized results are insignificant at p<0.10. n=39,654 sessions

Our results using the nested logit model are presented in Tables 7 and 8for the top and bottom level nests respectively. These results are consistent with the results presented above for the multinomial logit model: consumers are very sensitive to price (as evidenced by the coefficients

offers that appeared in the comparison tables. In contrast, during weeks 5 and 6 the BarnesandNoble offers did not appear in the tables, and thus our estimates of customer preferences remained accurate for the remaining choices.

on “lowest priced category,” price, and position in table), but still respond strongly to the presence of brand and retailer loyalty.[19]

Table 8: Nested Logit Model: Bottom Nests

	Branded Retailers	*Unbranded Retailers*
Price		
Total Price/Min Total Price	-5.735 (.246)	-1.841 (.066)
Position in Table		
First Price Listed	1.013 (.080)	1.296 (.096)
In First 10 Prices	1.054 (.076)	2.366 (.049)
Best Price In Nest	.634 (.068)	.897 (.095)
Delivery Time		
Delivery Average.	-.024 (.003)	-.028 (.002)
Delivery “N/A”	-.576 (.121)	-.534 (.043)
Retailer Brand		
Amazon.com	1.267 (.067)	
BarnesandNoble	.753 (.069)	
Borders	0	
A1Books		.130 (.054)
Kingbooks		-.381 (.050)
1Bookstreet		-.420 (.059)
AlphaCraze		.173 (.056)
AlphabetStreet		-.645 (.060)
Shopping.com		-.341 (.062)
Fat Brain		-.293 (.067)
Classbook.com		.267 (.075)
Books.com		-.548 (.064)
Other Retailers		0
Prior Choices		
Prior Last Click	.338 (.119)	.712 (.070)
Prior Click	*-.199 (.237)*	-.424 (.160)

* Standard Errors are listed in parenthesis. Italicized results are insignificant at p<.05. (Branded Retailer n=4,023, Unbranded Retailers n=11,480)

In addition the fit and predictive power of these models are quite good. Our hit rates for the nested logit results are slightly higher intra-temporally (.4880) and significantly higher inter-temporally (.4855) than those for the multinomial logit models reported above. The increase in inter-temporal hit rate reflects the fact that placing the branded retailers in a separate nest improves the predictions for branded retailers during week 10 when Borders’ share increases. The model still does not predict the increase in Borders share. However, because the nested logit models elasticity within nests, the actual shares for Amazon and BarnesandNoble fall within a

[19] Because the specifications in Table 12 control for different retailers (by construction) it is infeasible to use the same techniques presented in section 4.4 to compare coefficients between nests.

10% error bound of the predicted shares during week 10. Predicted and actual share for branded retailers under the nested logit model are shown in Appendix B. Because the share predictions for the unbranded retailers are similar to those shown in Appendix A, we suppress the graphs for these retailers. The similarity in the multinomial and nested logit results with regard to coefficients and predictions also provides confirmation that the IIA problem does not significantly impact our previous results.

One implication of the quality of our inter- and intra-temporal share predictions is that retailers may be able to use information gathered from Internet shopbots to create personalized prices for shopbot consumers. Shopbots could arrange to pass information regarding the consumer's prior search behavior and product characteristics for competing offers to retailers, allowing them to calculate a personalized price for this consumer to maximize their profits.

Using this information, the retailers could use the multinomial logit equation (equation 6) to calculate the probability that their offer would be chosen as a function of their price (P^*), their product characteristics (ϕ), the prices and product characteristics of competing offers (ϕ_{-1}, P^*_{-1}), and the consumer's characteristics (θ):

$$\mathrm{P}(P^*, \phi, P^*_{-1}, \phi_{-1}, \theta) \tag{22}$$

With this knowledge, the retailer could then choose a price to maximize their profit for this transaction:

$$\max_{P^*}[(P^* - c)\mathrm{P}(P^*, \phi, P^*_{-1}, \phi_{-1}, \theta)] \tag{23}$$

With an estimate of the annual frequency of the consumer's visits to the shopbot ($F(\theta)$) and the marginal loyalty advantage from being chosen on this purchase ($\Lambda(\theta)$), and a discount rate for future revenue (*i*), the retailer could instead maximize the net present value of being chosen in the current transaction:

$$\max_{P^*}\left[(P^* - c)\mathrm{P}(P^*, \phi, P^*_{-1}, \phi_{-1}, \theta) + \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathrm{P}(P^*, \phi, P^*_{-1}, \phi_{-1}, \theta) F(\theta) \Lambda(\theta)\right] \tag{24}$$

In implementing a personalized pricing system involving one or multiple retailers, the shopbot must be mindful of the overhead in processing time such a system would impose on their ability to return prices to their consumers and the privacy concerns of their consumers. Still, employing such a system would allow shopbots to build lock-in among their consumers and leverage their most important source of competitive advantage — knowledge of consumer behavior.

5. Conclusions

As Internet shopbot technologies mature, consumer behavior at shopbots will become an increasingly important topic for retailers, financial markets, and academic researchers.

Our findings demonstrate that while the market positions of branded retailers are substantially weakened at shopbots, brand name and retailer loyalty still strongly influence consumer behavior at Internet shopbots, giving such retailers a 3.1% and 6.8% price advantage respectively over their competitors. These results may derive from service quality differentiation, asymmetric market information regarding quality, or cognitive lock-in among consumers.

With regard to retailers, the reliability of our models compared to actual consumer behavior suggests that retailers may be able to use shopbot data to provide personalized prices to consumers.

For financial markets, our findings may help to focus the debate on the size and sustainability of market valuations for Internet retailers. Using Amazon.com as an example, our shopbot data indicate that the retailer maintains a 5.0% margin advantage over unbranded retailers and a 6.8% margin advantage among repeat visitors. Both of these statistics are likely to represent lower bounds on the actual margin advantages among their entire consumer base. A margin advantage of this magnitude, if sustainable and applicable across their entire product line, implies a very large capital value.[20] The relevant question then become whether companies such as

[20] For example, Amazon.com reports that 76% of their consumers are repeat visitors, giving them an average margin advantage of 10.2% on their customer base after combining our brand and loyalty results. Zack's Investment Research predicts that Amazon.com will grow by an average of 57.9% over the next 5 years. Amazon.com reports net revenue of $574 million for first quarter 2000 across all product categories. Assuming that Zack's growth projections hold, that growth stops after 5 years, and assuming a 5% interest rate, the net present value of Amazon.com's 10.2% margin advantage is over $40 billion.

Amazon.com can sustain current positions of competitive advantage, how much it will cost to sustain these positions, and whether they can transfer competitive advantage in one product category to other product categories to expand their revenue base.

Finally, for academic researchers, our results demonstrate the feasibility of using Internet shopping data to better understand consumer behavior in electronic markets. Future research in this regard may be able to extend these results to evaluate the cognitive processing costs of shopbot consumers, or to empirically analyze the application of personalized pricing strategies to shopbot consumers. Moreover, our results suggest that the quantity and quality of data available in Internet markets may introduce a revolution the analysis of consumer behavior rivaling that of the scanner data revolution in the 1980s.

References

Akaike, Hirotugu. 1973. Information Theory and an Extension of the Maximum Likelihood Principle. B. N. Petrov and F. Csaki, eds. *Second International Symposium on Information Theory*, Academiai Kiado, Budapest, 267-281.

Amemiya, Takeshi. 1978. On a two-step estimation of a multivariate logit model. *Journal of Econometrics* **8**(1) 13-21.

Bailey, Joseph P. 1998. Intermediation and Electronic Markets: Aggregation and Pricing in Internet Commerce. Ph.D. Thesis, Technology, Management and Policy, Massachusetts Institute of Technology, Cambridge, MA.

Bakos, J. Yannis. 1997. Reducing Buyer Search Costs: Implications for Electronic Marketplaces. *Management Science* **43**(12) 1613-1630.

Bakos, J. Yannis. 1998. The Emerging Role of Electronic Marketplaces on the Internet. *Communications of the ACM* **41**(8)

Bearse, Peter M., Hamparsum Bozdogan, Alan M. Schlottmann. 1997. Empirical econometric modelling of food consumption using a new informational complexity approach. *Journal of Applied Econometrics* **12**(5) 563-592.

Ben-Akiva, Moshe, Steven R. Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand.* The MIT Press, Cambridge, MA.

Bozdogan, Hamparsum. 1990. On the information-based measure of covariance complexity and its application to the evaluation of multivariate linear models. *Communication in Statistics — Theory and Methods* **19**(1) 221-278.

Brynjolfsson, Erik, Michael Smith. 2000. Frictionless Commerce? A Comparison of Internet and Conventional Retailers. *Management Science* **46**(4)563-585.

Bucklin, Randolph E., Sunil Gupta. 1992. Brand choice, purchase incidence, and segmentation: An integrated modeling approach. *Journal of Marketing Research* 29(2) 201-215.

Clay, Karen, Ramayya Krishnan, Eric Wolff, Danny Fernandes. 1999. Retail Strategies on the Web: Price and Non-price Competition in the Online Book Industry. Working Paper, Carnegie-Mellon University, Pittsburgh, PA.

Clemons, Eric K., Il-Horn Hann, Lorin M. Hitt. 1998. The Nature of Competition in Electronic Markets: An Empirical Investigation of Online Travel Agent Offerings. Working Paper, The Wharton School of the University of Pennsylvania, Philadelphia, PA.

de Figueiredo, John. 2000. Using strategic tools to generate profits in e-commerce. *Sloan Management Review* forthcoming.

Erdem, Tulin, Michael P. Keane. 1996. “Decision-Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets. *Marketing Science* **15**(1) 1-20.

Erdem, Tulin, Michael P. Keane, Baohong Sun. 1999. Missing price and coupon availability data in scanner panels: Correcting for the self-selection bias in choice model parameters. *Journal of Econometrics*. **89** 177-196.

Erdem, Tulin, Joffre Swait. 1998. Brand Equity as a Signaling Phenomenon. *Journal of Consumer Psychology* **7**(2) 131-157.

Fader, Peter S., Bruce G.S. Hardie. 1996. Modeling Consumer Choice Among SKUs. *Journal of Marketing Research* **33**(November) 442-452.

Goldberg, Pinelopi Koujianou. 1995. Product differentiation and oligopoly in international markets: The case of the U.S. automobile market. *Econometrica* **63**(4) 891-951.

Greenwald, Amy R., Jeffrey O. Kephart. 1999. Shopbots and Pricebots. *The Proceedings of International Joint Conference on Artificial Intelligence*.

Guadagni, Peter M., John D.C. Little. 1983. A Logit Model of Brand Choice Calibrated on Scanner Data. *Marketing Science* **2**(3) 203-238.

Guadagni, Peter M., John D.C. Little. 1998. When and What to Buy: A Nested Logit Model of Coffee Purchase. *Journal of Forecasting* **17** 303-324.

Hausman, Jerry A., David A. Wise. 1978. A conditional probit model for qualitative choice: Discrete decisions recognizing intedependence and heterogeneous preferences. *Econometrica* **46**(2) 403-426.

Johnson, Eric J., Steven Bellman, Gerald L. Lohse. 2000. “What makes a web site “sticky”? Cognitive lock in and the power law of practice. Working Paper, Columbia School of Business, Columbia University, New York, NY.

Kahneman, Daniel, Amos Tversky. 1979 Prospect theory: An analysis of decision under risk. *Econometrica* **47**(2) 263-292.

Kamakura, Wagner A., Gary J. Russell. 1989. A Probabilistic Choice Model for Market Segmentation and Elasticity Structure. *Journal of Marketing Research* **26**(November) 379-390.

McCulloch, Robert, Peter E. Rossi. 1994. An exact likelihood analysis of the multinomial probit model. *Journal of Econometrics* **64**(2) 207-240.

McFadden, Daniel. 1974. Conditional logit analysis of qualitative choice behavior. Zarembka, ed. *Frontiers in Econometrics*. Academic Press, New York, 105-142.

McFadden, Daniel. 1989. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica*. **57**(5) 995-1026.

Milgrom, Paul, John Roberts. 1986. Price and Advertising Signals of Product Quality. *The Journal of Political Economy*. **94**(4) 796-821.

Montgomery, Alan T. 1997. Creating micro-marketing pricing strategies using supermarket scanner data. *Marketing Science* **16**(4) 315-227.

Morwitz, Vicki, Eric A. Greenleaf, Eric J. Johnson. 1998. Divide and Prosper: Consumers' Reactions to Partitioned Prices. *Journal of Marketing Research*. **35**(November) 453-463.

Raftery, Adrian E. 1986. A note on Bayes factors for log-linear contingency table models with vague prior information. *Journal of the Royal Statistical Society: Series B (Methodological* **48**(2) 249-250.

Raftery, Adrian E. 1995. Bayesian model selection in social research, Marsden ed. *Sociological Methodology*. Blackwell Publishers, Oxford, 111-163.

Rossi, Peter E., Robert E. McCulloch, Greg M. Allenby. 1996. The value of purchase history data in target marketing. *Marketing Science* **15**(4) 321-340.

Schwarz, Gideon. 1978. Estimating the dimension of a model. *Annals of Statistics* **6**(2) 461-464.

Smith, Michael D.; Bailey, Joseph; Brynjolfsson, Erik. 2000. Understanding Digital Markets: Review and Assessment. E. Brynjolfsson, B. Kahin, eds. *Understanding the Digital Economy*, MIT Press, Cambridge, MA. Forthcoming.

Swait, Joffre, Jordan Louviere. 1993. The role of the scale parameter in the estimation and comparison of multinomial logit models. *Journal of Marketing Research* **30**(3) 305-314.

Varian, Hal R. 2000. Market Structure in the Network Age. E. Brynjolfsson, B. Kahin, eds. *Understanding the Digital Economy*, MIT Press, Cambridge, MA. Forthcoming.

Appendix A: Week-by-Week Predicted to Actual Choice Share, Multinomial Logit Model

Figure A.1: Amazon.com

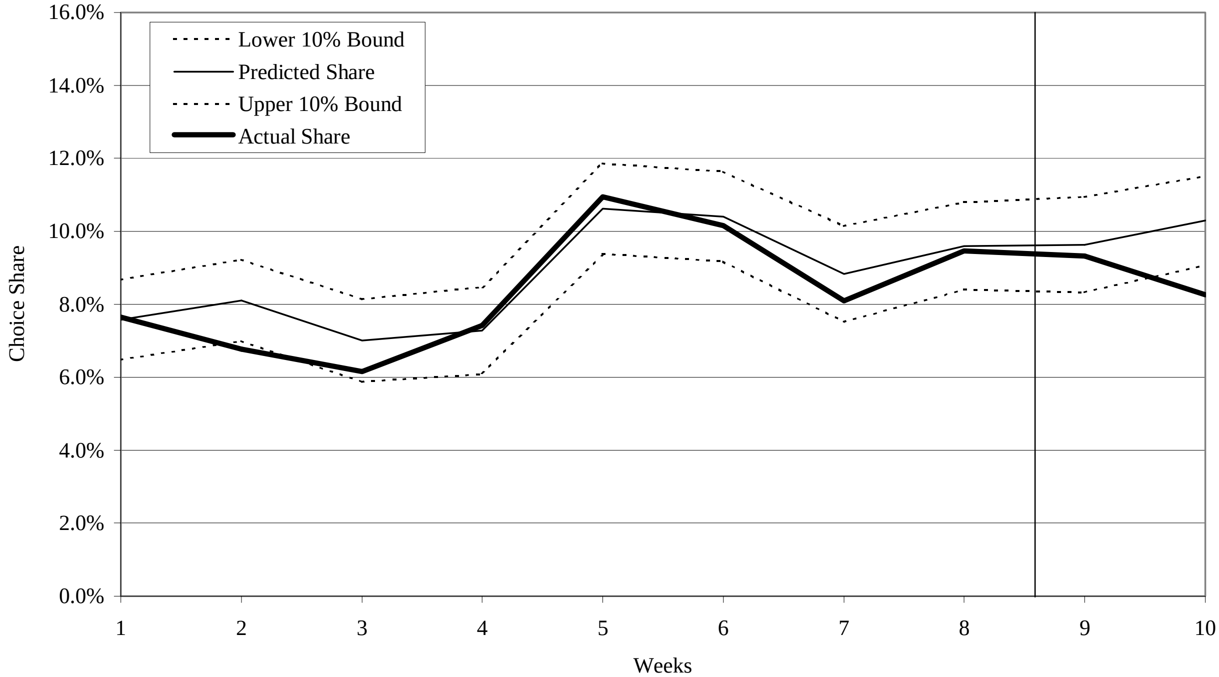


Figure A.2: BarnesandNoble.com

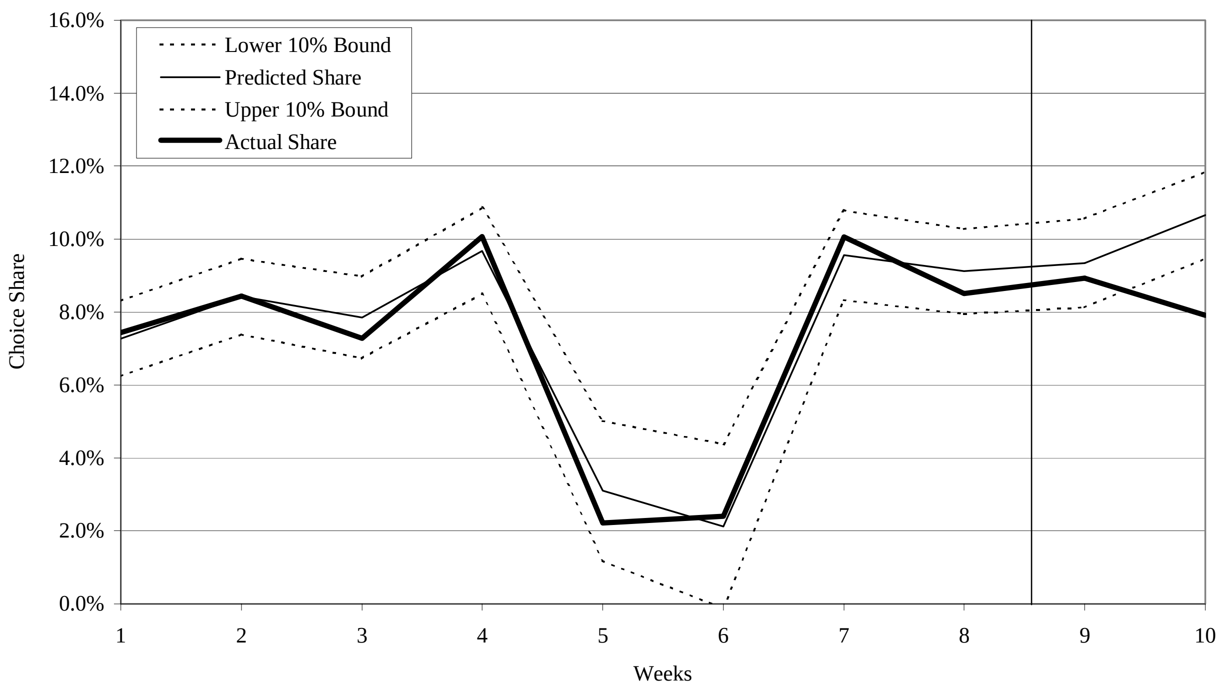


Figure A.3: Borders.com

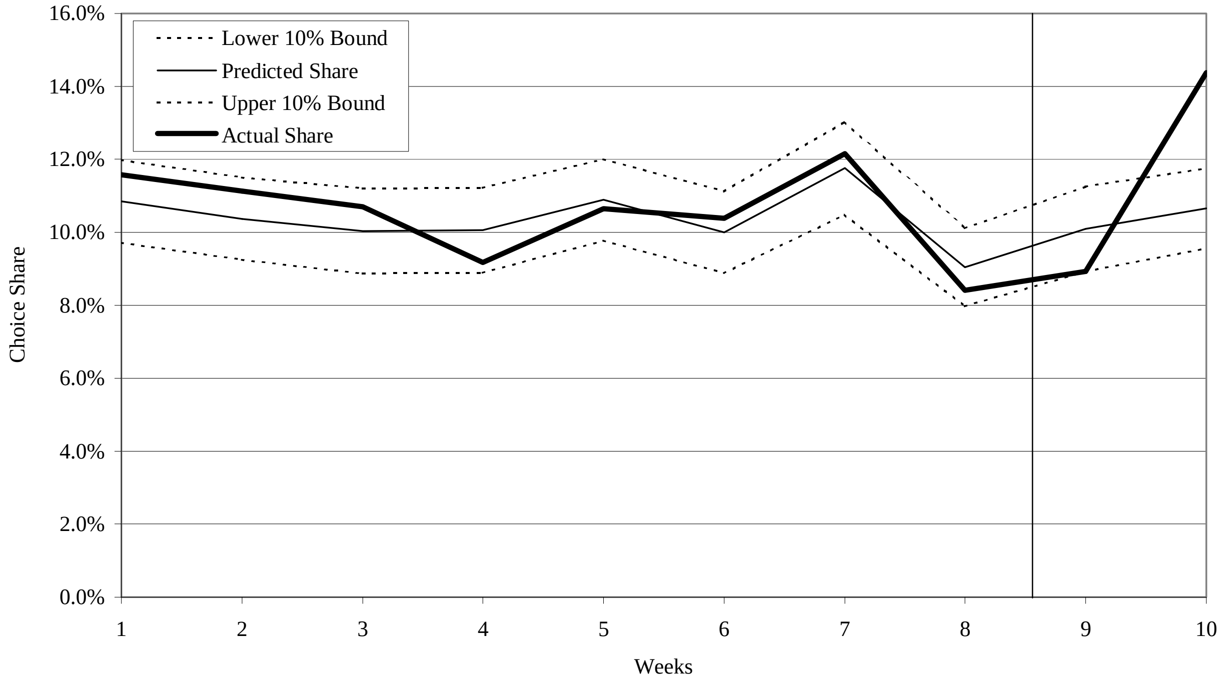


Figure A.4: A1Books.com

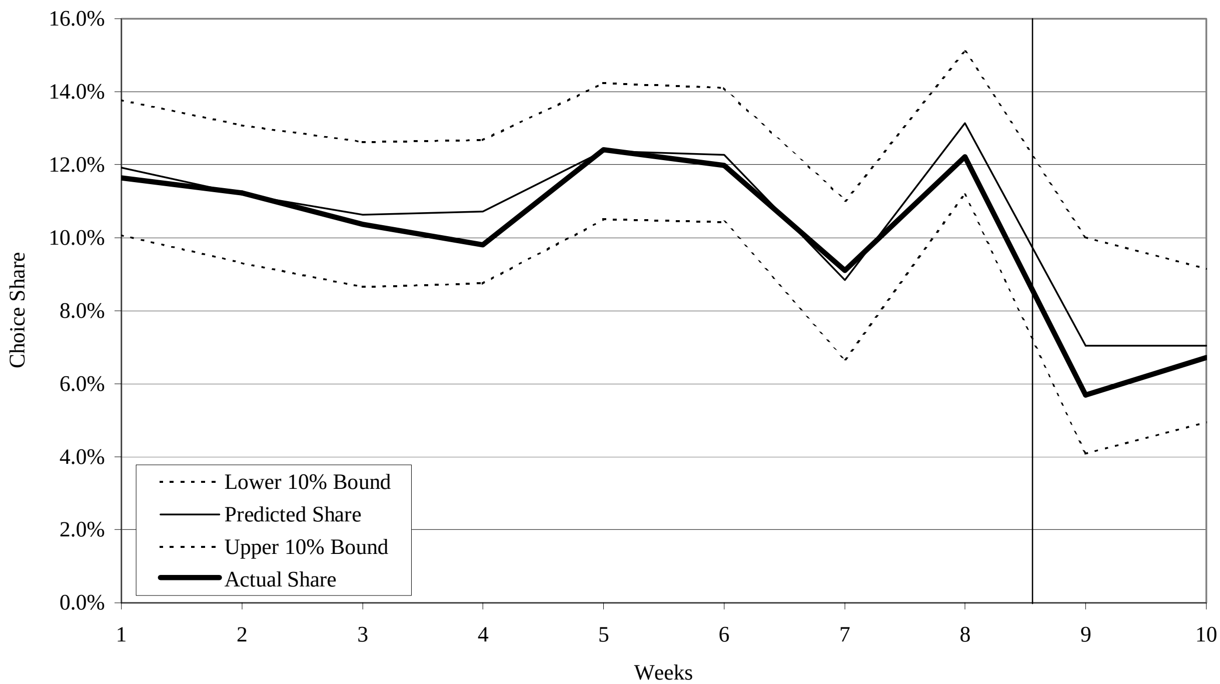


Figure A.5: Kingbooks.com

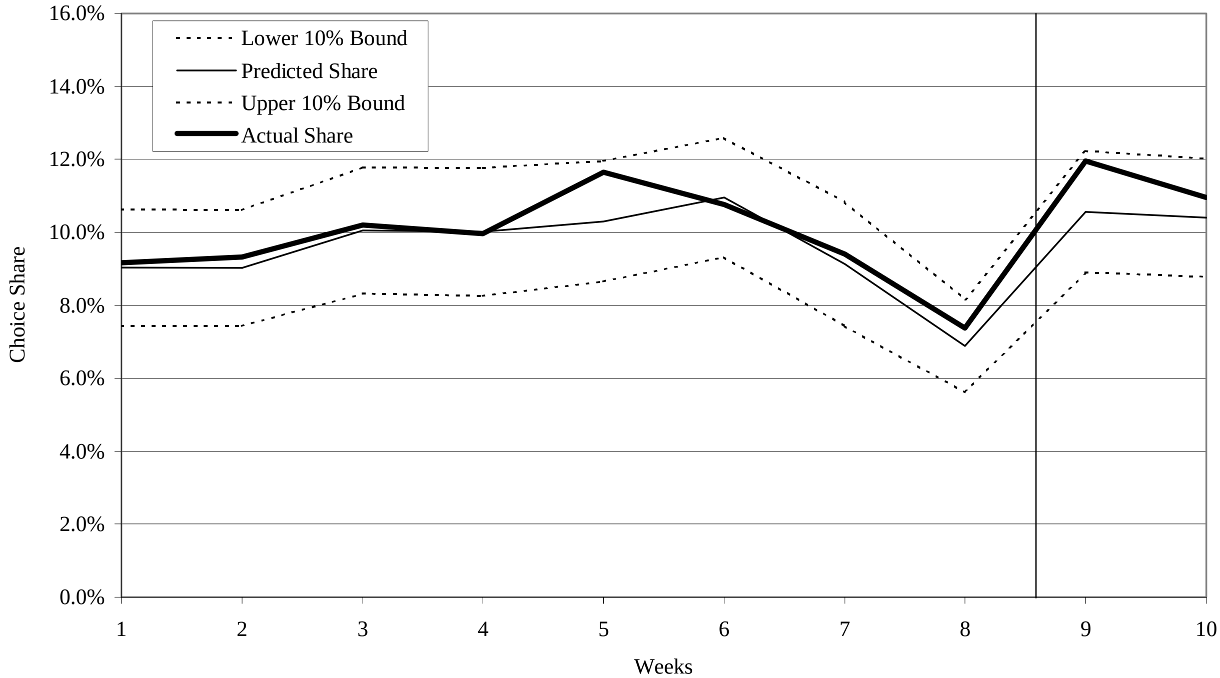


Figure A.6: 1Bookstreet.com

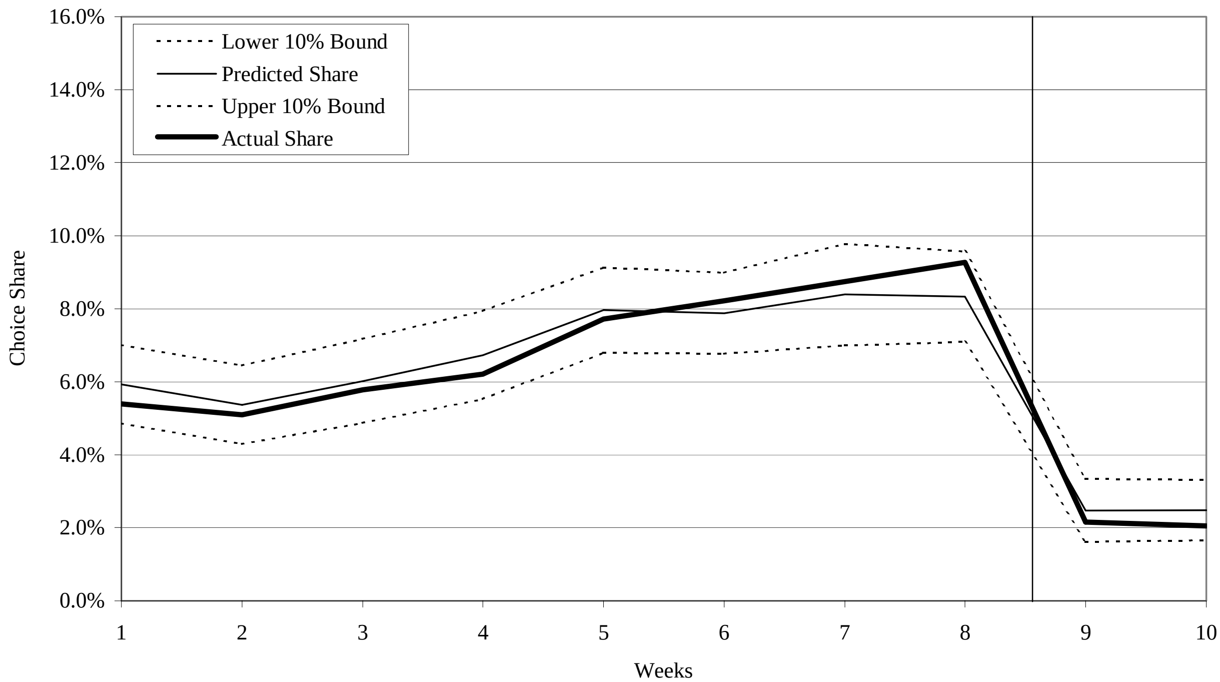


Figure A.7: AlphaCraze.com

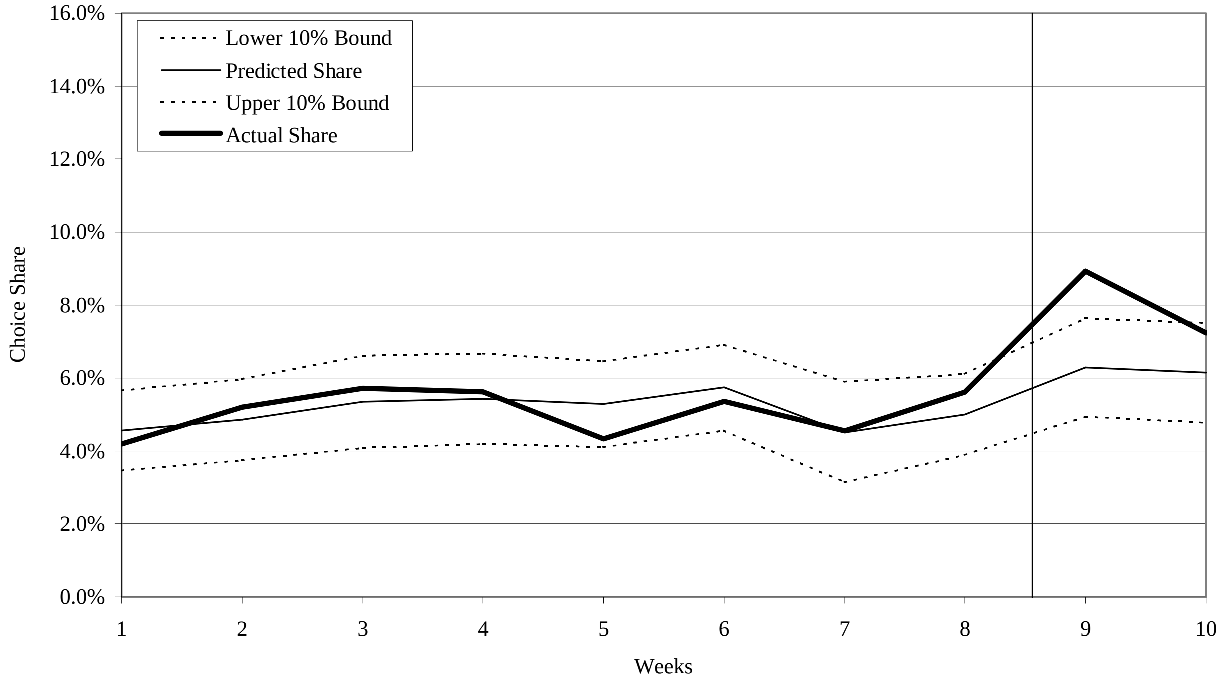


Figure A.8: Alphabetstreet.com

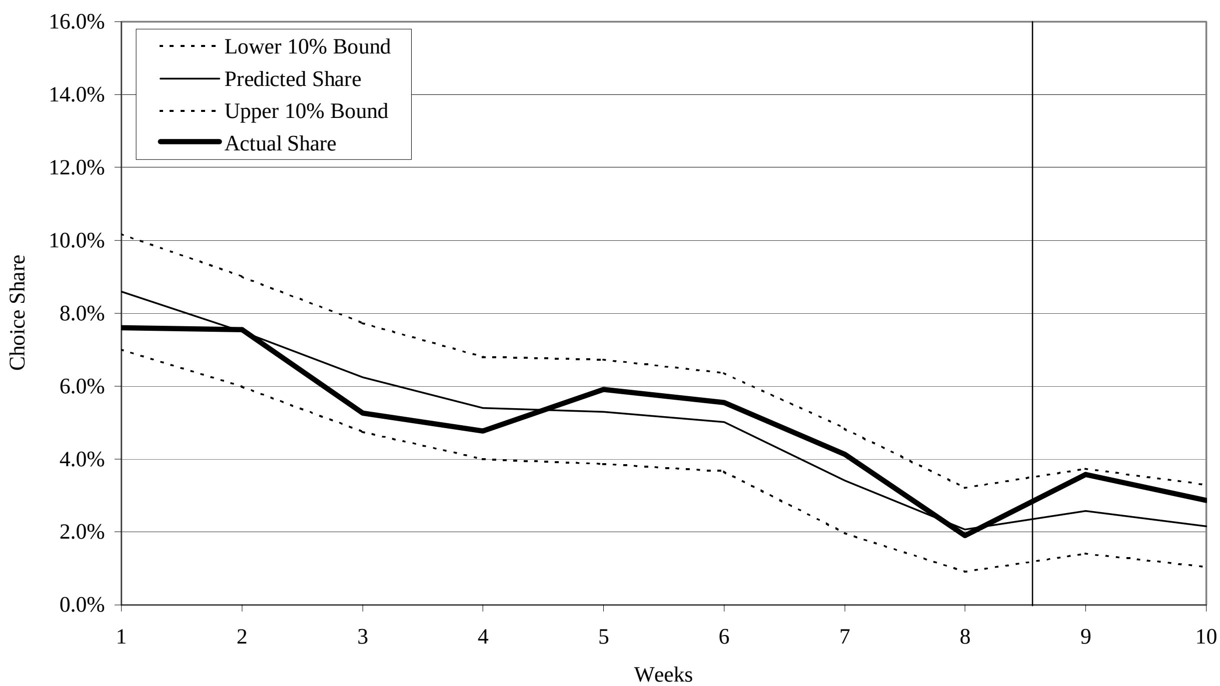


Figure A.9: Shopping.com

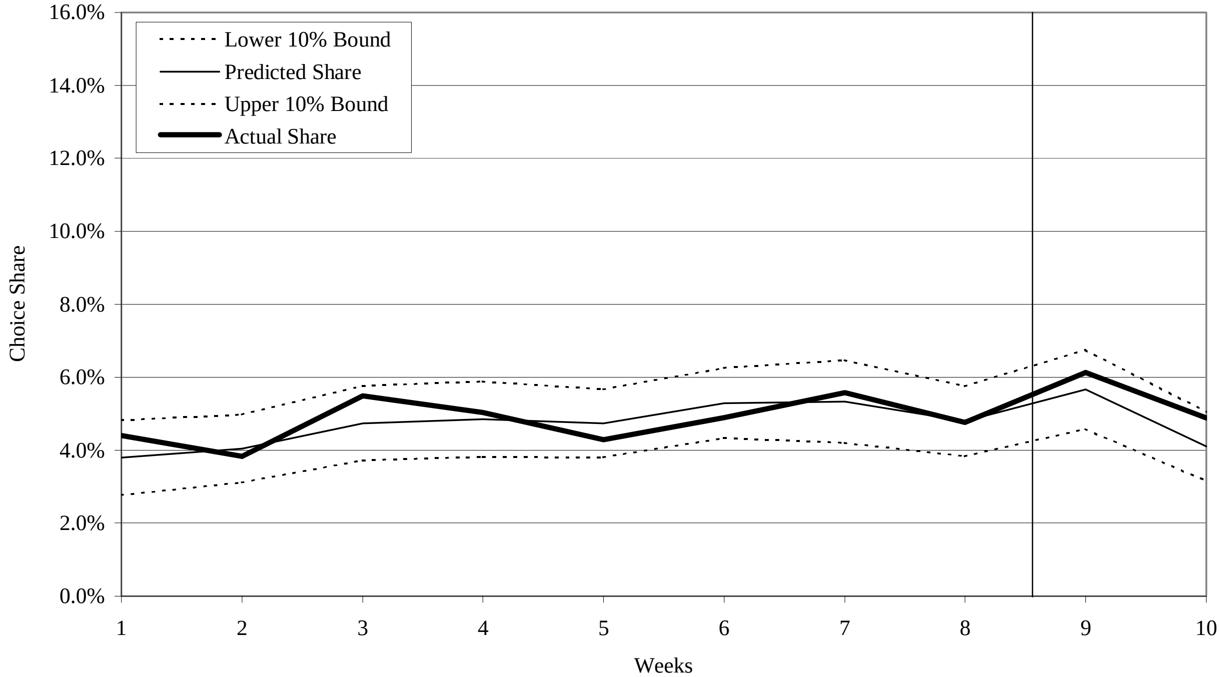


Figure A.10: FatBrain.com

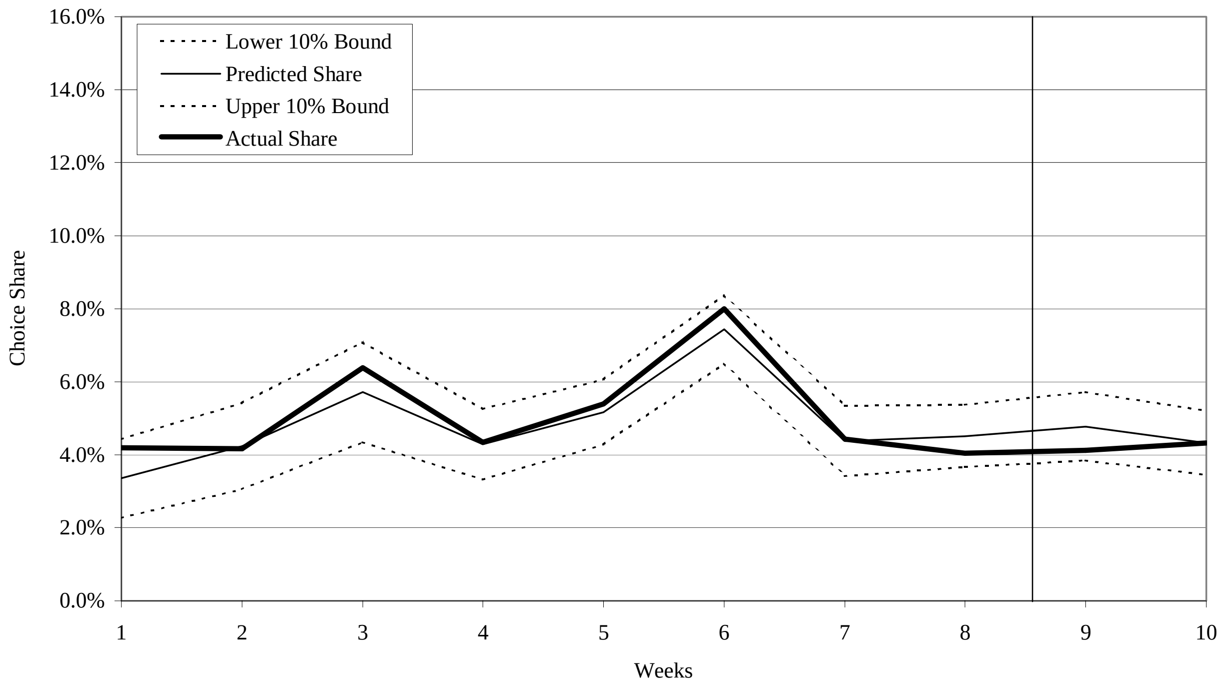


Figure A.11: Classbook.com

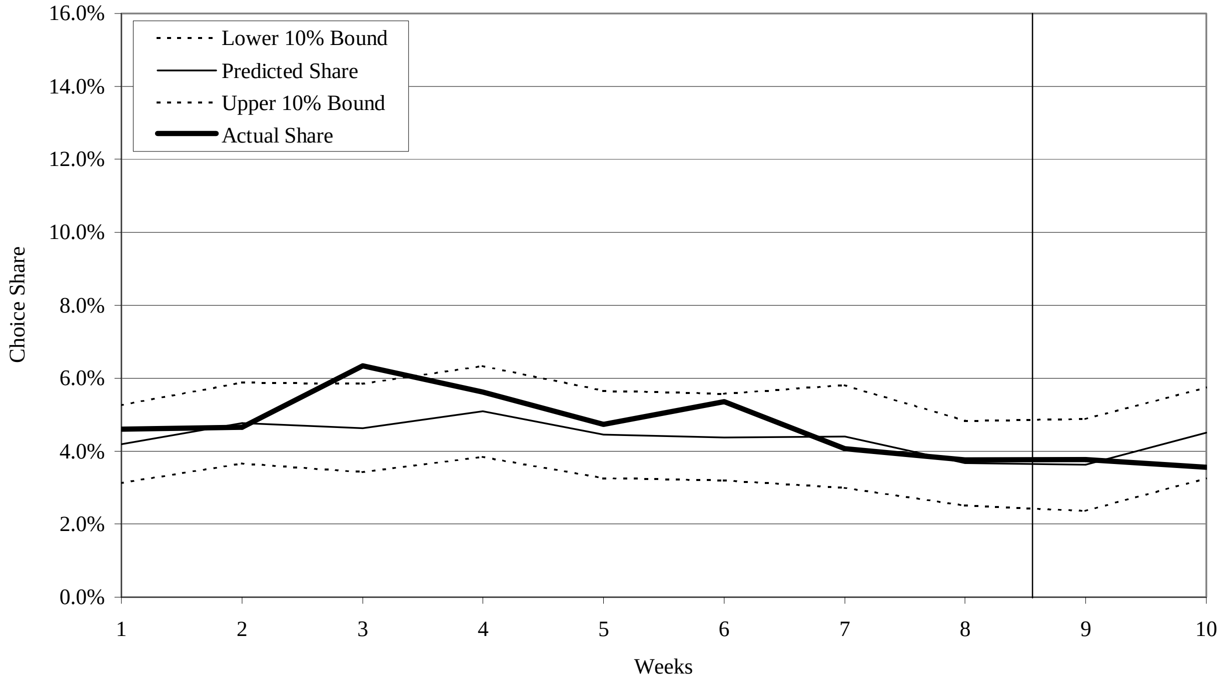


Figure A.12: Books.com

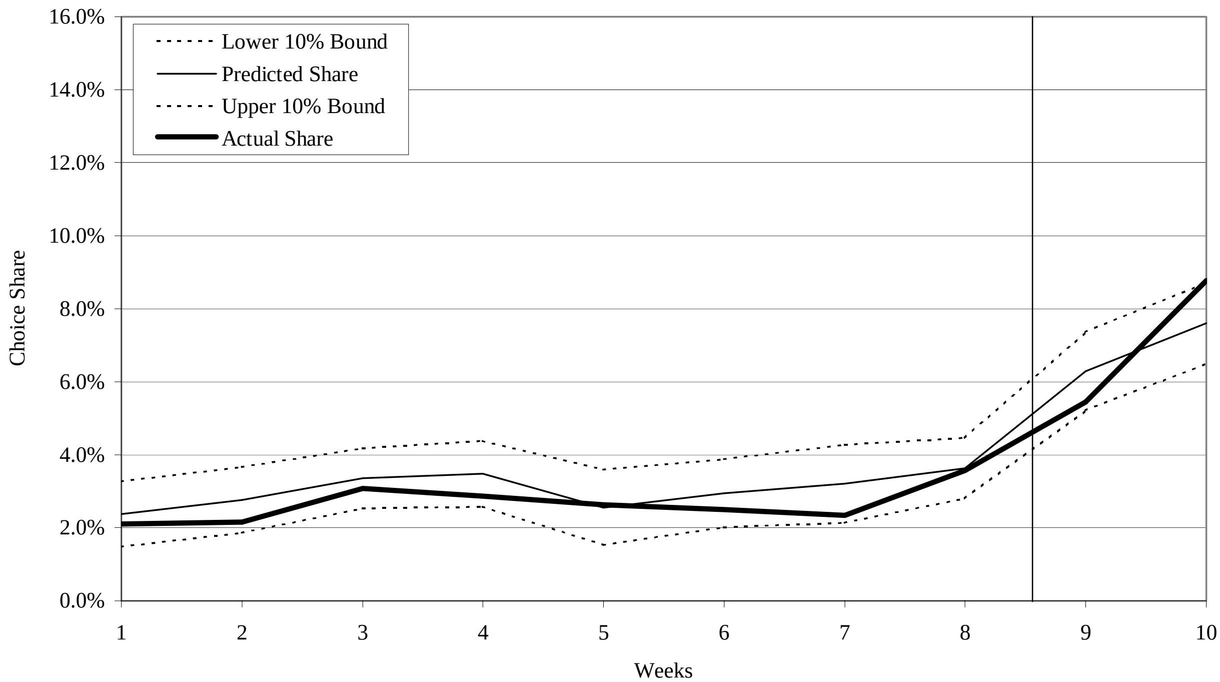
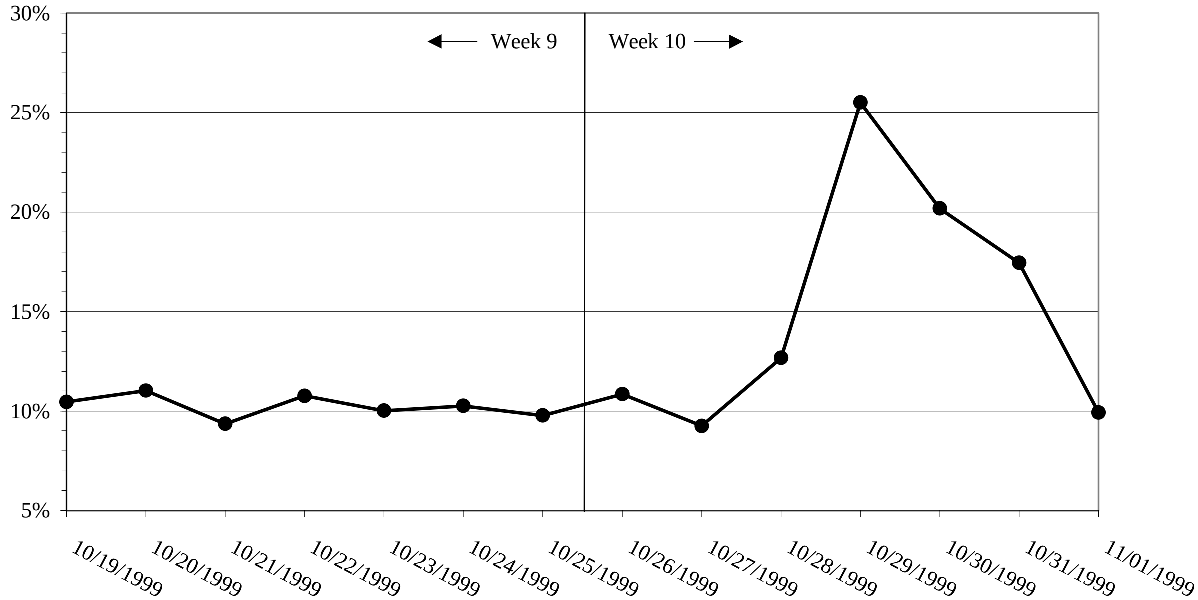


Figure A.13: Borders Last Click-Through Share — 10/19/99 - 11/1/99



Appendix B: Week-by-Week Predicted to Actual Choice Share, Branded Retailers, Nested Logit Model

Figure B.1: Amazon.com

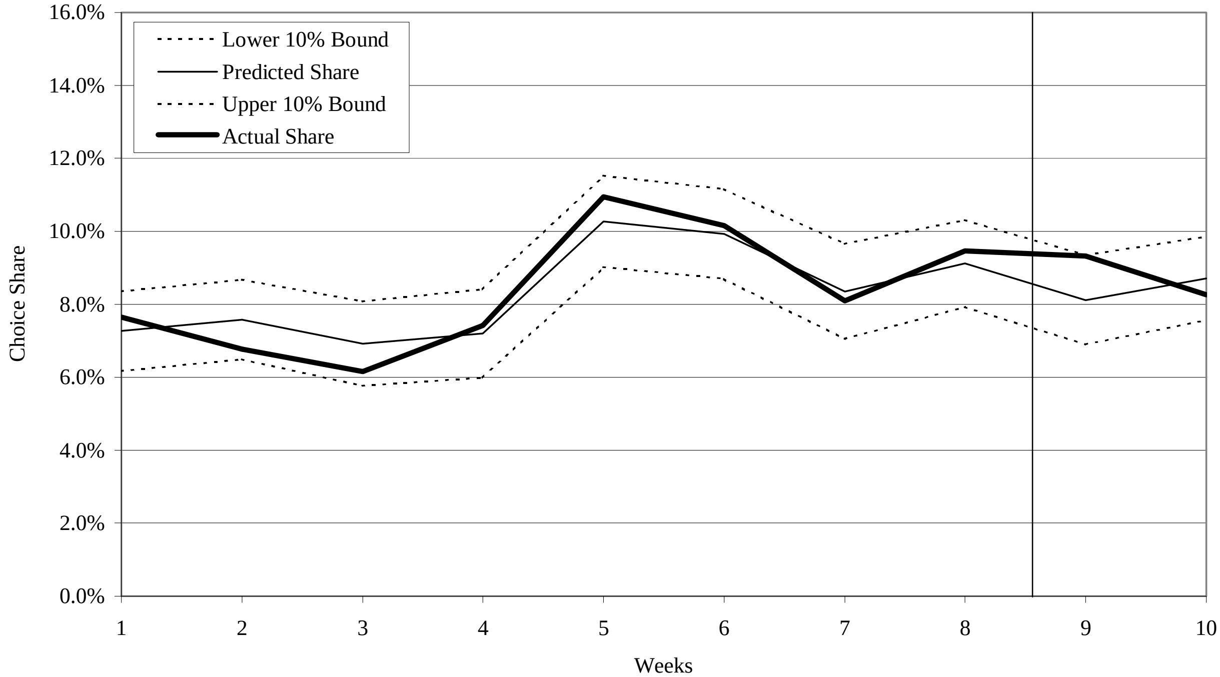


Figure B.2: BarnesandNoble.com

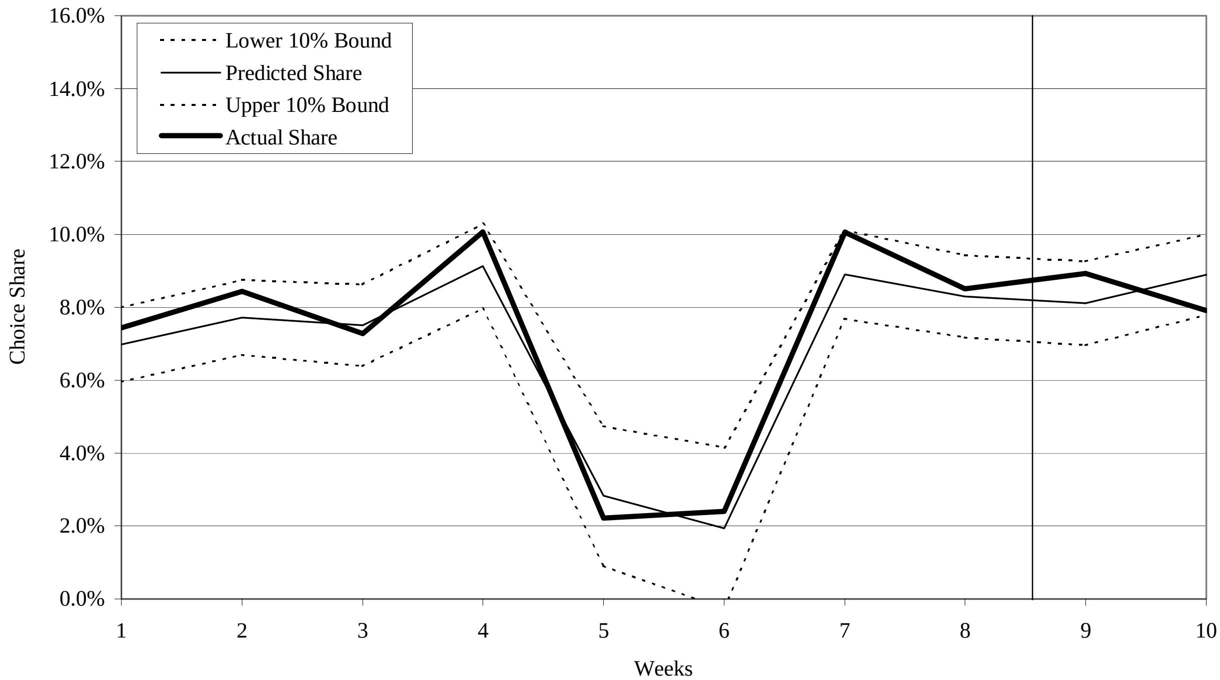
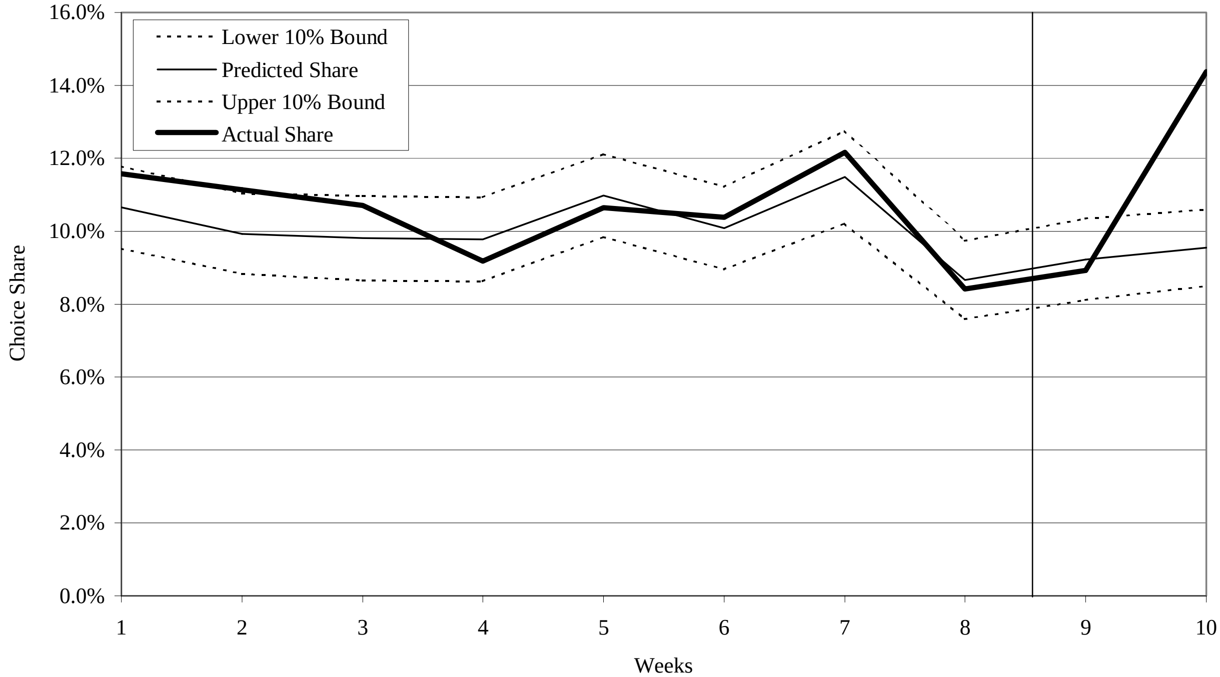


Figure B.3: Borders.com



Appendix C: Model Selection Criteria

This appendix presents several of the most common model selection criteria applied to multinomial logit models. As noted above, these criteria fall into two general categories: log likelihood-based measures and information theoretic measures. Significant criteria from each category are presented in turn below.

The most common model selection criterion is the likelihood ratio test. Likelihood ratio tests can be used to evaluate multiple restrictions on a model (e.g., Guadagni and Little 1983). Likelihood ratio tests in this setting are based on the observation that $2(\log(L(\hat{\theta}_A)) - \log(L(\hat{\theta}_B)) \sim \chi^2$ with degrees of freedom equal to the number of restrictions between model A and B.

Applied to our model, likelihood ratio tests reject at any reasonable confidence level the restrictions on specification 1 above with respect to all other specifications and on specification 2 with respect to specification 3. However, these tests are only applicable where one model can be expressed as a restricted subset of the second model. Therefore we cannot use likelihood ratio tests to compare specification 3 to specification 4, for example.

Another technique to choose among multinomial logit model specifications is to use a measure of fit analogous to R^2 in multivariate linear regressions. McFadden (1974) proposes to measure this value as

$$U^2 = 1 - \frac{\log L(\hat{\theta}^*)}{\log L(\hat{\theta}^0)} \tag{C.1}$$

where $L(\hat{\theta}^*)$ is the likelihood associated with the specification in question and $L(\hat{\theta}^0)$ is the likelihood of the null model (the constrained model excluding all regressors).

Ben-Akiva and Lerman (1985, p. 167) note that this measure will always (weakly) increase when new variables are added to the model whether or not these variables contribute usefully to explaining the data. Therefore, this measure does not adequately account for desired parsimony in the selected specification. For this reason, the Ben-Akiva and Lerman adjust McFadden's U^2 measure to penalize the addition of variables

$$\overline{U}^2 = 1 - \frac{\log L(\hat{\theta}^*) - k}{\log L(\hat{\theta}^0)} \tag{C.2}$$

where k is the number of independent variables in the model. Using either measure, the best model is the one with the largest U^2, corresponding to the model that explains the most variation in the data. Further, unlike the likelihood ratio presented above, these tests can be used to compare models that cannot be expressed as restricted subsets of each other.

A variety of model selection measures have been proposed based on concepts of information theory. The most well known of these measures, the Akaike Information Criterion or AIC (Akaike 1973) is specified as

$$AIC = \frac{-2\log L(\hat{\theta}) + 2P}{N} \tag{C.3}$$

where P is the number of parameters in the model (the number of independent variables plus the slope coefficient) and N is the number of observations. Intuitively, for models with better fit, $L(\hat{\theta})$ should increase and $-2\log L(\hat{\theta})$ should decrease. The $2P$ term will decrease with more parsimonious models. Thus, the "best" model minimizes the AIC criterion.

The Bayesian Information Criterion or BIC (Raferty 1986, Schwartz 1987) provides a similar measure, based on Bayesian statistical theory. In a Bayesian setting, we compare two models based on the ratio of their posterior probabilities. If Model 2 is preferred over Model 1 this odds ratio will be greater than 1. The posterior odds ratio of Model 2 to Model 1 can be written as

$$\frac{\mathbf{P}(M_2 \mid Data)}{\mathbf{P}(M_1 \mid Data)} = \frac{\mathbf{P}(Data \mid M_2)}{\mathbf{P}(Data \mid M_1)} \cdot \frac{\mathbf{P}(M_2)}{\mathbf{P}(M_1)} \tag{C.4}$$

where the first factor on the right hand side of the equation is called the Bayes factor for Model 2 against Model 1 and the second factor is the ratio of the prior probability for Model 2 against Model 1. In the general case where there is no prior probability for choosing Model 2 against Model 1, this ratio will be 1 and the posterior odds ratio will be equal to the Bayes factor. Unfortunately, calculating the Bayes factor is computationally prohibitive.

However, the Bayesian Information Criterion (BIC) presents a useful, and easily calculated, approximation to the Bayes Factor. BIC is defined as

$$BIC = -2\ln L(\hat{\theta}) - (N-k)\ln N \tag{C.5}$$

where $\hat{\theta}$ and N are defined as above, and k is the number of regressors. Relating this to the Bayes factor, it can be shown (Raftery 1995) that

$$2\ln\left(\frac{\mathbf{P}(Data \mid M_2)}{\mathbf{P}(Data \mid M_1)}\right) \approx BIC_1 - BIC_2 \,. \tag{C.6}$$

Thus, as with the AIC measure above, the best model is the model that minimizes BIC.

The information theoretic measure of complexity or ICOMP (Bozdogan 1990; Bearse, Bozdogan, Schlottmann 1997) provides an alternate model selection criteria. ICOMP uses the Fisher information matrix to measure (penalize) complexity in the model. The measure is defined as

$$ICOMP = -2\ln L(\hat{\theta}) - k\ln(tr(I^{-1}(\hat{\theta}))/k) - \ln\left|I^{-1}(\hat{\theta})\right| \tag{C.7}$$

where $I^{-1}(\hat{\theta})$ is the inverse Fisher information matrix. The advantage of ICOMP is that, instead of viewing complexity as arising from the number of parameters (e.g., $\overline{U}^2$, AIC, BIC), it evaluates model complexity from the correlation structure of the parameter estimates (through the inverse Fisher information matrix).