

# Consumer Experience and the Value of Search in the Online Textbook Market

Forrest Spence\*

May 27, 2015

## Abstract

This paper examines the effect of market experience on consumers' online search and purchasing decisions. I quantify this effect by estimating consumers' value of search in the online textbook market and how this value changes as they gain market experience. I then use these estimates to calculate the welfare gains to experience that result from changes in consumers' expected value of search. To conduct this analysis, I collect detailed individual-level purchasing and search data on 6,000 consumers in the University of North Carolina textbook market. I use these data to estimate a demand model that incorporates limited information and costly search as a function of past market experience. Search cost estimates imply that first-year students face a median search cost of \$13 in the online market relative to \$1 for more experienced consumers. Consumers' estimated returns to experience are approximately 5% of the purchase price. These results imply that policies that provide information to inexperienced consumers about market-level characteristics could significantly affect consumer and firm behavior, especially in markets with a large proportion of inexperienced consumers.

---

\*University of North Carolina - Chapel Hill; fspence@email.unc.edu. I thank my advisor, Brian McManus, and committee members Gary Biglaiser, Donna Gilleskie, James Roberts, and Helen Tauchen for helpful comments and encouragement throughout the course of this research project. I also thank Chris Cronin, Babur de los Santos, Liran Einav, Glenn Ellison, Robert Hammond, Matthew Harris, J.F. Houde, Clement Joubert, Myrto Kalouptsi, Robin Lee, Brett Matsumoto, Nick Papageorge, Klara Peter, Tiago Pires, Imke Reimers, Michelle Sovinsky, Andrew Sweeting, Chad Syverson, Matthjis Wildenbeest, Daniel Xu, and Andy Yates for their comments and suggestions. I am grateful for funding for this project provided through the Herbert Brown Mayo and Joseph Kampf Research Funds. All remaining errors are my own.

# 1 Introduction

First-time consumers in a market face several challenges. Information on the existence of retailers, the products they offer, and the prices and characteristics of these products are all costly to obtain. As consumers repeatedly participate in a market and interact with other market participants, they may acquire information about characteristics of the market that are common across all consumers. This paper investigates how consumers' decisions to search across retailers change as they learn about these market-level characteristics.

A considerable amount of research in the economics and marketing literatures focuses on how individuals learn about match-specific characteristics (e.g., Erdem and Keane, 1996; Akerberg, 2003), but little work has been done to understand how individuals learn about characteristics of a market as a whole. For instance, Crawford and Shum (2005) examine the importance of individuals' uncertainty about the efficacy of anti-ulcer medications (a match-specific characteristic). However, consumers may also face uncertainty about the availability of alternative medications, which retailers offer these medications, and the distribution of retailers' prices (market-level characteristics).

From a policy perspective, it is important to distinguish whether consumers are learning primarily about match-specific or market-level characteristics. Informational policies can be used to alleviate uncertainty about market-level characteristics, but will be ineffective at changing behavior in markets where the consumers primarily face uncertainty regarding match-specific characteristics. In order to evaluate the potential value of informational policies in a market, it is critical to understand if there are systematic differences in behavior across consumers that result from learning about market-level characteristics. If these differences exist, policies that provide consumers with information about market-level characteristics can increase the rate of learning for inexperienced consumers in a market and improve consumer outcomes. This may be particularly effective in markets where a large proportion of the consumers have made relatively few purchases (e.g., durable goods or health care markets).

Experience appears to play a large role in consumer decision-making in the college textbook market. First-year students (generally first-time participants) tend to purchase their textbooks at university bookstores much more frequently than upper-year students (experienced participants). All participants in the market are generally provided with information about the campus bookstore, but may significantly differ in the amount of information they

have about alternatives available outside of the campus bookstore. In particular, experienced consumers may have learned characteristics of the online market through repeated participation that changes their purchasing behavior.

As consumers gain information about the online market, this affects their decision to engage in search across retailers. The incentives to engage in search are broadly defined as consumers' value of search: consumers' beliefs about the benefits of search minus the cost of engaging in search. This paper contributes to the literatures on consumer search and learning by evaluating the level of heterogeneity in consumers' value of search, the degree to which this value varies by experience gained through repeated participation in a market, and the benefits consumers receive from increased information about characteristics of the online textbook market.

This paper makes a policy contribution by calculating the potential welfare gains that would result from eliminating the gap in consumers' expected value of search in the online market across experience levels. If experience leads to less uncertainty about the availability of online retailers, their reliability, and their prices, policies that provide information to consumers can increase the rate of learning over characteristics of the online market, thereby removing welfare gaps attributable to differences in experience. Policies that have similar effects have been enacted at several public and private universities such as NYU and the University of Texas; these campus bookstores use software designed by Verba, Inc. to provide students with prices from the campus bookstore and numerous online retailers.

Differences in search behavior across experience levels may be the result of consumers learning about characteristics of the market. However, in order to identify the difference in search behavior that is attributable to differences in consumers' expected value of search, it is necessary to rule out competing hypotheses for the observed difference in search and purchasing behavior. In the college textbook market, this difference could be due to a number of factors that are correlated with experience: who typically pays for the textbook, tastes for quality or retailers, previous online purchasing behavior, the types of courses taken, and beliefs about the value of textbooks among other things. I account for these factors by collecting detailed individual-level data on 6,000 consumers in the University of North Carolina – Chapel Hill (UNC) textbook market using online questionnaires.<sup>1</sup> These

---

<sup>1</sup>Each questionnaire was tailored to a specific course and textbook assignment. Instructors electronically distributed questionnaire links to students in their courses. Student participation was voluntary and encouraged through the opportunity to win cash prizes. This project received approval from the UNC IRB.

data span four semesters and more than 700 independent course-sections. Each observation contains the following information: whether and how the student obtained the assigned textbook, who paid or offered to pay for it, the student's online search behavior and past purchasing behavior, the student's price expectations, and his or her stated preferences regarding bookstore shopping and textbook conditions. In order to evaluate the benefits of search in the online market, I also scrape daily price data from 13 online textbook retailers across two academic calendar years. These data allow me to approximate the empirical distribution of prices across online retailers during the time period considered. I supplement these data with information provided by the university registrar and university bookstore.

In the data I find that inexperienced consumers are significantly less likely to search online retailers than their more experienced counterparts, despite the potential for substantial savings. The price of a new and used textbook from Amazon.com was approximately 30% lower on average than the price of the same textbook from the campus bookstore. This difference in prices was nearly identical for textbooks assigned to first-year students and upper-year students, but first-year students were 18% less likely to search online retailers. Reduced-form results suggest that search behavior is significantly different across experience levels even after controlling for differences in budgets, past experience purchasing online, reported tastes for online and bookstore shopping, and course characteristics.

In order to calculate the value of search across experience levels, I estimate a model of demand with limited information and costly search similar to Hortaçsu and Syverson (2004) and Hong and Shum (2006). In the model, consumers face a decision for one assigned textbook. They have full information about the price and characteristics of the campus bookstore alternatives (new, used, and rental), but face limited information about characteristics of textbooks sold in the online market. In particular, they face uncertainty over prices and their unobservable preference shocks for alternatives in the online market prior to making the search decision (although they know their joint distribution with certainty). Prices and preference shocks are revealed if the consumer pays a known search cost. This search cost is drawn from a distribution that depends on a number of observable factors including previous online shopping experience, experience in the textbook market, and reported confidence in the reliability of the online textbook market.

The majority of the learning literature identifies changes in information by assuming consumers' beliefs follow a Bayesian updating process. In contrast, I identify learning in a reduced-form manner similar to Akerberg (2001). I follow the the majority of other

structural models of consumer search in assuming that consumers have rational expectations for prices (i.e., consumers beliefs match the empirical distribution of online prices prior to making the search decision). Consumer learning is captured by reduced form parameters that shift the distribution of search costs. One advantage of this technique is that it avoids the endogeneity of consumers' decisions to acquire signals. The measure of experience used in this paper, semesters enrolled, is arguably exogenous to the decision to search an online retailer for an assigned textbook.<sup>2</sup> This is contrasted by a measure such as previous online purchases or searches, which are endogenously acquired based on beliefs about the value of search.<sup>3</sup>

The search cost and utility parameters are estimated using a modification of a routine proposed by Koulayev (2012). This estimation procedure takes advantage of a reservation utility rule to form choice probabilities used in simulated maximum likelihood estimation. Estimation is facilitated by the observation of both the search and the purchase decision, including the decision to search online retailers and still purchase from the bookstore.<sup>4</sup>

Estimates suggest that experience has a significant effect on the decision to search online retailers, and thus a significant effect on consumers' choice of retailer, even after controlling for consumer, textbook, and course differences. Median estimated search costs are \$13 for inexperienced consumers relative to \$1 for their more experienced counterparts. Counterfactual results suggest that eliminating the gap in search costs across experience levels leads to an average welfare gain of 5% of the price of a new textbook from the campus bookstore (\$8 for a typical textbook assignment).

Overall, these results imply that policies aimed at improving information in the textbook market could significantly benefit consumer welfare by encouraging more search across retailers. Moreover, the link between market information and search carries additional weight

---

<sup>2</sup>More specifically, as long as consumers are not making decisions about being enrolled in college based on their expectations about the value of searching for a textbook in the online market, then semesters enrolled is exogenous to the search decision for a particular textbook.

<sup>3</sup>The decision to acquire signals about the online market could be driven by selection (e.g. individuals that search more are also individuals with lower priors about the distribution of online prices), leading to biased estimates of beliefs and the rate of learning.

<sup>4</sup>A common problem in the search literature is dealing with the issue of endogenous choice set formation. I do not need to use methods to recover latent choice sets because these are observable in the data. The observation of individuals who search the online market and still choose a bookstore alternative also allows for the separate identification of the distribution of search costs versus heterogeneity in tastes for online and bookstore shopping.

in markets with positive externalities. Search not only reduces the expected price a consumer will pay, but also induces market participation for marginal consumers who were not willing to pay the pre-search price. In markets with positive externalities, such as textbooks (education) and alternative energy (pollution reduction), informational policies that alleviate uncertainty about market-level characteristics are a fiscally responsible way of increasing market participation, and hence welfare.

These results also have broad implications for firm behavior. Firms may be able to segment markets by experience and take advantage of informational differences to engage in price discrimination. For example, pricing and revision behavior may vary substantially for introductory textbooks which are primarily aimed at inexperienced consumers, than advanced undergraduate textbooks, which are aimed at more experienced consumers). Finally, these results have important implications for antitrust policy in markets with infrequent purchases. When evaluating whether a proposed merger is anticompetitive, antitrust authorities typically consider entry by new firms to be a mitigating factor. My analysis suggests that inexperienced consumers are less likely to search across unfamiliar retailers, representing a barrier to entry for new firms. These information frictions may therefore warrant stricter merger policy in markets where consumers tend to be inexperienced.

The following section provides an overview of research on demand estimation with limited information. Section 3 presents the empirical model. Section 4 presents the data and Section 5 provides parameter estimates and counterfactual results. Section 6 concludes.

## **2 Literature Review**

This paper draws from two well-developed literatures in economics: demand estimation incorporating constrained choice sets and learning. The following section describes the relationship between this paper and these literatures and elaborates on how this paper extends these literatures.

### **2.1 Constrained Choice Sets**

A critical assumption in a majority of the work done involving demand estimation is that consumers have full information on all product alternatives in a market. Although this assumption is realistic in some markets, this can lead to biased demand estimates in

markets where the number of product options is large or if a subset of consumers have not been exposed to various product alternatives. A large amount of work in marketing and economics has focused on how to incorporate limited information over product options into demand estimation. Particular interest has been paid to the determinants of these limited choice sets (e.g., advertising) and to what extent the assumption of full choice sets biases demand estimates. The empirical work done can be separated by two differing approaches: the consideration set approach and the estimation of structural consumer search models.<sup>5</sup> Both of these approaches relax the assumption that consumers have full information of all product alternatives and their characteristics, but differ substantially in their methodology. This paper develops and estimates a model that more closely resembles a structural model of consumer search, but a more reduced form consideration set model is discussed in the [Web Appendix](#) as a robustness check.

### 2.1.1 Consideration Set Approach

Incorporating consideration sets into demand estimation began in marketing with Howard and Sheth (1969) and Newell and Simon (1972).<sup>6</sup> The fundamental premise of the consideration set literature is that when confronted with a purchasing decision, consumers follow a two-step procedure. First, they narrow a full set of alternatives into a smaller, consideration set, and second, they make an optimal choice from this narrower set of alternatives.

In economics, Sovinsky Goeree (2008) uses this methodology to estimate the extent to which the assumption of full information in demand estimation biases estimates of product markups and elasticities in the U.S. personal computer industry. Goeree also examines to what extent advertising affects consumers' choice sets. This paper is similar to Goeree's in that it explores a potential determinant of consumers' information over product options: experience acquired through repeated participation in a market.

A common limitation of the consideration set approach is that it treats individuals' choice

---

<sup>5</sup>This is somewhat of a simplification. A few papers in marketing and economics more closely resemble the work done on estimating structural consumer search models than the majority of empirical work done incorporating consideration sets, but still reference the determination of choice sets as consideration set formation (e.g. Mehta et al. 2003; Pires 2013). In addition to these two approaches, there has been related work by Bruno and Vilcassim (2008) and Conlon and Mortimer (2012) focusing on how to incorporate limited product availability into demand estimation. Although these studies deal with constrained choice sets, they assume consumers have full information over all *available* alternatives.

<sup>6</sup>Van Nierop et al. (2010) provides a survey of the marketing literature on consideration set formation.

sets as being exogenously determined; for example, in Goeree’s framework, unobservable preferences for a Dell Inspiron are independent of any unobservables affecting the inclusion of a Dell Inspiron in an individual’s choice set. This paper relaxes this limitation by allowing the unobservables in the consideration phase to be correlated with the unobservables in the product choice phase.

### 2.1.2 Structural Models of Consumer Search

Another way of relaxing the assumption of exogenous choice sets is to estimate a more structural model that specifies a decision rule for how consumers choice sets are formed and estimates the parameters that underly the decision rule. This is the approach taken by the literature on the estimation of structural search models.

Despite the wealth of theoretical work done to explain how search costs or frictions affect equilibrium outcomes, no work had been done to structurally estimate consumers’ search costs until the early 2000’s.<sup>7</sup> The lack of empirical work on the structural side was primarily due to the difficulty of tracking and quantifying consumers’ search behavior. However, the increasing prevalence of internet shopping over the past decade has allowed for detailed data to be collected on consumers’ search behavior; with data on search behavior from various internet retailers and “shopbots”,<sup>8</sup> a number of recent papers have been written examining consumer search and price dispersion in online marketplaces (e.g., Hong and Shum, 2006).

The first work to empirically estimate a distribution of consumer search costs was Sorensen (2001). Sorensen uses store level pharmacy data on prices and quantities to estimate parameters of a search model, but because of the limitations of the data, he must make a number of relatively strong assumptions. Other studies have built upon this work to test different models of search (De los Santos et al., 2012), extend the BLP framework to incorporate endogenous choice sets (Moraga-González et al., 2009), and estimate consumer search costs from observed price dispersion (Hortaçsu and Syverson, 2004; Hong and Shum, 2006).

These models have two primary advantages over incorporating constrained choice sets

---

<sup>7</sup>Early empirical work focused on documenting price dispersion in markets with homogeneous goods such as cars and coal (Stigler, 1961), gasoline (Marvel, 1978), and a variety of other consumer goods (Pratt et al., 1979). Later empirical work focused on using reduced form methods to test results derived from theoretical search models (e.g. Carlson, 1980).

<sup>8</sup>Shopbots are websites devoted to tracking the lowest prices of specific goods from a number of online retailers and then linking consumers to these retailers.



through the more reduced form consideration set approach: they explicitly account for the endogeneity of what individuals' choose (or choose not) to gather information about and are able to incorporate uncertainty. For example, Koulayev (2012) uses data on online searching behavior for hotels to show how the assumption of full information leads to biased results in demand estimation for two reasons: i) because consumers have limited choice sets and ii) because these choice sets are endogenous. In other words, Koulayev shows that, even with knowledge of consumers' choice sets, if the endogenous nature of those choice sets is not accounted for, demand estimation estimates are still biased because consumers choose to gather information (and expand their choice sets) in a non-random manner.

## 2.2 Learning

Previous work on consumer learning has focused on how individuals learn about the specific characteristics of a product (e.g., Erdem and Keane, 1996; Crawford and Shum, 2005). For example, Crawford and Shum (2005), estimate a dynamic model of demand for prescription anti-ulcer medication that incorporates individuals' and doctors' uncertainty about the efficacy of various drugs. They show that initially there is considerable uncertainty over the curative and symptomatic effects of these drugs, but that as individuals repeatedly consume medication, learning takes place relatively quickly.

The majority of studies in the consumer learning literature assume that individuals' beliefs follow some form of Bayesian updating. A notable exception is Ackerberg (2001). Ackerberg examines the differential effects of advertising on first-time and repeat consumers by incorporating learning into a reduced form demand estimation framework. In his model, consumers learn about their preferences for yogurt indirectly through advertisement (informative effects) and directly through consumption; advertisement also may affect repeat consumers through prestige or image effects. This paper is similar to Ackerberg's in the reduced form manner in which learning is accounted for, but I allow for experience to affect individuals' choice sets in addition to the utility they receive from a product.

Two recent papers develop and estimate demand models that incorporate both learning and search: Koulayev (2013) and De los Santos et al. (2012). These papers investigate whether individuals update their beliefs within a search decision. For example, Koulayev uses data on prices and market shares of S&P 500 mutual funds to estimate a model of search that incorporates learning over the price distribution of funds. This paper differs with these

in that it investigates learning across consumers with varying experience levels, rather than focusing on learning within a purchasing decision.

### 3 Empirical Model

As students acquire experience in the college textbook market, their observed online search and purchasing behavior both change. Specifically, higher levels of experience are associated with a higher likelihood of searching online retailers, making an online purchase, and also choosing the outside option. To explore the effect of acquired experience on student textbook purchasing behavior, I develop a discrete choice model in which students make an optimal search and purchasing decision, conditional on an instructor's textbook assignment. Experience enters the model by affecting both the decision to search and the decision of which product alternative to choose, conditional on the search decision. This section first outlines individuals' choice sets and how limited information enters the model. Following this, I characterize an individual's optimal search and purchasing decision. I then describe the form of the alternative-specific utility function. The section concludes with a discussion of the simulated maximum likelihood technique used to estimate the parameters of the utility function and search cost distribution.

#### 3.1 Full Choice Set

For each course in which an individual is enrolled, individual  $i$  is modeled as facing a discrete choice from *at most* 7 alternatives that comprise the full choice set  $\Omega_F$ .<sup>9</sup>

---

<sup>9</sup>In practice, a used (or new) book may not be available from the bookstore at the time of purchase. Some assigned books are also not available from online retailers or offered as rentals from the campus bookstore. This exogenous variation in choice sets aids in identification and is accounted for in estimation.

$$\Omega_F = \left\{ \begin{array}{l} 0 \text{ Do not purchase a textbook (outside option)} \\ 1 \text{ Purchase a new textbook from a bookstore} \\ 2 \text{ Purchase a used textbook from a bookstore} \\ 3 \text{ Rent a textbook from a bookstore} \\ 4 \text{ Purchase a new textbook from an online retailer} \\ 5 \text{ Purchase a used textbook from an online retailer} \\ 6 \text{ Rent a textbook from an online retailer} \end{array} \right.$$

Alternatives are partitioned by type (new, used or rental<sup>10</sup>) and by retailer location (physical bookstore or online retailer). The choice to not purchase a textbook (the outside option) also includes the choice to use a reserve copy at the library, borrow a copy from another student, or purchase or rent a previous edition of the assigned textbook.

### 3.2 Limited Choice Set

A key feature of the college textbook market is that consumers are initially steered towards a common set of purchasing options: those offered at the university bookstore. Because of this, I model consumers as having full information about the bookstore options prior to the purchasing decision.<sup>11</sup> With this information in hand, consumers then make a decision to obtain information about alternatives available from online retailers,  $\Omega_O$ . If individual  $i$  decides not to search online retailers for information about the assigned textbook, her choice set is constrained to  $\Omega_L$ , otherwise she chooses the alternative with the maximal utility in the full choice set,  $\Omega_F$ :<sup>12</sup>

---

<sup>10</sup>Students are able to rent books for the duration of a semester at a lower price than the price to purchase a used book. If the book is not returned, the student is generally forced to pay the difference between the new and rental price. Also, the decision to rent is not separated into the decision to rent a new or used option primarily because the majority of online retailers do not make the quality of the rental option explicit.

<sup>11</sup>It is assumed that each student's choice set contains the bookstore options for a number of reasons: students are directed to the campus bookstore website in order to find the assigned book(s) for their courses, where prices are listed adjacent to the title and ISBN of the assigned book(s), information on the campus bookstore is presented at orientation and students receive periodical emails from the campus bookstore.

<sup>12</sup>The assumption that some individuals do not have the online options in their choice set does not necessarily imply that these individuals are unaware that textbooks can be purchased through online retailers. Instead, it implies that individuals who do not price the assigned textbook at an online retailer do not have

$$\Omega_F \left\{ \begin{array}{l} \Omega_L \left\{ \begin{array}{l} 0 \text{ Do not purchase a textbook (outside option)} \\ 1 \text{ Purchase a new textbook from a bookstore} \\ 2 \text{ Purchase a used textbook from a bookstore} \\ 3 \text{ Rent a textbook from a bookstore} \end{array} \right. \\ \Omega_O \left\{ \begin{array}{l} 4 \text{ Purchase a new textbook from an online retailer} \\ 5 \text{ Purchase a used textbook from an online retailer} \\ 6 \text{ Rent a textbook from an online retailer} \end{array} \right. \end{array} \right.$$

### 3.3 Search Decision

Consumers choose to search online retailers if the expected benefit of search outweighs an individual specific search cost. Let  $u_{ij}$  denote the alternative-specific utility individual  $i$  receives from choice  $j$ . Further, let  $U_{iL}^* = \max\{u_{ij} : j \in \Omega_L\}$ ,  $U_{iO}^* = \max\{u_{ij} : j \in \Omega_O\}$ , and  $U_{iF}^* = \max\{u_{ij} : j \in \Omega_F\}$  denote individual  $i$ 's utility from the alternative yielding the highest utility in the limited, online, and full choice sets respectively. If individuals are allowed to freely revisit the alternatives in the limited choice set (i.e. search with recall), the ex-post benefit of search to individual  $i$  is then,

$$U_{iF}^* - U_{iL}^* \tag{1}$$

or equivalently,

$$U_{iO}^* - U_{iL}^* \text{ if } U_{iO}^* > U_{iL}^* \tag{2}$$

$$0 \text{ if } U_{iO}^* < U_{iL}^* \tag{3}$$

Equation (1) states that the ex-post benefit of search is simply the maximal utility from all alternatives less the maximal utility from only the bookstore alternatives and the outside option (i.e. the limited choice set). Before considering the ex-ante benefit of search, it is helpful to rewrite Equation (1) into the two cases given by Equations (2) and (3). If the maximal utility from the online alternatives is greater than the maximal utility from the bookstore alternatives and the outside option, then the ex-post benefit of search is simply

---

the information necessary to purchase an online alternative and that these alternatives are not in their choice set.

the difference between the two. In the other case where the maximal utility from the online alternatives is less than the alternatives in the limited choice set, then the ex-post benefit of search is zero.

The maximal utility of the limited choice set,  $U_{iL}^*$ , is assumed to be known to the individual before search, but the individual faces uncertainty about characteristics of the online alternatives that search reveals. In particular, individuals face uncertainty about the prices and additional unobservable preference shocks of the online alternatives that is realized only after search.

The expected benefit of search is,

$$\int_{U_{iL}^*}^{\infty} (U_{iO}^*(p_o, \epsilon_o) - U_{iL}^*) dG(U_{iO}^*) \quad (4)$$

$G$  is the cumulative distribution function of individual  $i$ 's beliefs about the maximal utility from the online alternatives.  $U_{iO}^*$  is explicitly written as a function of  $p_o$  and  $\epsilon_o$  – vectors of the prices of the online alternatives and unobservable (to the individual and econometrician) preference shocks – the only random components in the above equation. The lower bound on the expected benefit from search is  $U_{iL}^*$  because the decision to search does not exclude options from an individual's choice set.<sup>13</sup>

Consumers' search costs are assumed to be known at the time of search, which yields the following search decision rule, where an individual searches if the expected benefit outweighs an individual specific search cost,  $c_i$ :

$$\int_{U_{iL}^*}^{\infty} (U_{iO}^* - U_{iL}^*) dG(U_{iO}^*) > c_i \quad (6)$$

Following the search decision, an individual chooses the alternative yielding the highest utility, conditional on that individual's choice set.

In order to estimate the model, assumptions need to be made on an individual's search costs,  $c_i$  and beliefs about the distribution of  $U_{iO}^*$ . Search costs,  $c_i$ , are parametrized as being drawn from a population lognormal distribution:

---

<sup>13</sup>Equation (4) can be rewritten as

$$\int_{p_o} \int_{\epsilon_o} (U_{iO}^*(p_o, \epsilon_o) - U_{iL}^*) \mathbb{1}[U_{iO}^* > U_{iL}^*] dH(\epsilon_o) dF(p_o) \quad (5)$$

where  $p_o$  is the vector of prices of the online alternatives and  $\epsilon_o$  is a vector of online alternative-specific unobservables.

$$c_i \sim \ln\mathcal{N}(\gamma'W_i, \theta^2) \tag{7}$$

Explicitly,  $W_i$  contains,

- Semesters and semesters squared terms.
- An indicator equal to one if individual  $i$  has an older sibling who attended college previously.
- An indicator equal to one if individual  $i$  has made *any* online purchase before.
- Stated preferences for confidence in online shopping.

Each covariate in  $W_i$  except for the semester indicators are assumed to be exogenous search cost shifters and are excluded from the specification of  $u_{ij}$ . For example, an individual who has previously made an online purchase may be more familiar with online shopping or may have a Paypal.com or Amazon.com account, reducing the search costs necessary to purchasing a textbook online. However, this is assumed to not affect how much an individual values a textbook from the internet versus a bookstore conditional on the search decision or the valuation of a textbook at the extensive margin. In other words, having previously made an online purchase is assumed to only affect the process of an individual searching.

Recall that individuals face uncertainty over the prices of the online alternatives (new, used, and rental) and also other unobservable factors realized after search. Unobservable demand shocks are assumed to be distributed Type 1 Extreme Value<sup>14</sup> and observed empirical price distributions are used to reflect consumers' beliefs about the price distribution of the online alternatives (i.e. rational expectations).<sup>15</sup>

---

<sup>14</sup>I hope to relax this assumption in the future by using a Generalized Extreme Value Distribution similar to that in Bresnahan et al. (1997)

<sup>15</sup>If consumers have upward biased beliefs about the distribution of prices from online textbook retailers, this will bias estimated search costs upward and price sensitivities towards zero (upwards biased price beliefs, high search costs, and low price sensitivities could all lead to lower levels of search). Matsumoto and Spence (2013) have found evidence that inexperienced consumers' price expectations are significantly higher than their more experienced counterparts using a novel dataset on consumers' subjective beliefs about the distribution of prices of textbooks from online retailers. Thus, a limitation of this paper is that any differences in beliefs across experience levels will cause search costs to be biased upwards for inexperienced consumers. Although this is a limitation, it does not diminish the counterfactuals of this paper. Instead, it simply does not allow differences in search costs across experience level to be differentiated from differences in price beliefs.

### 3.4 Alternative-Specific Utility

Student  $i$  receives utility from choice  $j$ ,

$$u_{ij} = \beta_j' X_{ij} + \alpha_i p_{ij} + \epsilon_{ij} \quad (8)$$

The utility from choice 0 (not purchasing a textbook) is normalized to  $u_{i0} = \epsilon_{i0}$ . For certain covariates in  $X_{ij}$ ,  $\beta_j = \beta_k \forall j, k \in \{1, 2, \dots, 6\}$ . These restrictions are discussed in the results section.

In order to control for differences across the characteristics specific to each alternative,  $X_{ij}$  contains the following:

- Indicators for new, used, or rental alternatives.
- Indicators for bookstore and online alternatives.
- A vector of stated preferences for quality,  $Z_{1i}$ , interacted with the new indicator.<sup>16</sup>
- A vector of stated preferences for bookstore shopping,  $Z_{2i}$ , interacted with the bookstore indicator.

$X_{ij}$  also contains the following individual, textbook, and course characteristics:

*Individual Characteristics:*

- Semesters and semesters squared terms.<sup>17</sup>
- Indicators for on-campus students, in-state students, and if the course is in the student's major field.

*Textbook Characteristics:*

- Years since revision/publication, number of pages, and textbook dimensions.
- Indicators for guaranteed buyback textbooks,<sup>18</sup> required books (as opposed to recommended), books with bundled materials, hardcover books, and UNC-custom editions or coursepacks.

*Course Characteristics:*

- Average SAT score of individuals in the course.

---

<sup>16</sup>A detailed description of how these stated preferences were collected and measured can be found in Section 4.1.

<sup>17</sup>In future versions of this paper, I plan to experiment with various functional forms of semesters enrolled.

<sup>18</sup>The campus bookstore offers consumers a guarantee that they will buy their textbook at a specified price for a small number of large enrollment courses.

- Total enrollment in the course.
- Indicators for upper level courses, lab courses, honors courses, courses in the spring, and if the instructor informed the individual of purchasing options before the semester began.

The covariates above are included to account for individual, course, and textbook characteristics that could affect purchasing and search behavior and may be correlated with market experience. For example, upper-level students may be observed to search online retailers more because the informal secondary market (e.g., student to student) is different for a typical upper-level course than a typical introductory course. This necessitates the inclusion of controls for upper-level and large enrollment courses.

Experience enters the utility function through the semesters and semesters squared term in  $X_{ij}$ . These terms reflect changes in how students value textbooks as they accrue experience.<sup>19</sup> For example, a first-semester student may have different beliefs before the semester begins about how much a textbook will be used during a semester. As the student progresses through college, she may refine that expectation, leading her to change her valuation of an assigned textbook. This change could also reflect differences in the value of the outside option as a student progresses through school. Consumers may acquire information about reserve copies at the library or the ability to purchase previous editions as they gain experience. If these options are preferable to purchasing/renting a current edition of the assigned textbook, this will be reflected in a change in the value of a textbook at the extensive margin.<sup>20</sup>

An additional covariate in  $X_{ij}$  measures the utility individuals receive from the continuation value of the textbook at the end of the semester.<sup>21</sup> This continuation value is given by the term  $\max\{r_{ij}, \bar{r}_{ij}\}$ . The expected resale value of choice  $j$  to student  $i$  at the end of the semester is given by  $r_{ij}$  (where the expectation is taken at the beginning of the semester), and  $\bar{r}_{ij}$  is student  $i$ 's expected reservation resale price (the lowest price a student expects she would be willing to sell her textbook for at the end of the semester).<sup>22</sup> If  $r_{ij} > \bar{r}_{ij}$

---

<sup>19</sup>Semesters are measured as the number of semesters a student has attended college *before* the semester in which the choice is made (i.e. the minimum value of  $sem_i = 0$ ).

<sup>20</sup>It should be noted that this reduced form measure of learning is not the focus of the paper. Experience enters the utility function due to the observed trend that upper-level students search online retailers more *and* choose the outside option more frequently than first-year students. This can only be rationalized if the outside option is becoming more attractive with experience.

<sup>21</sup>This covariate is excluded from the rental alternative as well as the outside option.

<sup>22</sup>Both  $r_{ij}$  and  $\bar{r}_{ij}$  are elicited through questions in the online questionnaire (i.e., both are data rather



the expected resale price is greater than the expected reservation price and the individual intends to sell the book at the end of the semester. If  $\bar{r}_{ij} > r_{ij}$ , the student expects that the value of the textbook to be greater than the resale price and plans to keep the textbook. Although the *max* operator does not fully characterize a individual's optimization problem with respect to the continuation value of the textbook at the end of the semester – it only depends on the first moments of these random variables – it provides a intuitively reasonable and tractable approximation.

Other terms in the utility function include the log price of choice  $j$ , given by  $p_{ij}$ . A random coefficient measuring price sensitivity is given by  $\alpha_i$ . Explicitly,

$$\alpha_i \sim \ln \mathcal{N}(\alpha_0 + \alpha_1 S_{1i} + \alpha_2 S_{2i}, \sigma^2) \quad (9)$$

This is similar to a traditional random coefficient in price sensitivity, but explicitly accounts for three scenarios: i) the student purchases the textbook with his or her own money, ii) the student receives a fixed payment from another person in order to purchase the textbook, and iii) the student receives a varying payment from another person. An example of the second case is another person or scholarship providing an individual with \$X for textbooks allowing the individual to keep any potentially remaining money. The third case corresponds to the case of another person completely covering the cost of textbooks regardless of the price. In order to allow for these potentially different price sensitivities,  $S_{1i}$  is an indicator variable equal to one if the student received a fixed payment from someone else. An additional indicator variable,  $S_{2i}$ , is equal to one if the student received a varying amount from another person to pay for the textbook. This coefficient is assumed to follow a log-normal distribution to allow for unobservable differences in price sensitivity with population variance  $\sigma^2$ . This flexible specification of  $\alpha_i$  allows individuals to have price sensitivities that vary for both systematic and idiosyncratic reasons.

### 3.5 Reservation Utility Rule

In order to write out the likelihoods of the observed product and search decisions, it is first helpful to characterize the decision rule in Equation (6) in terms of reservation utility in a manner proposed by Koulayev (2012). An individual will choose to search if the alternatives in the limited choice set provide utility below a threshold value,  $\bar{u}_i$ , at which the individual

---

than latent variables).

is indifferent between the decision to search and not to search. Explicitly, an individual will search if,

$$U_{iL}^* < \bar{u}_i \quad (10)$$

where the reservation utility,  $\bar{u}_i$ , is the utility level where the expected benefit from search is equivalent to the cost of search:

$$\int_{\bar{u}_i}^{\infty} (U_{iO}^* - \bar{u}_i) dG(U_{iO}^*) = c_i \quad (11)$$

The LHS of 11 can be approximated using simulation methods for any value of  $\bar{u}_i$  and deterministic portion of  $U_{iO}^*$ . Furthermore, for any value of  $c_i$ ,  $\bar{u}_i$  can be solved for by finding the root of  $\int_{\bar{u}_i}^{\infty} (U_{iO}^* - \bar{u}_i) dG(U_{iO}^*) - c_i$ .<sup>23</sup> Choice probabilities are formed based on these individual specific reservation utilities.

### 3.6 Likelihood Function

The individual likelihoods can be separated into the three following cases:

- The likelihood of an individual not searching and choosing the outside option or a bookstore option,  $j \in \{0, 1, 2, 3\}$ , is

$$Pr_{ij}^{NS} = Pr(u_{ij} > u_{ik}, \forall k \neq j \in \Omega_L) \cdot Pr(U_{iL}^* > \bar{u}_i) \quad (12)$$

- The likelihood of an individual searching and choosing the outside option or a bookstore option,  $j \in \{0, 1, 2, 3\}$  is

$$Pr_{ij}^S = Pr(u_{ij} > u_{ik}, \forall k \neq j \in \Omega_L) \cdot Pr(U_{iL}^* < \bar{u}_i) \cdot Pr(U_{iO}^* < U_{iL}^* | U_{iL}^* < \bar{u}_i) \quad (13)$$

- The likelihood of an individual searching and choosing an online or rental option,  $j \in \{4, 5, 6\}$  is

$$Pr_{ij}^S = Pr(u_{ij} > u_{ik}, \forall k \neq j \in \Omega_O) \cdot Pr(U_{iL}^* < \bar{u}_i) \cdot Pr(U_{iO}^* > U_{iL}^* | U_{iL}^* < \bar{u}_i) \quad (14)$$

where  $\Omega_O$  is the choice set composed of only the online and rental options.

Equation (12) states that the probability of an individual not searching and choosing either the outside option or a bookstore option is the choice probability conditional on

---

<sup>23</sup>A detailed algorithm for how  $\bar{u}_i$  is calculated is available in the [Web Appendix](#).

choosing from the limited choice set times the probability that the maximal utility from the limited choice set is greater than the individual's reservation utility.

Equations (13) and (14) include an additional term because the product choice conditional on choosing to search provides information about both  $U_{iL}^*$  and  $U_{iO}^*$ . The second term in equation (13) indicates that individual searched because the maximal element in  $U_{iL}^*$  was less than  $\bar{u}_i$ . Because the individual chose an element from  $\Omega_L$ , it must be the case that  $U_{iL}^* > U_{iO}^*$  (the third term in Equation (13)). The first term is simply the probability the individual chose alternative  $j$  in the limited choice set.

The first terms of Equations (12) - (14) have a closed form expression. The second terms of these equations have a closed form representation conditional on a value of  $\bar{u}_i$ . The third term of Equations (13) and (14) must be approximated using simulation methods.

Using these choice probabilities, the parameters of the alternative specific utility function and search cost distribution, collectively  $\Theta$ , can be estimated using simulated maximum likelihood estimation. The unconditional individual likelihood function is given by,<sup>24</sup>

$$\mathbb{L}_i(\Theta) = \int_c \int_{p_o} \left[ \prod_j (P_{ij}^S)^{d_{Sij}} (P_{ij}^{NS})^{d_{NSij}} \right] dH(c) dK(p_o) \quad (15)$$

for  $j \in \{0, \dots, 6\}$ . The distribution of search costs is given by the function  $H$ . The distribution of online prices is given by the function  $K$ . Individuals are assumed to know the empirical distribution of online prices (i.e., rational expectations). The indicator function,  $d_{Sij}$  ( $d_{NSij}$ ), is equal to one if the individual chooses to search (not search) and chooses alternative  $j$ . This individual likelihood can be approximated using simulation techniques. Each observation is treated as independent; the full likelihood function used in estimation is given by,

$$\mathbb{L}(\Theta) = \prod_i \mathbb{L}_i(\Theta) \quad (16)$$

The [Web Appendix](#) provides more detailed information on the form of the above choice probabilities, how the probabilities without closed forms are approximated, and the form of the likelihood function.

---

<sup>24</sup>The probability of not searching and choosing an online alternative,  $P_{ij}^{NS}$  for  $j \in \{4, 5, 6\}$ , is zero.

## 4 Data

### 4.1 Online Questionnaire Data

Micro-data were collected through online questionnaires sent from instructors to their students.<sup>25</sup> At the beginning of the fall and spring semesters for the 2011-2012 and 2012-2013 academic calendar years, instructors were sent an email containing survey directions and a link to an online survey. This email was then forwarded by the instructor to the students in his or her class. Each survey was individualized to the specific course; students were asked if and how they obtained the assigned textbook for the course, price information, which retailers they considered, details of their online search behavior, their price expectations of retailers not considered, past purchasing behavior, characteristics of the textbook they (may have) purchased, and questions gauging preferences for bookstore shopping and quality among other items.<sup>26</sup>

For the 2011-2012 and 2012-2013 academic calendar year, instructors of more than 700 course-sections forwarded an email containing an online survey link to their students.<sup>27</sup> This resulted in 8,279 students responding. Of the students who began the survey, 7,191 completed the survey: an 87% completion rate.<sup>28</sup>

The summary statistics that follow exclude non-undergraduate students (248 excluded), respondents whose textbooks were available for free through the UNC student stores because of scholarships offered through the university (315 excluded), or did not have information for other key variables (595 excluded).<sup>29</sup> These individuals are also excluded in estimation, leaving a sample of 6,033 students.

Table 1 summarizes the distribution of respondents' number of semesters enrolled. The number given in Table 1 includes the current semester of the respondent, as well as summer

---

<sup>25</sup>UNC IRB study #11-1177. The online surveys are conducted through the Odum Institute.

<sup>26</sup>The [Web Appendix](#) contains a detailed description of the online questionnaire.

<sup>27</sup>These surveys were distributed to instructors following the last day students were able to return textbooks to the campus bookstore for a full refund; this is roughly two weeks after the start of the semester. In an effort to bolster student response rates, respondents were entered in a drawing to win cash prizes.

<sup>28</sup>1,892 students responded in the fall of 2011, 2,493 students responded in the spring of 2012, 2,189 students responded in the fall of 2012, and 1,705 students responded in the spring of 2013.

<sup>29</sup>This last number is higher than it will be in the final version of the paper. Individuals who choose an "other" option for key questions are dropped in this version, but because they were prompted to write in responses, imputations can be made for the choices they made.

sessions and semesters enrolled at any other university.<sup>30</sup>

Table 1: Semesters Enrolled

| Semesters | Frequency | Percent | Cumulative |
|-----------|-----------|---------|------------|
| 1         | 821       | 13.6    | 13.6       |
| 2         | 789       | 13.1    | 26.7       |
| 3         | 614       | 10.2    | 36.9       |
| 4         | 782       | 13.0    | 49.8       |
| 5         | 681       | 11.3    | 61.1       |
| 6         | 705       | 11.7    | 72.8       |
| 7         | 617       | 10.2    | 83.0       |
| 8         | 552       | 09.1    | 92.2       |
| 9+        | 472       | 07.8    | 100        |
| Total     | 6,033     | -       | -          |

Table 2 summarizes the demographics of the respondents. This includes the percentage of students who were classified by the university as in-state students, live on campus, male, and were the first child in their immediate family to attend college. This table also includes the age of respondents and the percentage of students who reported that the course in which they were surveyed was in their major field. The total number of responses varies because response to these questions was optional.

Table 2: Student Demographics

|                       | N    | Mean | Std. Dev. | Min | Max |
|-----------------------|------|------|-----------|-----|-----|
| In-State              | 6028 | 0.79 | -         | 0   | 1   |
| On-Campus             | 6026 | 0.60 | -         | 0   | 1   |
| Male                  | 6022 | 0.38 | -         | 0   | 1   |
| First in College      | 6033 | 0.54 | -         | 0   | 1   |
| Course in Major Field | 6033 | 0.59 | -         | 0   | 1   |
| Age                   | 5978 | 19.9 | 2.1       | 16  | 61  |

<sup>30</sup>Data from the UNC registrar will be used to determine the representativeness of this sample. Details are available in the [Web Appendix](#)

Table 3 summarizes respondents' textbook purchasing decisions; students are separated into two categories: freshmen (defined as enrolled in two or fewer semesters) and upperclassmen. For freshmen, 84% of students decided to purchase the assigned textbook, 5.4% chose to rent the assigned textbook, 5.8% borrowed the textbook from another student, and a small percentage chose to use a copy from the library or to not use the textbook entirely. Upperclassmen tended to purchase the book less frequently, rent more frequently, and chose the other options at roughly the same rate.

Table 3: Students' Beginning of Semester Textbook Decisions

| Option                       | Freshmen  |         | Upperclassmen |         |
|------------------------------|-----------|---------|---------------|---------|
|                              | Frequency | Percent | Frequency     | Percent |
| Purchased the Textbook       | 1,353     | 84.0    | 3,330         | 75.3    |
| Rented the Textbook          | 87        | 05.4    | 450           | 10.2    |
| Borrowed the Textbook        | 94        | 05.8    | 262           | 05.9    |
| Used a Copy from the Library | 8         | 00.5    | 101           | 02.3    |
| Did not use the Textbook     | 68        | 04.2    | 280           | 06.3    |

Table 4 summarizes students' choice of retailers if they purchased the assigned textbook. Observations for courses that assigned textbooks which were not available at online retailers (UNC Coursepacks or Custom Editions) are excluded from this table. Conditional on being able to purchase the assigned textbook through an online retailer, the data show that upperclassmen were more likely to choose to purchase the assigned textbook through an online retailer than freshmen. Perhaps surprisingly, freshmen were more likely to purchase the assigned textbook from another student than upperclassmen. These purchases occurred almost exclusively in the spring semester and appear to be a result of the type of course the freshmen were taking (the supply of textbooks for introductory courses through other students is larger than for more specialized, core courses).

Table 4: Students' Retailer Choices

| Retailer                  | Freshmen  |         | Upperclassmen |         |
|---------------------------|-----------|---------|---------------|---------|
|                           | Frequency | Percent | Frequency     | Percent |
| UNC Student Stores        | 485       | 53.1    | 1,172         | 39.0    |
| Another College Bookstore | 4         | 00.4    | 61            | 02.0    |
| Online Retailer           | 359       | 39.3    | 1,641         | 54.7    |
| Another Student           | 65        | 07.1    | 127           | 04.2    |

Table 5 summarizes students choice of online retailer, conditional on purchasing or renting online. The majority of students purchased their textbook through Amazon.com, accounting for more than seven times the purchases of the next leading online retailer, Half.com. Roughly half of all rentals were done through Chegg.com (although this still represents a small proportion of the total number of online purchases and rentals). HTML source codes were collected for eight additional online retailers and two publishing websites, which comprised an additional 8.6% of sales and rentals. Price data were not gathered for 6.8% of the observed sales and rentals (e.g. alibris.com, bookbyte.com).

Table 5: Students' Online Retailer Choices - Purchased or Rented

| Choice of Online Retailer  | Frequency | Percent |
|----------------------------|-----------|---------|
| Amazon                     | 1,613     | 66.9    |
| Half (Ebay)                | 226       | 09.4    |
| Chegg                      | 203       | 08.4    |
| Other (Data Collected)     | 208       | 08.6    |
| Other (Data Not Collected) | 163       | 06.8    |

Table 6 contains information on whether students knew about the possibility of renting textbooks at the beginning of the semester and whether they priced the textbook at another physical bookstore, or an online retailer. Evidence that students learn about the availability of product options is seen in students' knowledge of rental options at the beginning of the semester; the vast majority of upperclassmen surveyed were knowledgeable of rental options when making their textbook decision, while a sizable fraction of freshmen were not.

This table also includes the proportion of students who had ever purchased *anything*

online or ever made an online sale. Nearly all students, freshmen and upperclassmen, had made an online purchase before, but less than half had ever made an online sale before.

Table 6: Students’ Purchasing Behavior

|                              | Freshmen<br>Percent | Upperclassmen<br>Percent |
|------------------------------|---------------------|--------------------------|
| Knowledge of Rental Options  | 74.9                | 93.1                     |
| Priced an Online Retailer    | 61.0                | 74.0                     |
| Ever Made an Online Purchase | 95.8                | 96.7                     |
| Ever Made an Online Sale     | 33.4                | 44.6                     |

Table 7 summarizes students’ stated preferences for aspects important to the decision to buy a new or used textbook and to deciding whether to buy from a physical bookstore or an online retailer.<sup>31</sup>

Table 7: Students’ Stated Preferences

| Stated Preferences                            | N    | Mean | Std. Dev. | Min | Max |
|---|------|------|-----------|-----|-----|
| New Rather than a Used Textbook               | 6033 | 1.77 | 2.25      | 0   | 10  |
| Textbook without Notes, etc.                  | 6033 | 4.00 | 3.09      | 0   | 10  |
| Seeing the Textbook Before Purchasing         | 6033 | 3.61 | 2.81      | 0   | 10  |
| Receiving the Textbook on the Day of Purchase | 6033 | 3.20 | 2.55      | 0   | 10  |
| Confidence in Online Shopping                 | 6033 | 7.36 | 1.81      | 0   | 10  |

These stated preferences include responses to the following questions important to the bookstore/online retailer decision:

- How important is seeing the condition of a textbook before purchasing it?
- How important is receiving a textbook on the day you purchase it?

<sup>31</sup>Respondents used a sliding scale to report their preferences, with response values ranging from 0 to 10. For example, for the question “How important to you is buying a new rather than a used textbook?”, the boundary at zero is labeled, “I am completely indifferent between buying a new and a used textbook.” The boundary at a response of ten is labeled, “I will only buy a new textbook.” In order to alleviate measurement error decimal responses to the nearest tenth were allowed. See the [Web Appendix](#) for a screenshot of these questions.



- How confident are you that you will be satisfied with an online purchase?

Questions gauging the importance of buying a new rather than a used textbook include the following:

- How important is purchasing a new rather than a used textbook?
- How important is purchasing a textbook without highlighting, dog-ears, or notes already in it?

Students reported placing low importance on purchasing a new rather than a used textbook. Slightly more importance was placed on purchasing a textbook without notes or highlighting in it, seeing the textbook before purchasing it, and receiving the textbook on the day of purchase. The majority of students reported being confident in being satisfied with an online purchase.

Table 8 reports the correlation between stated preferences.<sup>32</sup> Correlations between stated preferences are of the expected magnitudes and signs. Preferences for purchasing a new rather than a used textbook are positively correlated with preferences for purchasing a textbook with no notes or highlighting. These preferences are also correlated with seeing the book before purchasing it, but to a lesser degree. Preferences important for the bookstore/online retailer decision, the importance of receiving the textbook on the same day and seeing the textbook before purchasing it are also highly correlated.

As expected, confidence in online shopping is negatively correlated with preferences for seeing the textbook and receiving the textbook on the same day. Interestingly, freshmen seem to have stronger preferences for aspects of textbook purchasing that would lead to more new rather than used purchases and more bookstore than online purchases; this lends support to the need to include these measures when examining the effects of experience on purchasing behavior.

---

<sup>32</sup>P-values are indicated in parentheses.

Table 8: Correlation in Students' Stated Preferences

|                 | New vs.<br>Used | No<br>Notes     | Seeing<br>the Book | Same<br>Day     | Online<br>Confidence | Fresh.<br>Dummy |
|-----------------|-----------------|-----------------|--------------------|-----------------|----------------------|-----------------|
| New vs Used     | 1.00            | -               | -                  | -               | -                    | -               |
| No Notes        | 0.51<br>(0.00)  | 1.00            | -                  | -               | -                    | -               |
| Seeing the Book | 0.41<br>(0.00)  | 0.44<br>(0.00)  | 1.00               | -               | -                    | -               |
| Same Day        | 0.23<br>(0.00)  | 0.20<br>(0.00)  | 0.41<br>(0.00)     | 1.00            | -                    | -               |
| Online Conf.    | -0.06<br>(0.00) | -0.05<br>(0.00) | -0.20<br>(0.00)    | -0.20<br>(0.00) | 1.00                 | -               |
| Fresh. Dum.     | 0.05<br>(0.00)  | 0.07<br>(0.00)  | 0.11<br>(0.00)     | 0.04<br>(0.00)  | -0.09<br>(0.00)      | 1.00            |

Table 9 summarizes who paid for the assigned textbook. Less than half of all students paid for the assigned textbook out of pocket, with slightly more upperclassmen paying for the textbook than freshmen. Of the students who received money from another person to pay for the textbook, the majority received a payment that varied with the cost of the textbook (i.e. parents paid for the entire cost of textbooks, regardless of cost). A much smaller proportion received a fixed payment from another person to cover the cost of the textbook.

Table 9: Distribution of Payments for Textbooks

|                                  | Freshmen  |         | Upperclassmen |         |
|----------------------------------|-----------|---------|---------------|---------|
|                                  | Frequency | Percent | Frequency     | Percent |
| Self                             | 446       | 31.4    | 1,677         | 44.9    |
| Another person - Varying Payment | 853       | 60.0    | 1,785         | 47.7    |
| Another person - Fixed Payment   | 48        | 03.4    | 86            | 02.3    |
| Scholarship                      | 74        | 05.2    | 191           | 05.1    |

## 5 Results

This section first presents parameter estimates for the alternative specific utility function and search cost distribution. This is followed by a discussion of estimated elasticities. I conclude this section with a discussion of model fit and counterfactual results.

### 5.1 Parameter Estimates

Recall that search costs are distributed log-normal according to the following,

$$c_i \sim \ln\mathcal{N}(\gamma'W_i, \theta^2) \quad (17)$$

Table 10 presents estimates for the parameters that determine the mean and variance of the search cost distribution. More experienced consumers have significantly lower search costs, although the marginal reduction in search costs from an additional semester is decreasing. Consumers who have made a previous online purchase or have higher levels of stated confidence in online shopping also have significantly lower search costs. The parameters given below imply that the median search cost for a first year student is \$13.33 and \$1.34 for an upper-level student.<sup>33</sup>

Table 10: Search Cost Estimates

| Covariate             | Point Estimate | Standard Error |
|-----------------------|----------------|----------------|
| Constant              | 10.890         | (2.632)        |
| Online Confidence     | -0.894         | (.227)         |
| First in College      | -0.199         | (.146)         |
| Prev. Online Purchase | -5.478         | (1.54)         |
| Semesters             | -0.937         | (.078)         |
| Semesters Squared     | 0.0636         | (.002)         |
| $\theta^2$            | 5.34           | (.125)         |

Notes: N = 6033, Simulation Draws = 50

Bootstrapped Standard Errors in Parentheses: 250 Replications

<sup>33</sup>The variance of the search cost distribution is very high. Although median search costs are reasonable, mean search costs are relatively large. I plan to experiment with other distributions with positive support in the future (e.g., Gamma).

Table 11 presents the parameter estimates for the alternative-specific utility function. Consumers value textbooks less as they progress through school, for courses in their major field, and courses in the spring. Stated preferences are associated with greater valuation of new textbooks. Consumers paying for their textbook out of pocket are more price sensitive than consumers receiving fixed or varying payments.

Table 11: Utility Parameter Estimates

| Covariate                  | Point Estimate | Standard Error |
|----------------------------|----------------|----------------|
| Semesters                  | -0.286         | (.002)         |
| Semesters Squared          | 0.016          | (.000)         |
| Course in Major            | -0.106         | (.019)         |
| Spring Indicator           | -0.740         | (.017)         |
| Preference for New         | 0.087          | (.005)         |
| Preference for Quality     | 0.158          | (.005)         |
| $\alpha$ - Out of Pocket   | -0.560         | (.010)         |
| $\alpha$ - Fixed Payment   | -0.204         | (.031)         |
| $\alpha$ - Varying Payment | -0.309         | (.007)         |
| $\alpha$ - Scholarship     | -0.964         | (.008)         |
| Continuation Value         | 0.234          | (.040)         |

Notes: N = 6033, Simulation Draws = 50  
 Bootstrapped Standard Errors in Parentheses: 250 Replications

## 5.2 Elasticities

Table 12 presents simulated price elasticities. Demand for the bookstore alternatives is inelastic, with upper-level students (parentheses) having slightly more elastic demand than first-year students (brackets). This implies that bookstore profits would unambiguously increase from an increase in prices. This counterintuitive result is likely due to features of the textbook market. The first feature is that the campus bookstore is not necessarily profit maximizing. Bookstore residuals are used to support non-athletic scholarship and fellowship programs of the university.<sup>34</sup>

<sup>34</sup>In addition to this, there is already public sentiment against the rising costs of textbooks at the campus bookstore. In 2006, a UNC subcommittee was formed to review the rising costs of textbooks. It's possible

Demand is also inelastic for the online alternatives. This implies that if all online retailers and marketplace sellers raised their prices, overall profits would be increased. However, the price elasticity of demand for an individual seller on an online marketplace is likely much more elastic (relative to the entire market).

Because first-year students have higher search costs, they substitute into bookstore alternatives more frequently than upper-year students, while upper-level students are more likely to substitute into online alternatives. First year students are more likely to substitute into the outside option if bookstore prices increase, while the upper-year students are more likely to substitute into the outside option if the price of online alternatives increase.

---

that the bookstore does not charge the static profit maximizing price to reduce future objections about rising textbook prices (and potential negative publicity).

Table 12: Elasticity Estimates

|                  | Outside<br>Option | New/<br>Bkstore | Used/<br>Bkstore | New/<br>Online | Used/<br>Online | Rental  |
|------------------|-------------------|-----------------|------------------|----------------|-----------------|---------|
| Outside Option   | .                 | 0.09            | 0.08             | 0.06           | 0.09            | 0.04    |
|                  | .                 | [0.15]          | [0.10]           | [0.05]         | [0.08]          | [0.02]  |
|                  | .                 | (0.07)          | (0.07)           | (0.06)         | (0.09)          | (0.04)  |
| New / Bookstore  | .                 | -0.30           | 0.08             | 0.05           | 0.06            | 0.03    |
|                  | .                 | [-0.27]         | [0.10]           | [0.05]         | [0.05]          | [0.02]  |
|                  | .                 | (-0.31)         | (0.07)           | (0.05)         | (0.06)          | (0.03)  |
| Used / Bookstore | .                 | 0.08            | -0.29            | 0.04           | 0.06            | 0.03    |
|                  | .                 | [0.09]          | [-0.28]          | [0.04]         | [0.04]          | [0.02]  |
|                  | .                 | (0.06)          | (-0.30)          | (0.04)         | (0.08)          | (0.04)  |
| New / Online     | .                 | 0.06            | 0.05             | -0.33          | 0.11            | 0.05    |
|                  | .                 | [0.06]          | [0.08]           | [-0.31]        | [0.11]          | [0.02]  |
|                  | .                 | (0.06)          | (0.05)           | (-0.35)        | (0.11)          | (0.05)  |
| Used / Online    | .                 | 0.05            | 0.05             | 0.08           | -0.32           | 0.05    |
|                  | .                 | [0.06]          | [0.05]           | [0.07]         | [-0.29]         | [0.04]  |
|                  | .                 | (0.05)          | (0.05)           | (0.08)         | (-0.33)         | (0.06)  |
| Rental           | .                 | 0.06            | 0.07             | 0.07           | 0.15            | -0.41   |
|                  | .                 | [0.07]          | [0.07]           | [0.05]         | [0.12]          | [-0.35] |
|                  | .                 | (0.06)          | (0.07)           | (0.07)         | (0.15)          | (-0.44) |

Notes: N = 4234. Excludes textbook obs. not offered online and new/bkstr. price < \$40. Unbracketed results denote percentages for first-year and upper-level students. Brackets denote first-year students. Parentheses denote upper-level students. Elasticities are calculated by increasing the price of the column alternative by 10%.

### 5.3 Counterfactuals and Model Fit

The first two columns of Table 13 report observed choice probabilities and simulated choice probabilities. The model does a fairly good job of matching observed choice probabilities. All simulated choice probabilities are within three percentage points of the observed choice probabilities. The last row of Table 13 reports the actual percentage of individuals who search and the simulated percentage of individuals who search. Again, the model fits well,

only slightly under-predicting the amount of search for upper-level students.

The third column of Table 13 reports the simulated choice probabilities for a counterfactual where all consumers' search costs are drawn from the same distribution as first semester consumers (while holding their other observable characteristics constant). The fourth column presents the percentage change from column two to column three. Assigning upper-level students search costs for inexperienced consumers reduces online choices by roughly twenty percent, search by seventeen percent, and leads to an average welfare loss of \$5.15 for upper-level students.

The fifth and sixth columns are similar to the third and fourth columns, but correspond to a counterfactual in which all consumers' search costs are drawn from the same distribution as eighth semester students. In this counterfactual, inexperienced consumers decrease the frequency with which they choose the outside option by sixteen percent, reduce bookstore purchases by approximately sixteen percent, and increase online purchases/rentals by sixteen to nineteen percent. Search increases by twenty percent and the average welfare gain is \$8.49 for inexperienced consumers.

Table 13: Search Cost Counterfactuals and Model Fit

|                      | Actual<br>Percentage     | Baseline<br>Prediction (%) | 1 <sup>st</sup> Semester<br>Costs (%) | Percent<br>Change          | 8 <sup>th</sup> Semester<br>Costs (%) | Percent<br>Change         |
|----------------------|--------------------------|----------------------------|---------------------------------------|----------------------------|---------------------------------------|---------------------------|
| Outside<br>Option    | 17.2<br>[13.5]<br>(18.3) | 16.3<br>[11.0]<br>(17.8)   | 18.2<br>[11.4]<br>(20.1)              | 11.4<br>[3.4]<br>(12.9)    | 15.7<br>[9.3]<br>(17.5)               | -3.8<br>[-15.9]<br>(-1.5) |
| New/<br>Bookstore    | 19.4<br>[26.1]<br>(17.6) | 18.6<br>[24.1]<br>(17.1)   | 20.5<br>[24.6]<br>(19.4)              | 10.0<br>[2.2]<br>(13.3)    | 17.4<br>[19.8]<br>(16.8)              | -6.5<br>[-17.8]<br>(-1.6) |
| Used/<br>Bookstore   | 15.5<br>[17.6]<br>(13.9) | 18.7<br>[22.2]<br>(17.7)   | 20.6<br>[22.7]<br>(20.0)              | 10.1<br>[2.2]<br>(13.1)    | 17.7<br>[18.7]<br>(17.4)              | -5.6<br>[-15.6]<br>(-1.7) |
| New/<br>Online       | 12.1<br>[11.8]<br>(12.2) | 15.0<br>[14.5]<br>(15.2)   | 12.7<br>[13.9]<br>(12.2)              | -15.4<br>[-3.5]<br>(-19.8) | 15.8<br>[16.7]<br>(15.5)              | 5.5<br>[15.3]<br>(1.7)    |
| Used/<br>Online      | 24.3<br>[19.1]<br>(25.8) | 22.2<br>[20.1]<br>(22.8)   | 18.7<br>[19.4]<br>(18.2)              | -15.8<br>[-3.6]<br>(-20.3) | 23.3<br>[23.4]<br>(23.2)              | 5.5<br>[16.2]<br>(1.7)    |
| Rental               | 11.4<br>[8.2]<br>(12.3)  | 9.1<br>[8.2]<br>(9.4)      | 7.7<br>[7.9]<br>(7.5)                 | -15.6<br>[-3.1]<br>(-20.0) | 9.6<br>[9.5]<br>(9.6)                 | 5.6<br>[16.3]<br>(1.7)    |
| Search<br>Percentage | 73.5<br>[63.8]<br>(76.3) | 70.7<br>[63.3]<br>(72.8)   | 59.2<br>[61.1]<br>(58.1)              | -14.2<br>[-3.5]<br>(-17.5) | 74.9<br>[75.6]<br>(74.1)              | 5.9<br>[19.5]<br>(1.8)    |

Notes: N = 4234. Excludes individuals with textbooks not offered online and new bookstore price < \$40. Unbracketed results denote percentages for first-year and upper-level students. Brackets denote first-year students. Parentheses denote upper-level students.



## 6 Conclusion

This paper finds significant differences in consumers' search and purchasing behavior across levels of market experience. After controlling for consumer, course, and textbook characteristics that are correlated with experience, consumers' decision making is still markedly different across experience levels due to changes in consumers' expected value of search. These results imply that policies that provide information to consumers about characteristics of the online textbook market would yield significant benefits to inexperienced consumers. Moreover, these results carry additional weight in markets with positive externalities. I find that information leads to more search which induces market participation for marginal consumers who were not willing to pay the pre-search price. In markets with positive externalities, such as textbooks (education) and alternative energy (pollution reduction), informational policies that alleviate uncertainty about market-level characteristics are a fiscally responsible way of increasing market participation, and hence welfare.

This paper takes advantage of an exogenous measure of experience in the college textbook market (semesters enrolled), but there are tradeoffs associated with using this measure. I am able to capture average changes across consumers as gain experience in the market, but am not able to identify the precise signals that consumers are receiving. There are a number of potential signals that consumers receive in this market: past searches or interactions with other market participants (peer effects) are two possible sources. This paper is further limited in its identification of the characteristics of the market that consumers are learning about. In a separate paper with Brett Matsumoto, we find that inexperienced consumers have biased expectations about the price of textbooks in the online market, but that these expectations tend to converge to the empirical mean with experience (Matsumoto and Spence, 2013).

A natural followup question to this paper is whether firms are able to extract surplus from inexperienced consumers in the form of price discrimination or some other mechanism (e.g., revisionary behavior). Understanding firm responses to heterogeneity in consumer experience is important for policies related to both improving consumer outcomes and developing anti-trust policy. In markets such as durable goods markets, a large proportion of consumers will have made relatively few purchases. If inexperienced consumers are less likely to search unfamiliar or new retailers, than this represents a barrier to entry for new firms. These information frictions may therefore warrant stricter merger policy in markets where consumers tend to be inexperienced.

## References

- Akerberg, D. (2001). Empirically Distinguishing Informative and Prestige Effects of Advertising. *RAND Journal of Economics*, 32(2):316–333.
- Akerberg, D. (2003). Advertising, Learning, and Consumer Choice in Experience Good Markets: an Empirical Examination. *International Economic Review*, 44(3):1007–1040.
- Bresnahan, T., Stern, S., and Trajtenberg, M. (1997). Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the Late 1980s. *RAND Journal of Economics*, 28(0):S17–S44.
- Bruno, H. and Vilcassim, W. (2008). Structural Demand Estimation with Varying Product Availability. *Marketing Science*, 33(2).
- Carlson, J. and Pescatrice, D. (1980). Persistent Price Distributions. *Journal of Economics and Business*, 33(1):21–27.
- Conlon, C. and Mortimer, J. (2012). Demand Estimation Under Incomplete Product Availability. *Working Paper*.
- Crawford, G. and Shum, M. (2005). Uncertainty and Learning in Pharmaceutical Demand. *Econometrica*, 173(4):1137–1173.
- De los Santos, B. (2008). Consumer Search on the Internet. *Working Paper*.
- De los Santos, B., Hortacsu, A., and Wildenbeest, M. (2012). Testing Models of Consumer Search using Data on Web Browsing and Purchasing Behavior. *American Economic Review*, 102(6).
- Erdem, T. and Keane, M. (1996). Decision-making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets. *Marketing Science*, 15(1):1–20.
- Hong, G. and Shum, M. (2006). Using Price Distributions to Estimate Search Costs. *RAND Journal of Economics*, 37(2):257–275.

- Hortaçsu, A. and Syverson, C. (2004). Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds. *Quarterly Journal of Economics*, 119(2):403–456.
- Howard, J. and Sheth, J. (1969). *The Theory of Buyer Behavior*. John Wiley and Sons, New York.
- Koulayev, S. (2012). Estimating Demand in Search Markets: The Case of Online Hotel Bookings. *Federal Reserve Bank of Boston Working Paper No. 09-16*.
- Koulayev, S. (2013). Search with Dirichlet Priors: Estimation and Implications for Consumer Demand. *Journal of Business and Economic Statistics*.
- Marvel, H. (1978). Competition and Price Levels in the Retail Gasoline Market. *Review of Economics and Statistics*, 60(2):252–258.
- Matsumoto, B. and Spence, F. (2013). Price Beliefs and Experience: Do Consumers' Beliefs Converge to Empirical Distributions with Repeated Purchases? *Working Paper*.
- Mehta, N., Rajiv, S., and Srinivasan, K. (2003). Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation. *Marketing Science*, 22(1):58–84.
- Moraga-González, J., Sándor, Z., and Wildenbeest, M. (2009). Consumer Search and Prices in the Automobile Market. *Working Paper*.
- Newell, A. and Simon, H. (1972). *Human Problem Solving*. Prentice Hall, Englewood Cliffs, NJ.
- Pires, T. (2013). Consideration Sets in Storable Goods Markets. *Working Paper*.
- Pratt, J., Wise, D., and Zeckhauser, R. (1979). Price Differences in almost Competitive Markets. *Quarterly Journal of Economics*, 93(2):189–211.
- Sorensen, A. (2001). An Empirical Model of Heterogeneous Consumer Search for Retail Prescription Drugs. *NBER Working Paper 8548*.
- Sovinsky Goeree, M. (2008). Limited Information and Advertising in the U.S. Personal Computer Industry. *Econometrica*, 76(5):1017–1074.

Stigler, G. (1961). The Economics of Information. *Journal of Political Economy*, 69(3):213–225.

Van Nierop, E., Bronnenberg, B., Paap, R., Wedel, M., and Franses, P. (2010). Retrieving Unobserved Consideration Sets from Household Panel Data. *Journal of Marketing Research*, XLVII:63–74.