Price Partitioning and Consumer Rationality in Internet Retail Markets

Katharine Donovan Bodnar

Professor Andrew Sweeting, Faculty Advisor

Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University.

Duke University

Durham, North Carolina

2012

Acknowledgements

I would like to thank Professor Andrew Sweeting, without whom this project would not have been possible. I would also like to thank Professor Kent Kimbrough for his continued support and guidance on my work throughout the year. I would like to thank Chung-Ying Lee for his help with understanding the data set. And I would like to thank Fu Ouyang for his assistance with STATA.

Abstract

This paper seeks to further understand the bounds of consumer rationality and search on the Internet. Specifically this paper focuses on how consumers respond to partitioned prices when making their purchasing decisions. The goal of the paper is to determine if consumers are as sensitive to explicitly stated shipping prices, as they are to list prices, in an environment where items are sorted by list prices. After evaluating the data using a non-linear regression model, the results suggest that consumers do not weight partitioned prices (taxes or shipping prices) as much as they do list prices, contradicting the standard economic model about consumer rationality. The results imply that price partitioning is an effective obfuscation method that is allowing retailers to continue to maintain mark-ups and profit margins in Internet settings.

JEL classification: L1, L11, L81
Keywords: Shipping price, search, obfuscation, E-commerce, price partitioning, retail competition

1. Introduction

Since it's invention, the Internet has been a constant topic of media attention, public debate, and academic interest. The Internet allowed for the advent of E-Commerce, where consumers are able to purchase items online and have them shipped to them. E-Commerce has grown at around 20% annually for the past decade and has greatly affected the dynamics and competition in many different markets and industries. E-Commerce has taken an increasingly larger share of total retail commerce every quarter, reaching nearly 5% of total US commerce by the end of 2011 (US Department of Commerce 2012). In 2011, holiday shoppers planned to make over 30% of their purchases online. E-retail on Cyber Monday grew 22% over the 2010 to reach $1.25billion compared to a 5% growth in traditional retail on Black Friday (Tuttle 2012).

E-Commerce has grown extensively in a very short period of time leaving much unknown about the way consumers and retailers interact in this novel market setting. Many have asserted that the Internet facilitates the development of more competitive markets due to low search costs, low transaction costs and the large selection of buyers and sellers (Bakos 1997). Yet, Baye, Morgan and Sholten (2004) found that persistent price dispersion has decreased somewhat but not greatly. As technology has evolved to reduce consumers' search costs and provide more complete information, firms are finding methods to engage in obfuscation strategies to keep search costs high. These strategies have been shown to effectively keep mark-ups on retail goods at 10-15% (Ellison and Ellison 2008). This work relates to the broader topic of consumer information processing and how consumers weight the information they receive.

It is common practice in retail markets, traditional and Internet, for retailers to separately list the components of total price for an item. Often they will list the price for an item and separately list shipping prices, installation fees, insurance, and taxes. Most retailers online or on TV will advertise a price and then a separate shipping price will be listed on a later page or in fine print. In theory, the practice of dividing a price into these two components should have little effect on overall demand for a good. A perfectly informed and fully rational consumer will merely add together the two parts of a price to obtain the total out of pocket price for an item and then determine whether or not to buy based on this total price and their utility function. Yet many online search engines operate against this principle: items are ordered by list price of the item, not the total price that consumers will end up paying. Consumers must expend a high time cost to follow links and determine what the total price will be.

Consumers are known to be very price sensitive on search engines. Having the lowest posted price is associated with 60% more "clicks" than any other listing (Baye et al, 2006). Explicitly posting additional prices plays a clear role in changing consumer behavior: People are doubly sensitive to taxes when listings are ordered including taxes, and less responsive to taxes (relative to list price) when taxes are not explicitly listed at all. Ellison and Ellison (2008) found that consumers were less sensitive to changes in taxes than changes in list prices. Unlike many other studies, Ellison and Ellison used purchase data from a firm that lists items on a search engine. This is a much richer form of data to analyze because click data does not indicate an actual purchase and it is therefore impossible to determine if a purchase was actually made. Therefore I want to use this same purchase and search engine data used by Ellison and Ellison to determine how sensitive consumers are to explicitly stated shipping prices. Unlike past studies on explicitly stated taxes, the listings on Pricewatch.com are not ordered by their shipping included prices. Given two items on Pricewatch, one may be listed higher due to a lower listed "price" despite have large shipping price and therefore a larger overall price. Shipping price is listed directly next to the list price, making it as easy as possible for consumers to determine the total price, without explicitly stating it (see Appendix 2). In past studies shipping has been shrouded or put in fine print in the bottom of the page. Unlike taxes, shipping prices are explicitly listed on the website next to the list price so this likely effects consumer's ability to make price determinations. Shipping prices have been examined in an auction market setting but never in the setting of a search engine. I aim to quantify how sensitive consumers are to shipping costs relative to list prices and taxes on price search engines. This is an important contribution because I will quantify the extent to which consumers respond to explicitly partitioned prices (shipping) as compared to non explicitly stated partitioned prices (taxes) and list prices in a search engine setting.

In a broader context my work addresses the fundamental economic topic of how information affects consumers' decisions as well as the bounds on consumer rationality and search. Information and consumer search costs have long been a topic of debate in economics literature. It is believed the consumers will only invest in additional search costs (time) if they believe the price will be reduced by more than their search costs (Stigler). On Pricewatch, search costs are low and price savings exists, but it is unclear whether consumers pay attention to more information than just the simple order of list prices. With regards to consumer rationality in Ellison and Ellison (2008) taxes were found to matter but not in a one to one ratio to list prices. Yet past studies found consumers who are presented with listings sorted by tax inclusive prices are extremely sensitive to taxes. This suggests that there is a bound on consumer rationality and consumers ability to calculate costs. I aim to further specify where this boundary is. Do consumers only pay attention to the list order? Or are they able to make further simple calculations? Instead of having to determine the approximate tax and apply it to each price (if they live in that state), in my analysis consumers face two prices presented in equal clarity next to each other. If this is too much for consumers to calculate, then it will be a strong statement about the limits of consumers' rationality and ability to process all information presented.

2. Literature Review

Existing literature suggests that preferences are affected by whether a price is presented as one all-inclusive expense or partitioned into a set of mandatory charges. The literature has examined portioned prices in both traditional retail markets and online retail markets. In a laboratory setting modeling traditional retail markets, Markowitz (1998) found that partitioned prices reduce consumers' calculation of out of pocket expenses. Given two items of equal total price, the one with the partitioned prices seems less expensive to consumers. But, Betini and Watheiu (2008) found that partitioned prices focus attention on the characteristics of the goods that were partitioned and to undervalue the second price listed when prices are partitioned.

With respect to partitioned prices on the Internet, Hossain and Morgan (2006) studied consumer behavior on Ebay and found that shrouding shipping prices can reduce revenue if shipping costs are low and does not change revenue when shipping costs are high. Brynjolfsson and Smith (2001) examined consumers who visited Even-Better.com in 1999 and found that consumers are estimated to be twice as sensitive to differences in explicitly stated taxes as they are to differences in item prices. But in this setting the items were listed in order of their total tax inclusive price, which explains why consumers were easily able respond to these differences in taxes. They also found strong evidence that consumers prefer branded e-retailers over lesser known firms. The results are limited by the fact that they do not have quantity data, only click data; and thus, the quantity data was imputed by assuming that that consumers made a purchase from the e-retailer they visited last. This is unreasonable as consumers can visit many websites before making a purchase and often choose not to make a purchase at all.

Lewis (2006) uses data from an online grocer to examine partitioned prices and builds regression models, for consumer retention, consumer acquisition, repeat purchases and order size. The regressors in the model include price, branding, email coupons, shipping fee, the shipping penalty (measure by the different in shipping costs for large and medium and medium and small purchases). The evidence suggests that higher shipping fees reduce store (online website) traffic and that order size incentives (penalties) result in larger (reduced) order sizes. Although this is helpful to understand how partitioned prices effect consumer behavior, this analysis was done in a setting where consumers are buying multiple items (groceries) and does not represent evidence on how price sensitive consumers are for a given good and a given shipping price. Additionally the shipping price changed over time so consumers visiting the website were not able to make decisions from a selection of different shipping prices as they would on a search engine site.

Price search engines have received a whole separate body of literature. In price search engines there is strong incentive to cut prices to be the lowest price listed. The search engines list advertised items sorted by price from lowest to highest. Baye, Gatti, Kattuman and Morgan (2006) determined that a firm receives a significant share of the clicks when it offers the lowest price on a site. Using a dataset from Kelkoo.com they found that the lowest price receives 60% more clicks. They find that a firm receives 17% fewer clicks for every competitor listed above it. The data they had represented 40% of annual activity on Kelkoo, which is the biggest price search engine. Although the dataset is large, without quantity data they cannot determine what consumers actually purchased, it is unreasonable to assume that everyone who clicked on an item actually bought it. However their modeling of discrete choice behavior is helpful for this study.

Ellison and Ellison (2008) examined purchases of memory modules on Pricewatch.com and found that firms engage in obfuscation strategies and add-on techniques to incentivize consumers to purchase more. Consumers have very elastic demand when it comes to selecting an item from the Pricewatch listings. This is consistent with past research that asserts the advertising value of having an item listed as the lower price (and therefore higher up on list). The most popular obfuscation technique is to intentionally create an inferior quality good that can be offered at a low price. Firms offer a very low price for the low quality good to attract customers to their website. They then will use various techniques to convince consumers to upgrade to a higher quality product (with a higher mark up). Using the sales and cost data from a firm listing products on Pricewatch, they were able to calculate the markup on the different items. The mark up on the low quality goods (the ones that appear on the Pricewatch website) was very low but medium to high quality goods were able to achieve mark ups of 10-15%. This confirms that price search engines do not reduce firm profits as much as one would expect. They may lower search costs for certain types of goods, but firms engage in obfuscation strategies to maintain high search costs and mark ups on most of their goods.

Ellison and Ellison (2008) explore the same Pricewatch universe to examine how sensitive consumers are to taxes. Pricewatch displays items sorted by list price and does not display a tax inclusive price. The listing for each product does include the home state of the retailer. This allows fully informed and rational consumers to compute the overall (tax inclusive) price of each listing. However this is difficult for any consumer who has not memorized the tax rates in each state. They conclude that sales taxes are an important driver of e-retail activity. They use state-level regressions to show clearly that e-retail purchases are higher in states that levy higher sales taxes on traditional retail purchases (online purchases made from a retailer in a different state are tax free). Then they use discrete-choice analysis to find that consumers do not pay as much attention to differences in taxes as they do to differences in pre-tax prices when choosing between e-retailers. Since they only have limited quantity data, they build their model in a way that exploits the numerous price and rank changes made daily. Their model is used as a basis for analysis of shipping prices on consumer behavior in this paper. They conclude that taxes matter to consumers, though, and given how tightly distributed prices in this market are, they can have significant effects on consumer behavior. They also observe evidence that convinces them it is a tax effect and not some artifact of unobserved heterogeneity. They conclude by agreeing with previous assertions that applying sales taxes to e-retail sales could reduce e-retail demand by one-quarter or more. The body of literature relating to partitioned prices and behavior in a search engine setting build a strong foundation for my analysis. Important factors driving consumer behavior will be incorporated into my analysis. Yet to date there is no work examining explicitly stated shipping prices in a search engine setting and how these relative prices can affect consumer purchases. And that is where I seek to add to this body of literature.

More broadly, my work seeks to address the fundamental economic topic of information and consumer search costs. Beginning with Stigler's theoretical discussion of the economics of information (Stigler 1961), much economic literature has tried to identify and quantify the way in which consumers respond to search costs and the effect of search on market dynamics. Traditional economic models suggest there is a threshold to consumers search and ability to calculate costs. Consumers are willing to invest in an extra amount of search (as measured in time costs) if they believe it will result in a lower price (Stigler 1961). Bakos (1997) argues that as the Internet lowers search costs for prices and product information, consumers should be able to make rational decisions better and more easily. Yet as the Internet has offered low cost methods to compare prices, have consumers not become any better at using this information? Consumers go to Pricewatch to lower their search costs and have an easier purchasing decision. It is possible that they therefore overweight the most simple piece of information (the list price and ranking) to determine which good they wish to purchase. The two prices are presented in a clear way and if consumers cannot effectively add these two pieces of information together we might have two extremely conclusions: consumers are only able to effectively process one piece of information when comparing products (a single price or rank) and thus they underweight all additional pieces of price information; the internet has lowered search costs but also lowered consumers threshold for search and calculation.

3. Data

I will be using price data from Pricewatch.com and sales data from one computer company that lists its products on Pricewatch.com. Pricewatch.com is an Internet retail search engine that is popular with computer-savvy shoppers. Many small retailers list their products on Pricewatch and keep Pricewatch informed of their daily low prices. The firms are small, do little to no advertising and receive most of their customers through Pricewatch. Pricewatch sorts the items by list price (the price of a memory module). , with 12 items listed per page. Consumers then select an item, which links them to the website where they can make a purchase. The consumers shopping on this website are often small retailers or those building their own computers, and thus tend to frequent the website multiple times. Therefore although not the average consumer, this environment is valuable to study because if rationality is going to hold anywhere it should hold for these informed and computer savvy shoppers.

The products we are examining are memory modules, the part of a computer responsible for storing memory. Our data includes the sales of two different types of memory modules, 128MB PC100 and 128MB PC133. Both products are of similar quality and hold the same amount of memory (128MB); however PC100 is only compatible with older computers. PC133 is compatible with newer models of computers and also backwards compatible with computers that use PC100.

Potential consumers can choose product types, e.g. 128MB PC100, and will be given a list of participating retailers selling products in that category, sorted by price. The list with products also includes information on retailer location, shipping price, contact information, etc. An example of what a page on Pricewatch looks like can be found in Table 2 in the appendix. There is a fair amount of turnover and reordering of the price lists from day to day (and even from hour to hour in some periods). Over the course of the year there is a dramatic range in prices as shown in Chart 1. The minimum price listed is $20 for a 128MB module, whereas the maximum price is $131.

All of the price listing data was obtained by downloading the first and second pages from Pricewatch's memory module price lists on an hourly basis from May 2000 to May 2001. Although this was a decade ago, data on the Internet, particularly firm specific purchase data, is difficult to obtain for academic purchases, which explains why more recent data does not exist. Pricewatch.com is still in operation but its sorting mechanisms and design has changed since the collection of this data. Nevertheless, the environment we witness in our data is very similar to that of other retail search engines that exist today and so it is still a very worthwhile dataset to examine. Additionally this data is a very rich resource to help us answer fundamental questions about consumer search and decision making in retail markets.

Table 1 displays the average monthly low price on Pricewatch and the average price listed by our firm in a given month. As evident in the table, with the exception of June 2000, our firms price is usually close the lowest price on Pricewatch. Logically this also means the firms average rank is low (6.4) but it appears at multiple spots between positions 1 and 21 on the Pricewatch list during the time period we examine.

Chart 1: Summary Statistics for Firm Data

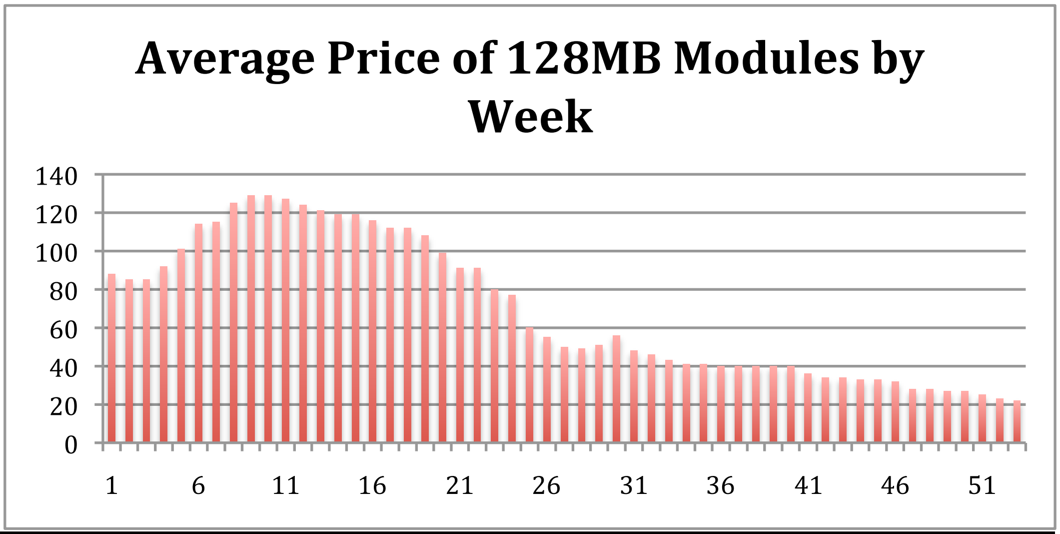

Table 1: Summary Statistics for Firm Data

Summary Statistics for 128MB PC100 Memory Modules			
Month	Average Lowest Price on Pricewatch	Our Firm's Average Price	Total Number of Orders Our Firm Received
Jun-00	$81.73	$105.03	246
Jul-00	$115.75	$118.12	381
Aug-00	$109.09	$113.81	173
Sep-00	$93.47	$99.73	217
Oct-00	$65.10	$69.06	448
Nov-00	$47.38	$51.29	643
Dec-00	$41.70	$43.14	601
Jan-01	$37.09	$39.02	642
Feb-01	$31.41	$34.92	748
Mar-01	$29.26	$32.85	494
Apr-01	$27.59	$29.18	810

The Ellison and Ellison (2008) analysis used Pricewatch.com data for 4 different types of memory modules. However the shipping prices are only available for the 128MB modules so that is what my analysis focuses on. I extracted the numerical shipping values from each listing on Pricewatch. Overall, shipping prices range from $0 to $16 dollars with a mean of $10 for PC100 modules and $9 for PC133 modules and a standard deviation of .4 and .5 respectively. A major shortcoming of Pricewatch is that the shipping prices are not listed in a standardized fashion. 14% of the listings for shipping prices are approximations (e.g "$6-7" or "$12 and up"). In these cases my analysis uses the lowest listed price for shipping. Shipping information also includes data on the type to shipping (ie. Fedex 2 day, Insured, or USPS) being used which I analyze using a dummy variable for higher quality shipping. Out of all the listings, 14% list a high quality service (FedEx, 2 day, UPS) and 9% list "Insured," with a small percentage listing as both.

Consumers can then click on any of the listed products, to be directed to a retailers website where they can complete the purchase. One retailer, which operates two different websites on Pricewatch, supplied sales data for the time period we are analyzing. This is an important asset to the analysis because it allows us to directly determine how the sales for a given firm is affected by the relative prices, shipping fees and other market conditions of other listings on Pricewatch at that time. Table 2 presents the summary statistics for prices and shipping prices in our dataset. Website 1 and Website 2 refer to the two websites operated by the firm.

Table 2.

Summary Statistics for Prices and Shipping Prices			
128MB PC100 Modules			
	Total Observations	Average Price (std dev)	Average Ship Price (std dev)
Other firms	202,879	64.34(35.5)	9.80 (1.81)
Website 1	9,452	59.69 (35.8)	10.66(.23)
Website 2	6,587	41.14 (21.08)	10.04 (.40)
Total	218,918	73.83 (35.5)	10.38(.48)
128MB PC133 Modules			
	Total Observations	Average Price (std dev)	Average Ship Price (std dev)
Other firms	187,144	70.03 (41.7)	9.91 (1.67)
Website 1	8,596	64.51(39.16)	10.67 (.21)
Website 2	5,701	42.91 (23.17)	10.09 (.43)
Total	201,441	66.303 (35.19)	9.93(.50)

Website two stopped selling modules for a period of time, resulting in the number of observations (times it appears on Pricewatch) relative to Website 1. The average prices are therefore deceptive, as Website 2 does not participate when the prices are extremely high. The firms vary both shipping and list prices during the period of analysis, although shipping prices do not vary greatly.

Shipping prices in our dataset are not positively correlated with list prices. There are some very unique instance in which listings with lower list prices also have lower shipping prices so that the order of the total (shipping inclusive) prices would be exactly the same as the ranking of the list prices. If this were the common trend throughout the data, it would be problematic as the total, all inclusive price, would be interpreted by a fully rational consumer the same way as the list price. In this data, there is a statistically significant and minute negative (-.0003) correlation between shipping prices and list prices.

The variation in shipping prices is somewhat limited. As can be seen in Table 2, the websites controlled by our firm do not vary their shipping prices much during the sample. Even though it would be useful if they varied more, we can still hope to identify the effect of shipping costs by looking at how the shipping costs and list prices of competitors affect our firm's sales..

4. Empirical Methodology

This analysis uses a methodology similar to the one employed by the previous Ellison and Ellison papers using the Pricewatch data. Ellison and Ellison (2008) employ demand estimation techniques in an unusual way to exploit the hourly variation in the data. In both of their papers they estimated discrete choice models that use the number of orders of a memory modules at a particular hour or day as a dependent variable. They found that preferences for home state purchases were significant; however consumer sensitivity to taxes was inconsistent: taxes were not statistically significant for 128MB PC100 modules, but were for the 128MB PC133 modules. They used this to conclude that consumers are sensitive to taxes to some degree but not as much as they are to list prices. Most of the literature on partitioned prices and online consumer demand uses discrete choice models. Usually discrete choice models use every firms' market shares, which often requires price and quantity data for every firm. But as discussed above, we only have the data for one firm (two websites) that list on Pricewatch.com. Ellison and Ellison (2008) were able to exploit the substantial intertemporal variation in the characteristics of the listings (there is an average of 4 list order changes per day) to complete their model. They examine how substitution between retailers and how the one firm's sales increase and decrease as rival firm's prices and characteristics change. Given that I am constrained by the same lack of complete data that they had, I aim to estimate a discrete choice model in a similar way.

In the tax sensitivity analysis Ellison and Ellison (2008) modeled the dependent variable as the number of units purchased at a given time in a given state. I employ the same definition of the dependent variable in my analysis. Although I am not focusing on the coefficient on taxes, state specific effects are significant and therefore necessary for my analysis. Most often the number of units purchased is 1, although it ranges between 0 and 4.

The demand for memory modules will be a result of an indirect utility function of the price, the ranking on the website, and time dependent demand variables. Therefore there are three main components to the discrete choice model that Ellison and Ellison, and in turn this analysis, follows. The first part of Ellison and Ellison model is an indirect utility function, which is modeled as:

$$\begin{aligned} u_{iksht} &= \beta_1(Price_{iht} + \beta_2 SalesTax_{isht}) + \beta_3 ShippingTime_{is} + \beta_4 HomeState_{is} \\ &+ \beta_5 NeighborState_{is} + \beta_6 SecondScreen_{iht} + \beta_7 SiteB_i + \epsilon_{ik}, \end{aligned}$$

In this model the utility for a given individual *k,* on website *i*, from state *s*, at hour *h* and day *t* for a purchase on Pricewatch is a function of price, tax, shipping time, whether the retailer is in their home or neighbor state, whether the item appears on the second page, which website it is purchased from, and an error. I aim to use a similar specification with a few changes. Ellison and Ellison use this model to estimate many parameters yet focus on β_2, which represents how sensitive consumers are to the sales tax relative to the list price. A value of one for β_2 would correspond to the traditional model in which consumers only care about their total expenditure, whereas a coefficient of zero would be that consumers are completely insensitive to taxes. In my analysis I will add shipping price and potentially a variable that captures shipping quality. My initial model for consumer's utility will be the following:

(1)

$$u_{iksht} = \beta_1(Price_{iht} + \beta_2 SalesTax_{isht} + \beta_3 ShippingPrice_{iht}) + \beta_4 HomeState_{is} + \beta_5 NeighborState_{is} + \beta_6 SecondScreen_{iht} + \beta_7 SiteB_i + \beta_8 ShippingQuality + \beta_9 ShippingTime_{is} + \varepsilon_{ik}$$

I will start by only adding Shipping Price to the model, so I can compare it directly to the Ellison and Ellison paper. My focus will be on the value of β_3. A value of 1 for β_3 would correspond with the standard rational model that consumers take the whole price into consideration. A value of 0 for β_3 would mean that consumers do not pay attention to shipping prices at all. And the initial model, I will adjust the model to see if I can find an even better fit for the data. Shipping time in the past model was found have a very small but negative effect; however this is possibly due to measurement error as it was calculated soley based on the zip codes of the retailer and purchaser. Additionally many listings offer specific quality (overnight, two day, Fed Ex, Insured) shipping services, meaning that the original analysis (Ellison and Ellsion 2008) used a simplified estimation for shipping time. Parsing out all of the individual shipping information would be difficult and potentially unsystematic so I keep the ShippingTime variable defined as it was in the original analysis, and add a quality variable which will be a dummy variable for a higher quality service (FedEx, insured, overnight etc). *HomeState* is a dummy for whether the retailer is in the state of a given purchaser. *NeighborState* attempts to capture if consumers have a preference for local purchases and could be compared to the coefficient for *HomeState*; however *NeighborState* was consistently found to be insignificant in previous analysis. *SecondScreen* is whether the price listing in on the second page of the Pricewatch list (places 13-24). This is important because it that extra effort of scrolling to a second screen increases the search and transaction costs for the consumer. In the past analysis it was found to have a negative coefficient and I expect the same result in my analysis. *SiteB* is a dummy variable for which of the two sites that our firm operates received the purchase. There was little conclusive evidence about consumers' sensitivity to taxes, but given that many factors in this analysis are state-dependant, it is essential to maintain my analysis separated by state and to include taxes in my analysis. Furthermore, this will allow me to compare the coefficient of indirectly stated prices (taxes) to that of directly stated prices (shipping prices).

As mentioned previously, in addition to the individual preferences, the discrete choice analysis also incorporates a general demand model. Given our lack of complete quantity data, we do not know the total purchases made on Pricewatch at every given time. Therefore a general demand model is built to estimate the number of purchasers at every given period we study. Ellison and Ellison (2008) developed the following function to model the number of purchasers N on Pricewatch.com from a given state *s*, at a given hour *h,* on a given day *t* as follows:

(2)

$$N_{sht} = \delta_s \overline{q}_h e^{\gamma_1 MinPrice_{ht} + \gamma_2 Weekend_t + \gamma_3 TimeTrend1_t + \ldots + \gamma_6 TimeTrend4_t} + \eta_{hst},$$

The demand is a function of state specific fixed effects δ, hour of day fixed effects q, and an exponential function that includes the minimum price listed on price watch *MinPrice*, a weekend dummy variable *Weekend*, 4 time trend dummy variables *TimeTrend* that allow for linear time trends of 90 days each and η_{sht} is an error term which is assume to have a mean of zero conditional on the right hand side variables. The state specific fixed effects are significant as previous studies found that states have certain attributes (higher internet usage, more computers per household, higher average income, higher average state taxes) that significantly effect the their demand for online purchases (Ellison and Ellison 2008). Hour of day fixed effects are also important, as demand should drastically decrease in the middle of the night. The *TimeTrend* dummy variables help to capture the unique trends in prices during the period of analysis: Prices listed on Pricewatch mostly decrease over time from June 2000 to May 2001, with the maximum reaching over $132 dollars and the minimum nearing $20 for the same item. I do not plan to make any changes to this portion of the model and I not believe that shipping fees would have an overall effect on general demand as they do not vary much over time.

An important factor to consider is endogeneity of prices, i.e. whether it is okay to use prices in the right hand side of our equation instead of other instruments. Ellison and Ellison (2008) asserts that endogeneity should not be of a concern because based on their interaction with the one firm, they believe firms have little information about demand shocks and little ability to determine what to do given these demand shocks. And even these firms did have information about demand shocks that are unaware of, they believe it is likely these demand shocks to do have a large effect on prices. Therefore it is unlikely the consumer behavior is affecting the prices that firms set.

The final part of the discrete choice analysis is the model that brings together the individual preferences and the general demand model. I plan to use a similar model to Ellison and Ellison (2008) for the final discrete choice analysis which is a logistic model defined as:

(3)

$$E(Q_{isht}|X_{sht}, N_{sht}) = N_{sht}\frac{e^{\beta X_{isht}}}{\sum_{j=1}^{24} e^{\beta X_{jsht}}}$$

I will be using a similar estimation method, nonlinear least squares, with hour-website-destination date sales as the dependent variable Q_{isht}. N_{sht} is all of the right hand side variables from the general demand model. The fraction in the model is our firm's market share. The X_{isht} in the numerator contains the characteristics of the consumer and module actually purchased (price, shipping price, tax, shipping quality) from our firm. The denominator represents all of the options that were available to the consumer, the top 24 listings on the website. Each term is calculated assuming that consumer had made the purchase from each of these listings (e.g, tax for an item is calculated using the state tax of the listed item and information about the state of the listing and the consumer who made the purchase). This gives us the total of what the consumer would have faced if they had decided to chose this particular listing instead of the one they did chose. Overall this gives us an idea of environment the consumer faced and allows the model to estimate what the key factors were in the decisions. Again, the model used in this analysis is not the standard discrete choice model. Given that we lack quantity data beyond our one firm, our model exploits the frequent price changes and movements on Pricewatch to build a model that captures the same effect as a standard discrete choice model would. The data set is extremely rich, thousands of purchases over our time period. The analysis was run by modifying the STATA code originally written for the Ellison and Ellison analysis (see appendix).

5. Results

Our model uses purchase and Pricewatch data for two different types of memory modules: 128MC PC100 and 128MB PC133. They are the same size and hold the same amount of memory. There are 793,950 observations and each observation is the hourly-state-website purchase of a memory module. Given that often the websites do not sell anything to a state on from a website in a given hour, the average quantity is low (.007 for PC100; .006 for PC133) with a minimum of 0 and a maximum of 4. Price represents the price charged by our website. The average price is $66 for PC100 and $73 for PC133. The price ranges from $21 to $131 for both with a significant downward trend overtime. This dramatic change is accounted for using the *TimeTrend* dummy variables in the regression. The minimum price is the minimum price listed on Pricewatch at the time a purchase is made, which is on average $62 for PC100 and $71 for PC133. This price is not the price that the customer paid if they made a purchase, but rather only what the lowest price listed was. The firm's rank is the average rank of our firm on the website (what position it was on the list, with 1 being the top of the list). Our firm's average rank was about 6, meaning that it was most often on the first page. The average tax that purchasers faced in their own state was 6.5%. All purchases bought memory modules from one state (as our firm only resides in one state), but the tax is what they faced had they purchased the product from a retailer (traditional or online) in their own state.

Table 3.

	Summary Statistics for 128MB PC100 Memory Modules on Pricewatch			
	Mean	Standard Deviation	Minimum	Maximum
Number of Orders	0.01	0.09	0.00	4.00
Our firm's price	66.24	34.37	21.00	123.00
Lowest Price	62.51	33.36	20.00	122.00
Our firm's rank	6.40	4.12	1.00	21.00
Quality	.2238	.1437	0	1
Tax	.064	.0165	0	.084

	Summary Statistics for 128MB PC133 Memory Modules on Pricewatch			
	Mean	Standard Deviation	Minimum	Maximum
Number of Orders	0.07	0.076	0.00	4.00
Our firm's price	73.8	36.57	21.00	136.00
Lowest Price	71.4	35.06	20.00	122.00
Our firm's rank	5.97	4.12	1.00	24.00
Quality	.318	.1437	0	1
Tax	.065	.0161	0	.084

In our dataset, the websites combined received about 5,700 orders for 128MB PC100 over the year and 4,075 orders for 128 MB PC133.

I ran 5 regressions for each type of memory module as a robustness check. Table 3 and Table 4 demonstrate the different regressions run for PC100 and PC133 memory modules respectively. Bolded coefficients are those that are significant at the 5% level. As the p-values on the table demonstrate, nearly all coefficients that were significant at the 5% level were also significant at the 1% level. All variations in regressions occurred in the individual demand variables, which are in the top portion of the table. Across all regressions, the coefficient for *Price* was about between -.5 and -.8 and statistically significant. *ShippingPrice* was statistically significant and positive, although essentially zero in magnitude. The *HomeState* dummy was positive and significant and the quality variable was negative and usually significant. The R^2 values were around .25 for PC100 modules and about .21 for PC133 modules. The preferred regression for each type of memory module is the middle regression in the table; this model includes *Price*, *Tax*, *ShippingPrice*, *HomeState*, and *Quality* in addition to the general demand variables. Subsequent discussion will focus on these two preferred regressions.

Table 4. Discrete Choice Analysis for 128MB PC 100 Purchases

Different regressions for purchases of 128 MB PC 100 memory modules					
Variables that effect individual demand					
Price	**-0.540**	**-0.565**	**-0.567**	**-0.568**	**-0.579**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Tax	0.059	0.048	.0550	0.042	0.0435
	(0.452)	(0.531)	(0.473)	(0.700)	(0.418)
Shipping Price	**0.0032**	**0.0030**	**0.0024**	**0.0025**	**0.0025**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Home State	**0.455**	**0.457**	**0.576**	0.152	
	(0.026)	(0.028)	(0.003)	(0.291)	
NeighborState	-0.084				
	(0.509)				
Quality		**-0.147**	**-0.145**	**-0.186**	**-0.188**
		(0.0001)	(0.0001)	(0.0001)	(0.0001)
Shipping Time	-0.040	-0.036			
	(0.123)	(0.130)			
Page 2	-1.391	-2.045	-1.92892	-3.849	-2.79823
	(0.046)	(0.149)	(0.131)	(0.651)	(0.149)
Variables that effect overall demand					
Weekend	**-0.424**	**-0.422**	**-0.422**	**-0.452**	**-0.453**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Minimum Price	**-0.033**	**-0.034**	**-0.034**	**-0.034**	**-0.034**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Timetrend 1	**0.016**	**0.017**	**0.017**	**0.017**	**0.017**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Timetrend 2	**0.037**	**-0.039**	**-0.039**	**-0.039**	**-0.039**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Timetrend 3	**0.017**	**0.018**	**0.018**	**0.018**	**0.018**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Timetrend 4	**-0.003**	-0.001	-0.0147	-0.001	-0.001
	(0.0001)	(0.095)	(0.058)	(0.068)	(0.03)
California				**2.810**	
				(0.0001)	
r2	**0.029**	**0.029**	**0.028**	**0.027**	**0.021**
bolded values are statistically significant at the 5% level					
P values are listed in parentheses					
State specific variables were included in regression but not shown as none yielded					

Table 5. Discrete Choice Analysis for 128MB PC133 Purchases

Different regressions for purchases of 128 MB PC 133 memory modules					
Variables that effect individual demand					
Price	**-0.789**	**-0.765**	**-0.760**	**-0.811**	**-0.808**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Tax	**0.346**	**0.335**	**0.340**	**0.328**	0.125
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.418)
Shipping Pric	**0.00455**	**0.00485**	**0.00489**	**0.00455**	0.00155
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.432)
Quality	**-0.210**		**-0.240**		
	(0.0001)		(0.001)		
Page 2	***-0.631***	**-0.438**	**-0.541**	**-0.608**	**-0.467**
	(0.020)	(0.020)	(0.008)	(0.009)	(0.009)
Shiping Time	-0.041			-0.065	
	(0.179)			(0.154)	
Home State	**1.479**	**1.539**	**1.532**	**1.539**	
	(0.0001)	(0.0001)	(0.0001)		
Variables that effect overall demand					
Minimum Price	**-0.027**	**-0.027**	**-0.026**	**-0.028**	**-0.027**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Weekend	**-0.412**	**-0.405**	**-0.407**	**-0.408**	**-0.404**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Timetrend 1	**0.019**	**0.017**	**0.018**	**0.019**	**0.018**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Timetrend 2	**-0.026**	**-0.025**	**-0.025**	**-0.026**	**-0.025**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Timetrend 3	0.001	0.002	0.000	0.002	0.002
		(0.873)	(0.767)	(0.437)	(0.326)
Timetrend 4	**-0.005**	**-0.004**	0.000	**-0.005**	**-0.004**
	(0.002)	(0.002)	(0.942)	(0.003)	(0.0001)
California					**1.803**
					(0.0001)
R^2	0.0215	0.0216	0.0216	0.0215	0.0175
State specific variables were included in regression but not shown as none yielded					
bolded values are statistically significant at the 5% level					

6. Discussion

My discussion focuses on the economic implications of the preferred model for each module. I will begin by discussing the coefficients related to price and taxes, before focusing on the shipping price and shipping quality. And then I will conclude with a brief discussion of state specific effects and the general demand variables in my model.

Price

The price coefficients for both types of PC128 memory modules are extremely significant, showing that consumers are very sensitive to small changes in price. As the price increases, consumers demand for these memory modules decreases. The own price elasticity for (holding all variables fixed at their sample means) for 128MC PC100 is -31 which means that every small change in price a firm makes (or when one firm is undercut by another) has a huge effect on the demand.

Taxes

The coefficients for taxes mirror those that were found in the original Ellison and Ellison analysis: they were statistically different from zero for some modules but not for others. Our analysis found that taxes were not statistically significant for PC100 modules. But for PC133 modules, they were statistically different from zero but not equal to one. Although unusual, the difference across modules is consistent with the Ellison and Ellison (2008) analysis on taxes. Overall, taxes were statistically different from 1 for both modules, meaning that consumers do not respond as much to taxes as they do to list prices.

Shipping Price

Shipping prices are the focus of our analysis. In our data shipping prices vary from $0 to $16 for both modules. The average shipping price for PC100 modules is 10.38 (std dev = .476) and for PC133 modules is 9.83 (std dev = .5). These standard deviations are small. But, the Pricewatch listings are extremely competitive, with sometimes only $2 separating the top 12 listings, therefore making small changes in shipping prices significant. Additionally there is a small but statistically significant negative correlation between list price and shipping price. This corresponds with pervious literature that asserts that retailers use higher shipping prices to compliment lower (more competitive) item prices and lower shipping prices to attract customers when item prices are higher. Based on average values, the shipping price is about 14% and 15% of modules prices (for PC100 an PC133 respectively). This is a significant percentage as it greatly exceeds the tax percent.

Table 6. Summary statistics for Shipping Prices and List prices

Module Prices				
Module Type	Average	Standard Deviation	Minimum	Maximum
128MB PC100	73.83	36.54	21	131
128MB PC133	66.30	35.20	21	123
Shipping Prices				
Module Type	Average	Standard Deviation	Minimum	Maximum
128MB PC100	10.385	0.477	0	15
128MB PC133	9.932	0.500	0	16
Shipping Prices as a Percent of List Price				
Module Type	Average	Standard Deviation	Minimum	Maximum
128MB PC100	14.08%	3.20%	6.45%	47.62%
128MB PC133	15.15%	3.30%	5.97%	51.28%

In the demand model, consumers are assumed to evaluate products on the basis of Price + β_2SalesTax + β_3ShippingPrice, where both *SalesTax* and *ShippingPrice* are measured in dollars. An estimate of one for each of the coefficients would correspond to the standard rational model in which consumers only care about their total expenditure. And a coefficient of zero would mean that consumers do not pay attention to sales tax or shipping prices when making their purchasing decisions. As evident in Table 3 and Table 4, all regression results thus far yielded a coefficient that is near zero. For both types of modules, the coefficient is statistically different from zero, however virtually zero in magnitude (.002). For PC100 modules, we cannot reject the hypothesis that the coefficients for Taxes and ShippingPrice are equal and equal to zero. However for the PC133 module, the coefficient for taxes is positive and statistically significant from zero, although statistically not equal to 1. The relatively small but significant coefficients for both models is important and is likely under representing the effect of shipping prices. It is important to remember that our model measures the effect of shipping, taxes, and prices by exploiting the variation in prices and the competitive environment on Pricewatch. And therefore variation is essential to capturing their effect in the model. In an ideal model, the shipping prices would be randomly distributed, and at the very least, not be clustered closely to the mean. Variation between shipping prices is much smaller than variation between tax amounts, which could explain the overall smaller coefficient on shipping prices in our model. What is also interesting is that the coefficient for shipping is positive and statistically significant across both modules, which is not true for the tax coefficient. Despite these potential explanations, overall the results suggest that consumer behavior on the Internet may not reflect the standard rational model and that consumers tend to ignore the smaller parts of partitioned prices regardless of whether they are explicitly (shipping prices) or implicitly (taxes) stated.

Shipping Quality

I hypothesized that higher shipping prices may be correlated with shipping quality and that in general, shipping quality is an important part of a consumers purchasing decision. Some listings indicated the type of shipping in addition to the price and so I created a quality variable that measures if the shipping is advertised as insured or of a high quality (Fed Ex expedited 2day, overnight etc). The quality variable is equal to one if the module includes indication of quality, insurance, or both. Overall this should signal a higher level of shipping quality to consumers, which economically would mean the value of the total package they are purchasing, is higher. Memory modules for a computer are significant investments but also delicate and therefore consumers probably want to ensure it is not damaged in transit. This also captures the idea that certain higher quality methods of shipping are likely more expensive. This variable was also used to take the place of the ShippingTime variable, which was defined in a way that may contain measurement error. The table below contains the summary statistics for the prevalence of high quality and insured shipping in our data set.

Table 7. Summary statistics for shipping listings

Characteristic	Prevalence in data
Insured Shipping	13.33%
High Quality Shipping (FedEx, 2day, UPS)	9.58%
Approximated Shipping Prices $6-8, $12 and up)	14.58%

The shipping price is positively correlated with shipping quality (p value < .05). When included in the model, the quality variable consistently has a statistically significant coefficient of about -.15 for PC100 modules, and -.23 for PC133 modules, signifying that modules with higher quality advertised shipping were less likely to be purchased; however, the inclusion of this variable did not increase the R^2 or adjusted R^2 values, nor did it increase the value of the *ShippingPrice* coefficient. Together the tendency for people to avoid high quality shipping and the correlation of that with higher shipping prices, suggests that purchasers may in fact be avoiding higher shipping prices. As the table displays, about 14% of the shipping prices are ranges ($6-8) or approximated values. This prevalence of estimation may discourage consumers from paying attention to the exact prices listed on Pricewatch. Instead, they may be using shipping quality information as a signal of higher shipping prices that may not be explicitly stated. For PC133 modules, the effect of quality (-.23) is nearly equivalent to the tax effect (-.25) on whether a consumer makes a purchase (recall the total effect of taxes on the purchase decision is $\beta_1 * \beta_2$). Past literature has shown that consumers are only willing to invest in time costs if they anticipate greater savings. Given the general lack of variation in shipping prices on Pricewatch, consumers may not find it worthwhile to invest the time to individually calculate total prices. Instead consumers may use the quality as a simple signal of a higher shipping price, which makes them less likely to purchase a product. So although our model does not capture a sizeable effect of shipping price, the quality variable may be demonstrating that consumers do pay attention to shipping prices and they do attempt to rationally and efficiently interpret the information presented to them.

So overall, consumers may associate significant price distinctions with different shipping qualities, whereas they are not able to make small numerical distinctions between listed shipping prices, as those prices are so close in value.

Economically our result that consumers pay little attention shipping prices can be extended to all forms of partitioned prices in a search engine setting. This is significant because it does reject the standard model that consumers evaluate the full price of an item; instead they appear to focus on the list price or the most basic piece of information the search engine gives them (a ranked order of list prices). In this case, the coefficient for shipping price is much smaller in magnitude than that for taxes although we cannot statistically reject the hypothesis that the coefficient for both taxes and shipping are the same. This result is important because it implies that the two forms of partitioned prices, both those that are implied (taxes) and those that are explicitly stated (shipping prices) have the similar and minimal effects on consumers overall calculations. Overall this result supports the notion that consumers use search engines to simplify their decision making tend to only pay attention to the list price component. Previous research found that lower shipping prices were associated with higher bidding for items in auctions on EBay (Einav 2011). My results in conjunction with the EBay results demonstrate the structure of the market and the pricing mechanisms significantly affects how consumers make purchasing decisions and how they respond to partitioned prices.

Geography

Geography affects our analysis in a few ways. The first is state specific affects. The model for our analysis is state-specific, meaning each observation is specific to website-state-hour-day. And therefore in the development of the general demand model

(Equation 2), we have a variable that captures any state specific effects. The first thing to note is the each model has variations in the purchases made from each state. Table 5 demonstrates the total purchases made from each state (the state is the home state of the consumer who made a purchase from our website). Especially in states where small purchases are made, the discrepancies could be a reason that leads to discrepancies in tax estimates. There are also discrepancies in the coefficients for state specific effects (see appendix). Some have coefficients that are positive for one type of modules but negative for another. This is concerning as state specific effects would not be expected to change between different types of memory modules. However these discrepancies usually have one estimate that is statistically not significant (likely due to a low number of purchases). California was not included in the original Ellison and Ellison analysis because "the fact that our retailer and most other retailers are located there would make demand different under reasonable departures from our assumptions" (Ellison and Ellison 2008). But, California purchasers do account for a significant percentage of purchases. So for robustness I included California in the regression analysis once (see Table 3 and Table 4), to confirm it did not significantly change our results for shipping price coefficients.

The next impact of geography is shipping time. The shipping time variable was included to capture the effect that shipping time has on consumers' decision to purchase an item. Theoretically the longer the shipping time the less likely a consumer would be to purchase an item if they value the time it takes to arrive. The coefficients for shipping time for each module is negative and about equal to about a 4% decrease in demand for every additional day or 20% for an item that needs to be shipped across the country. Relative to price however, this effect is not that large because consumers show a similar decrease in demand for a 50-cent price increase. The coefficient for Shipping Time was not significant, which is agreement with previous findings in the Pricewatch analysis (Ellison and Ellison 2008). This could be explained by the definition of the shipping time variable, which is only an estimated field (calculated using UPS estimates for shipping time based on zip codes reported by websites and consumers) and therefore contains significant measurement error. The ShippingTime variable was removed in later analysis and replaced by the quality variable

The other way geography enters our analysis is in the dummy variables for home state, to capture any additional preferences that consumers may have for purchasing an item from their home state. Previous analysis included a neighbor state dummy variable as well, but that was removed after it was found to be insignificant in all past analysis (Ellison and Ellison 2008) and our preliminary analysis (see appendix 1). The home state preferences were significant and positive indicating the individuals have strong preference for purchasing an item from their own state.

General Demand Variables

The weekend dummy, the minimum price variables, and the dummy for the second page were all significant and negative, meaning on the weekends demand is lower, and demand decreases as the lowest price on the website increases, and demand decreases for firms on the second page. This not only agrees with previous analysis but also makes economic sense.

The time trend dummy variables are used to capture specific effects in demand changes (due to the dramatic price changes) over the time period we examine. The coefficients on the time trend variables show growth and subsequent decline of Pricewatch during the year. The coefficient for time trend can be interpreted to show that overall demand was growing at just under two percent per day in the first three months of our data, or 62% per month. The growth rates for subsequent periods can be determined by adding the earlier coefficients. Growth rates for one period in each analysis are not significant which could be a result of the change in pricing trends that occurred around the 3rd and 4th quarter of our analysis. But this is something that is definitely worth further exploration.

Our analysis focused on the different factors that can affect consumer purchases on the internet, particularly whether consumers purchases in the face of partitioned prices reflect the standard rational model. Consumers were extremely sensitive to price, but not to taxes or shipping prices, which rejects the standard economic model of consumer rationality. State specific effects varied, higher quality shipping had zero to significant negative impact on purchases, and time trend variables captured the significant changes in prices and demand over the period we examined.

7. Conclusion

This paper aimed to examine consumer rationality on the internet. In particular I looked at how partitioned prices in a price sorted search engine affected consumer purchases. My analysis focused on explicitly stated partitioned prices to determine if consumers' purchasing decisions followed the standard rational economic model. The model took advantage of the intertemporal variation of prices and shipping prices on Pricewatch.com. The discrete-choice analysis found that consumers do not pay nearly as much attention to taxes and shipping prices as they do to list prices. Our results confirm that consumer decision making on these websites do not take into account partitioned prices regardless of whether they are explicitly (shipping prices) or implicitly (taxes) stated. Our conclusions agree with past literature, which demonstrates that consumers tend to underweight the second portion of a partitioned price. Together with analysis that found that consumers are extremely sensitive to shipping prices when items are sorted by shipping inclusive price, we see that consumers use these search engines to simplify their decision making and often ignore most other information besides the ranking of the item of the list.

Overall, our results suggest that the bounds on consumer rationality, search and cost computation are important. The evidence presented in this paper appears to contradict the notion that consumers base purchase decision on the total price on an item. Instead, it seems that when presented with partitioned prices, they give less weight to the shipping price, taxes, and other components of the total price than they do to the list price itself. The underweighting of partitioned prices in this Internet setting agrees with past literature that found underweighting of partitioned prices in traditional retail markets.

My research has three important implications for the design and marketing of products on the internet. First, consumers do not take into account small differences in prices but rather, appear to look for key signals (list order, shipping quality, state of purchase), to simplify their purchasing decision. Secondly, explicitly displaying the price information does not necessarily give consumers an easier ability to compute the total prices. And finally, partitioning of shipping prices, taxes, and other prices capable of partitioning can allow retailers to continue achieving high mark-ups and profits in a seemingly “more competitive” environment on the Internet.

Finally, this analysis is the second application of this type of discrete choice models applied to a dataset with limited quantity data. This could be a useful approach to expand the literature about competition on the Internet since this body of literature continues to lack of comprehensive data (especially quantity data) relative to other areas of competition analysis.

List of references

Baye, Michael, Rupert J. Gatti, Paul Kattuman, and John Morgan (2006). "Clicks, Discontinuities, and Firm Demand Online", Indiana University, University of Cambridge, and University of California-Berkeley

Baye, Michael, John Morgan, and Patrick Scholten (2004): "Price Dispersion in the Small and the Large: Evidence From an Internet Price Comparison Site," Journal of Industrial Economics, 52, 463–496.

Baye, Michael, John Morgan, and Patrick Scholten (2006). "Information, Search, and Price Dispersion," in T. Hendershott, ed. Handbook of Economics and Information Systems, Amsterdam: Elsevier.

Bakos (1997) "Reducing Buyer Search Costs: Implications for Electronic Marketplaces." Management Science, 43,12, 1676-1692

Bertini, M., and L. Wathieu (2008): "Attention Arousal Through Price Partitioning," Marketing Science, 27, 236-246

Brynjolfsson, Erik and Michael Smith (2001). "Consumer Decision-making at an Internet Shopbot," Journal of Industrial Economics, 49, 4, 541-558

Chetty, R., A. Loony, and K. Kroft (2008): "Salience and Taxation: Theory and Evidence," American Economic Review

Einav, Liran, Theresa Kuchler, Jonathan Levin and Neel Sundaresan. (2011). "Learning from Seller Experiments in Online Markets." NBER Working Paper No 17385.

Ellison, G. (2005): "A Model of Add-on Pricing," Quarterly Journal of Economics, 120, 585-637

Ellison, G and S Ellison (2004): "Search, Obfuscation, and Price Elasticities on the Internet," Working Paper 10570, NBER. [429,439,446]

Ellison, G., and S. Ellison (2007): "Search, Obfuscation, and Price Elasticities on the Internet," Econometrica

Ellison and Ellison (2004): "Tax Sensitivity and Home state preferences in Internet Purchasing" *The American Economic Journal.* 2009, 1,2, 53-71.

Hossain, Tanjim and John Morgan (2006): ". . . Plus Shipping and Handling: Revenue (Non)Equivalence in Field Experiments on eBay," Advances in Economic Analysis & Policy, 6, Article 3.

Levin, Jonathan. 2011. "The Economics of Internet Markets." NBER Working paper.16852

Morwitz, V.G., E.A. Greenleaf and E.J. Johnson (1998). "Divide and Prosper: Consumers' Reactions to Partitioned Prices," *Journal of Marketing Research*, 35 (November), 453–463.

Smith, M. and E. Brynjolfsson (2001). "Consumer Decision-making at an Internet Shopbot: Brand Still Matters," Journal of Industrial Economics, 49, 541-558.

Stigler, George (1961). "The Economics of Information." Journal of Political Economy, 69, 3 , 213-225

Tuttle (2011). "Holiday Shopping: Will Cyber Monday Outshine Black Friday?" Time Magazine Online. http://moneyland.time.com/2011/11/28/holiday-shopping-will-cyber-monday-outshine-black-friday/

Tyan, S. (2005): "The Effect of Shipping Costs on Bidder Entry and Seller Revenues in eBay Auctions," Senior Thesis, Department of Economics, Stanford University.

US Department of Commerce (2012). "Quarterly Retail E Commerce Sales." US Census Bureau News.

Xing, Xiaolin. (2010). "Can price dispersion be persistent in the internet markets?" 42.1927-1940.

Appendix

Table 1. Initial regression to compare with past analysis (Ellison and Ellison 2008)

Comparison of Regression for Purchases of 128MC PC100 Memory Modules				
Regression	**Bodnar 2012**		Ellison and Ellison 2008	
	Coefficient	**T-Value**	Coefficient	T-Value
PRICE	**-0.5404418**	**-60.95**	-0.56	64.17
SalesTax	**0.0593952**	**0.75**	0.05	0.59
HomeState	**0.4554824**	**2.23**	0.47	2.27
NeighborState	**-0.084164**	**-0.66**	-0.05	0.38
ShippingPrice	**0.0032297**	**9.49**		
ShippingTime	**-0.039846**	**-1.54**	-0.03	1.28
Page2	**-1.391047**	**-1.99**	-1.32	1.94
Weekend	**-0.4244648**	**-20.49**	-0.42	20.38
MinimumPrice	**-0.0328046**	**-13.72**	-0.03	14.09
T 1	**0.01632**	**10.49**	0.02	12.66
T2	**0.03693**	**10.39**	-0.04	11.54
T3	**0.01718**	**9.32**	0.02	10.36
T4	**-0.00318**	**5.06**	0	4
Observations		**793950**		793950
R^2		**0.03**		0.03

Notes: State and website dummy variables were included in the regression were excluded in this table for simplification and comparison purposes. T values were listed instead of P values, because Ellison and Ellison (2008) only reports T values.

Table 2. Screenshot of what a page on Pricewatch.com looks like at a random point during our analysis period.

BRAND	PRODUCT	DESCRIPTION	PRICE	SHIP	DATE/HR	DEALER/PHONE	ST	PART#
Generic	PRICE FOR ONLINE ORDERS ONLY - 128MB PC100 SDRAM DIMM - 8ns Gold leads	.- * LIMIT ONE - Easy installation - in stock	$ 68	**9.69 INSURED**	10/12/00 12:40:05 AM CST	Computer Craft Inc. 800-487-4910 727-327-7559 Online Ordering	FL	MEM-128-100PCT
Generic	ONLINE ORDERS ONLY - 128MB SDRAM PC100 16x64 168pin	- * LIMIT ONE	$ 69	**INSURED$9.95**	10/11/00 10:59:56 PM CST	Connect Computers 888-277-6287 949-367-0703 Online Ordering	CA	-
Generic	PRICE FOR ONLINE ORDER - 128MB PC100 SDRAM DIMM	- * LIMIT ONE - - InStock, 16x64-Gold Leads	$ 70	**10.75**	10/11/00 2:11:00 PM CST	1st Choice Memory 949-888-3810 -- P.O.'s accepted Online Ordering	CA	-
Generic	PRICE FOR ONLINE ORDER - 128mb True PC100 SDRAM EEPROM DIMM16x64 168pin 6ns/7ns/8ns Gold Leads	- * LIMIT ONE - in stock - with Lifetime Warranty	**$ 72**	**9.85**	10/10/00 11:30:39 AM CST	pcboost.com 800-382-6678 -- P.O.'s accepted Online Ordering	CA	-
Generic	IN STOCK, 128MB PC100 3.3volt unbuffered SDRAM Gold Lead 168 Pin, 7/8ns - with Lifetime warranty	- * LIMIT ONE Not compatible with E Machine	**$ 74**	**10.95- UPS INSURED**	10/11/00 12:44:00 PM CST	Memplus.com 877-918-6767 626-918-6767	CA	- 880060
Generic	PRICE FOR ONLINE ORDERS ONLY - 128MB True PC100 SDRAM DIMM - 8ns Gold - warranty	- * LIMIT ONE	**$ 74**	**10.25**	10/9/00 6:53:25 PM CST	Portatech 800-487-1327 .	CA	-
House Brand	128MB PC100 3.3volt SDRAM 168 Pin, 7/8ns - with LIFITIME WARRANTY	- * LIMIT ONE	**$ 74**	**10.50 FedEx**	10/11/00 10:20:23 AM CST	1st Compu Choice 800-345-8880 800-345-8880	OH	-
Generic	128MB 168Pin TRUE PC100 SDRAM - OEM 16X64	DIMM16x64 168pin 6ns/7ns/8ns Gold Leads	$ 75	**$10**	10/11/00 2:37:00 PM CST	Sunset Marketing, Inc. 800-397-5050 410-626-0211 -- P.O.'s accepted	MD	-
Generic	128MB 16x64 PC100 8ns SDRAM.	- * LIMIT ONE	$ 77	**10.90**	10/12/00 9:37:45 AM CST	PC COST 800-877-9442 847-690-0103 Online Ordering	IL	-
Generic	IN STOCK, PC100, 128MB, 168pins DIMM NonECC, - with Lifetime warranty	- * LIMIT 5	**$ 77**	**$10.95 & UP For UPS Ground**	10/9/00 5:11:10 PM CST	Augustus Technology, Inc 877-468-5181 909-468-1883 Online Ordering	CA	-
Generic	128MB PC100 8NS 16x64 SDRAM - one year warranty	- * LIMIT ONE	$ 78	**Ups Ground $10.62**	10/11/00 5:16:36 PM CST	Computer Super Sale 800-305-4930 847-640-9710 Online Ordering	IL	-
Generic	PRICE FOR ONLINE ORDERS ONLY - PC100 128MB NonBuffered, NonECC 16x64 SDRAM DIMM 3.3V 8ns mem module	- * LIMIT ONE - with lifetime warranty	**$ 78**	**10.95**	10/5/00 6:29:59 PM CST	Jazz Technology USA, LLC 888-485-8872 909-869-8859	CA	ME-GBP100128

Page 1 of 30
1 5 10 15 20 25 30 Next

Table 3. Summary statistics for orders placed by state

Orders Placed by Origin State							
128MB PC 133				128MB PC100			
postal		Freq.	Percent	Cum.	postal		Freq.	Percent	Cum.
AK	10	0.24	0.24	AK	31	0.54	0.54
AL	42	1.01	1.25	AL	82	1.43	1.98
AR	36	0.87	2.12	AR	54	0.94	2.92
AZ	113	2.72	4.83	AZ	153	2.68	5.60
CA	252	6.06	10.89	CA	342	5.98	11.58
CO	81	1.95	12.84	CO	132	2.31	13.89
CT	69	1.66	14.50	CT	72	1.26	15.15
DC	10	0.24	14.74	DC	7	0.12	15.27
DE	8	0.19	14.94	DE	16	0.28	15.55
FL	234	5.63	20.56	FL	352	6.16	21.71
GA	98	2.36	22.92	GA	170	2.97	24.69
HI	8	0.19	23.11	HI	20	0.35	25.03
IA	64	1.54	24.65	IA	89	1.56	26.59
ID	42	1.01	25.66	ID	38	0.66	27.26
IL	187	4.50	30.16	IL	257	4.50	31.75
IN	91	2.19	32.35	IN	136	2.38	34.13
KS	67	1.61	33.96	KS	87	1.52	35.65
KY	52	1.25	35.21	KY	76	1.33	36.98
LA	79	1.90	37.11	LA	127	2.22	39.21
MA	90	2.16	39.27	MA	138	2.41	41.62
MD	85	2.04	41.32	MD	93	1.63	43.25
ME	29	0.70	42.02	ME	39	0.68	43.93
MI	155	3.73	45.74	MI	187	3.27	47.20
MN	92	2.21	47.96	MN	135	2.36	49.56
MO	110	2.65	50.60	MO	135	2.36	51.92
MS	22	0.53	51.13	MS	37	0.65	52.57
MT	12	0.29	51.42	MT	22	0.38	52.96
NC	99	2.38	53.80	NC	135	2.36	55.32
ND	15	0.36	54.16	ND	28	0.49	55.81
NE	25	0.60	54.76	NE	20	0.35	56.16
NH	15	0.36	55.12	NH	34	0.59	56.75
NJ	83	2.00	57.12	NJ	124	2.17	58.92
NM	34	0.82	57.94	NM	57	1.00	59.92
NV	40	0.96	58.90	NV	56	0.98	60.90
NY	244	5.87	64.77	NY	309	5.41	66.31
OH	162	3.90	68.66	OH	194	3.39	69.70
OK	100	2.41	71.07	OK	99	1.73	71.43
OR	103	2.48	73.54	OR	120	2.10	73.53
PA	155	3.73	77.27	PA	237	4.15	77.68
RI	36	0.87	78.14	RI	47	0.82	78.50
SC	49	1.18	79.32	SC	63	1.10	79.60
SD	10	0.24	79.56	SD	23	0.40	80.00
TN	65	1.56	81.12	TN	125	2.19	82.19
TX	308	7.41	88.53	TX	406	7.10	89.29
UT	55	1.32	89.85	UT	49	0.86	90.15
VA	92	2.21	92.06	VA	141	2.47	92.62
VT	10	0.24	92.30	VT	10	0.17	92.79
WA	201	4.83	97.14	WA	255	4.46	97.25
WI	90	2.16	99.30	WI	121	2.12	99.37
WV	24	0.58	99.88	WV	20	0.35	99.72
WY	5	0.12	100.00	WY	16	0.28	100.00

Appendix 4. Regression estimates for state specific effects

PC 100 Modules

State	Coefficient	Std. Error	T Value	P Value
ST_FL |	3.007226	.1930641	15.58	0.000
ST_IL |	2.63302	.1931621	13.63	0.000
ST_OH |	2.375487	.1952965	12.16	0.000
ST_OR |	1.853614	.1996454	9.28	0.000
ST_PA |	2.525393	.1933564	13.06	0.000
ST_VA |	2.144028	.1957891	10.95	0.000
ST_WI |	1.908385	.1977363	9.65	0.000
ST_TX |	3.022677	.1917956	15.76	0.000
ST_AL |	1.379434	.2092294	6.59	0.000
ST_GA |	2.135127	.1952962	10.93	0.000
ST_AK |	.075484	.3659513	0.21	0.837
ST_AZ |	2.001224	.1964216	10.19	0.000
ST_AR |	1.200689	.2164797	5.55	0.000
ST_CO |	1.933575	.1971943	9.81	0.000
ST_CT |	.9584789	.2309397	4.15	0.000
ST_DE |	-.0133124	.3922112	-0.03	0.973
ST_DC |	-.3045006	.4965493	-0.61	0.540
ST_HI |	.0791657	.364636	0.22	0.828
ST_ID |	.9499836	.2313794	4.11	0.000
ST_IN |	1.950732	.1970998	9.90	0.000
ST_IA |	1.567356	.2036823	7.70	0.000
ST_KS |	1.345773	.2103696	6.40	0.000
ST_KY |	1.376506	.2093289	6.58	0.000
ST_LA |	1.774093	.1995122	8.89	0.000
ST_ME |	.7802408	.2460294	3.17	0.002
ST_MD |	1.660283	.2017164	8.23	0.000
ST_MA |	2.056617	.1959732	10.49	0.000
ST_MI |	2.28064	.1942457	11.74	0.000
ST_MN |	1.917485	.1974433	9.71	0.000
ST_MS |	.6180812	.2639801	2.34	0.019
ST_MO |	2.024596	.1962345	10.32	0.000
ST_MT |	-.1084294	.421844	-0.26	0.797
ST_NE |	-.035836	.3989563	-0.09	0.928
ST_NV |	1.162188	.2181988	5.33	0.000
ST_NH |	.4784637	.2835622	1.69	0.092
ST_NJ |	1.935169	.1972749	9.81	0.000
ST_NM |	1.176989	.2175333	5.41	0.000
ST_NY |	2.760336	.1924591	14.34	0.000
ST_NC |	1.866162	.198161	9.42	0.000
ST_ND |	.1826845	.3401809	0.54	0.591
ST_OK |	1.550486	.2041055	7.60	0.000
ST_RI |	.8260213	.2417221	3.42	0.001
ST_SC |	1.370689	.2095261	6.54	0.000
ST_SD |	.0277547	.3801413	0.07	0.942
ST_TN |	1.925854	.1973429	9.76	0.000
ST_UT |	.8981694	.2353534	3.82	0.000
ST_VT |	-1.035292	.9706873	-1.07	0.286
ST_WA |	2.545888	.19284	13.20	0.000
ST_WV |	.2190682	.3320091	0.66	0.509
ST_WY |	-.3845851	.5312409	-0.72	0.469

PC133 Modules

State	Coefficient	Std. Erro	T Value	P Valu
ST_FL |	1.385155	.2060182	6.72	0.000
ST_IL |	1.066425	.2059177	5.18	0.000
ST_OH |	1.124005	.2069483	5.43	0.000
ST_OR |	.4829834	.2115055	2.28	0.022
ST_PA |	.9667756	.207429	4.66	0.000
ST_VA |	.6511563	.2106542	3.09	0.002
ST_WI |	.492772	.2125029	2.32	0.020
ST_TX |	1.56085	.2041965	7.64	0.000
ST_AL |	-.3708951	.2439836	-1.52	0.128
ST_GA |	.4148945	.2124176	1.95	0.051
ST_AK |	-1.490698	.4571627	-3.26	0.001
ST_AZ |	.4786424	.2111982	2.27	0.023
ST_AR |	-.4294652	.2484613	-1.73	0.084
ST_CO |	.202522	.2168959	0.93	0.350
ST_CT |	.0149337	.2230474	0.07	0.947
ST_DE |	-2.034146	.7401807	-2.75	0.006
ST_DC |	-1.818893	.6088378	-2.99	0.003
ST_HI |	-1.971576	.6912391	-2.85	0.004
ST_ID |	-.3127058	.2395829	-1.31	0.192
ST_IN |	.2460169	.2159615	1.14	0.255
ST_IA |	.1431036	.218656	0.65	0.513
ST_KS |	.0533292	.2215595	0.24	0.810
ST_KY |	.028946	.222526	0.13	0.897
ST_LA |	.198477	.2170963	0.91	0.361
ST_ME |	-.654427	.2709438	-2.42	0.016
ST_MD |	.4752597	.2113968	2.25	0.025
ST_MA |	.3315575	.2140194	1.55	0.121
ST_MI |	.8885909	.2068956	4.29	0.000
ST_MN |	.3362218	.213843	1.57	0.116
ST_MS |	-1.071701	.3392409	-3.16	0.002
ST_MO |	.5911694	.2096788	2.82	0.005
ST_MT |	-2.238158	.8937287	-2.50	0.012
ST_NE |	-.9792697	.3203676	-3.06	0.002
ST_NV |	-.464276	.2509203	-1.85	0.064
ST_NH |	-1.63827	.5204328	-3.15	0.002
ST_NJ |	.338387	.213911	1.58	0.114
ST_NM |	-.7590538	.2839304	-2.67	0.008
ST_NY |	1.316814	.2048479	6.43	0.000
ST_NC |	.38236	.2130114	1.80	0.073
ST_ND |	-1.729568	.562413	-3.08	0.002
ST_OK |	.5856386	.2097492	2.79	0.005
ST_RI |	-.423362	.2480839	-1.71	0.088
ST_SC |	-.2319823	.2347805	-0.99	0.323
ST_SD |	-1.576389	.4937018	-3.19	0.001
ST_TN |	-.0097323	.223907	-0.04	0.965
ST_UT |	-.1371097	.2294182	-0.60	0.550
ST_VT |	-2.643276	1.324417	-2.00	0.046
ST_WA |	1.237309	.2049728	6.04	0.000
ST_WV |	-1.456228	.4480915	-3.25	0.001
ST_WY |	-2.656235	1.337583	-1.99	0.047

Appendix 6. Example of STATA Code for Discrete Choice Regression Model

```
/***************************************************************************/

/*****                                              *****/

/*****              Discrete Choice Model            *****/

/*****                                              *****/

/***************************************************************************/



/* estimates the discrete choice model                    */

/* of demand for the 128 MB pc 100 memory modules                */

/* program runs under stata 8.0

/* based on program written by for Ellison and Ellison analysis but editted to include shipping prices  for
KDB analysis                    */

capture clear
program drop _all
set more off
set mem 1g
capture log close
log using "f:/Output/KDBlog1.dta", replace
use "f:/Output/nldata123100merged.dta"
sort numdate h postal cnum
for num 1/24: gen byte bX=0
for num 1/24: replace bX=1 if cnum==2 & crank==X
for num 1/24: replace bX=1 if cnum[_n+1]==2 & crank[_n+1]==X & h==h[_n+1] &
numdate==numdate[_n+1]
drop totalPr*
generate cshippr = 0
foreach f in 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 {
replace cshippr = ShipPr`f' if crank==`f'
}
program define nlbase9
  version 7.0
  if "`1'"=="?" {
   global S_1 "HOME BOOST Taxwt PROX WKEND MINP PRICE SHIPPR PAGE2 SHIPTIME ST_FL
ST_IL ST_OH ST_OR ST_PA ST_VA ST_WI ST_TX ST_AL ST_GA ST_AK ST_AZ ST_AR ST_CO
ST_CT ST_DE ST_DC ST_HI ST_ID ST_IN ST_IA ST_KS ST_KY ST_LA ST_ME ST_MD ST_MA
ST_MI ST_MN ST_MS ST_MO ST_MT ST_NE ST_NV ST_NH ST_NJ ST_NM ST_NY ST_NC
ST_ND ST_OK ST_RI ST_SC ST_SD ST_TN ST_UT ST_VT ST_WA ST_WV ST_WY T1 T4 T7 T11"

   global HOME=0.477

   global Taxwt=0.050
   global PROX=-0.054
   global WKEND=-0.42
   global MINP=-0.0327
   global PRICE=-.566
global SHIPPR=-.1
```

```
global PAGE2=-1.14
global SHIPTIME=-0.035
global BOOST=-.169
global ST_FL=2.87
global ST_IL=2.49
global ST_OH=2.23
global ST_OR=1.70
global ST_PA=2.38
global ST_VA=2.02
global ST_WI=1.77
global ST_TX=2.86
global ST_AL=1.24
global ST_GA=2.00
global ST_AK=-0.09
global ST_AZ=1.84
global ST_AR=1.05
global ST_CO=1.78
global ST_CT=0.85
global ST_DE=-0.17
global ST_DC=-0.48
global ST_HI=-0.085
global ST_ID=0.77
global ST_IN=1.81
global ST_IA=1.43
global ST_KS=1.21
global ST_KY=1.24
global ST_LA=1.63
global ST_ME=0.66
global ST_MD=1.52
global ST_MA=1.91

global ST_MI=2.14

global ST_MN=1.77

global ST_MS=0.47

global ST_MO=1.89
global ST_MT=-0.25
global ST_NE=-0.20
global ST_NV=1.00
global ST_NH=0.34
global ST_NJ=1.79
global ST_NM=1.00
global ST_NY=2.63
global ST_NC=1.74
global ST_ND=0.05
global ST_OK=1.39
global ST_RI=0.69
global ST_SC=1.24
global ST_SD=-0.13
global ST_TN=1.78
global ST_UT=0.74
global ST_VT=-1.19
global ST_WA=2.38
global ST_WV=0.10
```

```
  global ST_WY=-0.56
  global T1=0.0185
  global T4=-0.040
  global T7=0.018
  global T11=-0.00236
  exit
}

tempvar m s

quietly gen double `m'=exp($ST_FL) if postal=="FL"

quietly replace `m'=exp($ST_IL) if postal=="IL"

quietly replace `m'=exp($ST_OH) if postal=="OH"

quietly replace `m'=exp($ST_OR) if postal=="OR"

quietly replace `m'=exp($ST_PA) if postal=="PA"

quietly replace `m'=exp($ST_VA) if postal=="VA"

quietly replace `m'=exp($ST_WI) if postal=="WI"

quietly replace `m'=exp($ST_TX) if postal=="TX"

quietly replace `m'=exp($ST_AL) if postal=="AL"

quietly replace `m'=exp($ST_GA) if postal=="GA"


quietly replace `m'=exp($ST_AK) if postal=="AK"

quietly replace `m'=exp($ST_AZ) if postal=="AZ"

quietly replace `m'=exp($ST_AR) if postal=="AR"

quietly replace `m'=exp($ST_CO) if postal=="CO"

quietly replace `m'=exp($ST_CT) if postal=="CT"

quietly replace `m'=exp($ST_DE) if postal=="DE"

quietly replace `m'=exp($ST_DC) if postal=="DC"

quietly replace `m'=exp($ST_HI) if postal=="HI"

quietly replace `m'=exp($ST_ID) if postal=="ID"

quietly replace `m'=exp($ST_IN) if postal=="IN"

quietly replace `m'=exp($ST_IA) if postal=="IA"

quietly replace `m'=exp($ST_KS) if postal=="KS"
```

```
quietly replace `m'=exp($ST_KY) if postal=="KY"
quietly replace `m'=exp($ST_LA) if postal=="LA"
quietly replace `m'=exp($ST_ME) if postal=="ME"
quietly replace `m'=exp($ST_MD) if postal=="MD"
quietly replace `m'=exp($ST_MA) if postal=="MA"
quietly replace `m'=exp($ST_MI) if postal=="MI"
quietly replace `m'=exp($ST_MN) if postal=="MN"
quietly replace `m'=exp($ST_MS) if postal=="MS"
quietly replace `m'=exp($ST_MO) if postal=="MO"
quietly replace `m'=exp($ST_MT) if postal=="MT"
quietly replace `m'=exp($ST_NE) if postal=="NE"
quietly replace `m'=exp($ST_NV) if postal=="NV"
quietly replace `m'=exp($ST_NH) if postal=="NH"
quietly replace `m'=exp($ST_NJ) if postal=="NJ"
quietly replace `m'=exp($ST_NM) if postal=="NM"
quietly replace `m'=exp($ST_NY) if postal=="NY"
quietly replace `m'=exp($ST_NC) if postal=="NC"
quietly replace `m'=exp($ST_ND) if postal=="ND"
quietly replace `m'=exp($ST_OK) if postal=="OK"
quietly replace `m'=exp($ST_RI) if postal=="RI"
quietly replace `m'=exp($ST_SC) if postal=="SC"
quietly replace `m'=exp($ST_SD) if postal=="SD"
quietly replace `m'=exp($ST_TN) if postal=="TN"
quietly replace `m'=exp($ST_UT) if postal=="UT"
quietly replace `m'=exp($ST_VT) if postal=="VT"
quietly replace `m'=exp($ST_WA) if postal=="WA"
quietly replace `m'=exp($ST_WV) if postal=="WV"
quietly replace `m'=exp($ST_WY) if postal=="WY"
```

```
  quietly replace `m'=`m'*hshare

  quietly replace `m'=`m'*exp($T1*t1+$T4*t4+$T7*t7+$T11*t11)

  quietly replace `m'=`m'*exp($MINP*price1)

  quietly replace `m'=`m'*exp($WKEND*weekend)




  quietly gen double
`s'=exp($HOME*home1+$PRICE*(price1*(1+$SHIPPR*ShipPr1+$Taxwt*tax*home1))+$SHIPTIME*sh
ip1+$PROX*prox1+$BOOST*b1)

  quietly replace
`s'=`s'+exp($HOME*home2+$PRICE*(price2*(1+$SHIPPR*ShipPr2+$Taxwt*tax*home2))+$SHIPTIME
*ship2+$PROX*prox2+$BOOST*b2)

  quietly replace
`s'=`s'+exp($HOME*home3+$PRICE*(price3*(1+$SHIPPR*ShipPr3+$Taxwt*tax*home3))+$SHIPTIME
*ship3+$PROX*prox3+$BOOST*b3)

  quietly replace
`s'=`s'+exp($HOME*home4+$PRICE*(price4*(1+$SHIPPR*ShipPr4+$Taxwt*tax*home4))+$SHIPTIME
*ship4+$PROX*prox4+$BOOST*b4)

  quietly replace
`s'=`s'+exp($HOME*home5+$PRICE*(price5*(1+$SHIPPR*ShipPr5+$Taxwt*tax*home5))+$SHIPTIME
*ship5+$PROX*prox5+$BOOST*b5)

  quietly replace
`s'=`s'+exp($HOME*home6+$PRICE*(price6*(1+$SHIPPR*ShipPr6+$Taxwt*tax*home6))+$SHIPTIME
*ship6+$PROX*prox6+$BOOST*b6)

  quietly replace
`s'=`s'+exp($HOME*home7+$PRICE*(price7*(1+$SHIPPR*ShipPr7+$Taxwt*tax*home7))+$SHIPTIME
*ship7+$PROX*prox7+$BOOST*b7)

  quietly replace
`s'=`s'+exp($HOME*home8+$PRICE*(price8*(1+$SHIPPR*ShipPr8+$Taxwt*tax*home8))+$SHIPTIME
*ship8+$PROX*prox8+$BOOST*b8)

  quietly replace
`s'=`s'+exp($HOME*home9+$PRICE*(price9*(1+$SHIPPR*ShipPr9+$Taxwt*tax*home9))+$SHIPTIME
*ship9+$PROX*prox9+$BOOST*b9)

  quietly replace
`s'=`s'+exp($HOME*home10+$PRICE*(price10*(1+$SHIPPR*ShipPr10+$Taxwt*tax*home10))+$SHIPT
IME*ship10+$PROX*prox10+$BOOST*b10)
```

```
  quietly replace
`s'=`s'+exp($HOME*home11+$PRICE*(price11*(1+$SHIPPR*ShipPr11+$Taxwt*tax*home11))+$SHIPT
IME*ship11+$PROX*prox11+$BOOST*b11)

  quietly replace
`s'=`s'+exp($HOME*home12+$PRICE*(price12*(1+$SHIPPR*ShipPr12+$Taxwt*tax*home12))+$SHIPT
IME*ship12+$PROX*prox12+$BOOST*b12)

quietly replace
`s'=`s'+exp($HOME*home13+$PAGE2+$PRICE*(price13*(1+$SHIPPR*ShipPr13+$Taxwt*tax*home13)
)+$SHIPTIME*ship13+$PROX*prox13+$BOOST*b13)

  quietly replace
`s'=`s'+exp($HOME*home14+$PAGE2+$PRICE*(price14*(1+$SHIPPR*ShipPr14+$Taxwt*tax*home14)
)+$SHIPTIME*ship14+$PROX*prox14+$BOOST*b14)

  quietly replace
`s'=`s'+exp($HOME*home15+$PAGE2+$PRICE*(price15*(1+$SHIPPR*ShipPr15+$Taxwt*tax*home15)
)+$SHIPTIME*ship15+$PROX*prox15+$BOOST*b15)

  quietly replace
`s'=`s'+exp($HOME*home16+$PAGE2+$PRICE*(price16*(1+$SHIPPR*ShipPr16+$Taxwt*tax*home16)
)+$SHIPTIME*ship16+$PROX*prox16+$BOOST*b16)

  quietly replace
`s'=`s'+exp($HOME*home17+$PAGE2+$PRICE*(price17*(1+$SHIPPR*ShipPr17+$Taxwt*tax*home17)
)+$SHIPTIME*ship17+$PROX*prox17+$BOOST*b17)

  quietly replace
`s'=`s'+exp($HOME*home18+$PAGE2+$PRICE*(price18*(1+$SHIPPR*ShipPr18+$Taxwt*tax*home18)
)+$SHIPTIME*ship18+$PROX*prox18+$BOOST*b18)

  quietly replace
`s'=`s'+exp($HOME*home19+$PAGE2+$PRICE*(price19*(1+$SHIPPR*ShipPr19+$Taxwt*tax*home19)
)+$SHIPTIME*ship19+$PROX*prox19+$BOOST*b19)

  quietly replace
`s'=`s'+exp($HOME*home20+$PAGE2+$PRICE*(price20*(1+$SHIPPR*ShipPr20+$Taxwt*tax*home20)
)+$SHIPTIME*ship20+$PROX*prox20+$BOOST*b20)

  quietly replace
`s'=`s'+exp($HOME*home21+$PAGE2+$PRICE*(price21*(1+$SHIPPR*ShipPr21+$Taxwt*tax*home21)
)+$SHIPTIME*ship21+$PROX*prox21+$BOOST*b21)

  quietly replace
`s'=`s'+exp($HOME*home22+$PAGE2+$PRICE*(price22*(1+$SHIPPR*ShipPr22+$Taxwt*tax*home22)
)+$SHIPTIME*ship22+$PROX*prox22+$BOOST*b22)

  quietly replace
`s'=`s'+exp($HOME*home23+$PAGE2+$PRICE*(price23*(1+$SHIPPR*ShipPr23+$Taxwt*tax*home23)
)+$SHIPTIME*ship23+$PROX*prox23+$BOOST*b23)

  quietly replace
`s'=`s'+exp($HOME*home24+$PAGE2+$PRICE*(price24*(1+$SHIPPR*ShipPr24+$Taxwt*tax*home24)
)+$SHIPTIME*ship24+$PROX*prox24+$BOOST*b24)
```

```
  quietly replace
`m'=`m'*exp($PRICE*(cprice*(1+$SHIPPR*cshippr+$Taxwt*tax*chome))+$SHIPTIME*cship)

  quietly replace `m'=`m'*exp($HOME*chome)
  quietly replace `m'=`m'*exp($PROX*cprox)
  quietly replace `m'=`m'*exp($BOOST) if cnum==2
  quietly replace `m'=`m'*exp($PAGE2) if crank > 12
  quietly replace `m'=`m'/`s'
  replace `1'=`m'
end

nl base9 norder if postal~="CA"


predict qpred
gen double npred=exp(_b[ST_FL]) if postal=="FL"
replace npred=exp(_b[ST_IL]) if postal=="IL"
replace npred=exp(_b[ST_OH]) if postal=="OH"
replace npred=exp(_b[ST_OR]) if postal=="OR"
replace npred=exp(_b[ST_PA]) if postal=="PA"
replace npred=exp(_b[ST_VA]) if postal=="VA"
replace npred=exp(_b[ST_WI]) if postal=="WI"
replace npred=exp(_b[ST_TX]) if postal=="TX"
replace npred=exp(_b[ST_AL]) if postal=="AL"
replace npred=exp(_b[ST_GA]) if postal=="GA"
replace npred=exp(_b[ST_AK]) if postal=="AK"
replace npred=exp(_b[ST_AZ]) if postal=="AZ"
replace npred=exp(_b[ST_AR]) if postal=="AR"
replace npred=exp(_b[ST_CO]) if postal=="CO"
replace npred=exp(_b[ST_CT]) if postal=="CT"
replace npred=exp(_b[ST_DE]) if postal=="DE"
replace npred=exp(_b[ST_DC]) if postal=="DC"
replace npred=exp(_b[ST_HI]) if postal=="HI"
replace npred=exp(_b[ST_ID]) if postal=="ID"
replace npred=exp(_b[ST_IN]) if postal=="IN"
replace npred=exp(_b[ST_IA]) if postal=="IA"
replace npred=exp(_b[ST_KS]) if postal=="KS"
replace npred=exp(_b[ST_KY]) if postal=="KY"
replace npred=exp(_b[ST_LA]) if postal=="LA"
replace npred=exp(_b[ST_ME]) if postal=="ME"
replace npred=exp(_b[ST_MD]) if postal=="MD"
replace npred=exp(_b[ST_MA]) if postal=="MA"
replace npred=exp(_b[ST_MI]) if postal=="MI"
replace npred=exp(_b[ST_MN]) if postal=="MN"
replace npred=exp(_b[ST_MS]) if postal=="MS"
replace npred=exp(_b[ST_MO]) if postal=="MO"
replace npred=exp(_b[ST_MT]) if postal=="MT"
replace npred=exp(_b[ST_NE]) if postal=="NE"
replace npred=exp(_b[ST_NV]) if postal=="NV"
replace npred=exp(_b[ST_NH]) if postal=="NH"
replace npred=exp(_b[ST_NJ]) if postal=="NJ"
replace npred=exp(_b[ST_NM]) if postal=="NM"
replace npred=exp(_b[ST_NY]) if postal=="NY"
replace npred=exp(_b[ST_NC]) if postal=="NC"
replace npred=exp(_b[ST_ND]) if postal=="ND"
replace npred=exp(_b[ST_OK]) if postal=="OK"
```

```
replace npred=exp(_b[ST_RI]) if postal=="RI"
replace npred=exp(_b[ST_SC]) if postal=="SC"
replace npred=exp(_b[ST_SD]) if postal=="SD"
replace npred=exp(_b[ST_TN]) if postal=="TN"
replace npred=exp(_b[ST_UT]) if postal=="UT"
replace npred=exp(_b[ST_VT]) if postal=="VT"
replace npred=exp(_b[ST_WA]) if postal=="WA"
replace npred=exp(_b[ST_WV]) if postal=="WV"
replace npred=exp(_b[ST_WY]) if postal=="WY"

replace npred=npred*hshare
replace npred=npred*exp(_b[T1]*t1+_b[T4]*t4+_b[T7]*t7+_b[T11]*t11)
replace npred=npred*exp(_b[MINP]*price1)
replace npred=npred*exp(_b[WKEND]*weekend)

gen share1=qpred/npred
gen share2=norder/npred
sum share1 share2 if postal~="CA"
sum share1 share2 if postal~="CA" & cnum==1
sum share1 share2 if postal~="CA" & cnum==2
sort localh
tab localh, sum(hshare)
for num 1/24: gen dtaxX=homeX*priceX*tax
egen mhome=rmean(home1-home24)
egen mprox=rmean(prox1-prox24)
egen mdtax=rmean(dtax1-dtax24)
sum norder cprice price1 crank mhome chome mprox cprox mdtax if postal~="CA"

log close
```