Tax Sensitivity and Home State Preferences in Internet Purchasing[1]

Glenn Ellison
MIT and NBER

and

Sara Fisher Ellison
MIT

Preliminary and Incomplete

January 2003

[1]We would like to thank Andrew Sweeting and Eric Moos for excellent research assistance, Patrick Goebel and Steve Ellison for assistance with data issues, and Austan Goolsbee and John Morgan for helpful comments. We also thank the NSF for financial support via grant SES-0219205. Finally, G. Ellison would like to thank the Center for Advanced Studies in the Behavioral Sciences and S.F. Ellison would like to thank the Hoover Institute.

1 Introduction

The growth of Internet commerce in the U.S. has been spurred by the *de facto* tax-free status of interstate purchases.[1] Internet taxation was hotly debated during the Internet boom when Internet advocates feared taxes could kill the most important infant industry in history and others feared that traditional retail might be completely wiped out. It remains hotly debated today (albeit with less hyperbole) as states look to make up revenue shortfalls caused by the current recession. An amendment that would have empowered states to collect use taxes from e-retailers if a sufficiently large group of states adopted a uniform tax code was narrowly defeated in 2001. A coalition of more than 30 states has now drafted such a code and will raise the issue again when the Internet Nondiscrimination Act comes up for renewal in late 2003 if not before. In the meantime, a number of states have stepped up efforts to collect "use taxes" from their residents, and California has led the fight to broaden the definition of nexus and force firms to collect use taxes.[2] The impact of such policies would depend on the extent to which consumers substitute between online and offline retailers and between different online retailers in response to tax differences.

There has been little academic work on sensitivity of online sales to taxation. The paper that has had the most substantial impact on the academic and nonacademic world is Goolsbee (2001).[3] It examines a 1997 survey in which 25,000 consumers were asked (among many other things) whether they had ever bought products online. Consumers living in states with higher sales tax rates are found to be more likely to have bought products online. The big-picture conclusion is that subjecting e-retailers to taxation could

[1] Forty five U.S. states levy sales taxes on traditional retail purchases. Each of these states also has laws assessing "use taxes" on purchases that its residents make from out-of-state firms. However, the Supreme Court ruled in *Quill vs. North Dakota* (1992) that absent new federal law, a state could not compel a firm without substantial physical "nexus" in that state to collect use taxes on its behalf. The 1998 Internet Tax Freedom Act makes explicit that web presence alone does not constitute nexus. While consumers are obligated to self-report use-tax liability, use taxes are rarely paid in practice. Note that states are able to collect sales taxes on e-retailers' in-state sales.

[2] Massachusetts, for example, added a line to its 2002 state income tax to allow consumers to report use tax obligations there. In September of 2002 the California Board of Equalization ruled that a promotion that involved Barnes & Noble.com paying Barnes and Noble Bookstores to distribute coupons and logo shopping bags was sufficient to established the nexus of the online retailer. A year earlier it has used the fact that items ordered at Borders.com could be returned to Borders bookstores to deny Borders.com's claim that it had no nexus.

[3] The work was presented to the Advisory Commission on Electronic Commerce created by the Internet Tax Freedom Act. A google search for "Goolsbee, Internet, tax" returns more than 1000 matches.

reduce online sales by 24%.[4] Another related study is Brynjolfsson and Smith (2001), which examines the behavior of a set of consumers who visited an Internet shopbot in 1999 using clickstream data. Its puzzling finding is that consumers are estimated to be twice as sensitive to differences in taxes as they are to differences in item prices. Also related is a small literature on the impact of sales tax changes on sales and retail location in border regions. See, for example, Fox (1986) and Walsh and Jones (1988).

In this paper we examine the impact of taxes on e-retail sales in the context of a particular segment of e-retail: a group of small firms selling computer parts via the price search engine, Pricewatch.com. The geographically distinguished price and quantity data available to us at the transaction level allow us to contribute to the literature in a number of ways. First, we provide additional evidence that the pattern identified in Goolsbee (2000) is is not an artifact of unobserved consumer heterogeneity. Second, our estimates are derived from actual sales data instead of consumer recollection and reflect quantities sold not just numbers of people who have bought something. Third, we examine both aggregate sales data and substitution between online retailers. And fourth, we attempt to separate tax effects from preferences that consumers may have for purhasing from firms in their home state. The particular elasticites we estimate obviously apply only to some e-retail segments: the firms we study are minimally differentiated, the consumers who find them are sophisticated Internet users, and, as we showed in Ellison and Ellison (2002), demand is extremely price sensitive. We feel, however, that some of our conclusions are likely to be much more broadly applicable.

This paper uses the same sources of data as Ellison and Ellison (2002). The quantity data come from a single firm that operates two different websites selling memory modules and other computer parts. For a period of approximately one year (May 2000-May 2001) we have data on each memory module sold including the time and date of the sale and the state to which the product was shipped. We combine this with price data we downloaded at hourly frequency over the same time period from Pricewatch.com. Most consumers buying

[4] The calculation assumes that the effect of subjecting e-retailers to taxation would be equivalent to eliminating sales taxes on offline purchases and implicitly that the same elasticity that describes whether consumers ever purchase in a consumer-weighted decision describes how demand changes in the quantity-weighted calculation.

from these websites learn of them by conducting a price search on Pricewatch.com. Our price data contain the twelve or twenty-four lowest prices listed on Pricewatch in each hour, the names of the websites offering each price, and the state in which each of the e-retailers is located.[5] The data allow two largely distinct ways to estimate the impact of sales taxes on Internet purchases. First, similarly to Goolsbee (2000), we can examine the degree to which state sales taxes affect the total sales made by the firm we study to each state. Second, similarly to Brynjolfsson and Smith (2001), we can examine the way in which differences in taxes assessed on purchases from various websites affect the substitution between websites.

Our first set of analyses is straightforward. In the simplest version we regress the total number of memory modules shipped to each state on state-level demographic controls and each state's sales tax rate. We find that our firm's sales are disproportionately high in states with high sales taxes. The magnitude of the effect is even larger than that reported in Goolsbee (2000): a one percentage point increase in a state's sales tax rate is estimated to increase online purchasing by the states' residents by seven to eight percent. Consumers are not randomly assigned to states, so a worry in any study of this form is that differences in purchasing patterns may be due to differences in consumer characteristics across states. For example, the results could be driven by the fact that California and Washington have high sales taxes in addition to populations inclined to use the Internet.[6]

To help assess whether the effects are due to tax rates or unobserved consumer characteristics, we do several things: we use cross-product comparisons to examine whether tax rates have a larger effect on sales of more expensive products; we make within-product cross-time comparisons to examine whether tax rates have a larger effect in the early part of the data when prices were much higher; and we examine how sales to California (on which our firm must assess sales tax) compare with sales to other states. The most striking result consistent with the hypothesis that we have found a tax effect is that our firm's sales to

[5] The firms advertising on Pricewatch are typically small firms and we assume that each has tax nexus only in one state.

[6] Despite the examples of California and Washington, sales taxes in the U.S. are, in fact, not positively correlated with the demographic controls for computer usage we employ. For example, Louisiana, Tennessee, Oklahoma, and Alabama each have both one of the eight highest average tax rates in the country and a below average fraction of households with home Internet access. Goolsbee casts doubt on the unobserved heterogeneity hypothesis in his data by using extensive household-level demographic controls, by including MSA dummies, and by showing that tax rates are not correlated with ownership of computers.

California customers are about 60 percent lower than what would we would have predicted (if these sales were also tax-free) using the data on sales in other states.

Our second set of analyses focuses on substitution between competing online retailers. Here we include state fixed effects, so we are not using any of the variation used in our first set of analyses to identify tax sensitivity. Instead, we are identifying tax sensitivity off of consumer subsitution patterns that result from variation in the home states of the firms they choose among. There is tremendous turnover in Pricewatch price rankings.[7] There is a lot of entry and exit in our sample, but the majority of the changes are probably due to firms trying to keep up with the substantial volatility in wholesale memory prices and to firms becoming dissatisfied with their position on the Pricewatch list as other firms jump above or below them.[8] The fact that we observe many different orderings of the firms on the Pricewatch list is potentially very informative about tax effects. Treating our firm's hourly sales to one particular state as the dependent variable, we can then, for example, examine whether our firm sells more or less memory to Pennsylvania customers when some of its close competitors (in pretax price) are Pennsylvania firms. The fact that different states have different tax rates allows us to separately identify a tax effect (assumed to be proportional to the tax owed) from a general preference for buying from in-state firms (assumed to be constant across states). For example, if we found that our firm's sales to Oregon are reduced when one of

the competing firms is in Oregon (which has no sales tax) and that its sales to Texas (which has a high sales tax) are higher when one of the competing firms is in Texas, this would suggest that consumers like buying from local firms but dislike paying taxes.

Our results vary somewhat from product to product. Among our conclusions are that consumers do not pay as much attention to sales tax differences as they do to price differences, that sales taxes nonetheless reduce sales by in-state firms substantially in some cases, and that apart from the tax disadvantage, consumers have a preference for buying from in-state firms.

[7] In the case of 128MB PC100 memory modules, for example, an average of three of the 24 lowest-priced firms will change prices each hour.

[8] One hundred thirty-eight websites appear in the 128MB PC100 top 24 list at some point during the year. Wholesale prices drop by about 65% from the start of the year to the end in addition to having high frequency volatility.

2 Data

For this study, we are focusing on four different types of memory modules, 128MB PC100, 128MB PC133, 256MB PC100, and 256MB PC133. In Ellison and Ellison (2002) we noted that memory modules also come in different "qualities" and examined the substitution between quality levels. In this paper we will simply focus on sales of generic memory modules. Most of the memory sold is of the low quality, and it is also the only quality level for which one can get price information from Pricewatch. We downloaded the first (or first and second) screens from Pricewatch's memory price lists on an hourly basis from from May 2000 to May 2001 (with some gaps). Our data on the 128MB modules include information on the twenty four lowest-priced websites listed on Pricewatch. The data on 256MB modules include information on the twelve lowest price websites. Pricewatch does not calculate sales taxes for consumers on these pages, but it does list the home state of each retailer so that a consumer who knew the tax rate in his home state (and understood that sales taxes will apply if and only if he or she buys from an in-state firm) could take sales tax differences into account. This feature of the Pricewatch rankings also allowed us to gather the state for each firm listing a product.

We obtained sales data for these products from an Internet retailer that gets most of its traffic from Pricewatch. It operates two different (but similar) websites, which typically have different prices for the products studied.[9] The sales data again cover May 2000 to May 2001 with some gaps. The raw data are at the level of the individual order and indicate from which website the customer made the purchase. In addition to specific information about the products ordered and prices paid, we have the shipping address for every order.

We also use a few state-level variables. The most important of these is the state's average sales tax rate. Sales tax rates vary by county and locality in many states. Our data are averages across the various jurisdictions within a state computed by a private firm. Our other state level variables come from Census Bureau datasets: the fraction of households with home Internet access as reported in a 2001 survey, the population of each

[9] Among the motivations for having multiple websites are that they may be given different looks and consumers may have heterogeneous reactions, that it allows the websites to be more specialized (which seems to be attractive to some consumers), that it facilitates experimentation, that it may help promote private-label branded products and that the firm may occupy multiple places on the Pricewatch screen.

state in the 2000 census, and the number of computer stores and gas stations reported in the 1997 Census of Retail Industries. For the state-level analysis, we compute the total sales of each size of memory module in each state aggregating over the full time period and over the two websites.[10]

Table 1 shows summary statistics for the state-level analysis. Each has 51 observations. The firm sells 204 128MB memory modules to a typical state over the course of the year. This ranges from a low of 19 in the District of Columbia to a high of 762 in Texas. Unit sales of 256MB memory modules are about half as large. The average sales tax rate is 5.7 percent. Four states have no state or local sales taxes. The percentage of households with home Internet access varies from a low of 40.6% in the District of Columbia to a high of 70.2% in New Hampshire. The average state has 230 computer stores. The ratio of computer stores to gas stations ranges from a low of 0.041 in West Virginia to a high of 0.184 in California. The mean price of a 128MB memory module is $70. The mean price of a 256MB memory module is $139. Although prices are not used in this state-level anaysis, they are relevant for our interpretation of the results. In particular, we would expect any estimated effect of tax rates from the 256MB modules to be quite a bit larger than estimates from the 128MB modules because the 256MB modules are approximately twice as expensive.

Unlike in the state-level regressions, for the individual-level discrete choice analysis, we can include many variables specific to the transaction, which could help explain the overall level of sales at a particular time or the consumers choice between possible products. These variables are listed in Table 2. The table reports summary statistics separately for 128MB PC100 and PC133 memory modules.[11] We have transactions data from two websites operated by our firm, so the unit of observation is hour-state-website. For instance, the number of 128MB PC100 modules we would expect one of the two websites run by our firm to sell in an hour to a state is 0.013.[12] *Price* is the price at which transactions with our

[10] Note that we aggregate over the PC100/PC133 distinction, but keep 128MB and 256MB memory chips separate. We do this because prices are very similar across the two speeds, but differ significantly between 128MB and 256MB modules, and we want to see if tax rate differences matter more for more expensive products.

[11] We have only estimated the individual-level model on 128MB modules so far but plan to add 256MB results in the next draft.

[12] We count a single order of multiple memory modules as having quantity one. For most of our time

firm took place whereas $MinPrice$ is the lowest market price for each hour. Along with $Weekend$, $MinPrice$ will help us predict overall market size. Note our firm has an average rank of around 6 for both types of modules.

The variables below the line are defined for each hour, each state, and each of the 24 websites listed. We know listed price, whether there is a match between the state of the observation and the state of the website, and the total sales tax that would be due for each potential transaction. Note that mean $SalesTax$ for potential transactions is only \$.16 or \$.18, depending on the type of module, but if we condition on $HomeState = 1$, the means are \$5.04 and \$5.14, respectively.

3 Aggregate sales

The first column of Table 3 presents coefficients obtained from estimating the regression

$$\begin{aligned}\log \frac{Quantity_s}{Population_s} &= \beta_0 + \beta_1 SalesTaxRate_s + \beta_2 California_s + \beta_3 InternetAccess_s \\ &\quad + \beta_4 \frac{ComputerStores_s}{GasStations_s} + \beta_5 \log(Population_s) + \epsilon_s\end{aligned}$$

with the log of per capita 128MB memory module sales as the dependent variable. Most of the estimated coefficients are highly significant.

The 7.4 estimated coefficient on $SalesTaxRate$ indicates that every percentage point increase in a state's sales tax is predicted to increase the quantity sold to that state by more than 7%. Goolsbee estimates the potential impact of applying sales taxes to Internet purchases by assuming that it would be equivalent to eliminating offline sales taxes. If we make the same assumption to evaluate the impact of use taxes, we would find that if a state with the mean tax rate applied a use tax, its residents would reduce their online purchases from our e-retailer by about 35%.

Suppose this estimated effect is a tax effect and not unobserved heterogeneity. What would we predict about our firm's sales to California? First, since our firm has no tax advantage relative to brick and mortar in California—its California customers must pay sales tax—we would expect its sales to be 42% less than would be predicted given California's demographics. (Our firm's California sales are taxed at 7.25%.) Second, our firm has

period, our firm limited purchases of memory modules to one per order.

a disadvantage relative to non-California firms when selling in California. Our regression does not speak to this effect, but we would expect it to be substantial. The estimated coefficient on $California$ is consistent with these observations. It indicates that sales to California customers are about 60% lower than we would have predicted given California's demographics. In addition, we think that the estimated coefficient on $California$ provides strong evidence that taxes do matter in consumer behavior. Even if the coefficient on $SalesTaxRate$ were due entirely to unobserved heterogeneity, the low sales in California would be due to the second effect, also a tax effect.

The coefficients on the control variables seem reasonable. Sales are higher in states where the fraction of residents with Internet access is higher. We cannot reject that the coefficient is one, which would correspond with sales being proportional to the number of people with home Internet access. The coefficient on the computer store-gas station ratio might have either sign: it reflects both interest in computers and the availability of computer parts at traditional retail stores. We find that the estimated coefficient is positive and significant, suggesting that this variable proxies for interest in computers. The population of a state may reflect the quality of retail offerings, consistent with its negative and significant estimated coefficient.

The second column of Table 3 presents coefficient estimates from a similar regression using 256MB memory sales rather than 128MB memory sales as the dependent variable. The coefficient estimates are quite similar to those in the first column. The fact that the coefficients on the tax rate and the California dummy are not larger than in the first column is unexpected given that the sales-weighted mean price of a 256MB memory module is about 60% higher than 128MB ones.

We do, however, investigate another specification where we see if the magnitude of the tax effect depends on the size of the purchase. Table 4 presents results obtained by estimating the above regressions on two subsamples each. The subsamples are constructed based on whether the lowest price listed on Pricewatch when each purchase was made was in the indicated ranges. For both sizes of chip, the low range is near the end of our data period and the high range is near the beginning, mostly in the summer of 2000. In each case our estimates are that the coefficient on the sales tax rate is substantially larger in the period when the items are more expensive. This is further evidence that the cross-state patterns are due to tax effects rather than some sort of unobserved heterogeneity.

4 Substitution between e-retailers

In the section we exploit the turnover in Pricewatch rankings to assess the extent to which consumers pay attention to tax differences when choosing between online retailers. We also offer evidence on whether consumers (absent tax considerations) prefer buying from local e-retailers. Intuitively, we exploit the fact that the sales by our websites to consumers in a given state depend on how attractive it is relative to other websites listed on Pricewatch, and that a ranking of websites by their mean attractiveness would often be dramatically different if consumers did or did not take tax differences into account. The churning of the Pricewatch list gives us the opportunity to see our websites' demand when its price is slightly above and below those of firms from many different states. The fact that we can observe these competitive effects in both high-tax and low-tax states and at different times when prices are higher and lower allows us to separately estimate both a tax effect and a home-state preference (assumed to be state-independent and independent of the item price).

Figure 1 gives two examples of Pricewatch rankings. It shows the twelve e-retailers listed on the first screen of Pricewatch's 128MB PC100 memory page at 9am and 11am on August 1, 2000. Two of the e-retailers made price changes between these two times. Coast to Coast Memory of New Jersey, which offered the lowest price of $112 at 9am, raised its price sufficiently so as to disappear from the top twelve by 11am. UpgradePlanet.com of Virginia, which was on the second page of the 9am list at $128, reduced its price to $111 and took over the top slot. The first three columns show information presented on Pricewatch: the websites' names, the states in which they are located and their prices. In the fourth through the sixth columns we've added additional information consumers would have to compute for themselves: the tax-inclusive prices that customers in New Jersey, Virginia, and California, respectively, would pay if they purchased from each of the e-retailers.

Suppose for the sake of this discussion that our sales data included sales by Connect

Computers.[13] At 9am Connect Computers' tax-inclusive price for New Jersey residents is lower than that of any other website. At 11am Connect Comptuters' tax-inclusive price for New Jersey residents is only the second lowest. Accordingly, if consumers pay attention to sales taxes we would expect Connect Computers' sales into New Jersey (controlling for hour effects, etc.) to be higher at 9am than at 11am. Similarly, we would expect its sales into Virginia to be higher at 11am than at 9am.

The fact that we will observe Connect Computers' sales when it is in many different positions relative to its competitors makes it easy to precisely estimate consumers' reactions to price differences between Connect Computors and its competitors. If we did not want to allow for the possibility that consumers might have a preference for buying from an in-state firm, then using data on either the extent to which Connect is making more sales in Virginia at 11am than at 9am or the extent to which Connect is making fewer sales in New Jersey at 11am than at 9am would allow us to estimate whether consumers fully account for tax differences, ignore tax differences, or do something in between. Two sources of variation allow us to do the same allowing for a fixed home state preference: we can take advantage of the fact that different states have different tax rates and identify the tax effect off the difference between the Virginia increase and the New Jersey decrease; and we can exploit the fact that as prices change over time, the number of dollars in taxes that New Jersey and Virginia residents have to pay to buy from their local firms is changing. The home state preference itself is easily identified by sales levels. For example, if consumers fully account for tax differences, then Connect Computers' sales to Virginia residents would in expectation be the same at 9am and 11am if buying from an in-state firm gave consumers an extra $4 worth of utility, and we can estimate how much bigger or smaller the utility benefit is by looking at how much higher or lower sales to Virginia are at 11am relative to 9am.

Let N_{sht} be the number of consumers in state s purchasing a particular type of memory module in hour h of day t from the twenty-four (or twelve for 256MB modules) websites whose prices we observe. Assume that consumer k's utility if he purchases from website i

[13] Connect Computers is, in fact, not one of the websites from which we have data.

Information on Pricewatch			Price	Price	Price
Website	State	Price	into NJ	into VA	into CA
Pricewatch ranking at 9:01am EDT					
Coast-to-Coast Memory	NJ	112	118.72	112	112
Connect Computers	CA	113	113	113	121.64
Computer Craft	FL	114	114	114	114
Advanced PCBoost	CA	115	115	115	123.80
1st Choice Memory	CA	116	116	116	124.87
Jazz Technology	CA	117	117	117	125.95
Memplus.com	CA	117	117	117	125.95
Portatech	CA	119	119	119	128.10
Augustus Technology	CA	120	120	120	129.18
EconoPC	IL	120	120	120	120
Advanced Vision	CA	121	121	121	130.26
Computer Super Sale	IL	122	122	122	122
Pricewatch ranking at 11:01am EDT					
UpgradePlanet.com	VA	111	111	115.99	111
Connect Computers	CA	113	113	113	121.64
Computer Craft	FL	114	114	114	114
Advanced PCBoost	CA	115	115	115	123.80
1st Choice Memory	CA	116	116	116	124.87
Jazz Technology	CA	117	117	117	125.95
Memplus.com	CA	117	117	117	125.95
Portatech	CA	119	119	119	128.10
Augustus Technology	CA	120	120	120	129.18
EconoPC	IL	120	120	120	120
Advanced Vision	CA	121	121	121	130.26
Computer Super Sale	IL	122	122	122	122

Figure 1: Sample Pricewatch rankings: 128MB PC100 memory modules on August 1, 2000

is

$$u_{iksht} = \beta_1(Price_{iht} + \beta_2 SalesTax_{isht}) + \beta_3 HomeState_{is} + \beta_4 SecondScreen_{iht} + \epsilon_{ik},$$

where $SalesTax$ is the dollar sales tax due on the purchase and ϵ_{ik} is an independent logit random variable. Note that $\beta_2 = 0$ corresponds to consumers ignoring sales taxes and $\beta_2 = 1$ corresponds to consumers treating tax differences exactly as they do differences in item prices. Magnitudes of the home-state preference, β_3, can be interpreted in dollar terms relative to the coefficient on price or in terms of the magnitude of the effect on quantities sold. Writing X_{sht} for the vector of attributes on the right hand side of this expression, we have the familiar logit formula for the number of consumers in state s buying from website i:

$$E(Q_{isht}|X_{sht}) = N_{sht}\frac{e^{\beta_1(Price_{iht}+\beta_2 SalesTax_{isht})+\beta_3 HomeState_{is}+\beta_4 SecondScreen_{iht}}}{\sum_{j=1}^{24} e^{\beta_1(Price_{jht}+\beta_2 SalesTax_{jsht})+\beta_3 HomeState_{js}+\beta_4 SecondScreen_{jht}}}$$

Our dataset only contains the number of consumers purchasing from two particular websites, not all websites. To estimate the model we assume that the unobserved number of consumers satisfies

$$E(N_{sht}|X_{sht}, Z_{sht}) = \delta_s \overline{q}_h e^{\gamma_1 TimeTrend1_t + \ldots + \gamma_4 TimeTrend4_t + \gamma_5 Weekend_t + \gamma_6 MinPrice_{ht}},$$

where δ_s is a state-fixed effect to be estimated, $\overline{q}_h$ is the fraction of all sales in the sample which occur in hour h, the $TimeTrend$ variables allow for linear time trends with slopes changing every ninety days, $Weekend_t$ is a dummy for whether day t is a Saturday or Sunday, $Minprice_{ht}$ is the lowest price listed on Pricewatch (which we think of as being very close to the wholesale price of memory), and Z_{sht} is the set of additional variables on the right hand side of this model.

Note that the $MinPrice$ variable is being used to control both for the increase in overall demand for memory modules caused by a decrease in their price, and for any substitution between e-retailers and traditional retailers due to dollar sales taxes being lower when prices are lower. It also may reflect an endogenous relationship: wholesale prices for memory may rise in response to demand shocks. It could also reflect additional nonlinearities in the time trends. Accordingly, we will think of $MinPrice$ simply as a control variable and not try to interpret the estimates in terms of demand elasticities or in terms of an effect of sales taxes on the aggregate share of online versus offline purchasing.

Note also that we are implicitly assuming that the aggregate sales by Pricewatch e-retailers in any given state, N_{sht}, is not affected by the variables on the right hand side of the utility function, *i.e.* the states in which the e-retailers are located and the difference between the n^{th} loweset price and the lowest price do not cause consumers to substitute to the outside good (buying from a traditional retailer or not buying). In the case of price differences, we think this is reasonable because the prices on Pricewatch are almost always tightly bunched. In the case of state locations we think the assumption is reasonable when modeling demand in any particular state (other than California) because having more than one or two e-retailers on the list being from that state is extremely rare. (The majority of the e-retailers listed on Pricewatch are based in California.) We would expect, for example, that whether zero or one or two firms are from Florida would have little impact on the maximum utility that a Florida consumer can get by buying through Pricewatch, and hence feel that it is reasonable to ignore any substitution to the outside good that it might cause.

We feel that it is reasonable to think of the turnover in the Pricewatch rankings as exogenous and not driven by any information the firms have about a particular hour on a particular day being a good time to have the third or seventh lowest price. Accordingly, we will simply estimate the model via nonlinear least squares. The model could in principal be estimated as a very large dataset using sales to each of the fifty states in each of the approximately 7900 hours by each of the two websites as separate observations. Some states, however, rarely make any purchases. Other states have a nontrivial number of purchases, but rarely or never have an in-state firm listed on Pricewatch. Data from such states could help us to estimate the coefficient on price, for instance (which is easy to estimate precisely), but will not otherwise contribute to estimating tax sensitivities because consumers in these states can purchase from any websites on the list without paying salex tax. For this reason, we decided to reduce the computational burden by carrying out our analysis on a smaller dataset containing hourly sales by our two websites in each of ten states: Alabama, Florida, Georgia, Illinois, Ohio, Oregon, Pennsylvania, Texas, Virginia, and Wisconsin.

Table 5 presents coefficient estimates obtained by estimating the model on sales of each speed of 128MB memory modules in the the non-California states. The price coefficient indicates that (as we reported in Ellison and Ellison (2002)) demand is extremely price sensitive. The estimates imply that a one dollar increase in the price charged by one of our websites (holding all variables fixed at their sample means) reduces demand by 53%.[14] The β_2 coefficient on the $SalesTax$ variable is 0.03 in the PC100 regression and 0.33 in the PC133 regression. The latter is significantly different from zero.[15] Both are very clearly different from one. We conclude in contrast to Brynjolfsson and Smith that consumers pay less attention to tax differences than to price differences.

Given the extreme price sensitivity of demand, the fact that consumers pay less attention to tax differences than to price differences does not imply that taxes do not have a large impact on quantities sold. For example, using the prices listed in the top panel of Figure 1, the coefficient estimates from the PC133 regression imply that if Florida was able to require that all out-of-state firms collect use taxes, then Computer Craft's share of online sales to Florida residents would increase by almost 350%.[16] We would, however, also expect that requiring out-of-state firms to collect use taxes would reduce the total number of Florida residents buying memory online. As noted above, we have not tried to capture this substitution in this model. Our state-level analysis suggested that collecting use taxes might reduce online puchases in Florida 38%. Hence, the net effect on Computer Craft of instituting use taxes might be to raise its Florida sales by something closer to 150% to 200%.

Given that Florida has not been able to require that out-of-state firms collect use taxes, applying sales taxes to in-state sales causes Florida consumers to buy from out-of-state firms and will bring in limited tax revenue. One could then imagine that Florida might want to consider a modification to its tax policy that has not been debated in many legislatures: eliminating sales-taxes on online sales made by in-state firms.[17] Again taking the prices

[14]This calculation assumes our website has the sixth lowest price charging its average price. Also, the prices of the other firms are set to the sample means for firms with those ranks.

[15]The standard errors in this draft do not allow for correlations over time or across states. The t-statistics therefore overstate the true significance of the results.

[16]The prices in Figure 1 are for PC100 memory modules, but PC133 module prices are usually similar to PC100 prices.

[17]Hal Varian has advocated eliminating all sales taxes as a way to resolve online-offline conflicts in "Eco-

from the top panel of Figure 1 as an example, the estimates imply that making Computer Craft's in-state sales tax free would increase sales of PC133 memory by $71 for every dollar in tax revenue that is lost.

The home state coefficient is positive in both regressions and significant in the PC133 regression. The magnitudes of the coefficients in the PC133 regression imply that consumers are willing to pay $2.01 more to buy from an in-state firm. Because consumers are estimated to weight taxes only one-third as heavily as prices, our estimates imply that an in-state firm will make more sales in a state with a 6% sales tax than an out-of-state firm charging the same price on products that cost less than $100. Prices are below $100 for most of the observations in our dataset.

5 Conclusion

Using two distinct identification strategies, one exploiting variation in state-level tax rates and demographics and one exploiting variation in prices and taxes that individuals face, we find evidence of significant, often large, tax effects. A few caveats apply, though. First, our analysis was performed in a market with extremely elastic demand and relatively little price dispersion at the bottom of the price distribution. We would expect tax impacts typically to be smaller in markets where small price changes do not produce large changes in ranks and large demand responses. Our state-level results on the tax impacts on overall demand, however, would probably not vary as much depending on specifics of the market.

Second, the way in which prices and taxes are presented by the price search engine we study may be important. In particular, while Pricewatch gives consumers the information they need to compute tax-inclusive prices for all of the products (assuming they know the tax rate in their home state), it does not actually compute the tax-inclusive prices for them. Brynjolfsson and Smith's result suggesting that consumers were *more* sensitive to taxes and shipping rates than to prices was obtained by studying behavior of consumers using a price search engine that displayed prices inclusive of taxes and shipping. Hossain and Morgan (2002) also indicates that such framing could be important in these contexts. In a series of field experiments performed on eBay, they find that bidders are not sensitive

nomic Scene: Forget Net Taxes. Forget Sales Taxes Altogether," *New York Times,* March 8, 2001.

to shipping charges (as long as they fall in some reasonable range). Similar to our setting, shipping charges are available to bidders on eBay, but prices inclusive of shipping are not conspicuously displayed.

Finally, we think the estimation of a home-state preference is an important addition to the literature. First, in conversations with e-retailers, it is clear that they believe that such a preference sometimes exists. Second, if one does exist, it is important to allow for it so that tax effects can be properly estimated. If, for instance, consumers have a home-state preference but strongly dislike paying taxes, the net effect may appear to be a mild dislike of paying taxes. This distinction would be important in predicting consumer behavior under different tax scenarios. Finally, knowing the magnitude of a possible home-state preference could be important in predicting the market structure of e-commerce. Strong home-state preferences, for instance, would tend to support more geographically dispersed e-retail and presumably less concentrated markets overall.

References

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

Ellison, Glenn and Sara Fisher Ellison (2001): "Search, Obfuscation, and Price Elasticities on the Internet," mimeo, MIT.

Fox, William (1986): "Tax Structure and the Location of Economic Activity Along State Border," *National Tax Journal,* Vol. 39, n. 4, pp. 387-401.

Goolsbee, Austan (2000): "In a World without Borders: The Impact of Taxes on Internet Commerce," *Quarterly Journal of Economics,* Vol. 115, n. 2, pp 561-576.

Hossain, Tanjim and John Morgan (2002): "A Test of the Revenue Equivalence Theorem using Field Experiments on eBay," mimeo, Berkeley.

Walsh, Michael and Jonathan Jones (1988), "More Evidence on the 'Border Tax' Effect: The Case of West Virginia," *National Tax Journal,* Vol. 41, n. 2, pp. 261-265.

Variable	Mean	St.Dev	Min	Max
$Quantity128$	203.5	176.0	19.0	762.0
$Quantity256$	85.6	84.3	5.0	391.0
$SalesTaxRate$	0.057	0.021	0.000	0.084
$InternetAccess$	0.566	0.067	0.406	0.702
$\frac{ComputerStores}{GasStations}$	0.092	0.034	0.041	0.184
$\log(Population)$	15.02	1.04	13.11	17.33

Table 1: Summary statistics for state-level regressions

Variable	Mean	St.Dev	Min	Max
128MB PC100				
$Quantity$	0.013	0.119	0	4
$Price$	66.24	34.4	21	123
$MinPrice$	62.51	33.4	20	122
$Rank$	6.4	4.1	1	21
$SecondScreen$	0.11	0.31	0	1
$Weekend$	0.29	0.45	0	1
$Price$	69.25	35.0	20	131
$HomeState$	0.033	0.18	0	1
$SalesTax$	0.16	1.04	0	10.19
Number of Observations on *Quantity*: 158790				
128MB PC133				
$Quantity$	0.011	0.104	0	2
$Price$	73.82	36.6	21	131
$MinPrice$	71.50	37.1	20	131
$Rank$	6.0	4.4	1	24
$SecondScreen$	0.10	0.30	0	1
$Weekend$	0.28	0.45	0	1
$Price$	78.16	39.3	20	145
$HomeState$	0.035	0.18	0	1
$SalesTax$	0.18	1.10	0	11.09
Number of Observations on *Quantity*: 141470				

Table 2: Summary statistics for individual-level regressions

	Set of Products	
	128MB	256MB
$SalesTaxRate$	7.40	8.28
	(3.06)	(3.24)
$California$	-0.94	-0.81
	(2.92)	(2.37)
$InternetAccess$	1.54	0.74
	(1.96)	(0.88)
$\frac{ComputerStores}{GasStations}$	3.63	5.95
	(2.12)	(3.29)
$Log(Population)$	-0.15	-0.10
	(3.28)	(2.14)
Observations	51	51
R^2	0.43	0.36

Note: t-statistics in parentheses.

Table 3: State-level regressions

	Product/Price Range			
Product:	128MB		256MB	
Item Price:	\$20-\$40	\$100+	\$40-\$60	\$150+
$SalesTaxRate$	5.25	10.41	5.99	9.85
	(1.93)	(3.81)	(1.60)	(1.90)
$California$	-0.87	-1.41	-0.63	-0.59
	(2.41)	(3.89)	(1.27)	(0.87)
$InternetAccess$	1.15	2.85	-0.54	3.88
	(0.89)	(2.98)	(0.44)	(2.15)
$\frac{ComputerStores}{GasStations}$	2.10	6.25	6.59	5.35
	(1.09)	(3.18)	(2.48)	(1.45)
$Log(Population)$	-0.11	-0.09	-0.18	-0.14
	(2.08)	(1.61)	(2.60)	(1.31)

Note: t-statistics in parentheses.

Table 4: State-level regressions on different subsamples

Product:	128MB PC100	128MB PC133
Website characteristics		
Price	-0.54	-0.78
	(34.47)	(28.66)
SalesTax	0.03	0.33
	(0.23)	(2.73)
HomeState	0.44	1.59
	(1.43)	(4.91)
SecondScreen	-1.70	-0.64
	(1.02)	(1.45)
Number of Potential Consumers		
Weekend	-0.44	-0.35
	(11.28)	(7.62)
MinPrice	-0.03	-0.03
	(7.19)	(6.69)
Observations	158790	141470
R^2	0.04	0.03
Home-state preference dollar equivalent	\$0.81	\$2.01
Item price at which sales tax offsets home preference for PA customers	\$419.24	\$96.69

Note: Dependent variables are number of distinct customers in each of ten states ordering from each of websites A and B in each of approximately 7900 hours.
Regressions also contain state dummies and time-trends with different slopes for each 90 day period.

Table 5: Substitution between e-retailers: the discrete-choice model of hourly sales of 128MB memory modules in ten states