**WORKING PAPER SERIES**

# Advertising in the U.S. Personal Computer Industry

Michelle S. Goeree

WP 2004-09

500 East 9th Street
Claremont, CA 91711
Phone: (909) 607-3203
Fax: (909) 621-8249

# Advertising in the US Personal Computer Industry

Michelle Sovinsky Goeree [1]

March, 2004

(First Version June 2002)

**Abstract**

Traditional models of consumer choice assume consumers are aware of all products for sale. This assumption is questionable, especially when applied to markets characterized by a high degree of change, such as the personal computer (PC) industry. I present an empirical discrete-choice model of limited information on the part of consumers, where advertising influences the set of products from which consumers choose to purchase. Multi-product firms choose prices and advertising in each medium to maximize their profits. I apply the model to the US PC market, in which advertising expenditures are over $2 billion annually.

The estimation technique incorporates macro and micro data from three sources. Estimated median industry markups are 19% over production costs. The high industry markups are explained in part by the fact that consumers know only some of the products for sale. Indeed estimates from traditional consumer choice models predict median markups of one-fourth this magnitude. I find that product-specific demand curves are biased towards being too elastic under traditional models of consumer choice. The estimates suggest that PC firms use advertising media to target high-income households, that there are returns to scope in group advertising, and that word-of-mouth or experience plays a role in informing consumers. The top firms engage in higher than average advertising and earn higher than average markups.

JEL Classification: L15, D12, D21, M37, L63

Keywords: Advertising, information, discrete choice models, product differentiation, personal computer industry

---

[1] This paper is based on various chapters from my dissertation. Special thanks to my dissertation advisors, Steven Stern and Simon Anderson, for their guidance. This research has benefited from discussions with Jacob Goeree, Aviv Nevo, Matthew Shum, Greg Crawford, Dan Ackerberg, Jeroen Hinloopen and seminar participants at Warwick, Virginia, Leuven, Amsterdam, Tilburg CentER, Edinburgh, Arizona, Claremont McKenna, and the EARIE meetings in Helsinki. I thank the co-editor and three anonymous referees for useful comments and suggestions. I am grateful to Gartner Inc. for making the data available, and to Sandra Lahtinen for her willingness to provide time and assistance. Financial support from the University of Virginia's Bankard Fund for Political Economy is gratefully acknowledged. Address for correspondence: University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands (email: m.s.goeree@uva.nl)

*"He who has a product to sell*
*And goes and whispers in a well*
*Is not so apt to get the dollars*
*As one who climbs up a tree and hollers"*
- Author unknown

# 1 Introduction

In 1998 over 36 million PCs were sold in the US, generating over $62 billion in sales – $2 billion of which was spent on advertising. The PC industry is one in which products change rapidly, with a total of 200 new products introduced by the top 15 firms every year (Gartner Inc., 1999). Due to the large number of PCs available and the frequency with which new products are brought into the market, consumers are unlikely to be aware of all products for sale. More generally, it is reasonable to suspect consumers have limited information regarding the products available for purchase in many industries. Price elasticities calculated under the assumption of full information may be misleading, which could lead to incorrect conclusions regarding the nature of competition. The goal of this study is to examine the effects of limited choice set information on price and advertising demand elasticities and markups, which have implications for the sources of market power and the role of advertising in the US PC industry.

I develop and estimate a model in which consumers are not assumed to know all products that are available when they make their purchase decision. Advertising influences the set of products from which consumers choose to purchase, where the probability a consumer knows a product is a function of advertising and consumer attributes. Allowing for heterogeneity in consumer's choice sets yields more realistic estimates of substitution patterns between goods. The results suggest that (i) firms benefit from limited information on part of consumers and (ii) that assuming full information may result in incorrect conclusions regarding the nature of competition. For example, estimated median markups over marginal costs in the PC industry are 19%. The high markups are explained in part by the fact that consumers know only some of the products for sale, which results in a less competitive environment. Assuming consumers are aware of all products generates inconsistent estimates of product-specific demand curves that are biased towards being too elastic. Indeed estimates from traditional consumer choice models predict median markups of one quarter of the magnitude predicted by a model of limited information.

Estimation of the structural model is complicated by the lack of individual-level purchase and advertising exposure data for the PC industry. I use three primary data sets in estimation. One is from Gartner Inc.'s Dataquest and includes product-level market shares and other product characteristics. The second is from Leading National Advertisers (LNA) and includes national advertising expenditures across media. The final data set is from Simmons Market Research and includes consumer-level purchases across manufacturers, consumer characteristics, and media exposure information. I exploit the information in these data to estimate a model which allows for three important sources of consumer heterogeneity: choice sets, tastes, and advertising media exposure.

Petrin (2002) shows how combining aggregate data with data that links average consumer attributes to product attributes allows one to obtain more precise estimates. The approach I take is similar in that I augment market share data with data relating consumers to the characteristics of the products they purchase. It differs in that the individual level data I have connect consumers to firms from which they purchased, thus associating consumer attributes and average *product* attributes (across firms). I combine the manufacturer-choice data with aggregate product-level data to obtain more precise estimates of the parameters of the taste distribution.

Additionally, I show how to combine aggregate advertising data with information relating consumer attributes to media exposure. This new technique allows one to obtain a more precise picture of how advertising media exposure and demand are related (relative to using only aggregate data). Together, both parts of the data augmentation methodology enable one to control for heterogeneity in tastes *and* advertising exposure across households – thereby permitting a model which allows for individual heterogeneity in choice sets while having limited information connecting consumers to purchases and advertising.

There are a number of recent structural studies of advertising utilizing individual-level purchase and advertising exposure data. Erdem and Keane (1996) (hereafter EK) estimate models where consumers learn (in a Bayesian manner) about the quality of laundry detergents through past experience and advertising exposures. Experience and advertisements relay noisy information about product quality which consumers use to update their expectations. Ackerberg (2003) estimates a model in the spirit of Milgrom and Roberts (1986) by allowing advertising for Yoplait yogurt to have indirectly informative (signaling) and uninformative (prestige) effects. He posits that advertising will affect consumers who have had prior experience with Yoplait differently than those who have not tried the product. Consumers who know the product (experienced consumers) should not be affected by exposures to informative advertising. Alternatively, the prestige effects of advertising should affect both inexperienced and experienced consumers. He finds a large and significant signaling effect of advertising and an insignificant prestige effect. Anand and Shachar (2001) examine the market for network television programs and the effects of preview advertising on channel viewing decisions. They extend the model of EK to incorporate additional sources of information (such as word-of-mouth), which are modeled as noisy signals as well. They also allow advertising to have a persuasive effect, by affecting utility directly.[1] They find that exposure to advertising is informative and results in an improved match of consumers to products. Shum (forthcoming) also examines how the effect of advertising differs across households. In his model, brand loyal behavior is the cause of different responses by consumers to advertising. He employs a micro-level panel dataset on weekly household

[1] Becker and Murphy (1993) present a model in which firms use advertising to suggest consumption of their product is prestigious. Rational consumers prefer to consume more heavily advertised goods. When a consumer buys a product, she is also buying an image. In these models, uninformative advertising directly effects utility.

purchases of breakfast cereals and matches this to aggregated advertising data.

Grossman and Shapiro (1984) (hereafter GS) present a theoretical model of informative advertising. Theirs is a circular model of spatially differentiated products in which advertising messages provide consumers with information about product availability and price.[2] The research presented here is loosely related to GS in that (i) consumers are heterogenous and seek to purchase the product that gives them the highest utility and (ii) advertising conveys information about the existence and attributes of differentiated products. However, the empirical model differs greatly from GS along several dimensions. Most importantly, I wish to allow for a more flexible model of product differentiation and hence abandon the circular city framework. The empirical model in this work is a discrete choice model of product differentiation (see Anderson, de Palma, and Thisse 1989, 1992). In addition, unlike GS, advertising is not the only source of information for consumers. In this study, consumers may be informed if there is no advertising where the degree to which they are informed depends on their characteristics. Finally, I do not observe individual specific advertising messages, which is central to the GS framework. For these reasons, in addition to the necessity of having an empirically tractable model, the "information technology" (and the resulting market shares) presented in this work differ greatly in form from those in GS.

I have an additional challenge in modeling choice set heterogeneity across consumers; I do not observe which of the possible $2^J$ choice sets the consumer faces (where $J$ are the number of products). Chiang, Chib, and Narasimhan (1999) use individual-level purchase and advertising data to estimate a model that allows for "consideration set" heterogeneity across consumers. A consideration set is a (potential) subset of the possible $2^J$ choice sets. The consideration set they take to the data is obtained by making assumptions regarding a consumer's decision making process. For example, a consumer's consideration set may be the set of all previously purchased brands. Due to the stable nature of their industry of

[2] Grossman and Shapiro (1984) extended the model of Butters (1977), who considers a market of homogenous goods.

interest (the ketchup industry), the authors provide evidence that a consumer's consideration set doesn't change over time. Under this assumption, they are able to eliminate choice sets which do not contain all previously purchased brands.[3] In addition, the scanner data they use contains information on four main brands in the industry, which eases computational burden significantly since there are not a large number of products to contend with.

The PC industry is much different than the ketchup industry, it is rapidly changing and there are over 200 products introduced by the top firms alone each year. Therefore, I use a different approach than Chiang, *et. al.* to model consumer choice set heterogeneity. Due to the large number of products in the PC industry, it is not feasible to calculate all possible purchase probabilities for each product corresponding to each possible choice set. Instead I simulate a choice set for each individual in each period and construct an importance sampler to smooth the simulated choice probabilities. The simulator for the market share is the average over individuals of these smoothed choice probabilities. In addition, the distribution of consumer tastes is an empirical one, which also makes simulation of the market shares necessary. The simulator is discussed in more detail in section 4.4 and appendix B.[4]

Estimation is in two parts. First, I use consumer-level data on media exposure (from Simmons) to estimate media-specific parameters that measure how exposure varies with consumer attributes. Given the nature of the data, these parameters are estimated by ordered response maximum likelihood. In the second stage, I simultaneously estimate the remainder of the parameters using generalized method of moments. There are four "sets" of moments. The first utilizes aggregate product-level data (from Gartner) to fit observed market shares to those implied by the model. The second arises from the firm's profit maximizing choices of prices. These two sets of moments are similar to those in Berry,

---

[3] Consideration set formation is modeled by taking the power set of the available products and assigning a probability mass on each subset. This probability mass has a Dirichlet distribution across the consumers with known parameters. A subset that doesn't contain one of the previously purchased products receives a weight of zero.

[4] I utilize antithetic acceleration to reduce the variance due to simulation.

Levinsohn, and Pakes (1995) (hereafter BLP). The third set of moments arises from the firm's optimal choices of advertising medium (using data from LNA). These require care in constructing because some firms find it optimal not to advertise in some media. The fourth set of moments exploits the micro-level data on manufacturer choice (from Simmons).

As is common in this literature, one must address the issue of endogeneity of prices and, in this setting, advertising levels. While previous studies use product specific dummies to correct for the endogeneity of advertising, I form exogenous instruments that are approximations to the optimal instruments to correct for the correlation between unobserved quality and advertising or price.

To summarize, my research differs from previous studies in a few fundamental ways. First, I explicitly model the effect of advertising when consumers have limited information about the choice set available to them. Previous studies have modeled advertising as part of the consumer's utility function or as a noisy signal of product attributes, where consumers are assumed to know all products for sale. Secondly, prior studies utilize individual-level data on both consumption and advertising exposures. I don't have access to such detailed data for the PC industry, yet I am still able to determine the influence of advertising on an individual's choice set, where this influence differs across consumers. Thirdly, I allow the effect of advertising to differ across media. Due to data limitations, previous studies only considered television advertising. Finally, I model firms' decisions with regard to pricing and advertising choices across media, which allows me to examine the additional markup firms earn as a result of limited consumer information. I compare these markups to those predicted by traditional consumer choice models.

The rest of the paper is organized as follows. In the next section, I discuss the data. I develop the model in section 3 and present the estimation technique in section 4. The results from preliminary regressions and from the full model are discussed in sections 5 and 6 respectively. I present specification tests and conclusions in the final sections 7 and 8.

# 2 Data

## 2.1 Product Level Data

The product-level data were provided by Gartner Inc.'s Dataquest and consist of quarterly shipments and dollar sales of all PCs sold between 1996 and 1998.[5] The majority of firms sell to the home market as well as to businesses, educational institutions, and the government. The Dataquest data detail sales across sectors and, since the focus of this research is on consumer purchasing behavior, I use the home market data to estimate the model.[6] Sales in the home market comprise over 30% of all PCs sold.

The PC industry is concentrated, with the top five firms accounting for over 63% (61%) of the dollar (unit) home market share on average. In addition, over 80% of PC sales to the home market sector are from the top 10 manufacturers. Table 1 shows home market shares of the leading manufacturers. The major market players did not change over the period of the data. I restrict my attention to the top 10 firms (based on home market share) and to 5 others.[7] These 15 "included" firms account for over 85% (83%) of the dollar (unit) home market share on average.

There is substantial product differentiation in the industry. PCs are differentiated along many dimensions such as processor type and speed, hard drive space, form factor (desktop, laptop, etc.), RAM, etc. Data limitations prevent me from including all product characteristics. Gartner collects information on five main attributes of each of the PCs: the manufacturer (e.g. Dell), the brand (e.g. Latitude LX), the form factor (e.g. desktop), the CPU type (e.g. Pentium II), and the CPU speed. I define a model as a manufacturer, brand, CPU type, CPU speed, form factor combination. Even though I do not have data

[5] I constructed a price variable by dividing dollar sales by the number of units sold, which was deflated using the Consumer Price Index from BLS.

[6] Since firms' profits depend on sales to all sectors, I use the non-home sector data when developing the supply side of the problem in section 3.3.

[7] This enables me to make full use of consumer-level manufacturer purchase data (from Simmons), which is discussed shortly. The included manufacturers are Acer, Apple, AST, AT&T, Compaq, Dell, DEC, Epson, Gateway, Hewlett-Packard, IBM, Micron, NEC, Packard-Bell, and Texas Instruments.

on some product attributes, the richness of the Dataquest data still allows for a very narrow model definition. For example, Compaq Armada 3xxx Pentium 150/166 laptop and Compaq Armada 4xxx Pentium 150/166 laptop are two separate models, as are an Apple Power Macintosh Power PC 604 180/200 desktop and deskside.

Treating a model/quarter as an observation, the total sample size is 2112.[8] These 2112 observations represent 723 distinct models sold in the home market. The majority of the PCs offered to home consumers were desk PCs, about 71%, and over 83% of the processors were Pentium-based (either Pentium, Pentium II, or Pentium Pro). The number of models offered by each firm varied. Compaq had the largest selection with 138 different choices, while Texas Instruments offered only five. On average, each firm in my sample offered a particular model for 3 quarters. The "modal" PC offered by each firm was a desktop with a Pentium processor having an average speed of 220 MHz.

The potential market size is given by the number of US households in a given period, as reported by the Census Bureau. Market shares are computed as unit sales of each model divided by the market size. The market share of the outside good is one minus the share of the inside goods.

## 2.2 Advertising Data

The product level sales data are combined with advertising data as reported in Competitive Media Reporting's *LNA/ Multi-Media* publication. These data consist of quarterly advertising expenditures across 10 media. From the 10 media, I construct four main media categories: newspaper, magazine, television (TV) and radio.[9]

Unlike the product-level data, the advertising data are not broken down by sector (e.g.

[8] This is the sample size after eliminating observations with negligible quarterly market shares.

[9] The "magazine" medium also includes Sunday magazines. The "television" medium encompasses all programs shown on network, spot, cable or syndicated TV. The "radio" medium encompasses network and spot radio advertising. In addition, I include outdoor advertising in the radio medium. Outdoor advertising represents a very small fraction of expenditures (on average less than 0.3%). There are too many zero observations for outdoor advertising to use it separately, and so I choose to add it to the radio medium.

home, business, etc). Competitive Media Reporting (CMR) reports total advertising expenditures on all computers and computer systems. The CMR measure includes advertising for non-PCs intended for the business, government or education markets (such as mainframe servers and unix workstations).[10] Fortunately, CMR categorizes advertising across products, which in some instances, allows me to isolate certain expenditures as for non-home computers. For example, expenditures are sometimes generally reported (e.g. IBM various computers) and other times are more detailed (e.g. IBM RS/6000 Server). Since some expenditures are generally reported, it is not possible to construct a measure that consists solely of advertising for the home PC market. As a result the advertising measure used in this research includes some expenditures on non-PC systems used in the business, education, or governmental sectors.

Total advertising expenditures in the computer industry have grown from $1.4 billion in 1995 to over $2.3 billion in 1999 (an average annual rate of close to 13%). Table 2 provides a summary for the leading manufacturers. There is much variation in advertising expenditures across firms. Notably, fifty percent of the industry expenditures are by IBM, resulting in an (total) ad-to-sales ratio of over 19 percent (compared to the industry ratio of 3%).

IBM is the largest computer manufacturer in the world. The large advertising expenditures by IBM, relative to other firms in the industry, may be due to their position in the non-PC category of the computer industry. To examine whether IBM's sizable advertising expenditures are a result of its non-PC interests (servers, mainframes, unix workstations, etc.), I allow the position of the firm in the non-PC sector to affect the marginal revenue of advertising from the non-home sectors.

In the PC industry, it is common for firms to advertise products simultaneously in groups. For example, in 1996 one of Compaq's advertising campaigns involved all Presario brand PCs (of which there are 12). I have to make some assumptions about the informativeness of group

[10] However, advertising expenditures on computer components and accessories (such as printers) are itemized separately.

advertising. One possibility is that it provides as much information about the products in the group as product-specific advertising. However, if group advertising were as effective as product advertising, we would observe only group advertising (because this would be the most efficient use of resources). An alternative possibility is that group advertising is not informative about the products in the group; it merely informs the consumer about the manufacturer. If this were the case, we should observe either firm-level (the largest possible group) or product-specific advertising.

In reality, firms use a combination of product-specific and group advertising (with groups of varying sizes). I need a measure of advertising expenditures by product that incorporates *all* advertising done for the product. I construct "effective" product advertising expenditures by adding observed product-specific expenditures to a weighted average of all group expenditures for that product, where the weights are estimated. To be more precise, let $\mathcal{G}_j$ be the set of all possible product groups that include product $j$ (I suppress the time subscript). Let $ad_{\mathcal{H}}$ be (observed) total advertising expenditures for group $\mathcal{H} \in \mathcal{G}_j$, where the average expenditure per product in the group is

$$\overline{ad}_{\mathcal{H}} \equiv \frac{ad_{\mathcal{H}}}{|\mathcal{H}|}$$

Then "effective" advertising expenditures for product $j$[11] are given by

$$ad_j = \sum_{\mathcal{H} \in \mathcal{G}_j} \gamma \overline{ad}_{\mathcal{H}} + \pi \overline{ad}_{\mathcal{H}}^2 \tag{1}$$

where the sum is over the different groups that include product $j$.[12] If there is only one product in the group (i.e. it is product-specific), I restrict $\gamma$ to unity and $\pi$ to zero. Notice that this specification allows for increasing or decreasing returns to group advertising, where $\gamma$ and $\pi$ are parameters to be estimated.

---

[11] I call these "effective" product advertising expenditures to indicate they are constructed from observed group and product-specific advertising.

[12] To get an idea of the level of detail in the data: in the first quarter of 1998, there were 18 group advertisements for Apple computers. The groups advertised ranged from "various computers" to "PowerBook" to "Macintosh Power PC G3 Portable" (the later being a specific model). In this quarter the Apple Macintosh Power PC G3 Portable computer belonged to 7 different product groups.

## 2.3 Consumer Level Data

The consumer level data come from the *Survey of Media and Markets* conducted by Simmons Market Research Bureau. Simmons collects data on consumers' media habits, product usage, and demographic characteristics from about 20,000 households annually. I use two years of the survey from 1996-1997 (data from 1998 were not publicly available).[13] Households in the Simmons data contain at least one respondent 18 years or older. Descriptive statistics for the overall population and the sample used in estimation are given in Table 3.[14]

Ideally, one would have data on an individual's product purchase, demographics, and exposure to product advertising. Unfortunately, micro-level purchase and advertising exposure data are not available for the PC industry. While the Simmons data are not ideal, they do contain information that allows me to link demographics with purchases and to control for heterogeneity across households in advertising media exposure. I use these data in combination with the macro market share (from Gartner) and advertising data (from LNA) to obtain more precise parameter estimates.

Simmons collects information on PC ownership, including whether the individual purchased a PC in the past year and the PC manufacturer. Approximately 11% of the households purchased a PC in the last 12 months. However, the Simmons data are not detailed to the product level. Respondents were not asked any specifics regarding their PC other than the manufacturer. In addition, only 15 manufacturers were listed separately. I use the Simmons data to construct moments relating individual purchases and demographic attributes to product attributes. The micro-moments are valuable when used in conjunction with the macro-level product data. The strategy of using both micro-data and macro-data in estimation follows recent work by BLP (2004) and Petrin (2002). I discuss the construction of the micro-moments in more detail in section 4.2.

[13] To reduce the sample to a manageable size, I select respondents randomly from each year. The final sample size is 13,400.

[14] The Simmons survey oversamples in large metropolitan areas, however this oversampling causes no estimation bias because residential location is treated as exogenous.

Relying on aggregate advertising data has the drawback that there is no observed variation in advertising across households. In addition to purchase information, the Simmons respondents were asked about their media habits. They were ranked according to how often they viewed TV programs, read newspapers, etc. relative to others in the surveyed population. I use the self-reported media exposure information to control for heterogeneity in advertising media exposure across households. The Simmons demographic and media exposure data are combined with (separate) information on market shares and product characteristics, which provides a more precise picture of how media exposure and demand are related. Table 4 details media exposure across households. I have information on the ranges of answers given by the respondents, but the survey reports only the quintile to which the consumer belongs. I discuss the specific way in which the media-exposure data are used in section 4.1.

I also construct BLP-type macro-moments. I use the *Consumer Population Survey (CPS)* data to define the distribution of consumer characteristics for use in the macro-moments. I use the CPS data because the Simmons data are available only for the first two years of the three for which I have product and advertising level data whereas the CPS is available over all the years of interest. I use the CPS data in constructing the macro-moments[15] and the Simmons data in constructing the micro-moments, which are discussed in more detail in section 4.2.

# 3 Economic Model

The primitives of the model are product attributes, consumer preferences, and the notion of equilibrium. I assume the econometrican observes price, advertising choices across media, and quantities sold by each firm. The structural estimation strategy requires me to specify

[15] I drew a sample of 3,000 individuals from the March CPS for each year. Quarterly income data were constructed from annual data and were deflated using the Consumer Price Index from BLS. A few households reported an annual income below $5000. These households were dropped from the sample. Examination of the Simmons data indicate that purchases were made only by households with annual income greater than $5000, therefore eliminating very low income households should not affect the group of interest.

a model of consumer choice, derive the implied relationships among choice probabilities, specify firm behavior, and estimate the parameters of the model.[16] The econometric technique I employ follows those found in recent studies of differentiated products, such as Berry, Levinsohn, and Pakes (1995, 1998), Nevo (2000), and Petrin (2002).

## 3.1 Utility and Demand

Individual $i = 1, ..., I$ chooses from $j = 1, ..., J$ products at time $t = 1, ..., T$. A product pertains to a specific PC model defined as a manufacturer-brand- CPU type-CPU speed-form factor combination. The characteristics of product $j$ are represented by $(x_j, p_{jt}, \xi_{jt})$. Observable attributes are represented by $p_{jt}$, the price, and the vector $x_j$ which consists of CPU speed (MHz), a laptop dummy variable, a Pentium dummy variable, firm fixed effects, and a constant.[17] Attributes that are unobserved by the econometrician but known to the consumer are represented by $\xi_{jt}$. The income of individual $i$ at time $t$ is given by $y_{it}$. The indirect utility consumer $i$ obtains from product $j$ at time $t$ is given by

$$u_{ijt} = \alpha \ln(y_{it} - p_{jt}) + x_j' \beta_{it} + \xi_{jt} + \epsilon_{ijt} \tag{2}$$

where $\beta_{it}$ are individual specific components, and $\epsilon_{ijt}$ is a mean zero stochastic term, which represents idiosyncratic individual preferences and is assumed to be independently and identically distributed across products and consumers.[18]

[16] The model presented in this paper is static. The primary reason for choosing a static model is lack of adequate data. To properly examine dynamics, consumer-level purchasing and advertising exposure data are necessary. Modeling choices as static has certain implications. For one, it does not capture the long-term effects associated with advertising, such as brand building. While branding is an important issue, the majority of PC firms have not changed over the period and most had been in existence for a number of years prior to the start of the data. This suggests these firms would not have as much need to establish a brand-image as to spread information about new products. There are advantages to a static framework. I am able to focus on the influence of advertising on the choice set absent the additional structure and complications of a dynamic setting. Also, the nature of advertising in the PC industry lends itself to a static framework. Products change rapidly. Advertising today informs about products today, and its effects on future information provision are minimal, if not zero. That is, advertising from last year is not informative this year since the same products are no longer for sale.

[17] I do not include brand fixed effects as there are over 200 brands.

[18] Note that the indirect utility in equation (2) can be derived from a Cobb-Douglas utility function (see BLP). I have chosen this specification for indirect utility to allow for wealth effects. See also Petrin (2002)

To allow for correlation between product choices and consumer characteristics, I employ a random coefficient model. Let

$$\beta_{it} = \beta + \Pi D_{it} + \Sigma \nu_i, \qquad \nu_i \sim N(0, I_k) \tag{3}$$

where $\beta$ are the mean preferences for observable attributes of the good excluding price, $D_{it}$ are observed consumer attributes, $\Pi$ are coefficients measuring how tastes vary with attributes, and $\Sigma$ is a scaling matrix. A consumer's taste for a product characteristic may depend on characteristics not observed by the econometrician, as captured by $\nu_i$. I assume that the $\nu_i$ are independently normally distributed across the population with mean zero and variance to be estimated. The distribution of consumer characteristics is an empirical one given by the CPS (see section 2.3).[19]

Equations (2) and (3) can be rewritten as

$$u_{ijt} = \delta_{jt} + \mu_{ijt} + \epsilon_{ijt}, \tag{4}$$

where $\delta_{jt} = x_j'\beta + \xi_{jt}$ captures the base level of utility every consumer derives from product $j$ and the composite random shock, $\mu_{ijt} + \epsilon_{ijt}$,[20] captures heterogeneity in consumers' tastes for product attributes where

$$\mu_{ijt} = \alpha \ln(y_{it} - p_{jt}) + x_j{}'(\Pi D_{it} + \nu_i).$$

One can use market-level data to estimate the parameters of the taste distribution. However, I have additional information linking consumers to the products they purchase, which I use to augment the market level data to obtain more precise estimates of the parameters of the taste distribution.

who uses a more general functional form which includes wealth effects.

[19] The demographic variables include measures of age, household size, income, sex, and race.

[20] Since choices of an individual are invariant to multiplication of utility by a person-specific constant, I have fixed the standard deviation of the $\epsilon_{ijt}$. In theory, I could estimate an unrestricted variance-covariance matrix, however in practice this is not feasible given there are over 2000 products (it implies estimating $J(J-1)/2$ parameters).

The consumers may decide not to purchase any of the goods; instead they may decide to purchase the "outside" good. The outside good might include nonpurchase, purchase of a used PC, or purchase of a new PC from a firm not included in the "15 included firms." The presence of an outside good allows for lower market sales in response to a market wide uniform price increase. The indirect utility from the outside option is

$$u_{i0t} = \alpha \ln(y_{it}) + \xi_{0t} + \epsilon_{i0t},$$

where the price of the outside good is normalized to zero. Since I cannot identify relative levels of utility, the mean utility of one good is not identified. Thus, I normalize $\xi_{0t}$ to zero.

I assume a consumer purchases at most one good per period.[21] The consumer chooses the good which provides the highest utility, $U$, from all the goods in his choice set. First I consider the case in which the consumer has full-information regarding the products for sale. Consumer $i$ will purchase good $j$ at time $t$ if and only if

$$U_{jt} \geq U_{rt}, \forall r \neq j.$$

Define the set of variables that results in the purchase of good $j : R_{jt} \equiv \{(y_{it}, D_{it}, \nu_i, \epsilon_{ijt}) : U_{ijt} \geq U_{irt} \quad \forall r \neq j\}$. Assuming ties occur with zero probability, the probability that $(y_{it}, D_{it}, \nu_i, \epsilon_{ijt})$ fall into the region $R_{jt}$ (i.e. the market share of product $j$) is

$$s_{jt} = \int_{R_{jt}} dF(y, D, \nu, \epsilon) = \int_{R_{jt}} dF_{y,D}(y, D) dF_\nu(\nu) dF_\epsilon(\epsilon) \tag{5}$$

where $F(\cdot)$ denotes the respective known distribution functions. To derive the market share of product $j$, I integrate over the observed joint distribution of $(y_{it}, D_{it})$ and the assumed distribution of $(\nu_i, \epsilon_{ijt})$ in the population, where the second equation follows from independence assumptions. In order to obtain simple expressions for choice probabilities, I assume the $\epsilon$ are distributed i.i.d. type I extreme value. Therefore the probability that product $j$ is chosen conditional on $(y, D, \nu)$ is given by the multinomial logit (MNL) choice probability.

[21] This assumption may be unwarranted for some products for which multiple purchase is common. However it is not unreasonable to restrict a consumer to purchase one computer per quarter. Hendel (1999) examines purchases of PCs by businesses and presents a multiple-choice model of PC purchases.

## 3.2 Information Technology

In my model, the probability that consumer $i$ purchases product $j$ will depend upon the probability she is aware of product $j$, which products are competing with $j$ (that is which products are in her choice set) and, given her choice set, the probability she would buy product $j$. The firm's advertisement alerts the consumer to the product's existence and thereby increases the probability that the product is in the consumer's choice set.[22]

Let $\mathcal{C}_j$ be the set of all possible choice sets that include product $j$. Assuming consumers are aware of the outside option with probability one, and that the $\epsilon$ are distributed i.i.d. type I extreme value, the (conditional) probability that consumer $i$ purchases product $j$ is given by

$$s_{ijt} = \sum_{\mathcal{S}\in\mathcal{C}_j} \prod_{l\in\mathcal{S}} \phi_{ilt} \prod_{k\notin\mathcal{S}} (1-\phi_{ikt}) \frac{\exp\{\delta_{jt}+\mu_{ijt}\}}{{y_{it}}^{\alpha} + \sum_{r\in\mathcal{S}} \exp\{\delta_{rt}+\mu_{irt}\}} \tag{6}$$

where $\phi_{ijt}$ is the probability consumer $i$ is informed about product $j$, the outside sum is over the different choice sets that include product $j$, and the ${y_{it}}^{\alpha}$ in the denominator is from the presence of an outside good. Advertising affects demand through the information technology function, $\phi_{ijt}$, which describes the effectiveness of advertising at informing consumers.

The information technology is modeled as a function of product $j$'s advertising by medium, observed consumer characteristics, and unobserved idiosyncratic consumer-advertising-medium-specific effects. The information technology function for consumer $i$ is given by

$$\phi_{ijt}(\theta_\phi, \Upsilon) = \frac{\exp(\tau_{ijt})}{1+\exp(\tau_{ijt})} \tag{7}$$

$$\tau_{ijt} = \widetilde{D}'_{it}\lambda + a'_{jt}(\varphi + \rho a_{jt} + \Psi_f + \Upsilon D^s_{it} + \kappa_i) \quad \ln\kappa_i \sim N(0, I_m)$$

[22] This is not a model of advertising content, but of product existence. That is, once a consumer is aware of the existence of a product she is also aware of the product's attributes. As noted in the introduction GS present a theoretical model of informative advertising in a circular city framwork, however the market share specification I develop in this section differs greatly from that in GS.

where $\theta_\phi = \{\lambda, \gamma, \pi, \varphi, \rho, \Psi\}$.[23] The number of advertisements for product $j$ is broken down across four media and represented by the $m(=4)$ dimensional vector, $a_{jt}$. The four media are magazines, newspapers, television, and radio.[24] The $\widetilde{d}$ dimensional parameter vector $\lambda$ measures the fraction of consumers of type $\widetilde{D}_{it}$ who are informed without seeing any advertising ($\widetilde{D}$ is a subset of all consumer characteristics, $D$);[25] $\varphi$ measures advertising's effectiveness; $\rho$ captures the decreasing or increasing effectiveness of advertising; and $\Psi$ are firm fixed effects. These capture how variation in advertising effectiveness varies across firms.[26]

Ideally, one would have data on an individual's exposure to product $j$'s advertising (measured as number of messages). Unfortunately, micro-level advertising exposure data are not available for the PC industry. However, I control for heterogeneity in advertising exposure across households (as it is related to observables) by taking advantage of self-reported media exposure information (including the amount of time spent watching TV, reading magazines, etc.) from the Simmons survey. The matrix $\Upsilon_{m\times d}$ captures how advertising media's effectiveness varies by observed consumer characteristics from the Simmon's survey, $D^s$.[27]

In addition I include a stochastic consumer-medium-specific term ($\kappa_{im}$) in the information technology specification. The $\kappa_{im}$ are unobserved consumer heterogeneity with regard to advertising medium effectiveness, these include consumer attributes that may influence the effectiveness of medium $m$ at informing the consumer, but that aren't picked up by observed

[23] The reason for separating $\Upsilon$ from the other parameters in the information technology function will become clear when I discuss estimation.

[24] The number of advertisements are advertising expenditures, $ad_{jm}$, divided by the weighted average price of an advertisement in medium $m$. Recall, from section 2.2 that $ad_j$ is a weighted sum of model specific and group advertising where the weights, $\gamma, \pi$, are to be estimated.

[25] The subset of consumers characteristics ($\widetilde{D}$) consist of a constant, and dummy variables for high school graduate, whether income is below \$60,000, and whether income is above \$100,000 (that is, $\widetilde{d} = 4$). I include fixed effects only for those firms that offered a product every quarter.

[26] The advertising technology depends upon own product advertising only. Implicit in this specification is the assumption that product specific or group advertising for product $r \neq j$ provides no information about product $j$.

[27] See Table 9 for a list of attributes included in $D^s$.

demographic characteristics.

As advertising increases, the information technology approaches one but it is non-zero even for zero advertisements. The latter property comes from the presumption that a fraction of consumers are informed even if there is no advertising, that is $\phi(a=0) > 0$, which allows for positive demand when no advertising occurs. The magnitude of the probability that a consumer is informed when no advertising occurs is determined by $\widetilde{D}'_{it}\lambda$.[28]

Since I assumed the unobservables have known, independent distributions, I integrate over them and the joint distribution of income and consumer attributes in the population. This yields the market share for product $j$

$$s_{jt}(p,a) = \int s_{ijt} dF(y,D) dF(\nu) dF(\kappa) \tag{8}$$

where $s_{ijt}$ is given in equation (6). Market share is a function of prices and advertising of all products. When all firms advertise more and the information probability for all products approaches one, market share approaches the standard full information choice probability. Naturally, the smaller is the information technology the smaller is product market share. Demand for product $j$ at time $t$ is $M_t s_{jt}$, where $M_t$ is the market size given by the number of households in the US.

## 3.3 Firm Behavior

I assume there are $F$ firms in an oligopolistically competitive industry and that they are non-cooperative, Bertrand-Nash competitors. Each firm produces a subset of the $J$ products, $\mathcal{J}_f$. Suppressing time notation, the profits of firm $f$ are

$$\sum_{j\in\mathcal{J}_f}(p_j - mc_j) M s_j(p,a) + \sum_{j\in\mathcal{J}_f}\Pi_j^{nh}(p^{nh},a) - \sum_m mc_{jm}^{\text{ad}}(\sum_{j\in\mathcal{J}_f} a_{jm}) - \mathcal{C}_f \tag{9}$$

[28] There are a number of reasons why individuals may be informed, even in the event that they haven't seen an advertisement. They may have received information by word-of-mouth, experience with the product, or exposure to other non-advertising media coverage (i.e. read a magazine article). See Anand and Shachar (2001).

where $s_j$ is the vector of home market shares for product $j$ given in (8), which is a function of the prices and advertising levels of all products; $mc_j$ is the marginal cost of production; $\Pi_j^{nh}$ is the gross profit (before advertising) from sales to the non-home sectors; $mc_{jm}^{\mathrm{ad}}$ is the marginal cost of advertising in medium $m$; $a_{jm}$ is the number of medium $m$ advertisements; and $\mathcal{C}_f$ are fixed costs of production.

Following the approach taken in BLP, I assume marginal costs are log-linear and composed of unobserved and observed cost characteristics, $\omega_j$ and $w_j$ respectively. I expect unobserved cost characteristics to be correlated with $\xi_j$, that is PCs with high unobserved quality might be more expensive to produce. I account for the correlation between $\omega$ and $\xi$ in estimation. The (log) marginal cost function is given by

$$\ln(mc_j) = w_j'\eta + \omega_j \tag{10}$$

where $\eta$ is a vector of parameters.

Similarly, I assume marginal costs of advertising are composed of observed components, $w_{jm}^{ad}$ (such as the average price of an advertisement),[29] and unobserved product-specific components, $\upsilon_j$. The (log) marginal cost of advertising in medium $m$ is given by

$$\ln(mc_{jm}^{\mathrm{ad}}) = w_{jm}^{ad\prime}\psi + \upsilon_j \qquad \upsilon_j \sim N(0, I_m) \tag{11}$$

where $\psi$ is to be estimated.

Given their products and the advertising, prices, and attributes of competing products, firms choose prices and advertising media levels simultaneously to maximize profits. Product attributes that affect demand $(x_j, \xi_j)$ and those that affect marginal costs $(w_j, \omega_j, w_{jm}^{ad}, \upsilon_j)$, which will be discussed shortly, are treated as exogenous to the firm's pricing and advertising decisions. Since PC firms may sell to non-home sectors (such as the business, education,

[29] The LNA data consist of advertising expenditures across 10 media, from which I construct 4 main media groups. I observe the distribution of firm spending across the original 10 media. The quarterly average ad price in media group $m$ is a weighted average of the ad prices in the original categories comprising the media group. The weights are firm specific and are determined by the distribution of the firms advertising across the original media.

and government sectors), some discussion is necessary regarding optimal choices of prices and advertising levels. Constant marginal costs imply pricing decisions are independent across sectors.[30] Therefore, any product sold in the home market sector will have prices that satisfy the following first order conditions

$$s_j(p,a) + \sum_{r \in \mathcal{J}_f} (p_r - mc_r) \frac{\partial s_r(p,a)}{\partial p_j} = 0 \tag{12}$$

However, an advertisement intended to reach a home consumer may affect sales in other sectors. Optimal advertising choices must equate the marginal revenue of an additional advertisement in all sectors with the marginal cost. Optimal advertising medium choices $a_{jm}$ must therefore satisfy

$$M \sum_{r \in \mathcal{J}_f} (p_r - mc_r) \frac{\partial s_r(p,a)}{\partial a_{jm}} + mr_j^{nh}(p^{nh}) = mc_{jm}^{\mathrm{ad}} \tag{13}$$

where $mr^{nh}(p^{nh};\theta^{nh})$ is the marginal revenue of advertising in non-home market sectors. I assume $mr^{nh}$ is a linear combination of price in the non-home sector, $p_j^{nh}$, and the vector $x_j^{nh}$ which consists of CPU-speed, non-PC firm sales[31] , a constant, and parameters to be estimated ($\theta^{nh}$). Specifically, $mr_j^{nh} = \theta_p^{nh} p_j^{nh} + x_j^{nh\prime}\theta_x^{nh}$.[32] For ease of exposition, let $\eta_{\mathrm{AD}} = \{vec(\psi), vec(\theta_{nh})\}$ be the vector of parameters associated with advertising medium choices.

[30] There are reasons to believe that pricing decisions may not be independent across sectors. For instance, if the price of a particular laptop is lower in the business sector, a consumer might buy the laptop from their business account for use at home. I abstract away from this problem for two reasons. First, identification of a model which includes pricing decisions across all sectors would require much richer data than I have on the non-home sectors. Second, education, government, and business purchases usually involve more than one computer. Hence, one should allow for multiple purchases per period in the non-home sector, which greatly complicates the model (see Hendel, 1999). While the assumptions on firm behavior that I impose imply independent pricing decisions, the parameter estimates that I obtain are sensible. In addition, goodness-of-fit tests suggest the model fits the data reasonably well. Hence, the results suggest that the model does not do a poor job of describing the PC industry even when pricing decisions are independent across sectors.

[31] Non-PC sales are constructed by subtracting quarterly PC sales from quarterly total manufacturer sales (as recorded in firm quarterly reports). Therefore "non-home sales" include sales of computer systems such as mainframes, servers, and unix workstations.

[32] Ideally, one would construct $mr^{nh}$ in a structural framework analogous to that used to construct the marginal revenue of advertising in the home market sector. However, identification of the parameters would require much richer data than I have. In addition, one should allow multiple purchases per period in the non-home sector, which greatly complicates the model (see Hendel, 1999).

# 4 The Estimation Technique

The parameters of the model associated with the demand-side are $\Upsilon$, $\beta$, and $\theta = \{\alpha, \Sigma, \Pi, \theta_\phi\}$ and the supply-side parameters are $\eta$ and $\eta_{AD}$. Under the assumption that the observed data are the equilibrium outcomes, I estimate the parameters of the model in two stages.

First, I use individual-level data on media exposure (from Simmons) to estimate the media-specific parameters, $\Upsilon$, that measure how exposure varies with observable consumer attributes. These parameters are estimated separately using maximum likelihood.

Then the remainder of the parameters, $\{\beta, \theta, \eta, \eta_{AD}\}$, are estimated simultaneously by generalized method of simulated moments (GMM). There are four "sets" of moments:

**(i)** Moments arising from the demand side, which fit the model's predictions for product $j$'s market share to its observed market shares

**(ii)** Moments arising from firm's pricing decisions, which express an orthogonality between the cost side unobservable $\omega_j$ and appropriate instruments

**(iii)** Moments arising from the firms advertising media choices, which express an orthogonality between the advertising residuals (constructed so as to allow for corner solutions) and the instruments

**(iv)** Moments arising from individual-level data on manufacturer choice, which match the model's predictions for the probability of a purchase from firm $f$ (conditional on consumer and product characteristics) to observed purchases from the Simmons data

Before discussing the moments in more detail, I outline the method used to control for heterogeneity in household media exposure, which results in the first-stage estimate of $\Upsilon$.

## 4.1 Individual Media Exposure

The information technology is a function of consumer attributes, product attributes, and parameters and is given by

$$\phi_{ijt}(\theta_\phi, \Upsilon) = \frac{\exp(\tau_{ijt})}{1 + \exp(\tau_{ijt})}$$

$$\tau_{ijt} = \widetilde{D}'_{it}\lambda + a'_{jt}(\varphi + \rho a_{jt} + \Psi_f + \Upsilon D^s_{it} + \kappa_i) \quad \ln \kappa_i \sim N(0, I_m)$$

I have data on respondents' exposure to media from Simmons, which I use to first estimate the $\Upsilon$ parameters by maximum likelihood.[33] This subsection concerns the first stage of the estimation procedure. Once I obtain an estimate of $\Upsilon$, the value of these parameters remain constant for the remainder of the estimation process. For ease of exposition, I suppress the time subscript below.

Recall that the Simmons survey reports only the quintile to which the consumer belongs, therefore I construct an ordered-response likelihood function which I use to obtain an estimate of $\Upsilon$. Let $Y_{im}$ be the amount of exposure individual $i$ has to medium $m$ and let $D^s_i$ be a $dsim(= 13)$ dimensional vector of observed attributes for individual $i$ from the Simmons survey where

$$Y_{im} = D^{s\prime}_i \Upsilon_m + \varepsilon_{im}$$

and $\varepsilon_{im}$ is a mean zero stochastic term with an i.i.d. standard normal distribution. Defining quintile one as the highest, the consumer belongs to the $q$th quintile in medium $m$ if $c_{qm} <$

[33] This two-step estimation procedure can be motivated as follows. The probability that consumer $i$ is informed about product $j$ can be written as

$$1 - \prod_m [1 - \Pr(\mathcal{I}_{ijm} \mid a_{jm}, D_i)]$$

where $\Pr(\mathcal{I}_{ijm} \mid a_{jm}, D_i)$ is the probability that $i$ is informed about product $j$ through medium $m$. Using Bayes' law, this probability can be written as $\Pr(\mathcal{I}_{ijm} \mid a_{jm}, D_i, \mathcal{M}_{im} = 1) \Pr(\mathcal{M}_{im} = 1 \mid a_{jm}, D_i)$ where $\mathcal{M}_{im} = 1$ if $i$ is exposed to medium $m$ and 0 otherwise (with $\Pr(\mathcal{I}_{ijm} \mid a_{jm}, D_i, \mathcal{M}_{im} = 0) = 0$). Assuming consumer characteristics explain exposure to advertising media but do not affect the probability of being informed conditional on media exposure, $\Pr(\mathcal{I}_{ijm} \mid a_{jm}, D_i)$ can be written as $\Pr(\mathcal{I}_{ijm} \mid a_{jm}, \mathcal{M}_{im} = 1) \Pr(\mathcal{M}_{im} = 1 \mid a_{jm}, D_i)$. Since the first probability in the product is not a function of consumer characteristics, the $\Upsilon$ parameters can be estimated separately.

$Y_{im} < c_{(q-1)m}$, where $c_{qm}$ and $\Upsilon_m$ are parameters. Let $Z_{iqm}$ be an indicator function equal to one if $i$'s level of exposure falls in quintile $q$, which yields

$$\Pr(Z_{iqm} = 1) = \Phi\left(c_{(q-1)m} - D_i^{s\prime}\Upsilon_m\right) - \Phi\left(c_{qm} - D_i^{s\prime}\Upsilon_m\right)$$

where $\Phi$ is the cumulative standard normal. The likelihood function for medium $m$ is

$$\pounds_m = \prod_i^N \prod_q^5 \left[\Phi\left(c_{(q-1)m} - D_i^{s\prime}\Upsilon_m\right) - \Phi\left(c_{qm} - D_i^{s\prime}\Upsilon_m\right)\right]^{Z_{iqm}}$$

I estimate the *dsim*-dimensional parameter vector $\Upsilon_m$ separately for each medium. The estimates remain constant throughout the remainder of the estimation procedure. There are two implicit assumptions in this process. The first is that the error terms are independent across media and the second that the errors associated with estimating $\Upsilon$ are independent of the errors associated with the rest of the model.

## 4.2 The Moments

I use individual and product level data to fit moments predicted by the model to their data analogs. First, I discuss the derivation of demand and marginal cost unobservables used in the first two sets of moments. The derivation of these moments follows the algorithm used in BLP. Then I discuss the role of corner solutions and the resulting (third set of ) moments associated with advertising media choices. Finally, I explain the role of the individual-level data in constructing the fourth set of micro moments.

**BLP-Type Moments** Following BLP, I restrict the model predictions for product $j$'s market share to match the observed market shares.

$$S_t^{obs} - s_t(\delta, \theta) = 0, \tag{14}$$

where $S_t^{obs}$ and $s_t$ are the vectors of observed and predicted market shares, respectively. I solve for the mean utility vector $\delta(S, \theta)$ that is the implicit solution to (14). I substitute

$\delta(S,\theta)$ for $\delta$ when calculating the models predictions for the moments discussed in the remainder of this section.[34] Using $\delta(S,\theta)$, I solve for the demand side unobservable used in the first moment

$$\xi_{jt} = \delta_{jt}(S,\theta) - x_j'\beta. \tag{15}$$

In vector form, the $J$ first-order conditions from (12) are

$$s - \Delta(p - mc) = 0$$

where $\Delta_{j,r} = -\frac{\partial s_r}{\partial p_j} I_{j,r}$ with $I_{j,r}$ an indicator function equal to one when $j$ and $r$ are produced by the same firm and zero otherwise. These FOC's imply marginal costs given by

$$mc = p - \Delta^{-1}s \tag{16}$$

Following in the tradition of the new empirical IO (Bresnahan, 1989), I use the estimates of the demand system to compute the marginal costs implied by equation (10). The production cost side unobservable used in the second moment is obtained by combining equations (16) together with equation (10) and rewriting as follows

$$\omega = \ln(p - \Delta^{-1}s) - w'\eta \tag{17}$$

where $\Delta$ and $s$ are functions $\theta$ and $\delta$.

**Advertising Moments (Corner Solutions)** The third set of moments arises from firm's optimal advertising media choices. If products were advertised in all media, construction of these moments would be more straightforward. However, some firms choose not to advertise

[34] I use a contraction mapping suggested by BLP to compute the vector $\delta(S,\theta)$. Under mild conditions on the distribution of consumer tastes Berry(1994) shows that there is a unique relationship between the choice of the vector $\delta$ and the vector of market shares predicted by the model, $s(\delta,\theta)$.

some products in some media. To construct moments that allow for corner solutions, I use a method proposed by Gourieroux *et al.*(1987).

The method is best illustrated by considering a simple example. Let $y_i^*$ denote the latent variable, where $y_i^* = x_i\beta + u_i$. We observe the latent variable if $y_i^* \geq 0$, and 0 otherwise. The errors, $u_i(\beta)$, are linked with the latent variables $y_i^*$. Since they depend on unobserved variables, the errors cannot be used in constructing moments. Gourieroux *et al.* suggest an alternative method: replace the errors by their best prediction conditional on the observable variables, $E[u_i(\beta) \mid y_i]$, and use these to construct moments.

In this paper, the latent variables are optimal advertising levels and, due to non-linearities, the application is more complicated, but the technique is the same. Let $a_{jm}^*$ denote the latent variable. We observe

$$a_{jm} = \begin{cases} a_{jm}^* & \text{if } \partial\Pi_j/\partial a_{jm} \mid_{a_{jm}=a_{jm}^*} = 0 \\ 0 & \text{if } \partial\Pi_j/\partial a_{jm} \mid_{a_{jm}=0} < 0 \end{cases}$$

where $\Pi_j$ is the profit associated with product $j$ given in equation (9). For ease of exposition rewrite the advertising medium FOC (equation 13) as

$$\ln\left(h_{jm}(a_{jm})\right) - w_{jm}^{ad\prime}\psi = \upsilon_{jm} \tag{18}$$

where $h_{jm}$ is the marginal revenue associated with advertising medium $m$ (the left-hand side of equation 13). The latent variable, $a_{jm}^*$, is the implicit solution to equation (18). In this context the errors, $\upsilon_{jm}$, cannot be used since they depend on the latent variable $a_{jm}^*$. Instead, I use the best prediction of the errors conditional on the observed level of advertising to construct moments.

Following the method proposed by Gourieroux *et al.*(1987), to allow for corner solutions I construct moments arising from a tobit likelihood function. Using the marginal costs associated with advertising media (equation 11) and the interior first order conditions (equation

13), the a tobit likelihood function is given by

$$\pounds = \prod_{j:a_{jm}>0} \frac{1}{\sigma_\upsilon} \phi \left( \frac{\ln(h_{jm}) - w_{jm}^{ad\prime}\psi}{\sigma_\upsilon} \right) \prod_{j:a_{jm}\leq 0} 1 - \Phi \left( \frac{\ln(h_{jm}) - w_{jm}^{ad\prime}\psi}{\sigma_\upsilon} \right)$$

where $\phi$ is the standard normal pdf and $\Phi$ is the cumulative standard normal. Due to computational complexity I must be parsimonious in my choice of parameters, therefore in estimation I normalize the value of $\sigma_\nu = 1$. Hence, the corresponding log likelihood function is

$$\sum_j 1(a_{jm} > 0) \ln \phi(\widetilde{h}_{jm}) + 1(a_{jm} = 0) \ln \left(1 - \Phi(\widetilde{h}_{jm})\right)$$

The generalized residual for the $j$th observation is

$$\widetilde{\upsilon}_{jm}(\widehat{\Theta}) = E[\upsilon_{jm}(\widehat{\Theta}) \mid a_{jm}]$$

where $\Theta$ are the parameters associated with equation (18) and $\widehat{\Theta}$ its maximum likelihood estimator. The generalized residual is then

$$\widetilde{\upsilon}_{jm} = \widetilde{h}_{jm} 1(a_{jm} > 0) - \frac{\phi(\widetilde{h}_{jm})}{1 - \Phi(\widetilde{h}_{jm})} 1(a_{jm} = 0)$$

The (third set of ) moments express an orthogonality between the generalized residuals $\widetilde{\upsilon}_{jm}(\widehat{\Theta})$ and the instruments. For instance, the $\Theta$ that solves

$$\frac{1}{J} \sum_j \frac{\partial \widetilde{h}_{jm}}{\partial \Theta} \widetilde{\upsilon}_{jm} = 0$$

is the MOM estimator, where $\frac{\partial \widetilde{h}_{jm}}{\partial \Theta}$ are the appropriate instruments. Let $\mathcal{T}(\delta, mc, \theta, \eta_{\mathrm{AD}})$ be the vector formed by stacking the residuals over media and over products. I use the resulting sample moments to construct a GMM estimator.

**Micro Moments** The Simmons *Survey of Media and Markets* consists of a sample of consumers, their characteristics, and their manufacturer purchase information. Up until this point the Simmons data have been used only to estimate the $\Upsilon$, in calculating the other moments demographic characteristics were given by the CPS. The micro-moments are constructed using individual level purchase data from the Simmons survey. Therefore the demographic characteristics for the micro-moments are not given by the CPS, but are linked directly to purchases as detailed in the Simmons data (denoted $D^s$). Since these moments are constructed from individual purchase data one can think of them as being particularly informative about the parameters of the taste distribution ($\Pi$ and $\Sigma$).

Recall the survey reports only the firm from which the PC was purchased. Let $B_i$ be a $F \times 1$ vector of firm choices for individual $i$. Let $b_i$ be a realization of $B_i$ where $b_{if} = 1$ if a brand produced by firm $f$ was chosen. Define the residual as the difference between the vector of observed choices in the data and the model prediction given $(\delta, \theta)$. The residual for individual $i$, denoted $G_i(\delta, \theta)$, can be written as

$$G_i(\delta, \theta) = b_i - E_{\nu,\kappa} E[B_i \mid D_i^s, \delta, \theta] \tag{20}$$

For example, the element of $E_{\nu,\kappa} E[B_i \mid D_i^s, \delta, \theta]$ corresponding to firm 2 for consumer $i$ is

$$\sum_{j \in \mathcal{J}_2} \int \sum_{\mathcal{S} \in \mathcal{C}_j} \prod_{l \in \mathcal{S}} \phi_{ilt} \prod_{k \notin \mathcal{S}} (1 - \phi_{ikt}) \frac{\exp\{\delta_{jt} + \mu_{ijt}\}}{y_{it}{}^{\alpha} + \sum_{r \in \mathcal{S}} \exp\{\delta_{rt} + \mu_{irt}\}} dF_{\nu}(\nu) dF_{\kappa}(\kappa)$$

where the first summand is over all the products sold by firm 2, the integral is over the assumed distributions of unobserved consumer attributes ($\nu$ and $\kappa$), and the second summand is over all the different choice sets that include product $j$. The population restriction for the micro moment is $E[G_i(\delta, \theta) \mid (X, \xi)] = 0$. Let $G(\delta, \theta)$ be the vector formed by stacking the residuals $G_i(\delta, \theta)$ over individuals.[35]

## 4.3 Identification

In this section, I present an informal discussion of how variation in the data identifies the parameters of the model. I begin with the parameters of the demand side. Recall that associated with each PC is a mean utility, $\delta_j = x_j' \beta + \xi_j$, which is chosen to match observed

[35] The Simmons sample is annual so in estimation the outermost summand is over all products sold by each firm over the course of the year.

and predicted market shares. Heterogeneity in consumers tastes around the mean is given by $\alpha \ln(y_{it} - p_{jt}) + x_j{}'(\Pi D_{it} + \Sigma \nu_i)$. If consumers were identical, then all variation in sales would be driven by variation in product attributes. Variation in product market shares corresponding to variation in the observable attributes of those products (such as CPU speed) is used to identify the parameters of mean utility ($\beta$). However, while a PC may have attributes that are preferred by many consumers (high $\beta$'s), the PC may also have attributes that appeal to certain types of consumers. For instance, if children like to play PC video games, then consumers from large households may place a higher valuation on CPU speed relative to smaller households.

Identification of the parameters of the taste distribution $(\Sigma, \Pi)$ relies on information on how consumers substitute. There are two issues that merit attention. First, new product introductions are common in the PC industry.[36] Variation of this sort is helpful particularly for identification of the variance in tastes. Since the distribution of unobserved consumer heterogeneity is fixed over time, variation in sales patterns over time as the choice sets change allows for identification of $\Sigma$. Second, I augment the market level data with individual level data on manufacturer choice. The extra information in the micro data helps with identification. That is, it allows variation in choices to mirror variation in tastes for product attributes. Correlation between $x_j D_i$ and choices identifies the $\Pi$ parameters.

The information technology ($\phi$) describes the effectiveness of an ad at informing consumers. Most of the variation in $\phi$ (and the induced variation in market shares) is due to variation in advertising. The model should incorporate the impact of media-specific advertising on $\phi$ and its differing effect across households ($\Upsilon$). One major drawback of the aggregate advertising data is that I don't observe variation across households. However, observed variation in market shares corresponding to variation in household ad media exposure is necessary to identify the $\Upsilon$ matrix. Fortunately, the individual-level data contain useful information on media exposure across households. Using the Simmons data, and

[36] I assume that this is exogenous variation as is common in the literature.

taking media exposure as a proxy for ad exposure, I estimate the $\Upsilon$ separately by maximum likelihood. Covariation in observable consumer characteristics and choices of media exposure identifies the $\Upsilon$. The use of the Simmon's data allows me to side-step the need for observed ad heterogeneity across households.

Variation in sales corresponding to variation in advertising identifies the remainder of the parameters of $\phi$. Returns to scale in media advertising ($\rho_m$) are identified by covariation in sales with the variance of $a_{jm}$. The relative effectiveness of advertising in different media ($\varphi_m$) is identified by any variation in sales due to variation in $a_{jm}$, which is constant across households (and not explained by returns to scale). Conditioning on $\Upsilon D^s$, any extra variation across individuals due to extra $a_{jm}$ that is not explained by the model is captured in $\kappa_{im}$ : unobserved heterogeneity with regard to ad medium effectiveness. The parameter $\Psi_f$ is identified by additional variation in firm market shares due to advertising that is not explained by other elements in $\phi$. That is, it captures the fact that some firms are more effective at informing consumers through advertising, where this "extra" effectiveness is independent of media or households. Finally, the parameters on group advertising ($\gamma$ and $\pi$) are identified by covariation in average medium advertising per product ($\overline{ad}_m$) over time and by functional form assumptions (relative to what a logit specification would imply for the parameter values in the case of the outside good). The other parameters of the information technology function which do not interact with advertising ($\lambda$) are separately identified from $\Pi$ due to non-linearities in the model.

Variation in prices and markets shares corresponding to variation in observed cost characteristics identifies the corresponding cost characteristics' effect on production costs. Covariation in advertising prices, advertising and the generalized residuals identifies the effect of advertising prices on advertising costs.

## 4.4 The GMM Estimator

I use generalized method of simulated moments to find the parameter values that minimize the objective function

$$\Lambda' Z A^{-1} Z' \Lambda$$

where $A$ is an appropriate weighting matrix which is a consistent estimate of $E[Z'\Lambda\Lambda' Z]$ and $Z$ are instruments orthogonal to the composite error term $\Lambda$. Specifically if $Z_\xi, Z_\omega, Z_{ad}, Z_{micro}$ are the respective instruments for each disturbance/residual the sample moments are

$$Z'\Lambda = \begin{bmatrix} \frac{1}{J}\sum_{j=1}^{J} Z_{\xi,j}\xi_j(\delta,\beta) \\ \frac{1}{J}\sum_{j=1}^{J} Z_{\omega,j}\omega_j(\delta,\theta,\eta) \\ \frac{1}{J}\sum_{j=1}^{m*J} Z_{ad,j}\mathcal{T}_j\left(\delta,\theta,\eta_{AD}\right) \\ \frac{1}{N}\sum_{i=1}^{N} Z_{micro,i}G_i(\delta,\theta) \end{bmatrix}$$

where $Z_{\xi,j}$ is column $j$ of $Z_\xi$. Joint estimation of the parameters takes into account the cross-equation restrictions on the parameters that affect both demand and supply, which yields more efficient estimates. This comes at the cost of increased computation time since joint estimation requires a non-linear search over all the parameters of the model. As in Nevo (2000), I restrict the non-linear search to a subset of the parameters $\Omega = \{\theta, \eta_{AD}\}$, reducing the searching time. This restriction is possible since the first-order conditions with respect to $\beta$ and $\eta$ can be expressed in terms of $\theta$.[37]

**Instruments** The estimation method outlined above requires instruments that are correlated with specific functions of the observed data, but not correlated with the disturbances. That is, appropriate instruments satisfy (i) $E[\Lambda \mid Z] = 0$ and (ii) $Z'E[\Lambda_\Omega]/NM$ converges to a nonsingular matrix where $NM$ is the sample size and $\Lambda_\Omega = \partial\Lambda/\partial\Omega$. The discussion

[37] I provide the details in Appendix A.

surrounding the instruments used in estimation proceeds as follows. First, I discuss assumptions I make regarding the relationship between unobservables and observables and the implications of these assumptions, then I discuss endogeneity issues, the form of optimal instruments, and finally the exogenous instruments I use, which are an approximation to optimal instruments.

A common assumption made in the literature regards the relationship between unobservable and observable characteristics. Namely, the unobservables associated with demand and pricing (evaluated at the true value of the parameters) are assumed to be mean independent of a vector of observable product characteristics and cost shifters, $(x, w)$ :

$$E\left[\xi_j \mid (x, w)\right] = E\left[\omega_j \mid (x, w)\right] = 0 \tag{21}$$

In choosing which variables to include in the conditioning vector I look for observables that shift demand and cost functions. As can be seen from pricing first-order conditions in equation (12), the optimal price for product $j$ depends upon characteristics of all of the products offered. Products which face more competition (due to more rivals offering similar products) will tend to have lower markups relative to more differentiated products. Similarly from the advertising FOCs in equation (13) we see that advertising for product $j$ also depends in part on the markup for product $j$. A firm will tend to advertise a product more the more they make on the sale of the product. The optimal price and advertising for product $j$ depends upon the characteristics, prices, and advertising of all products marketed. Thus the optimal instruments associated with product $j$ will include functions of attributes and cost shifters of all other products.

It is important to note that the moment restriction in equation (21) has nontrivial economic implications. Indeed while some product characteristics may be uncorrelated with unobserved attributes (such as CPU speed), others such as prices and advertising are likely to be correlated with unobserved attributes. Recall that the econometrician does not observe the product characteristic $\xi_j$ or the cost characteristic $\omega_j$, but market participants do. For example, $\xi_j$ could represent unobserved product quality, which firms observe and base pricing and product advertising decisions upon. This leads to endogeneity problems since prices and advertising choices are most likely functions of unobserved characteristics. If the correlation between unobserved characteristics ($\xi_j$ and $\omega_j$) and prices and advertising is ignored, it will lead to biased estimates. For example, if price is positively correlated with unobserved quality, price coefficients (in absolute value) will be understated (as preliminary estimates indicate, see section (5)). Whereas if advertising is positively correlated with quality, advertising's effect will be overstated.

Berry (1994) was the first to discuss the implementation of instrumental variables methods to correct for endogeneity between unobserved characteristics and prices, and BLP provide an estimation technique. My model and estimation strategy is in this spirit but is adapted to correct for advertising endogeneity.

Given the mean independence assumption above and some additional regularity conditions, Chamberlain (1987) shows the optimal instrument for any disturbance-parameter pair is the expected value of the derivative of the disturbance with respect to the parameters (evaluated at the true value of the parameters, $\Omega_0$). That is, the optimal instrument for each parameter is one that places more weight on the disturbances that are most sensitive to changes in the parameter value (at $\Omega = \Omega_0$). Consider the optimal instruments for the market-level disturbances $\xi$ and $\omega$. These are

$$E\left[\frac{\partial \xi_j(\Omega_0)}{\partial \Omega} \mid (x, w)\right] \quad \text{and} \quad E\left[\frac{\partial \omega_j(\Omega_0)}{\partial \Omega} \mid (x, w)\right] \tag{22}$$

evaluated at a consistent estimate for $\Omega_0$.[38] Product characteristics are optimal instruments for the demand side $\beta$ parameters, and likewise cost characteristics are optimal instruments for the $\eta$ cost parameters. However, the optimal instruments for the other parameters are functions of either prices or advertising. Due to endogeneity of price and advertising the expectations given in equation (22) will be correlated with the disturbances, that is (i) is

[38] For completeness, the optimal instruments for the advertising and micro-residuals are $E\left[\frac{\partial \widetilde{h}(\Omega_0)}{\partial \Omega} \mid (x, w, w^{ad})\right]$ and $E\left[\frac{\partial G(\Omega_0)}{\partial \Omega} \mid (x, \xi)\right]$.

violated. Similarly, the optimal instruments for the advertising and micro residuals are functions of price and advertising.

Since the derivatives are functions of advertising and prices, to calculate the optimal instruments I would have to calculate the pricing and advertising equilibrium for different $\{\xi_j, \omega_j\}$ sequences, compute the derivatives at equilibrium values, and integrate out over the distribution of the $\{\xi_j, \omega_j\}$ sequences. This is computationally demanding and perhaps more importantly, requires additional assumptions on the joint distribution $(\xi, \omega)$.

Instead of computing the optimal instruments, I form approximations to them. The approach I take is in the spirit of BLP(1999). While the optimal instruments are the expectation of the derivative of the disturbance, the approximations are formed by evaluating the derivatives at the expected value of the unobservables (i.e. at $\xi = \omega = 0$). The estimate of the instruments will be biased since the derivatives evaluated at the expected values are not the expected value of the derivatives. However the approximations are functions of exogenous data, and are constructed such that they are highly correlated with the relevant functions of prices and advertising. Hence the exogenous instruments will be consistent estimates of the optimal instruments.[39]

The method used to construct the exogenous instruments is as follows:

**(i)** Construct initial instruments for prices ($\widehat{p}_{int}$) and advertising.[40]

**(ii)** Use the initial instruments to obtain an initial estimate of the parameters, $\widehat{\Omega}$.

**(iii)** Construct estimates of $\delta$, $mc$,and $mc^{ad}$. I used $\widehat{\delta} = x\widehat{\beta}$ , $\ln(\widehat{mc}) = w\widehat{\eta}$, and $\ln(\widehat{mc}^{ad}) = w^{ad}\widehat{\psi}$.

**(iv)** Solve the first-order conditions for equilibrium advertising, $\widehat{a}$, as a function of $\widehat{\Omega}, \widehat{\delta}, \widehat{mc}$,

[39] BLP(1995) show that variables that shift markups are valid instruments in models of differentiated products markets. One could also use a series approximation as in BLP to construct exogenous instruments. I chose to use the approximation method outlined above since it is more closely tied to the model.

[40] I constructed a distance variable based on observables and used kernel estimates for prices and advertising as the initial instruments. BLP-type instruments would also work for prices.

$\widehat{mc}^{ad}$, $\widehat{p}_{int}$ and $x$.

**(v)** Solve the first-order conditions of the model for equilibrium prices, $\widehat{p}$, as a function of $\widehat{\Omega}, \widehat{\delta}, \widehat{mc}$, $\widehat{a}$, and $x$.

**(vi)** These imply a value for predicted market shares, $\widehat{s}$, which is a function of $\widehat{\Omega}, \widehat{p}, \widehat{\delta}$, $\widehat{a}$ and $x$.

**(vii)** Calculate the required disturbance-parameter pair derivatives.[41]

**(viii)** Repeat steps (iv)-(vii) where each time the new $\widehat{p}_{int}$ is replaced by the $\widehat{p}$ found from the previous round.

**(iv)** Form approximations to the optimal instruments by taking the average of the exogenous derivatives found in step (vii)

**Simulation** There are two separate reasons why I must simulate market shares. First, as is common in many papers using random coefficient models of demand, I use an empirical distribution to define the distribution of consumer demographics. As a result there is no analytical solution for predicted market shares, even if one assumes consumers are aware of all products for sale. Second, consumers may not know of products for sale, but I don't observe the choice set they face when making a purchase decision. Due to the large number of products in the PC industry, it is not feasible to calculate all possible purchase probabilities for each product corresponding to each of the $2^J$ possible choice sets. Instead I simulate a choice set for each individual in each period and construct an importance sampler to smooth the simulated choice probabilities.

A general outline of the simulation technique follows (for more detail see Appendix B). I sample a set of "individuals" where each individual consists of $(v_{i1}, \ldots, v_{ik})$ taste parameters drawn from a multivariate normal; demographic characteristics, $(y_i, D_{i1}, \ldots, D_{id})$ drawn

[41] These are $\frac{\partial \widehat{\xi}_j(\widehat{p},\widehat{a},\widehat{s},\widehat{\delta},x,\widehat{\Omega})}{\partial \Omega}$, $\frac{\partial \widehat{\omega}_j(\widehat{p},\widehat{a},\widehat{s},\widehat{\delta},\widehat{mc},x,\widehat{\Omega})}{\partial \Omega}$, $\frac{\partial \widehat{\widehat{h}}(\widehat{p},\widehat{a},\widehat{s},\widehat{\delta},\widehat{\Omega})}{\partial \Omega}$ and $\frac{\partial \widehat{G}(\widehat{p},\widehat{a},\widehat{s},\widehat{\delta},\widehat{\Omega})}{\partial \Omega}$.

from the CPS in the case of the macro-moments and data in the case of the micro-moments; and unobserved advertising medium effectiveness draws $(\kappa_{i1}, \ldots, \kappa_{im})$ from a multivariate log normal.[42]

For the macro moments, for each individual, I draw $J$ uniform random variables. For a given value of the parameters, I compute the probability each consumer is informed about each product $(\phi_{ij})$. Then I construct a choice set for each individual by comparing her vector of $\phi_i$'s with her uniform draws and compute choice probabilities. I construct an importance sampler by using the initial choice set weight to smooth the simulated choice probabilities.[43] The simulator for the market share is the average over individuals of these smoothed choice probabilities. The process is similar for the micro moments, but I take $R$ draws for each individual. I construct a simulator for individual product choice probabilities which is the average over the $R$ draws. I construct individual firm choice probabilities by summing over the products offered by each firm.

**The Estimation Algorithm and Properties of the Estimator** In summary, in the first stage I calculate the $\Upsilon$ parameters by maximum likelihood and hold them constant for the duration of estimation procedure. In the second stage, I employ the following estimation algorithm. Given a value of the parameters, $\Omega$ (1) Compute (via simulation) the market shares implied by the model (see equation 6) (2) Solve for the vector $\delta$ that equates simulated market shares and observed shares (3) Calculate $\beta$ and compute the vector of demand unobservables $\xi$ (see equation 15) (4) Calculate $\eta$ and compute the cost side unobservable, $\omega$ (see equation 17) (5) Compute the ad residual implied by the tobit likelihood function $\mathcal{T}$ (6) Compute (via simulation) the firm purchase probabilities implied by the model (7) Calculate the micro moment residual (see equation 20) (8) Calculate the

[42] I sample 3000 individuals each year from the March CPS for the macro moments. For the micro moments, I sample 6700 individuals each year from the 20,000 surveyed.

[43] The initial choice set weight is the product over all the $\phi$'s for products in the choice set (computed at the initial value of the parameters) multiplied by the product of $(1 - \phi)$ for all the products not in the choice set.

instruments and interact them with the macro disturbances and micro residuals. Hold the instruments fixed at these values for the duration of the estimation. (9) Search for the parameter values that minimize the objective function

$$\widehat{\Lambda}' Z A^{-1} Z' \widehat{\Lambda}$$

where $\widehat{\Lambda}$ is the composite error term resulting from simulated moments. If the parameters don't minimize the moments (according to some criteria) make a new guess of the parameters. Repeat until moments are close to zero.[44]

Using the results of Pakes and Pollard (1989), this estimator is consistent and asymptotically normal. As the number of psuedo random draws used in simulation $R \to \infty$, the method of simulated moments covariance matrix approaches the method of moments covariance matrix. To reduce the variance associated with simulation, I employ antithetic acceleration as described in the simulation literature (for an overview of simulation techniques see Stern, 1997 and 2000). Geweke (1988) shows if antithetic acceleration is implemented during simulation then the loss in precision is of order $1/N$ (where $N$ are the number of observations), which requires no adjustment to the asymptotic covariance matrix. The reported (asymptotic) standard errors are derived from the inverse of the simulated information matrix which allows for possible heteroskedasticity.[45]

# 5 Preliminary Analysis

Before estimating the full-model, I conduct a series of logit and probit regressions, which allows me to examine in a simple framework how advertising impacts demand and supply and guides the choice of variables to include in the structural analysis.[46]

---

[44] I use BHHH derivative based optimization routine to obtain parameter estimates. To obtain a smooth simulator (necessary in order to use derivative based optimization algorithms), I construct an importance sampler. This sampler uses the initial choice set weight to smooth choice probabilities.

[45] The reported standard errors do not included additional variance due to simulation error.

[46] While reduced-form estimation is computationally easy, structural analysis has many advantages. It provides estimates that are invariant to changes in policy or competitive factors. Another advantage is

First I estimate a series of probit models of the decision to purchase a PC using the individual-level Simmons data (see Table 5). I started by allowing for many explanatory variables, including interactions between consumer attributes, education and income splines, and media exposure variables.[47] I found the consumer attributes which matter most are age, education, and marital status. Household income and size also significantly affect the probability of purchase, although including the presence and/or number of kids does not improve the fit.

The estimates in the first two columns suggest media exposure effects the decision to buy a PC, controlling for observed consumer covariates.[48] Results from a likelihood ratio test (columns three and four) suggest that exposure to TV and magazine media impact the decision the most.[49] Results with no media exposure variables (final two columns) indicate that media exposure does matter. Indeed, we can reject the hypothesis that media exposure has no effect on PC purchase at a smaller than 0.01 significance level.

I next estimate a multinomial logit model of demand to study the effects of advertising on product choice using all datasets. Due to data restrictions I estimate a model of manufacturer choice. The probability consumer $i$ purchases a PC produced by firm $f$, $P_{if}$, is given by

$$P_{if} = \frac{\exp\{\beta_f X_{if}\}}{1 + \sum_{r=1}^{F} \exp\{\beta_r X_{ir}\}}$$

where $X_f$ includes product characteristics (such as average weighted price, average weighted CPU speed, proportion of pentium processors, and proportion of laptops), manufacturer

that it allows one to specify the effects of advertising. If advertising affects a consumers choice set, we would expected changes in consumer behavior as advertising changes. This effect of a limited choice set is not captured in non-structural models since it is not possible to be specific about how advertising affects demand. Also, we would expect changes in firm behavior as variables relating to advertising change, which will have an impact on markups and prices.

[47] Table available on request.

[48] There may be unobserved consumer attributes which influence media exposure. I account for this possibility in the structural model, which allows for unobserved consumer heterogeneity in media effectiveness in the information technology (these are the $\kappa_i$ from section 3.2).

[49] The final rows present the results from likelihood ratio tests. We cannot reject the hypothesis that all other media have no impact on purchase probabilities.

characteristics (such as advertising expenditures), and manufacturer specific parameters, $\beta_f$.

Selected results are given in Table 6.[50] In the first specification (column one), the only explanatory variable is price. All coefficient estimates are positive and significant (except for Epson). The most obvious explanation is that prices are correlated with quality: it appears consumers prefer a higher price, when most likely they prefer a higher quality product Even after including CPU type, CPU speed, and laptop as explanatory variables, the majority of the price coefficients are still positive. This suggests there are product attributes (in addition to those mentioned) that are positively correlated with prices.

In a second specification (column two), I include total advertising expenditures as an explanatory variable. The specification with advertising fits the data better than those without advertising. Indeed, it fits better than the specification with product characteristics included even though there are fewer explanatory variables. Without specifying how advertising affects demand, the coefficient estimates indicate that advertising may be correlated with higher quality. This obtains from comparing the estimates in the first two columns, the price coefficients in the specification with advertising are much smaller. It seems that advertising may be picking up some of the effect of unobserved product attributes. As is common in the empirical IO literature, I will account for the possibility that unobserved attributes are correlated with prices and correct for the possible correlation with advertising.

I also ran specifications broken down across advertising media. The results are given in the final column of Table 6. I found advertising's effect on manufacturer choice differs across media. The coefficient estimates indicate that advertising in magazines and newspapers has a positive (and significant) effect on choice for almost all PC manufacturers. Recall, consumer level probit estimates suggest exposure to TV and magazine media mattered in

[50] I estimate the parameters for two separate definitions of the 'outside' good. Under the first specification the outside good encompasses no purchase of a new PC. Under the second, it encompasses no purchase of one of the 15 inside options. The results did not differ greatly, so I report the results under the second more broad definition of the outside good.

the decision of whether to purchase a PC. Finally, I find that, after controlling for observed consumer covariates, advertising still influences the decision of manufacturer choice.

# 6 Structural Estimation Results

I use results from the preliminary regressions to guide the choice of which consumer and product variables to include. While there are many viable characteristics, I must be parsimonious in my choice, due to the computational complexity of the model. Included product characteristics ($x_j$) are CPU speed (MHz), a dummy for whether the computer has a Pentium chip, a dummy for whether the computer is a laptop, firm level fixed effects, and a constant. CPU speed and Pentium are measures of computational speed and laptop is a measure of convenience. In all tables, the (asymptotic) standard errors are given in parentheses.

The structural results are broken down into three categories. First, I discuss the importance of product differentiation and the substitution patterns present in the PC market. Next, I discuss the importance of advertising: its influence on consumers, and implications for firm behavior and the resulting nature of competition. Finally, I discuss results which highlight the importance of information in the PC industry and contrast these with results from different base-line models in which consumers are completely informed.

**Differentiation and Substitution Patterns** Not surprisingly, the results indicate there is much variation across consumers with respect to product attributes. (Recall from the model that the marginal utility of product attributes varies across consumers.) I estimate the means and the standard deviations of the taste distribution for CPU speed, Pentium, and laptop. The mean coefficients ($\beta$) are given in the first column, first panel in Table 7. Estimates of heterogeneity around these means are presented in the next few columns. The means of CPU speed and laptop are positive and significant. The coefficient on the interaction of CPU speed with household size is significant, while the other coefficients on interactions with demographics ($\Pi$) are insignificant. The results imply that the product characteristics CPU speed and laptop have a significant positive effect on the distribution of the utilities. In addition, the marginal valuation for CPU speed is increasing in household size (3.9). This result is intuitive since children often use the PC to play games (which require higher CPU speeds).

None of the coefficients for the Pentium dummy are significant. This is a somewhat surprising result and suggests that once you control for CPU speed (and other product characteristics) consumers don't place extra value on whether the chip is a Pentium. During the period considered in this study 80% of PCs had a Pentium chip; the AMD chip was not yet a strong market contender. In that light, the results may not be so surprising.

The non-random coefficient results are also presented in the first panel. The coefficient on $\ln(y-p)$ is of the expected sign and is highly significant (1.2). Firm fixed effect estimates indicate that the marginal valuation for a product is higher if it is produced by Apple, Dell, IBM or Packard Bell. This could capture prestige-effects of owning a computer produced by one of top firms (Apple, IBM, and Packard Bell). Apple operates on a different platform, so Apple fixed effects could reflect the extra valuation consumers on average place on the Apple platform. Finally, they could capture extra valuation consumers place on enhanced services offered by the firms (for instance Dell is known for its excellent consumer service) or other reputational effects.

The estimated parameters have important implications for pricing behavior and markups. Table 8 presents a sample from 1998 of own- and cross-price elasticities of demand. Elasticities are computed by multiplying the numerical derivative of estimated demand by price and dividing by actual sales. The table shows all negative elements on the diagonal. Consistent with oligopolistic conduct, the results indicate that the products are priced in the elastic portion of the demand curve. The substitution patterns implied by these elasticities are intuitive. The results show that products are more sensitive to changes in prices of computers with similar characteristics. For example, Apple computers are most sensitive to changes in the prices of other Apple computers, implying there is less substitution across platforms. Among PC's that have a windows operating system, form factor plays a strong role in driving substitution patterns. For example, Compaq Presario laptop is most sensitive to changes in prices of other laptops rather than to changes in other Compaq non-laptop computers. These patterns are consistent across the data.

**Differentiation and Advertising** The effect of advertising also varies across consumers. The effect of advertising on a consumers information set is determined by the information technology function. I estimate some of the parameters of the information technology separately before the rest of the model ($\Upsilon$). The parameter estimates, given in Table 9, measure how medium exposure varies with observed demographic characteristics. The coefficients can be used to proxy for effectiveness of ads in reaching consumers through various media. The results indicate that exposure varies across media and households. The parameter estimates suggest magazine advertisements are most effective at reaching mature, high income individuals and the effectiveness is increasing in household size. Newspaper advertising is most effective at reaching married individuals above the age of 30 who have a high income. However, newspaper advertising is less likely to reach a family the larger is their household size (-0.03). Hence, newspaper advertising targeted at large households would not be effective in increasing the probability of being informed for this particular cohort. Perhaps not surprisingly, TV advertising is the most effective media for reaching low-income households. Television advertising is also effective at reaching mature, married individuals, although not as effective as newspaper advertising. Most of the advertising in the PC industry is in magazines; this suggests PC manufacturers are targeting high-income households.

Some products are advertised in groups while others are advertised individually. The coefficient estimates on group advertising ($\gamma$) and group advertising squared ($\pi$) are given in the second panel of Table 7. These (unrestricted) estimates predict that we will observe both group advertising and product-specific advertising, which is supported by the data.

The estimate on advertising squared (0.09) indicates there are economies of scope in group advertising. Specifically, the estimates imply that if average group ad expenditures ($\overline{ad}$) for a particular product group are above a threshold level of \$1.4 million per quarter[51] (either the advertising expenditures for a group are high or the groups are small) the firm will find it worthwhile to engage in group advertising to capitalize on the returns to scope. To put this into context, in the first quarter of 1998 Apple's advertising strategy involved 18 group advertisements. The parameter estimates suggest we would observe 18 group advertisements only if Apple's home-sector advertising budget was at least \$25 million. As we see from Table 2, Apple spent over \$180 million in advertising in 1998, and more than 25\$ million of that was in the first quarter – consistent with the model prediction.

The information technology coefficients are presented in the third panel of Table 7. Consumers may be informed without seeing any advertisements. The coefficient estimate for income less than \$60,000 (dummy variable for low income), 0.48, indicates that low income individuals are likely to be informed about 38% of the products without seeing any advertisement. Having a high income is not significantly different from having a middle income in terms of being informed. This is perhaps not surprising since low income individuals are likely to have lower opportunity costs of time and thus more time to search for information about available products. The coefficient estimate for high school graduate, implies that the probability of being informed without seeing any advertising is higher for high-school graduates relative to non-high school graduates. These results suggest that other means of information provision, such as word-of-mouth or experience play a role in informing consumers in this market.

The results also indicate that there are decreasing returns to advertising in the tv (-0.03) and newspapers and magazines (-0.02) media, but that they are decreasing at a faster rate for tv advertising. Estimates of firm fixed effects interacted with advertising ($\Psi$) indicate that some firms are more effective at informing consumers through advertising. Most notably,

[51] The ad threshold is $(1-\gamma)/\pi$.

ads by Compaq, Gateway, IBM and Packard Bell are significantly more effective, which could arise due to differences in advertising techniques.

Estimated advertising elasticities of demand indicate that, for some firms, advertising one product has negative effects on other products sold by that firm, but it is less negative than for some of the rival products. Table 10 presents a sample from 1997. Each semi-elasticity gives the percentage change in the market share of the row computer associated with a \$1000 increase in the advertising of the column computer. For instance, a \$1000 increase in advertising for Apple Macintosh computers results in a decreased market share of around 0.1% for Compaq Presario brand computers, but has very little effect on the market share for Apple PowerBook computers.[52] In contrast, an increase in advertising for Dell Latitude has a large effect (relative to increase in own market share of less than 0.01%) on the market share for Dell Dimension computers (decline of 0.04%).

The cost and non-home sector estimates are given in Table 11. The marginal cost of production estimates are given in the first panel. Most of the coefficients ($\eta$) are of the expected sign and are significantly different from zero. The estimates indicate marginal costs are declining over time and increases in CPU speed or offering a laptop increase marginal costs. The only variable with an unexpected sign is Pentium (-0.39), the negative sign indicates that PCs with a Pentium chip are cheaper to produce. The coefficient estimates for log marginal cost of advertising is given in the second panel of Table 11. The coefficient on the (log) price of advertising ($\psi$) is highly significant, and indicates that there are not many product-specific cost characteristics that affect the cost of advertising.

The parameter estimates associated with non-home sector marginal revenue are given in the bottom panel. All coefficients are positive and significant. Recall, from the discussion in section 2.2, that 50% of industry advertising expenditures are by IBM. My conjecture that IBM's advertising expenditures were so large (relative to other firms) was due to non-

[52] The diagonal elements report the increase in market share from own-advertising. For example, an increase of \$1000 for advertising on Dell Dimension results in an increased market share of 0.03%.

PC related enterprises seems to be supported by the data. I included non-PC sales in the non-home marginal revenue specification (last panel of Table 11) to adjust for the fact that the measure of advertising includes that for non-PCs. Indeed the coefficient on non-PC sales (5) is significant (although only at the 10% level) and positive. However, the interaction term between IBM and advertising in the information technology function (0.62) indicates that advertising by IBM is still more effective relative to some other firms after controlling for non-PC enterprises. If the IBM fixed effect in the information technology were not significantly different from zero than I would have concluded that IBMs presence in the non-PC sector fully explained their large advertising expenditures.

To gain more insight into the advertising choices of firms and to learn more about the competitive nature of the PC industry, I use estimated demand to infer marginal costs and markups. Summary statistics are given in Table 12. The median markup charged by PC firms is 19% over marginal costs of production and 10% over per unit production and (estimated) advertising costs. As can be seen from the first two rows, the top firms have higher than average markups and engage in higher than average advertising relative to the total industry. Indeed, the non-top firms average median markups is a much lower 14% with an ad-to-sales ratio of about 2%. The final column shows that even after controlling for the fact that the top firms advertise more, they continue to earn higher than average markups.

The bottom portion of the table gives detailed information for the top firms. Firms advertising choices are determined by their markup and their advertising elasticity of demand, as can be seen from examining the advertising FOCs given in equation (13). IBM has one of the highest ad-to-sales ratios. The advertising demand elasticities for IBM are not more sensitive to advertising relative to other top firms, however, IBM markups are higher than average. The results indicate that IBM is advertising more than the average non-top firm because they earn more per product than the average non-top firm. Compaq, on the other hand, has one of the highest markup margins, 24%, but still advertises less than average (although not less than the average non-top firm). As expected Compaq's demand is less sensitive to advertising relative to other firms in the industry, which is the driving factor in their advertising decision. In addition, the table shows that Gateway has the highest median price of the top firms, but earns lower than average markups. The lower markups are due to higher than average costs, as reflected in a higher than average cost unobservable ($\omega$), suggesting they are not as cost-effective in making their computers.

**Information** High industry markups are explained in part by the fact that consumers know only some of the products for sale. If all consumers had full information (the assumption made in the literature to date) the market would look very different. Table 13 presents the markups resulting from a model of limited information, to those predicted by traditional consumer choice models. I estimated a benchmark BLP model for the baseline model of comparision.[53] Estimating the BLP model allows me to examine the additional markup firms earn as a result of limited consumer information. The estimates indicate industry median markups would be 5% under full information, one-fourth the magnitude of those under limited information.

The bottom rows of Table 13 present markup comparisions broken down by top firms, with some representative products for each firm. The model of limited information suggests there is a larger markup gap between the top firms and the industry average relative to the prediction under full information. Not surprisingly, the firm with the largest negative percentage change in markups is IBM, the one that spends the most on advertising currently. The extent to which a firm in can exercise market power depends on the elasticity of its demand curve. A comparision of estimated firm price elasticities for the top firms under both models is given in Table 14. The greater the number of competitors or the larger the cross-elasticity of demand with the products of other firms, the greater the elasticity of the

[53] More accurately, I estimate a BLP model with micro-moments. Since my focus is on examining the effect of advertising, I include the micro-moments in estimating the BLP model to obtain as precise estimates of the parameters of the taste distribution as possible given the data (see Petrin for more detail). The parameter estimates are given in Appendix C.

firm's demand curve and the less its market power.

The model of full information presents an image of an industry that is quite competitive, and indicates markups are similar across the top firms. In addition, demand is very sensitive to price changes and cross-elasticites imply the products are somewhat substitutable. However, if we remove the full information assumption the industry looks very different. Firms have much more market power as evidenced by the elasticities given along the diagonal in the top panel. Also cross-price elasticities in the top panel indicate products are not as substitutable. This is intuitive, if consumer know of fewer products than firms effectively face fewer competitors resulting in a less competitive industry. The results suggest that traditional models of full-information yield estimates for product specific elasticities that are biased towards being too elastic. Hence industry analysts could reach incorrect conclusions regarding the nature of competition in an industry, if they use elasticities and markups based on models of full information.

# 7 Specification Tests

I examined the robustness of the informative advertising model by conducting a number of goodness-of-fit tests. First I tested whether all moments were satisfied. This test is conditional on all the assumptions of the model and therefore tests the overidentifying moment restrictions together with all the functional form and distributional assumptions. By construction the objective function

$$\Lambda' Z\{\text{ Est. Asy. } Var[Z'\Lambda]\} Z'\Lambda$$

is a Wald statistic and is distributed chi-squared with degrees of freedom equal to the number of moment restrictions less the number of parameters. However, the test is stringent and generally rejects for large samples. It is not surprising then, given the large sample size and stylized nature of the model, that the model is rejected by the data.

I conducted additional goodness-of-fit tests focused on various aspects of the model. To conduct the tests I partitioned the region in which the response variables (and in some cases covariates) lie into disjoint cells.[54] I calculated the quadratic form based on the difference between the observed number of outcomes in each cell and the expected number (given the observed covariates). If the model is correct, the normalized quadratic form converges in distribution to a chi-square random variable as the sample size increases.

Formal tests were not able to reject the null that the predicted values for market shares are the same as the observed values.[55] I also constructed test statistics based on the average value of shares that fall into specified cells. Again, the test statistic is below the 10% level of significance critical value: the null hypotheses is not rejected. Examination of the cells indicate that the model does a good job of predicting average market shares across cells. Finally, controlling for product characteristics, the model does a good job of predicting average market shares across cells. However, the model tends to miss more among non-Pentium based products.

In addition, I compare the informative model predictions to those from two other baseline models: one in which there is no role for advertising and one in which advertising takes on an uninformative role. The first model is the BLP model (with micro moments) discussed in the section above. This model can be viewed as a version of the limited information model where the information technology is restricted to be one for every product. The second model is one in which consumers are assumed to know all products for sale, but advertising affects the utility function directly, this model I refer to as the uninformative model.[56]

I would prefer to be able to test the relative fit of the models parametrically. Unfortunately a formal test of non-nested hypotheses (Vuong, 1989) would require additional

[54] These tests are based on those presented in Andrews(1988). The predetermined number of cells are centered at the mean of the response variable with a width proportional to its standard deviation.

[55] The test statistic is distributed chi-squared with 7 degrees of freedom, and the realized value of 4.8 is below the 10% level of significance critical value of 12. While the model fits well, it misses more among lower market share products.

[56] The parameter estimates for the BLP baseline model and the uninformative advertising model can be found in Appendix C.

assumptions on the distribution of the error terms. While the data suggests no natural assumptions on the error distributions, there are some ways to view the results of the model to highlight the strengths and weaknesses of the fit of the informative model relative to the other models. For instance, both the informative and uninformative models predict a threshold level of average group advertising expenditures above which products will be advertised in groups and below which they will be advertised individually. Therefore we should never observe expenditures on group advertisements below this level, nor product-specific expenditures above the threshold level. The informative and uninformative models predict different threshold levels, and these predictions are presented in the second panel of Table 15. The informative model misses about 4 percent of the time, while the uninformative model misses twice as much, 8.2%. Most of the misses for both models are among Apple products (7%, informative and 8% uninformative), while both models' predictions match the data for HP and Packard Bell. In addition, both models miss more among television advertisements (3.8%, informative and 5.5%, uninformative). The fact that the uninformative model fits worse in this dimension is not surprising, since the uninformative model predicts a higher threshold level of 1.66 million, so we expect to observe a larger percentage of group expenditures below the predicted level. However, it is surprising that the informative model does no worse than the uninformative model regarding the proportion of product-specific expenditures above the predicted level. Both models miss less than 1% on average, with all the misses coming among Apple and Compaq products. This anecdotal evidence suggests at the very least that the informative model fits no worse than the uninformative model regarding advertising expenditures.

Another dimension along which the models can be compared regards the role of unobserved product attributes. Recall the mean utility is a function of observed and unobserved product attributes. In all models, the mean utility levels are chosen such that predicted market shares match observed market shares. While there is no explicit role for advertising in the BLP model, one can interpret the unobserved product heterogeneity terms ( $\xi_j$) as containing product advertising. In the model of limited information, a product with little advertising is unlikely to be in many consumer's choice sets and will have a low market share. In the BLP model, a small market share would be explained by a low value for $\xi_j$.[57]

Under the BLP approach advertising is not in the model and is captured only indirectly through the $\xi_j$, whereas with the other models advertising enters explicitly either through utility or through the information technology. Using the parameter estimates, I restricted $\xi_j$ to zero and recalculated the predicted market shares for each of the three models. These "psuedo"predicted market shares are presented in the first panel of Table 15. These give insight into the importance of unobserved product attributes in each model as well as indicating how well the model fits market shares based solely on observables and the form of the model. The results for the BLP model are presented in the last column. The BLP model's predicted psuedo shares do not come within 10% of the observed market shares for any of the top firms. This is not so surprising as the $\xi$ play a larger role in the BLP model relative to the other models. The informative model fits the market shares better than the uninformative model, for Gateway and HP the psuedo market shares are within 5% of observed shares, and for Compaq, IBM, and Packard Bell the psuedo shares are within 10%. The uninformative model comes within 5% of the observed market shares for Gateway, and within 10% for HP and Packard Bell. Neither model of advertising predicts the Apple market shares within 10%, this is perhaps not so surprising given that the firm for which advertising misses the most is Apple as discussed previously. These results suggest the informative advertising model of limited information does a good job of predicting advertising and market shares in the PC industry, relative to models in which consumers are assumed to be aware of all products.

Modeling advertising as affecting a consumer's choice set, requires significant computation time since the choice sets for each consumer must be simulated. To test if the benefits of simulating choice sets where worth the costs of increased computation time, I performed a

[57] I thank an anonymous referee for this point.

monte-carlo experiment with a simplified version of the model. Consider a market consisting of two products and one outside good. Denote the probability consumers are aware of a product by $\phi_j$. A simplified version of the market share of product 1 as presented in this study is given by

$$s_1 = \phi_1(1-\phi_2)\frac{D_1}{1+D_1} + \phi_1\phi_2\frac{D_1}{1+D_1+D_2}$$

where $D_j$ represents $\exp(\delta_j)$, the mean utility from product $j$. The market share is defined analogously for product 2. Also, define version of market share which would not require simulating choice sets

$$s_1^* = \frac{\phi_1 D_1}{1+\phi_1 D_1} + \frac{\phi_1 D_1}{1+\phi_1 D_1+\phi_2 D_2}$$

I calculated the values of $s_j$ and $s_j^*$ for different values of $\phi$ and $D$. The monte-carlo experiments indicated that the value of $s_j^*$ was within 5% of the value of $s_j$ only 2% of the time.

Notice also that the specification for $s_j^*$ is not separately identifiable from a model in which advertising enters the utility function directly (or a model in which advertising is included in $\xi_j$). This obtains by defining $\phi^* = \ln(\phi)$ and $D = \exp(\delta + \phi^*)$. These results suggests two things. The more computationally demanding model presented in this study cannot be replaced by a simplified version. Secondly, advertising which influences consumers' choice sets has very different effects from that which shifts demand directly through utility. That is the standard BLP model and models in which advertising are one of the observed product attributes are not observationally equivalent to the model presented in this research.

# 8 Conclusions

In markets characterized by rapid change, such as the PC industry, consumers are unlikely to be aware of all products for sale when making their purchase decision. In this paper, I develop and estimate a structural model of demand and supply which incorporates limited consumer information. Multi-product firms may choose to provide information to consumers about the existence of their product through informative advertising. The empirical model allows for three important sources of consumer heterogeneity: in tastes, choice sets, and advertising exposure. However, as is common in many industries, individual level purchase and advertising exposure data are not available. I show how to use an auxiliary dataset on media exposure together with aggregate advertising data to incorporate heterogeneity across consumers with respect to advertising effectiveness.

This study adds to the existing literature in that it (i) provides a structural model of informative advertising where advertising directly affects the consumers choice set, (ii) develops a simulator to deal with limited information and the large number of possible choice sets facing a consumer, (iii) incorporates firm behavior with regard to prices and advertising choices across media (allowing for corner solutions in the econometric model), (iv) shows how to use data on media exposure to incorporate heterogeneity across consumers with regard to advertising effectiveness, (v) and develops a technique to deal with the existence of group advertising. The results explain variation in behavior across firms with regard to advertising expenditures in general and medium choices in particular. I find that there are economies of scope in group advertising and some firms find it worthwhile to engage in group advertising to capitalize on the increasing returns. Estimated advertising elasticities indicate that, for some firms, advertising one product has a negative effect on other products sold by that firm, but it is less negative than for most of the rival products.

The economic importance of this study may be summarized as follows. The estimates indicate that products are priced in the elastic portion of the demand curve, consistent with oligopolistic conduct. Median markups in the PC industry are high: 19% over production costs, with the top firms earning higher than average markups and engaging in higher than average advertising. Furthermore, the results suggest firms are using advertising media to target high-income households. The high industry markups are explained in part by the fact that consumers know only some of the products for sale. Indeed estimates from traditional consumer choice models predict median markups of one-fourth this magnitude. These findings indicate that ignoring the consequences of limited information, and hence the strategic role of informative advertising, may yield misleading conclusions regarding the degree of competition in the market. Considering the effects of informative advertising are of particular importance when conducting policy analysis in industries characterized by rapid change. Assuming consumers are aware of all products generates estimates of product-specific demand curves that are biased towards being too elastic. As a result, antitrust authorities may reach different conclusions regarding the welfare implications of mergers relative to traditional models of full information[58] , if they consider limited information on the part of consumers.

[58] Inconsistent demand side estimates yield inconsistent estimates of profit changes (especially when supply side information is also limited). Under the assumptions of full information, the effects of consumer welfare (as measured as the area under the Hicksian or Marshallian demand curves) will thus be understated. See Goeree (2003) for a start on this topic.

# Tables

| Manufacturer | Percentage Unit Share | | | Percentage Dollar Share | | |
|---|---|---|---|---|---|---|
| | 1996 | 1997 | 1998 | 1996 | 1997 | 1998 |
| Acer | 6.02 | 5.89 | 4.42 | 6.20 | 6.02 | 4.37 |
| Apple | 7.39 | 5.00 | 6.97 | 6.66 | 5.79 | 9.16 |
| Compaq | 12.18 | 18.81 | 17.53 | 11.89 | 16.29 | 16.43 |
| Dell | 2.22 | 2.42 | 1.92 | 2.46 | 2.87 | 2.57 |
| Gateway | 7.80 | 11.13 | 13.36 | 8.94 | 11.77 | 16.43 |
| Hewlett-Packard | 4.07 | 5.37 | 10.16 | 4.02 | 5.52 | 10.05 |
| IBM | 8.06 | 7.01 | 7.75 | 8.49 | 7.42 | 6.85 |
| NEC | 3.13 | | | 3.22 | | |
| Packard Bell | 26.83 | | | 23.48 | | |
| Packard Bell-NEC | | 23.23 | 17.78 | | 21.02 | 16.33 |
| 15 included | 85.31 | 83.15 | 81.08 | 83.61 | 82.34 | 83.88 |
| Total Home Sales | 7,736 | 9,217 | 11,343 | $16,529 | $18,610 | $17,673 |

Notes: Others in the "15 included" are AST, ATT(NCR), DEC, Epson, Micron, and Texas Instr.
Total unit sales are in thousands, dollars in millions. In 1997 three mergers occurred:
Packard Bell, NEC, and ZDS; Acer and Texas Instr.; Gateway and Advanced Logic Research

**Table 1:** Home Market Shares of Leading Manufacturers

| Manufacturer | Advertising Expenditures | Total Market Share | Ad to Sales Ratio | Ad$ per Market Share Point |
|---|---|---|---|---|
| Apple | $181 | 8.88% | 4.90% | $20.37 |
| Compaq | $240 | 16.10% | 2.56% | $14.91 |
| Dell | $227 | 16.02% | 2.28% | $14.17 |
| Gateway | $358 | 15.07% | 5.99% | $23.75 |
| Hewlett-Packard | $466 | 9.62% | 10.28% | $48.44 |
| **IBM** | **$1,079** | **7.30%** | **19.55%** | **$147.82** |
| Total for PC market | $2,068 | | 3.34% | |

Note: Dollars are in millions. Total market share is dollar market share for all sectors (home, business, education, and government).

**Table 2:** 1998 Advertising Expenditures for Selected Manufacturers

| Variable Description | Sample | | Population | |
|---|---|---|---|---|
| | Mean | Std. Dev. | Mean | Std. Dev. |
| male | 0.663 | 0.474 | 0.661 | 0.473 |
| white | 0.881 | 0.324 | 0.881 | 0.324 |
| age (years) | 47.381 | 15.676 | 46.866 | 15.129 |
| midage (=1 if 30<age<50) | 0.443 | 0.497 | 0.449 | 0.497 |
| education (years) | 13.980 | 2.543 | 13.998 | 2.347 |
| married | 0.564 | 0.496 | 0.572 | 0.495 |
| household size | 2.633 | 1.429 | 2.631 | 1.428 |
| employed | 0.695 | 0.460 | 0.693 | 0.461 |
| income ($) | 56745.33 | 45246.23 | 56340.40 | 44464.85 |
| inclow (=1 if income<$60,000) | 0.667 | 0.471 | 0.669 | 0.471 |
| inchigh (=1 if income>$100,000) | 0.107 | 0.309 | 0.106 | 0.308 |
| own pc (=1 if own a PC) | 0.466 | 0.499 | 0.470 | 0.499 |
| pcnew (=1 if PC bought in last 12 months) | 0.113 | 0.317 | 0.112 | 0.316 |

Notes: Unless units are specified variable is a dummy. Number of observations in survey is 39,931. Sample size is 13,400.

**Table 3:** Selected Consumer Attributes

| Variable Description | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|
| cable (=1 if receive cable) | 0.749 | 0.434 | 0 | 1 |
| hours cable (per day) | 3.607 | 2.201 | 0 | 7 |
| hours non-cable (per day) | 3.003 | 2.105 | 0 | 6.2 |
| hours radio (per day) | 2.554 | 2.244 | 0 | 6.5 |
| magazine (=1 if read last quarter) | 0.954 | 0.170 | 0 | 1 |
| number magazines (read last quarter) | 6.870 | 6.141 | 0 | 95 |
| weekend newspaper (=1 if read last quarter) | 0.819 | 0.318 | 0 | 1 |
| weekday newspaper (=1 if read last quarter) | 0.574 | 0.346 | 0 | 1 |

Notes: These summary statistics are based on reports published by Simmons Market Research.

**Table 4:** Media Exposure

| | Dependent Variable: Purchased PC in Last 12 Months | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Explanatory Variable | Coefficient | | Std. Error | Coefficient | | Std. Error | Coefficient | | Std. Error |
| Constant | -1.5549 | ** | (0.1399) | -1.5133 | ** | (0.1376) | -1.4907 | ** | (0.1383) |
| age | 0.0141 | ** | (0.0058) | 0.0140 | ** | (0.0058) | 0.0132 | ** | (.0058) |
| age squared | -0.0002 | ** | (0.0001) | -0.0002 | ** | (0.0001) | -0.0002 | ** | (.00006 |
| edusp (education if <11) | -0.0585 | ** | (0.0075) | -0.0588 | ** | (0.0075) | -0.0609 | ** | (.0074) |
| eduhs (=1 if highest edu 12 years) | -0.3427 | ** | (0.0503) | -0.3441 | ** | (0.0502) | -0.3579 | ** | (.0500) |
| eduad (=1 if highest edu 1-3 college) | -0.1735 | ** | (0.0466) | -0.1715 | ** | (0.0465) | -0.1838 | ** | (.0463) |
| edubs (=1 if highest edu college grad) | -0.1028 | ** | (0.0398) | -0.1008 | ** | (0.0398) | -0.1023 | ** | (.0396) |
| married (=1 if married) | 0.1082 | ** | (0.0307) | 0.1067 | ** | (0.0306) | 0.1036 | ** | (.0304) |
| hh size (household size) | 0.0660 | ** | (0.0093) | 0.0660 | ** | (0.0093) | 0.063 | ** | (.0092) |
| inclow (=1 if income<$60,000) | -0.1436 | ** | (0.0305) | -0.1438 | ** | (0.0303) | -0.1586 | ** | (.0301) |
| inchigh (=1 if income>$100,000) | 0.1067 | ** | (0.0406) | 0.1093 | ** | (0.0405) | 0.1042 | ** | (.0403) |
| malewh (=1 if male and white) | 0.0834 | ** | (0.0283) | 0.0828 | ** | (0.0283) | 0.0927 | ** | (.0282) |
| mag 1 (=1 if magazine quintile=1) | -0.0383 | | (0.0325) | -0.0338 | | (0.0321) | | | |
| mag 2 (=1 if magazine quintile=2) | 0.0482 | | (0.0306) | 0.0497 | * | (0.0304) | | | |
| np 1 (=1 if newspaper quintile=1) | 0.0176 | | (0.0308) | | | | | | |
| np 2 (=1 if newspaper quintile=2) | -0.0059 | | (0.0334) | | | | | | |
| tv 1 (=1 if television quintile=1) | -0.1264 | ** | (0.0627) | -0.1240 | ** | (0.0626) | | | |
| tv 2 (=1 if television quintile=2) | -0.0664 | ** | (0.0314) | -0.0657 | ** | (0.0314) | | | |
| radio 1 (=1 if radio quintile=1) | 0.0856 | | (0.0549) | | | | | | |
| radio 2 (=1 if radio quintile=2) | 0.0116 | | (0.0264) | | | | | | |
| Log Likelihood | -6479 | | | -6481 | | | -6536 | | |
| Likelihood Ratio Test Statistic | | | | -4.7 | | | -114.6 | | |
| Prob>Test Statistic | | | | 0.4538 | | | 0.0000 | | |

Note: These results use the complete Simmons data set; sample size 20,100. The first specification is the unrestricted model to which I compare the other specifications. ** indicates significant at the 5% level; * significant at the 10% level.

**Table 5:** Probit Estimates of Purchase Probabilities

| Choice | Variable | Coefficient | Std. Error | Coefficient | Std. Error | Coefficient | Std. Error |
|---|---|---|---|---|---|---|---|
| Apple | price | 0.0005 ** | (0.0000) | 0.0003 ** | (0.0000) | 0.0005 ** | (0.0000) |
| | total advertising | | | 0.0135 ** | (0.0002) | | |
| | newspaper advertising | | | | | 0.0951 ** | (0.0026) |
| | magazine advertising | | | | | 0.0148 ** | (0.0018) |
| | television advertising | | | | | -0.0001 | (0.0004) |
| Compaq | price | 0.0006 ** | (0.0000) | 0.0004 ** | (0.0000) | 0.0007 ** | (0.0000) |
| | total advertising | | | 0.0134 ** | (0.0002) | | |
| | newspaper advertising | | | | | 0.0969 ** | (0.0026) |
| | magazine advertising | | | | | 0.0139 ** | (0.0018) |
| | television advertising | | | | | -0.0159 ** | (0.0005) |
| Dell | price | 0.0007 ** | (0.0000) | 0.0006 ** | (0.0000) | 0.0009 ** | (0.0000) |
| | total advertising | | | 0.0133 ** | (0.0002) | | |
| | newspaper advertising | | | | | 0.0963 ** | (0.0026) |
| | magazine advertising | | | | | 0.0141 ** | (0.0018) |
| | television advertising | | | | | -0.0122 ** | (0.0007) |
| Gateway | price | 0.0009 ** | (0.0000) | 0.0009 ** | (0.0000) | 0.0009 ** | (0.0000) |
| | total advertising | | | -0.0074 ** | (0.0007) | | |
| | newspaper advertising | | | | | 0.0564 ** | (0.0049) |
| | magazine advertising | | | | | -0.0153 | (0.0183) |
| | television advertising | | | | | -0.0088 ** | (0.0008) |
| HP | price | 0.0005 ** | (0.0000) | 0.0002 ** | (0.0000) | 0.0005 ** | (0.0000) |
| | total advertising | | | 0.0128 ** | (0.0002) | | |
| | newspaper advertising | | | | | 0.0961 ** | (0.0026) |
| | magazine advertising | | | | | 0.0116 ** | (0.0018) |
| | television advertising | | | | | -0.0029 ** | (0.0004) |
| IBM | price | 0.0003 ** | (0.0000) | 0.0000 * | (0.0000) | 0.0003 ** | (0.0000) |
| | total advertising | | | 0.0135 ** | (0.0002) | | |
| | newspaper advertising | | | | | 0.0963 ** | (0.0026) |
| | magazine advertising | | | | | 0.0146 ** | (0.0018) |
| | television advertising | | | | | 0.0012 ** | (0.0004) |
| Log Likelihood | | -79,052 | | -57,408 | | -54,343 | |

Notes: ** indicates significant at the 5% level; *significant at the 10% level. The base group is Texas Instruments and estimation is for all 15 included firms.

**Table 6:** Preliminary Estimates of Manufacturer Choice

| Variable | Means | Standard Deviation | Coefficient estimates for interactions | | | |
|---|---|---|---|---|---|---|
| **utility coefficients** | | | | | | |
| | | | **interactions with demographic variables** | | | |
| | | | household size | income > $100,000 | 30<age<50 | white male |
| cpu speed (MHz) | 9.9490 ** | 0.1390 ** | 3.9209 ** | -- | -- | -- |
| | (1.1262) | (0.0283) | (0.8774) | | | |
| pentium | 0.2499 | 0.2978 | -- | 0.0744 | -- | -- |
| | (5.8061) | (0.7234) | | (3.3086) | | |
| laptop | 3.7080 ** | 1.0712 | -- | -- | 1.3537 | 4.4313 |
| | (1.6484) | (0.8553) | | | (5.4335) | (3.0218) |
| constant | -12.2812 ** | | | | | |
| | (4.5332) | | | | | |
| ln(income-price) | 1.2260 ** | | | | | |
| | (0.1026) | | | | | |
| acer | 3.5451 | | | | | |
| | (25.4209) | | | | | |
| apple | 3.1367 ** | | | | | |
| | (1.2641) | | | | | |
| compaq | 3.8394 | | | | | |
| | (4.8206) | | | | | |
| dell | 2.4035 ** | | | | | |
| | (0.1744) | | | | | |
| gateway | 4.0235 | | | | | |
| | (8.2857) | | | | | |
| hewlett packard | 0.5294 | | | | | |
| | (3.3285) | | | | | |
| ibm | 2.2635 ** | | | | | |
| | (1.0363) | | | | | |
| micron | 1.7562 | | | | | |
| | (4.5577) | | | | | |
| packard bell | 4.0766 ** | | | | | |
| | (1.6603) | | | | | |
| **advertising coefficients** | | | | | | |
| group advertising | 0.8706 ** | | | | | |
| | (0.2265) | | | | | |
| (group advertising)$^2$ | 0.0918 ** | | | | | |
| | (0.0007) | | | | | |

| Variable | Means | **interactions with total advertising** | | |
|---|---|---|---|---|
| **information technology coefficients** | | | | See table 9 for the coefficients on interactions between advertising and demographic variables |
| constant | -1.5700 ** | acer | 0.5824 | |
| | (0.0872) | | (118.4350) | |
| high school graduate | 0.5504 ** | apple | 0.3644 * | |
| | (0.0187) | | (0.2066) | |
| income < $60,000 | 0.4818 ** | compaq | 0.6244 ** | |
| | (0.1048) | | (0.1607) | |
| income > $100,000 | 0.4512 | dell | 0.5827 * | |
| | (0.3926) | | (0.3533) | |
| | | gateway | 0.9046 ** | |
| np and mag advertising | 1.0317 | | (0.1385) | |
| | (2.1168) | hp | 0.8049 | |
| tv advertising | 1.0626 ** | | (23.7230) | |
| | (0.1418) | ibm | 0.6215 ** | |
| | | | (0.1383) | |
| (np and mag advertising)$^2$ | -0.0212 * | micron | 0.7384 | |
| | (0.0110) | | (3.7633) | |
| (tv advertising)$^2$ | -0.0346 ** | packard bell | 0.6550 ** | |
| | (0.0169) | | (0.0861) | |

Notes: ** indicates significant at the 5% level; * significant at the 10% level. Standard errors are given in parentheses.

**Table 7:** Structural Estimates (for Home Sector)

| | Apple Power Mac | Apple PowerBook* | Compaq Presario* | Compaq Presario | Dell Dimension | HP Pavilion | HP Omnibook* | IBM Aptiva | IBM Thinkpad* |
|---|---|---|---|---|---|---|---|---|---|
| Power Mac | -11.2544 | 0.0749 | 0.0151 | 0.0178 | 0.0180 | 0.0182 | 0.0206 | 0.0146 | 0.0234 |
| PowerBook* | 0.0987 | -12.7893 | 0.0122 | 0.0144 | 0.0162 | 0.0147 | 0.0173 | 0.0117 | 0.0188 |
| Presario* | 0.0136 | 0.0168 | -6.0864 | 0.0315 | 0.0266 | 0.0318 | 0.0369 | 0.0255 | 0.0409 |
| Presario | 0.0251 | 0.0301 | 0.0495 | -7.1969 | 0.0347 | 0.0298 | 0.0346 | 0.0239 | 0.0334 |
| Dimension | 0.0135 | 0.0169 | 0.0268 | 0.0319 | -8.0341 | 0.0322 | 0.0280 | 0.0259 | 0.0311 |
| Pavilion | 0.0268 | 0.0330 | 0.0261 | 0.0311 | 0.0349 | -5.5801 | 0.0361 | 0.0252 | 0.0241 |
| Omnibook* | 0.0166 | 0.0203 | 0.0328 | 0.0194 | 0.0219 | 0.0197 | -5.6129 | 0.0158 | 0.0505 |
| Aptiva | 0.0239 | 0.0296 | 0.0307 | 0.0367 | 0.0309 | 0.0278 | 0.0323 | -5.8874 | 0.0359 |
| Thinkpad* | 0.0098 | 0.0122 | 0.0198 | 0.0231 | 0.0257 | 0.0239 | 0.0338 | 0.0187 | -9.0513 |

Notes: Cell entries *i,j* where *i* ,indexes row and *j* column, give the percentage change in market share of brand *I,* with a one percentage change in the price of j. Each entry represents the average of the elasticities from 1998. A * indicates a laptop.

**Table 8:** A Sample from 1998 of Estimated Price Elasticities

| | Media | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Magazine | | Newspaper | | Television | | Radio | |
| Variable | Coefficient | Std. Error | Coefficient | Std. Error | Coefficient | Std. Error | Coefficient | Std. Error |
| constant | -1.1543 ** | (0.0385) | -1.0420 ** | (0.0378) | -0.8799 ** | (0.0392) | -1.7820 ** | (0.0412) |
| midage | -0.0425 * | (0.0234) | 0.1988 ** | (0.0242) | 0.0173 | (0.0236) | -0.0301 | (0.0241) |
| mature (age>50) | 0.0033 | (0.0239) | 0.5284 ** | (0.0243) | 0.1894 ** | (0.0238) | -0.2484 ** | (0.0242) |
| married | -0.0239 | (0.0182) | 0.1869 ** | (0.0184) | 0.0729 ** | (0.0182) | -0.0107 | (0.0187) |
| hh size | 0.0391 ** | (0.0057) | -0.0348 ** | (0.0060) | 0.0173 ** | (0.0058) | 0.0119 * | (0.0061) |
| inclow | -0.1958 ** | (0.0195) | -0.2520 ** | (0.0197) | 0.1118 ** | (0.0201) | -0.1197 ** | (0.0205) |
| inchigh | 0.1565 ** | (0.0294) | 0.1355 ** | (0.0283) | -0.0196 | (0.0305) | 0.0736 ** | (0.0304) |
| malewh | -0.0791 ** | (0.0170) | 0.0066 | (0.0171) | -0.0193 | (0.0169) | 0.0072 | (0.0173) |
| eduhs | -0.1219 ** | (0.0254) | -0.3354 ** | (0.0251) | 0.2682 ** | (0.0256) | 0.0839 ** | (0.0260) |
| eduad | 0.0185 | (0.0265) | -0.1538 ** | (0.0261) | 0.2541 ** | (0.0269) | 0.1313 ** | (0.0273) |
| edubs | -0.0276 | (0.0260) | -0.0504 ** | (0.0253) | 0.1317 ** | (0.0265) | 0.1020 ** | (0.0269) |
| edusp | -0.0302 ** | (0.0034) | -0.0660 ** | (0.0035) | 0.0315 ** | (0.0033) | -0.0134 ** | (0.0034) |
| Log Likelihood | | -31983 | | -31087 | | -31052 | | -28597 |

Notes: Estimates include time dummies. ** indicates significant at the 5% level; * significant at the 10% level.

**Table 9:** Likelihood Estimates of Media Exposure

| | Apple Macintosh | Apple PowerBook* | Compaq Presario* | Compaq Presario | Dell Dimension | Dell Latitude* | HP Omnibook* | IBM Thinkpad* |
|---|---|---|---|---|---|---|---|---|
| Macintosh | 0.0212 | -0.0102 | -0.0216 | -0.0202 | -0.0206 | -0.0262 | -0.0206 | -0.0141 |
| PowerBook* | -0.0097 | 0.0076 | -0.0092 | -0.0093 | -0.0098 | -0.0098 | -0.0094 | -0.0069 |
| Presario* | -0.1011 | -0.0867 | 0.0935 | -0.0944 | -0.0925 | -0.0938 | -0.0954 | -0.0571 |
| Presario | -0.1050 | -0.1002 | -0.0130 | 0.1025 | -0.1006 | -0.1028 | -0.1028 | -0.0790 |
| Dimension | -0.0297 | -0.0368 | -0.0382 | -0.0334 | 0.0330 | -0.0304 | -0.0338 | -0.0214 |
| Latitude* | -0.0081 | -0.0096 | -0.0093 | -0.0106 | -0.0094 | 0.0095 | -0.0091 | -0.0074 |
| OmniBook* | -0.0046 | -0.0037 | -0.0045 | -0.0043 | -0.0049 | -0.0041 | 0.0047 | -0.0038 |
| Thinkpad* | -0.0157 | -0.0153 | -0.0129 | -0.0183 | -0.0151 | -0.0153 | -0.0083 | 0.0108 |

Notes: Cell entries *i,j* where *i* ,indexes row and *j* column, give the percentage change in market share of *i* . with a $1000 increase in the advertising of j. * indicates the PC is a laptop.

**Table 10:** A Sample from 1997 of Estimated Advertising Semi-Elasticities

| Variable | Coefficient | Standard Error |
|---|---|---|
| **ln marginal cost of production** | | |
| constant | 10.0230 ** | (0.1784) |
| ln(cpu speed) | 0.4932 ** | (0.0062) |
| pentium dummy | -0.3978 ** | (0.0926) |
| laptop dummy | 1.3054 ** | (0.2691) |
| quarterly trend | -0.1320 ** | (0.0214) |
| **ln marginal cost of advertising** | | |
| constant | 7.0356 ** | (1.9230) |
| ln(price of advertising) | 1.0004 ** | (0.0002) |
| **non-home sector marginal revenue** | | |
| constant | 2.4053 * | (1.2869) |
| non-home sector price | 1.0350 ** | (0.0749) |
| cpu speed | 0.0198 ** | (0.0028) |
| non-pc sales | 5.2707 * | (2.8202) |

Notes: ** indicates significant at the 5% level; * significant at the 10% level.

**Table 11:** Structural Cost and Non-Home Sector Estimates

| | Median Price | Ad to Sales Ratio | Median Percentage Markup over Marginal Costs | including ad costs |
|---|---|---|---|---|
| Total Industry | $2,239 | 3.34% | 19% | 10% |
| Top 6 firm | $2,172 | 8.66% | 22% | 12% |
| Apple | $1,859 | 4.90% | 19% | 10% |
| Compaq | $2,070 | 2.56% | 24% | 16% |
| Gateway | $2,767 | 5.99% | 12% | 9% |
| Hewlett Packard | $2,203 | 10.28% | 17% | 11% |
| IBM | $2,565 | 19.55% | 17% | 10% |
| Packard Bell | $2,075 | 19.55% | 18% | 12% |

Note: Ad to sales ratios are from LNA and include ad and sales across all sectors. Percentage markups are the median (price-marginal costs)/price across all products. The last column is percentage total markups per unit after including advertising. These are determined from estimated markups and estimated effective product advertising in the home sector.

**Table 12:** Summary Statistics for Prices, Advertising, and Markups

| | Median Percentage Markup | | |
|---|---|---|---|
| | Under Partial Information | Under Full Information | Change in Markups |
| Total industry | 19% | 5% | -73% |
| Apple | | 2.5% | -86% |
| iMac | 22.2% | 3.1% | |
| Power Mac | 13.7% | 2.0% | |
| PowerBook Duo* | 15.9% | 2.0% | |
| Compaq | | 7.0% | -70% |
| Armada* | 42.3% | 3.5% | |
| Presario | 18.1% | 2.6% | |
| Presario* | 15.3% | 2.0% | |
| ProLinea | 23.4% | 7.0% | |
| Gateway | | 1.7% | -85% |
| Gateway Desk Series | 12.9% | 1.9% | |
| Gateway Portable Series | 8.1% | 1.5% | |
| HP | | 4.5% | -73% |
| OmniBook* | 15.7% | 8.4% | |
| Pavilion | 21.8% | 3.1% | |
| Vectra | 15.3% | 6.8% | |
| IBM | | 2.0% | -88% |
| Aptiva | 16.0% | 2.3% | |
| Thinkpad* | 11.9% | 2.0% | |
| IBM PC | 23.2% | 2.3% | |
| Packard Bell | | 3.0% | -83% |
| NEC Versa* | 11.1% | 1.6% | |
| NEC Desk Series | 17.7% | 2.5% | |

Notes: Percentage markups are defined as (price-marginal cost)/price. Full information is the traditional model in which consumers know all products; under partial information the choice set is estimated. * indicates that computers are laptops.

**Table 13:** Estimated Percentage Markups

| | Apple | Compaq | Gateway | HP | IBM | Packard Bell |
|---|---|---|---|---|---|---|
| **under partial information** | | | | | | |
| Apple | -10.1720 | 0.0576 | 0.0199 | 0.0249 | 0.0178 | 0.0241 |
| Compaq | 0.0404 | -6.6810 | 0.0416 | 0.0516 | 0.0370 | 0.0499 |
| Gateway | 0.0396 | 0.1184 | -10.1177 | 0.0511 | 0.0366 | 0.0494 |
| Hewlett-Packard | 0.0434 | 0.1288 | 0.0447 | -6.2509 | 0.0399 | 0.0538 |
| IBM | 0.0296 | 0.0870 | 0.0304 | 0.0377 | -7.5830 | 0.0364 |
| Packard Bell | 0.0312 | 0.2808 | 0.0543 | 0.0676 | 0.0484 | -7.4831 |
| **under full information (blp benchmark)** | | | | | | |
| Apple | -29.2337 | 0.1901 | 0.0448 | 0.1036 | 0.0726 | 0.0735 |
| Compaq | 0.0188 | -34.1364 | 0.0237 | 0.0548 | 0.0384 | 0.0389 |
| Gateway | 0.0378 | 0.2024 | -34.6917 | 0.1103 | 0.0773 | 0.0782 |
| Hewlett-Packard | 0.0032 | 0.0171 | 0.0040 | -37.9312 | 0.0066 | 0.0067 |
| IBM | 0.1014 | 0.5431 | 0.1277 | 0.2963 | -32.5824 | 0.2102 |
| Packard Bell | 0.0707 | 0.3793 | 0.0895 | 0.2066 | 0.1446 | -35.8697 |

Notes: Cell entries *i,j* where *i* ,indexes row and *j* column, give the percentage change in market share of brand *I,* with a one percentage change in the price of j. Each entry represents median elasticities over all products for 1998.

**Table 14:** Estimated Firm Price Elasticities (1998)

| Response Variable | Observed | Prediction for different models of advertising: Informative | | Uninformative | | No role |
|---|---|---|---|---|---|---|
| **average annual percent market shares** | | | | | | |
| Apple | 6.45% | 8.68% | | 8.96% | | 5.15% |
| Compaq | 16.17% | 17.48% | * | 17.98% | | 19.74% |
| Gateway | 10.76% | 11.32% | ** | 10.99% | ** | 13.07% |
| HP | 6.53% | 6.31% | ** | 5.99% | * | 1.98% |
| IBM | 7.60% | 8.40% | * | 8.59% | | 9.38% |
| Packard Bell | 22.61% | 20.60% | * | 24.34% | * | 27.41% |
| **group and product-specific advertising** | | | | | | |
| Predicted threshold value (in millions) | | 1.41 | | 1.66 | | not applicable |
| **percent group expenditures below predicted threshold value** | | | | | | |
| All products | | 4.2% | | 8.2% | | |
| Apple | | 6.9% | | 8.2% | | |
| Compaq | | 4.4% | | 4.4% | | |
| Gateway | | 1.3% | | 2.6% | | |
| HP | | 0.0% | | 0.0% | | |
| IBM | | 1.3% | | 3.8% | | |
| Packard Bell | | 0.0% | | 0.0% | | |
| Newspaper | | 0.2% | | 0.8% | | |
| Magazine | | 0.2% | | 0.9% | | |
| Television | | 3.8% | | 5.5% | | |
| Radio | | 0.9% | | 0.9% | | |
| **percent product-specific expenditures above predicted threshold value** | | | | | | |
| All products | | 0.8% | | 0.8% | | |
| Apple | | 0.9% | | 0.9% | | |
| Compaq | | 0.9% | | 0.9% | | |
| Gateway | | 0.0% | | 0.0% | | |
| HP | | 0.0% | | 0.0% | | |
| IBM | | 0.0% | | 0.0% | | |
| Packard Bell | | 0.0% | | 0.0% | | |
| Newspaper | | 0.0% | | 0.0% | | |
| Magazine | | 0.8% | | 0.8% | | |
| Television | | 0.1% | | 0.1% | | |
| Radio | | 0.8% | | 0.8% | | |

Notes: Predicted market shares are evaluated at parameter estimates with unobserved product attributes restricted to zero.
** indicates that predicted values within 5% of the observed value * within 10% of the true value
level. Predicted group advertising expenditures threshold value is in millions. Advertising expenditures are computed using equation (1) evaluated at the optimal parameter values. Firm percentages are calculated as percent of product/medium advertising by that firm.

**Table 15:** Goodness of Fit Comparisons

# Appendices

## A Restricting Nonlinear Parameter Search

The non-linear search over the parameters $\{\theta,\ \beta,\ \eta, \eta_{\text{AD}}\}$ for the objective function given by

$$\Lambda' Z A^{-1} Z' \Lambda$$

can be restricted to a search over the parameters $\Omega = \{\theta,\ \eta_{\text{AD}}\}$ because $\beta,\ \eta$ are uniquely determined by the choice of $\theta$. This can be seen by rewriting the first-order conditions that are functions of $\beta$, and $\eta$ as follows

$$\begin{bmatrix} \delta(\theta) - x'\beta \\ \\ \ln(p - \Delta(\delta,\theta)^{-1}s) - w'\eta \end{bmatrix} =$$

$$\begin{bmatrix} \delta(\theta) \\ \\ \ln(p - \Delta(\delta,\theta)^{-1}s) \end{bmatrix} - \begin{bmatrix} x & 0 \\ \\ 0 & w \end{bmatrix} \begin{bmatrix} \beta \\ \\ \eta \end{bmatrix} \equiv Y - Q \begin{bmatrix} \beta \\ \\ \eta \end{bmatrix}$$

The values of $\beta$ and $\eta$ that minimize the objective function are

$$\left[\ \beta \quad \eta\ \right]' = [Q'Z^{*}A^{*-1}Z^{*\prime}Q]^{-1}Q'Z^{*}A^{*-1}Z^{*\prime}Y(\theta)$$

which are determined by the choice of $\theta$, where $Z^{*}A^{*-1}Z^{*\prime}$ are the portions of the instrument and weighting matrices corresponding to the moments arising from the demand and pricing first-order conditions.

## B Simulation Details

A general outline for simulation follows, I omit the time subscript for clarity. First prepare random draws, which, once drawn, do not change throughout estimation.

1. In the case of the macro moments,

   (a) Draw $i = 1, \ldots, ns$ consumers from the joint distribution of characteristics and income given by the CPS, $F(D, y)$, and corresponding draws from multivariate normal distribution of unobservable consumer characteristics, $F(\nu)$, one for each product characteristic, call these $\nu_{ik}$ (where I drew a sample of 3000 for each year, $ns = 9000$)

   (b) Draw log normal variables one for each medium combination, call these $\kappa_{im}$.(where $m = 4$)

(c) Draw uniform random variables one for each product-individual pair, call these $u_{ij}$.

2. For the micro moments

   (a) For each Simmons consumer $i = 1, \ldots, ncons$ draw $R$ times from multivariate normal distribution of unobservable consumer characteristics, $F(\nu)$, one for each product characteristic, call these $\nu_{ikr}$. (where $ncons = 13400$)

   (b) Draw $R$ uniform random variables for each product-individual combination, call these $u_{ijr}$.

   (c) Draw $R$ log normal variables one for each medium-individual combination, call these $\kappa_{imr}$.

3. Choose an initial value of the parameters $\theta_0$

4. For the macro-moments, do for $i = 1, \ldots, ns$

   (a) Calculate $\phi_{ij}(\theta)$ for each product $j = 1, ..., J$ for each period

$$\phi_{ij}(\theta) = \frac{\exp(\tau_{ij})}{1 + \exp(\tau_{ij})}$$

$$\tau_{ij} = \sum_d \widetilde{D_{id}}' \lambda_d + \sum_m \varphi_m a_{jm} + \sum_m \rho_m a_{jm}^2 + +\Psi_f \sum_m a_{jm} + \sum_m \sum_d \Upsilon_{md} D_{id}^s a_{jm} + \sum_m a_{jm} \kappa_{im}$$

   (b) Given $\phi_{ij}(\theta)$ and $u_{ij}$ construct a $J$ dimensional Bernoulli vector, $b_i(\theta)$. This defines the choice set $\mathcal{S}'$, where the $j$th element is determined according to

$$b_{ij} = \begin{cases} 1 & \text{if } \phi_{ij}(\theta) > u_{ij} \\ 0 & \text{if } \phi_{ij}(\theta) \leq u_{ij} \end{cases}$$

   Define $b_i^0$ to be the Bernoulli vector generated from the initial choice of parameters, $\theta_0$.

   (c) Calculate

$$P_{ij}(\theta) = \frac{\exp\{\delta_j + \mu_{ij}\}}{y_i{}^{\alpha} + \Sigma_{k:b_{i,k}^0=1} \exp\{\delta_k + \mu_{ik}\}}$$

   where $\mu_{ij}$ is value of $\alpha \ln(y_i - p_j) + \sum_k x_{jk}(\sigma_k \nu_{ik} + \sum_d \Pi_{kd} D_{id})$ given the $i$th draw and $\theta$.

(d) Calculate

$$s_{ij}(\theta) = \Pi_{l \in \mathcal{S}} \phi_{il} \Pi_{k \notin \mathcal{S}} (1 - \phi_{ik}) \frac{P_{ij}(\theta)}{\phi_i^0(\theta_0)}$$

where $\phi_i^0(\theta_0)$ is the value of $\Pi_{l \in \mathcal{S}_0} \phi_{il} \Pi_{k \notin \mathcal{S}_0} (1 - \phi_{ik})$ using the initial value of the parameters and the initial choice set. During estimation the parameter values will be updated so the simulated product over the $\phi_{ij}$ will differ from the initial $\phi_i^0(\theta_0)$ in all but the first simulation.

5. Calculate the simulator for the market share

$$\widehat{s_j} = \frac{1}{ns} \sum_i s_{ij}$$

6. For the micro-moments: For each consumer, $i = 1, \ldots, ncons$, calculate $\overline{\tau_{ij}}$

$$\overline{\tau_{ij}} = \sum_d \widetilde{D_{id}^s}' \lambda_d + \sum_m \varphi_m a_{jm} + \sum_m \rho_m a_{jm}^2 + \Psi_f \sum_m a_{jm} + \sum_m \sum_d \Upsilon_{md} D_{id}^s a_{jm}$$

do for $r = 1, \ldots, R$ draws

(a) Calculate $\phi_{ijr}(\theta)$

$$\phi_{ijr}(\theta) = \frac{\exp(\tau_{ijr})}{1 + \exp(\tau_{ijr})}$$

$$\tau_{ijr} = \overline{\tau_{ij}} + \sum_m a'_{jm} \kappa_{imr}$$

(b) Given $\phi_{ijr}(\theta)$ and $u_{ijr}$ construct a $J$ dimensional Bernoulli vector, $b_{ir}(\theta)$. This defines the choice set $\mathcal{S}_r$ for the $r$th loop, where the $j$th element is determined according to

$$b_{ijr} = \begin{cases} 1 & \text{if } \phi_{ijr}(\theta) > u_{ijr} \\ 0 & \text{if } \phi_{ijr}(\theta) \leq u_{ijr} \end{cases}$$

Define $b_{ir}^0$ to be the Bernoulli vector generated from the initial choice of parameters, $\theta_0$.

(c) Calculate

$$P_{ijr}(\theta) = \frac{\exp\{\delta_j + \mu_{ijr}\}}{y_i^{\alpha} + \Sigma_{k:b_{ir,k}^0=1} \exp\{\delta_k + \mu_{ikr}\}}$$

where $\mu_{ijr}$ is value of $\alpha \ln(y_i - p_j) + \sum_k x_{jk}(\sigma_k \nu_{ikr} + \sum_d \Pi_{kd} D_{id}^s)$ given the $r$th draw and $\theta$.

(d) Calculate

$$s_{ijr}(\theta) = \Pi_{l\in\mathcal{S}_r}\phi_{il}\Pi_{k\notin\mathcal{S}_r}\left(1-\phi_{ik}\right)\frac{P_{ijr}(\theta)}{\phi_{ir}^0\left(\theta_0\right)}$$

where $\phi_{ir}^0(\theta_0)$ is the value of $\Pi_{l\in\mathcal{S}_r}\phi_{il}\Pi_{k\notin\mathcal{S}_r}\left(1-\phi_{ik}\right)$ using the initial choice set evaluated at the initial value of the parameters, $b_{ir}^0$.

7. Calculate the simulator for the choice probability

$$\widehat{s_{ij}} = \frac{1}{R}\sum_r s_{ijr}$$

The firm choice probability (used in the micro moments) is

$$\widehat{B}_{if} = \sum_{j\in\mathcal{J}_f}\widehat{s_{ij}}$$

# C Full Information Parameter Estimates

| Variable | | Standard Deviation | Interactions with demographic variables: household size | income > $100,000 | 30<age<50 | white male |
|---|---|---|---|---|---|---|
| ln(income-price) | 1.1980 ** | | | | | |
| | (0.5130) | | | | | |
| constant | -32.4815 ** | | | | | |
| | (13.5997) | | | | | |
| cpu speed (MHz) | 12.1745 ** | 0.2878 ** | 0.6967 ** | -- | -- | -- |
| | (2.2525) | (0.0566) | (0.2925) | | | |
| pentium | 2.2631 | 0.7168 ** | -- | 0.7495 * | -- | -- |
| | (2.9031) | (0.3617) | | (0.3893) | | |
| laptop | 3.0241 * | 0.3158 ** | -- | -- | -0.2052 | 0.3913 * |
| | (0.8242) | (0.1425) | | | (0.5434) | (0.2015) |
| acer | 2.2559 | | | | | |
| | (12.7105) | | | | | |
| apple | 7.3454 ** | | | | | |
| | (0.6321) | | | | | |
| compaq | 8.7814 ** | | | | | |
| | (3.2137) | | | | | |
| dell | 1.2345 | | | | | |
| | (0.6980) | | | | | |
| gateway | 9.9450 * | | | | | |
| | (5.1786) | | | | | |
| hewlett packard | 4.5117 * | | | | | |
| | (2.3775) | | | | | |
| ibm | 6.1112 ** | | | | | |
| | (0.6909) | | | | | |
| micron | 1.1279 | | | | | |
| | (2.2789) | | | | | |
| packard bell | 6.6300 * | | | | | |
| | (3.3207) | | | | | |

Notes: ** indicates significant at the 5% level; * significant at the 10% level. Standard errors are given in parentheses.

**Appendix C Table 1:** BLP Model Demand Estimates

| Variable | | Interactions with demographic variables: Standard Deviation | household size | income < $60,-000 | income > $100,000 | 30<age<50 | high school graduate | white male |
|---|---|---|---|---|---|---|---|---|
| ln(income-price) | 1.3962 ** | | | | | | | |
| | (0.6839) | | | | | | | |
| constant | -16.3836 ** | | | | | | | |
| | (6.7999) | | | | | | | |
| cpu speed (MHz) | 18.5052 ** | 0.5352 ** | 0.9336 ** | -- | -- | -- | -- | -- |
| | (4.5050) | (0.2262) | (0.4387) | | | | | |
| pentium | 4.3071 | 0.0649 ** | -- | -- | -1.9431 | -- | -- | -- |
| | (8.7092) | (0.0289) | | | (1.6543) | | | |
| laptop | -1.8485 * | 0.1562 ** | -- | -- | -- | -2.8122 | -- | 1.5265 |
| | (0.9696) | (0.0778) | | | | (2.7168) | | (1.5109) |
| magazine | 7.5328 ** | | | | | | | |
| | (3.1603) | | | | | | | |
| newspaper | -0.0726 | | | | | | | |
| | (0.4387) | | | | | | | |
| radio | -5.3824 | | | | | | | |
| | (2.8625) | | | | | | | |
| television | 2.6127 * | 0.0880 | 0.0382 | 0.0152 ** | 0.0021 | -0.0177 * | 0.0290 | -0.0724 |
| | (1.5094) | (0.0792) | (0.1580) | (0.0074) | (0.0055) | (0.0094) | (0.1509) | (0.0439) |
| magazine and newspaper | | 0.6122 * | -0.6658 * | -0.1630 ** | -0.0248 | 0.7535 | 0.2328 | -0.8555 |
| | | (0.3167) | (0.3187) | (0.3178) | (0.0125) | (0.6232) | (0.8034) | (0.7299) |
| **firm fixed effects** | | | | | | | | |
| acer | 2.6190 | | | | | | | |
| | (12.7105) | | | | | | | |
| apple | 7.1964 ** | | | | | | | |
| | (3.1603) | | | | | | | |
| compaq | 3.9684 | | | | | | | |
| | (2.4103) | | | | | | | |
| dell | -3.5496 | | | | | | | |
| | (2.6175) | | | | | | | |
| gateway | 4.0329 | | | | | | | |
| | (4.1429) | | | | | | | |
| hewlett packard | -5.6777 * | | | | | | | |
| | (2.9198) | | | | | | | |
| ibm | 3.8068 ** | | | | | | | |
| | (1.5545) | | | | | | | |
| micron | 6.1322 | | | | | | | |
| | (5.4693) | | | | | | | |
| packard bell | -2.8169 * | | | | | | | |
| | (1.5094) | | | | | | | |
| **advertising coefficients** | | | | | | | | |
| group advertising | (0.9456) ** | | | | | | | |
| | 0.4530 | | | | | | | |
| (group advertising)$^2$ | (0.0328) ** | | | | | | | |
| | 0.0160 | | | | | | | |

Notes: ** indicates significant at the 5% level; * significant at the 10% level. Standard errors are given in parentheses.

**Appendix C Table 2:** Uninformative Model Demand Estimates

| Variable | BLP model Coefficient | | BLP model Standard Error | Uninformative Model Coefficient | | Uninformative Model Standard Error |
|---|---|---|---|---|---|---|
| **ln marginal cost of production** | | | | | | |
| constant | 12.6836 | ** | (0.3503) | 7.5037 | ** | (0.7005) |
| ln(cpu speed) | 1.2788 | * | (0.6788) | 0.2486 | ** | (0.0185) |
| pentium | -0.8888 | ** | (0.1854) | -0.4403 | ** | (0.2039) |
| laptop | 0.5078 | ** | (0.1347) | 1.1417 | ** | (0.5387) |
| quarterly trend | -0.1009 | ** | (0.0432) | -0.1874 | ** | (0.0886) |
| **ln marginal cost of advertising** | | | | | | |
| constant | | | | 4.6904 | ** | (2.3076) |
| price of advertising | | | | 1.0000 | ** | (0.0197) |
| **non-home sector marginal revenue** | | | | | | |
| constant | | | | 1.2943 | | (1.1699) |
| non-home sector price | | | | 1.0252 | ** | (0.1648) |
| cpu speed | | | | 0.0169 | ** | (0.0083) |
| non-pc sales | | | | 5.1320 | * | (2.6860) |

Notes: ** indicates significant at the 5% level; * significant at the 10% level.

**Appendix C Table 3:** Supply Side Estimates

# References


Ackerberg, Daniel (2003) "Advertising, Learning, and Consumer Choice in Experience Goods Markets: A Structural Empirical Examination," *International Economic Review* 44:1007-1040

Anand, Bharat and Ron Shachar (2001) "Advertising, the Matchmaker," Harvard Business School Working Paper No. 02-057.

Anderson, Simon, A. de Palma, and J.F. Thisse (1989) "Demand for Differentiated Products, Discrete Choice Models, and the Characteristics Approach," *Review of Economic Studies* 56: 21-35.

Anderson, Simon, A. de Palma, and J.F. Thisse (1992) *Discrete Choice Theory of Product Differentiation*, Cambridge: MIT Press.

Andrews, Donald (1988) "Chi-square Diagnostic Tests for Econometric Models," *Journal of Econometrics* 37: 135-156.

Becker, Gary and Kevin Murphy (1993) "A Simple Theory of Advertising as a Good or Bad," *Quarterly Journal of Economics* 942-964.

Berkovec, James and Steven Stern (1991) "Job Exit Behavior of Older Men," *Econometrica* 59(1): 189-210

Berry, Steven (1994) "Estimating Discrete Choice Models of Product Differentiation," *Rand Journal of Economics* 25(2): 242-262.

Berry, Steven, James Levinsohn and Ariel Pakes (1995) "Automobile Prices in Market Equilibrium," *Econometrica* 63(4): 841-890.

Berry, Steven, James Levinsohn and Ariel Pakes (2004) "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy* 112(1,1):68-105.

Berry, Steven, James Levinsohn and Ariel Pakes (1999) "Voluntary Export Restraints on Automobiles: Evaluating a Trade Policy," *American Economic Review* 89(3): 400-430.

Butters, G. (1977) "Equilibrium Distribution of Prices and Advertising," *Review of Economic Studies* 44: 465-492.

Chamberlain, G. (1987) "Asymptotic Efficiency in Estimation with Conditional Moment Restrictions," *Journal of Econometrics* 34: 305-344.

Chiang, Jeongwen, Siddhartha Chib and Chakravarthi Narasimhan (1999) "Markov Chain Monte Carlo and Models of Consideration Set and Parameter Heterogeneity," *Journal of Econometrics* 89: 223-248.

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

Geweke, John (1988) "Antithetic Acceleration of Monte Carlo Integration in Bayesian Inference." *Journal of Econometrics* 38: 73-89.

Goeree, Michelle S. (2002) "Informative Advertising and the US Personal Computer Market: A Structural Empirical Examination," Ph.D. Dissertation, University of Virginia.

Goeree, Michelle S. (2003) "Was Mr. Hewlett Right? Mergers, Advertising, and the PC Industry," mimeo, University of Amsterdam.

Gourieroux, Christian, *et al.* (1987) "Generalized Residuals," *Journal of Econometrics*, 34: 5-32.

Grossman, G. and Carl Shapiro (1984) "Informative Advertising with Differentiated Products," *Review of Economic Studies* 51: 63-82.

Hendel, I. (1999) "Estimating Multiple-Discrete Choice Models: An Application to Computerization Returns," *Review of Economic Studies* 66(2): 423-46

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

Nevo, Aviv (2000) “A Practitioner’s Guide to Estimation of Random Coefficients Logit Models of Demand,” *Journal of Economics and Management Strategy*

Pakes, Ariel and David Pollard (1989) “Simulation and the Asymptotics of Optimization Estimators,” *Econometrica* 57(5): 1027-1057.

Petrin, Amil (2002) “Quantifying the Benefits of New Products: The Case of the Minivan,” *Journal of Political Economy* 110(4):705-729.

Shum, Matthew “Does Advertising Overcome Brand Loyalty? Evidence from the Breakfast Cereals Market,” forthcoming, *Journal of Economics and Management Strategy*

Stern, Steven (1997) “Simulation-Based Estimation,” *Journal of Economic Literature* 35: 2006-2039.

Stern, Steven (2000) “Simulation Based Inference in Econometrics: Motivation and Methods,” in *Simulation-Based Inference in Econometrics: Methods and Applications*, ed., Mariano, Weeks, and Schuermann. Cambridge: University Press.

Vuong, Quang H. (1989) ”Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses,” *Econometrica* 57: 307-333.