TECHNICAL WORKING PAPER SERIES

ESTIMATING MULTIPLE-DISCRETE CHOICE MODELS: AN APPLICATION TO COMPUTERIZATION RETURNS

Igal Hendel

Technical Working Paper No. 168

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 1994

Most of this work was done while I was at Yale University. I would like to thank their hospitality, in particular I am grateful to Ariel Pakes, Moshe Buchinsky and Steve Berry for their invaluable advice, comments and useful conversations. Thanks to Zvi Griliches for his support and advice. I am grateful to Manuel Trajtenberg who provided the motivation for studying the topic and provided access to the data. Thanks to Eli Berman, Gary Chamberlain and John Sutton for useful comments and to Ariel Hendel of Standard Microsystems Corporation for technical advice. Financial support from the Sloan Foundation Grant to the NBER and the Sloan Doctoral Dissertation Fellowship are gratefully acknowledged. The errors are all mine. This paper is part of NBER's research program in Productivity. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Technical Working Paper #168
October 1994

ESTIMATING MULTIPLE-DISCRETE CHOICE MODELS: AN APPLICATION TO COMPUTERIZATION RETURNS

ABSTRACT

This paper develops a multiple-discrete choice model for the analysis of demand of differentiated products. Users maximize profits by choosing the number of units of each brand they purchase. Multiple-unit as well as multiple-brand purchases are allowed. These two features distinguish this model from classical discrete choice models which consider only a single choice among mutually exclusive alternatives. Model parameters are estimated using the simulated method of moments technique. Both requirements - microfoundations and estimability - are imposed in order to exploit the available micro level data on personal computer purchases. The estimated demand structure is used to assess welfare gains from computerization and technological innovation in peripherals industries. The estimated return on investment in computers is 90%. Moreover, a 10% increase in the performance to price ratio of microprocessors leads to a 4% gain in the estimated end user surplus.

Igal Hendel
Department of Economics
Fisher Hall
Princeton University
Princeton, NJ 08544
and NBER

SECTION 1 - INTRODUCTION

The personal computer (PC) market is of high interest to industrial organization economists because of its size, continual technological change, strategic interactions among players and above all because of the far reaching implications of PCs on the economy as a whole. PC demand presents distinguishing features difficult to capture. PCs are a differentiated durable good, in constant evolution and whose demand presents multiple-discrete choices. That is, users can hold multiple-units as well as multiple-brands at the same time. In this paper I present: first, a model that captures those distinguishing features and second, a method of estimating a demand structure consistent with the microfundations of the model. This method is developed to exploit micro-data on PC purchases by firms in order to assess the welfare implications of the computer revolution.

The main contributions of the paper are two: First, to develop a framework that generalizes discrete choice models. Classical discrete choice models enable one choice among mutually exclusive alternatives. The proposed framework enables users to demand multiple-units as well as multiple-brands; that is, it adds the quantity dimension and makes alternatives non-mutually exclusive. This model can be applied to other economically interesting cases of multiple-discrete choices, e.g., the composition of airline fleets, car holdings per household, an individual's clothing choices or the field of the faculty in an economics department. Although pervasive, multiple-discrete problems have received little attention in the literature due to their complexity. Second, the proposed framework together with new econometric machinery - simulated method of moments, Pakes and Pollard (1989) and McFadden (1989) - and a new micro-data set on PC purchases enables me to estimate the demand structure at the micro-level. As explained later in the paper micro-data and microfundations are essential for reliable answers.

There are many questions surrounding the PC market. The central one is the current controversy in the economic literature about the "computerization puzzle." This puzzle stems from the empirical finding of no effect of computerization in productivity and profitability (e.g., Griliches and Siegel (1991) and Morrison and Berndt (1990)). The basic ingredient in addressing this question is the demand structure, i.e., revealing preferences. Furthermore, knowing preferences enables us to answer questions about the gain from technological improvement of peripherals, e.g., in the microprocessors, how do we reflect quality changes in price indices or how do PC suppliers locate and price in the attribute space. My aim is to present a demand estimation approach that enables me to address those questions as precisely as possible.

In a previous work, Hendel (1992), I studied the computerization puzzle by estimating a production function for the manufacturing industry at the 4 digit SIC level. Contrary to previous findings, which were mostly performed at higher levels of data aggregation, I showed that a computing capability index, when carefully calculated, enters significantly into the production function.[1] Pointing on that more reliable and disaggregated data solves the puzzle. In this paper, I take an additional step exploiting even more dissagregated data in order to get a more accurate assessment of the contribution of PCs to welfare. In order to do this, I will present a framework that stems from profit maximization behavior and that is estimable at the micro level. Furthermore, this framework enables not only to perform welfare assessments, but also to explain demand patterns and the relation between productivity and attributes, that other approaches would not.

The approach taken is to build a model that resembles the distinguishing features of PC demand. The behavioral model predicts PC purchases for each firm, given its characteristics. Predicted behavior is fitted to actual purchases. I exploit the moment conditions defined by the difference between observed and predicted behavior to obtain estimates using the generalized method of moments. Given the computational complexity of the model, particularly dictated by the multiple-discretness, I simulate, rather than integrate out, expected behavior. That is, I use the simulated method of moments.

Section 2 describes the model, which assumes that each establishment (i.e., sampled firms in this data set) has a given number of potential tasks that can be performed by PCs. The number of such tasks relates, naturally, to the establishment's characteristics such as size and sector. Through this novel concept of task is that the model captures the first distinguishing feature of the demand in question, i.e., the non-mutually exclusiveness nature of choices. Moreover, I assume that firms' profit functions are concave in the number of PCs used. This assumption implies that any number of units may be purchased. This gives rise to the second feature of the model, i.e., the quantity dimension. These two features necessitate the particular approach taken here to address the PC market, characterized by multiple-discrete choices.

The characteristics approach, Lancaster (1979), in which each brand/model is identified by a bundle of attributes, is particularly useful for the following reasons. First, it allows me to precisely define each good, and the ways in which they differ from each other. Second, it makes it possible to identify the sources of welfare gains from product innovations by estimating the valuation of each attribute as a function of the characteristics of the users. Finally, it simplifies the estimation by reducing the dimensionality of the parameter vector of interest, because differences (as well as elasticities of substitution) between goods can be completely explained in terms of the various attributes.

The model considered here belongs to the class of random utility/discrete choice models,[2] which is extensively used in the current industrial organization literature. It represents consumer preferences over products as a function of individual consumer characteristics and of the attributes of the products. Quality is assumed to be subjectively perceived by users; that is, users have their own perceptions about brand quality. Those perceptions, being subjective, may differ among users. Clearly, perceptions are known by the user but not by us. This creates a potential source of bias in demand estimation, as quality and prices are normally thought to be positively correlated. Therefore, treating the quality dimension properly is of double interest: (a) for overcoming the explained econometric problems and (b) for estimating suppliers' positions on the vertical dimension (i.e., quality).

The proposed framework resembles the discrete/continuous problem analyzed by McFadden and Dubin (1984) and Hanemann (1984). A few differences are worth noting: First, in this paper, I allow for many alternatives to be chosen at the same time. Second, the number of units purchased is discrete, whereas it was continuous in the previous studies. In addition, the estimation procedures available today (i.e., simulation procedures) allow for richer assumptions regarding the distribution of the unobservables as well as for a richer behavioral model. In a related work, Manski and Sherman (1980) estimate demand for cars at the micro-level, allowing for up to two cars to be purchased by each household. Their main focus was on the relationship between different purchases by the same household. The error term was assumed to be distributed as an extreme value, hence the decision of how many units and what brands to purchase could be nested. Manski and Sherman then proceeded to estimate demand conditional on the number of units purchased, that is, estimating the model for those households that purchased one car and those purchasing two cars separately. Their problem is conceptually similar to ours but their solution cannot be extended to our case as the number of PCs purchased ranges from zero to several thousands, rather than from one to two. Train, McFadden and Ben-Akiva (1987) analyse demand for local telephone services using an approach similar to Manski and Sherman's. Their feasible set is much more complex. They assume an extreme value distribution for the error term, which enables them to overcome the choice set complexity by sampling alternatives. My framework differs from theirs in that I do not treat alternatives as mutually exclusive and distributional assumptions are not restrictive at all. In particular, the framework allows for the inclusion of random coefficients, making demand substitution patterns much more reasonable (see Berry, Levinsohn and Pakes, 1992).

Once the goal of estimating demand is achieved, I proceed to assess the welfare implications of computerization in Section 5. My estimates show that return on investment on PCs by the banking industry was 90%. This framework allows, among other things, for estimation of productivity gains from different PC attributes. An increase of 10% in the performance to price ratio in the microprocessor industry translates into a 4% increase of end user surplus.

The rest of the paper is structured as follows: Section 2 presents the model underlying the multiple-discrete choice problem. Section 3 discusses the estimation approach. The data is presented and explained in Section 4. Section 5 presents the empirical results. Conclusions are presented in Section 6.

Section 2 - The Model

2.1. Motivating Facts

Table I presents the joint distribution of the number of PCs and the types of PCs held by 8000 firms in the United States. From this table we learn: First, that multiple-discretness is an essential feature of PC demand. Hence any serious attempt of demand estimation should conform to this main feature. Disregarding this characteristic would imply not using all the available information and thereby distorting the results. Second, aggregation over firms can be misleading, as different firms behave very differently. In particular, a firm purchasing 2000 units is not equivalent to 2000 firms purchasing one PC. Finally, the table makes very clear the complexity of the choice set - as any combination of types in any number of units are feasible. These are the main reasons for presenting a framework that: enables estimation at the micro level and is consistent with the essential features of demand.

2.2. Basic Assumptions

In order to treat the differentiated product market in a manageable way, I adopt the characteristics approach, Lancaster (1979), assuming each PC represents a different bundle of characteristics in an M-dimensional space.

These attributes consist of three different types. First, I distinguish between N built-in (MHZ, wordsize, etc.) and M-N added (hard drives, monitors, etc.) attributes. I focus only on the built-in attributes because M-N attributes are not essential to the make/model as they can be modified by the end user. I implicitly assume that the user maximizes his/her utility/profits with respect to the added-in attributes of each computer. Hence, the profit function, to be defined later, is in a sense an indirect one, already optimized in some other dimensions. There are two reasons for this simplifying assumption: (a) to reduce the dimensionality of the problem to the essential attributes; and (b) to fit the available data (the data describes individual purchases without any information about the added-in attributes).

Furthermore, I distinguish one of the N built-in attributes from all the other N-1. This Nth dimension, which I simply call "quality", includes quality control, service, reliability, software compatibility, etc. The main reason for distinguishing that attribute from the other N-1 dimensions is that the researcher in general knows the clock speed, size and ram of different brands (i.e., I observe the N-1 dimensions) but has no available data on reliability, quality control or software compatibility. Hence, this quality dimension is perceived by the users but is unknown to us. This is a common reason for endogeneity bias in oligopolistic equilibrium models, even if micro data is available (see Berry (1993)).

The particular model used here is a random utility (more precisely, random profit) one. Random utility models arise when one assumes that although the purchaser's payoff function is deterministic and known to him or her, it contains some components which are unobservable to the econometrician and hence can be treated as random variables. In our particular case, they represent firms' subjective tastes and perceptions about which I have no information. I denote these random components on firms f's payoffs as

$$A_f = [a_{f_1}, \ldots, a_{f_{(N+I-1)}}] \qquad for \quad f=1,2,\ldots F$$

That is, A_f is a vector with N+I-1 entries, the first I entries represent the subjective quality perceptions about the I different brands and the last N-1 entries are the valuations for the built-in attributes. Note that each user is allowed his/her own valuation of the built-in attributes, as well as his/her own subjective perception about the location of every brand in the quality dimension. I assume a particular distribution for individual tastes (i.e., random coefficients). Note that if the firm has more than one potential task, then A_f becomes a matrix rather than a vector. Every row in matrix A_f represents random coefficients for a specific task. This distinguishing feature of the model allows firms to have different valuations of attributes for different tasks. Let us call the random coefficients in task j (i.e. row j of the matrix A_f) A_{fj}.

While one might be interested in the estimation of those valuations, the data would allow only the estimation of mean valuations, conditional on establishment characteristics. That is, I look for the mean in willingness to pay for different attributes and also for the mean perception of the quality of different brands conditioned, for instance, on the establishment's sector or size. The mean valuation for the N-1 built-in attributes enables me to assess the contribution of technological progress to welfare. See Section 5 for estimated assessments. The estimated J quality perceptions, represent brands' locations in the vertical dimension.

Keeping that framework in mind, let us move to the details of purchasers' behavior. Let D_f represent the available demographic information about purchaser f (an establishment in our sample) (f=1,2,3 ...F). The demographic information includes size, sector, labor composition, establishment function, other information processing equipment (IPE). That is, D_f contains all the observed individual information, while A_f represents the set of (random) coefficients for firm f. The pair $<D_f, A_f>$ completely describes firm f and hence determines its behavior.

A novel feature of the model in question is the following assumption. Each of the establishments may use PCs to perform J_f different tasks. I assume that the number of tasks[3] that can potentially be performed, J_f, is a function of the demographics, that is, different sectors find a different number of potential uses for Pcs. That is:

$$J_f = \Gamma(D_f) . \qquad (1)$$

Given the number of potential tasks to be performed with PCs, I assume that establishment f maximizes profits by choosing X_j ($j=1,..,J_f$), the vector of PC inputs in task j:

$$Max \quad \phi^f(\pi_1^f, \ldots, \pi_{J_f}^f, X_1, \ldots, X_{J_f}) \tag{2}$$

$$where \quad \pi_j^f = \pi(A_j, X_j, D_f) \tag{3}$$

That is, each firm f maximizes an overall profit function whose arguments are the returns from using PCs in each one of the potential tasks, π_j^f, and the PC input levels. The profit functions at the task level depend on the demographics, D_f, the vector of random coefficients corresponding to task j, A_j, as well as X_j (the vector that specifies how many units of each brand of PCs are used in task j). The maximization is over the number and brands of PCs used as inputs in each task j.

Notice the difference between potential and actual tasks performed with PCs. The former is generated by some function of the demographics while the latter stems from profit maximization by each establishment given J_f (potential tasks), the demographics and the random coefficients. In theory, an establishment might find it unprofitable to spend money on PCs for any of its potential tasks.

Let us assume a simple specification for both the overall and the task level profit function. Let the overall profit function be:

$$\Phi^f(\pi_1^f, \ldots, \pi_{J_f}^f, X_1, \ldots, X_{J_f}) = \sum_1^{J_f} \pi_j^f(A_j, X_j, D_f) \quad . \tag{4}$$

This specification assumes no inter-task effects (network externalities). An inter-task effect may be empirically tested later. This assumption, as will become clear later, drastically simplifies the problem, making the estimation algorithm tractable. However, notice that the general specification, allowing for network externalities, may be tractable in other situations such as airlines' fleet composition or the number of cars per household, as the choice sets there are much simpler. For instance, more than four cars per household is barely observed, in contrast to the 4000 PCs that some establishments in this sample hold.

Furthermore, let the return function at the task level be:

$$\pi_j^f(A_j, D_f, X_j) = (\textstyle\sum_{i=1}^{I} W_{ijf} X_{ijf})^{\alpha} S(D_f) - \sum_{i=1}^{I} P_i X_{ijf} \tag{5}$$

where X_{ijf} is the number of PCs of type i used by establishment f in task j; $S(D_f)$ is a return shifter, a function of the demographics that captures the effect of the latter on the scale of purchases; P_i is the price of a PC of brand i; A_j represents f's random coefficients for task j; and W_{ijf} is firm f's valuation or weight of a PC of brand i in task j. W_{ijf} is a function of firm f's random coefficients in task j, A_j, as well as the attributes of the PC brand i. Let us assume the following functional form for W_{ijf}:

$$W_{ijf} = (A_{fj} \cdot C_i)^{m(D_f)} \tag{6}$$

where A_{fj} represents f's random coefficients in task j; C_i is the $\underline{N-1+I}$ vector of brand i's attributes, whose first $\underline{I}$ entries are zeroes and a 1 in position i, and the last $\underline{N-1}$ entries are the built-in attributes of brand i. Hence, the dot product inside the parenthesis represents the sum of f's valuations for the $\underline{N-1}$ objective (observable) attributes of PC of type i plus f's perception of brand i's unobservable quality; and $m(D_f)$ is f's taste for quality. Notice, I am using the terms type and brand interchangeably.

Notice that it is precisely in W_{ijf} that the random coefficients play their role. In this specification m_f is allowed to depend on D_f to test which demographics influence the taste for quality; that is, which of the firm's demographic characteristics can be tied the vertical dimension. In general, the interactions between demographics and attributes could be introduced here by making the random coefficients' distribution a functions of D_f. For instance, I will allow for some attributes' valuation mean to vary across sectors.

My interest lies in estimating the following: (1) the parameters of the structure imposed, (2) the mean quality perception facing each PC supplier in order to reveal his position in the vertical dimension axis, and (3) the marginal valuation for the different attributes. Those parameters reveal all the relevant information about the demand structure.

2.3. BEHAVIORAL IMPLICATIONS OF THE MODEL

Let us return to the specification proposed and analyze its implications. First notice the additivity of the overall profit function in π^f_j. The purpose of this assumption is to simplify the choice problem to a small number of independent decisions. That is, optimal behavior by firm f becomes simply to maximizes $\pi_j^f(A,D,X)$ over X_j; that is, decides what PC brands and how many units to use in task j, independent of all other potential tasks.

Second, notice the linearity of π^f_j in X_{ijf}. The decision process, of maximizing profits at each task j, proceeds as follows: any combination of brands is feasible, but since X_{ijf} appears linearly in both the terms composing π^f_j, in (5), just one type of PC is actually used per task. Hence, the decision process collapses to comparing maximum (over the number of units used) profits achieved by using each of the available brands. For every firm I can define a latent vector variable:

$$\pi^*_j = (\pi^*_{j1}, \ldots . \pi^*_{jI}) \tag{7}$$

where

$$\pi^*_{ji} = Max_X \quad \pi_j(D,A,X) \quad s.t.\ X \text{ is of brand } i.$$

That is, the latent profits π^*_{ji} are maximum profits obtained in task j if only brand i is available. Hence, the problem for the firm is to find the latent profits for each brand and then compare those latent profit levels. Thus, firm f chooses brand i' if

$$\pi^*_{ji'} = Max(\pi^*_{j1},\ldots,\pi^*_{jI}) \tag{9}$$

If purchases were not restricted to integer units, the optimal purchase of brand i would solve the following first order condition:

$$\alpha \, . \, W^{\alpha}_{ijf} \, . \, X^{\alpha-1}_{ijf} \, . \, S_f - P_i = 0 \tag{10}$$

Which implies that optimal level of purchase is:

$$X^*_{ijf} = \left(\frac{\alpha \, . \, W^{\alpha}_{ijf} \, . \, S_f}{P_i}\right)^{\frac{1}{1-\alpha}} \tag{11}$$

Plugging optimal purchases back into the profit function, we can get π^* for the continuous choice case:

$$\pi^*_{ij} = (1-\alpha) \, . \, S_f^{\frac{1}{1-\alpha}} \, . \left(\frac{W_{ijf} \, . \, \alpha}{P_i}\right)^{\frac{\alpha}{1-\alpha}} \tag{12}$$

The last equation shows profits are monotonically increasing in the performance-to-price ratio, W/P. Hence, the chosen brand is clearly the one with the highest performance-to-price ratio, W_{ijf}/P_i .

When only integer units are available, the chosen brand may not maximize the ratio W/P. Imagine brand i' as having the maximal W/P; however, the optimal purchase might be 3.5 units. If that choice is not feasible, alternative i" may become more profitable than i' when purchases are restricted to integer values. The estimation procedure takes care of this integer problem, in the way explained at the end of Section 3.

Third, provided $0 < \alpha < 1$, an interior solution for the number of units purchased is assured.

Fourth, notice that by allowing $A_f(1,\ldots I)$ to differ across establishments, I allow for a horizontal dimension (i.e., at the same price, different users may prefer different PC brands). On the other hand, notice that by allowing m_f to vary across firms, I allow a vertical dimension[4] in the sense that two users agreeing on the attributes of PCs may purchase different types if their tastes for quality parameters, m_f, change the above-mentioned W/P ratios.

In order to understand the way in which m_f affects a firm's decision, imagine a continuum of goods, each one with a different quality level, and assume that all consumers agree in their perceptions, i.e., they agree about product locations on the quality dimension W. Now assume that price, P(W), is an increasing and convex function of W (otherwise, goods in some region are not purchased; by deleting them, the remaining price function becomes increasing and convex) then, as shown in Appendix 2, each consumer prefers the good with a quality such that

$$\eta_{PW} \equiv \frac{\partial P}{\partial W} / \frac{W}{P} = m_f \tag{13}$$

By assumption $P'(W) > 0$ and $P''(W) > 0$, hence, the elasticity of P with respect to W increases in W; then, consumers with higher m prefer better quality, and hence higher priced goods.

A very important feature of the model, that determines its usefulness in dealing with the different dimensions of multiple-discrete behavior, is present in the different roles played in the model by $S(D_f)$ and $m(D_f)$. The former -- a shifter function capturing the effect of demographics on the size of purchases -- as can be seen in the first order condition (10), affects the number of units purchased. But as the indirect profit function (12) indicates, it has no effect on the discrete (brand) choice as it affects all of the alternatives equally. On the other hand, the latter -- the taste for quality $m(D_f)$-- affects the specific choice among different brands as (12) shows. This explains how both functions of the demographics can be identified in practice. They can, as they capture different dimensions of behavior. One determines the number of units purchased and the other the brands chosen.

As mentioned above, there is a dynamic aspect to computer purchase. I am assuming that there is no adjustment cost of investment[5] and that a resale market exists,[6] hence the relevant alternative cost of using a PC is the rental price (of buying it today, keeping it for say one year and reselling it at the used market price).

To summarize: the behavioral model provides precise firm behavior as a function of all the relevant individual information, that is, $X^*(D_f,A_f,P)$. Moreover, the behavior predicted by the model resembles the observed multiple-discretness features of PC demand.

2.4. Advantages of this approach over Alternative modelling strategies

An alternative approach to the one suggested here in estimating demand is to posit a functional form on the market level demand system; however this approach is problematic, as explained in Bresnahan (1989), because of the number of parameters to be estimated (as many cross-price elasticities are involved as the square of the number of goods). Furthermore, that approach is inappropriate for dealing with innovations (new goods and new attributes), an essential feature of the market to be analyzed here. Moreover, that approach would only exploit aggregate data and not the available micro data. Alternatively, as extensively used in the recent industrial organization literature, one can take as primitives a system which represents consumer preferences over products as a function of those product attributes as well as of consumer characteristics. This approach which is the one taken here, enables us to reduce the dimensionality of the parameters to estimate, and to treat new products consistently.

The framework presented in this paper is developed with a particular goal in mind: to exploit the available micro level data. An important question is then: What is to be gained from using micro level rather than aggregate data? First, more information is used. Second, in contrast to the automobile market where it is not unreasonable to assume that every consumer purchases at most one unit of the good, the PC case presents an extra dimension, i.e., the number of units purchased. This makes it unconvincing to pose an individual utility function and then aggregate it to the market level without use of those micro explanatory variables that determine the size dimension. It would mislead the welfare analysis since it does not distinguish between a consumer purchasing m units from m consumers buying one unit each. Finally as a result of using micro data, additional questions such as Who benefits from different technological innovations? can be answered. Hence, the extra advantage of our approach is enabling to use micro data, which in turn conveys more of the information contained in the data.

As stated in the introduction, the main aim is to estimate demand, i.e., to reveal the parameters of the structural model just described. I will now present the estimation procedure and data.

SECTION 3 - ESTIMATION PROCEDURE

The estimation procedure is, in principle, the standard one. That is, those parameter values that minimize some metric function between observed and predicted behavior were chosen as estimates. In our particular case, because of the complexity of the choice set and the process generating the potential number of tasks, I am unable to compute predicted behavior. The difficulty appears in integrating out over the random coefficients as well as over the task generating process. These sources of randomness represent the error terms from our standpoint. The approach taken is simulation.

Assume that the random coefficients A_t are distributed normally with a mean and variance to be estimated. Furthermore, I have to specify $\Gamma(D_t)$, the relationship between demographics and J_t (total number of potential tasks). Assume that J_t is distributed Poisson with the mean being a function of the establishment characteristics. That is,

$$J_t \sim \text{Poisson}(\Gamma(D_t))$$

The model described above and the distributional assumptions enable to predict behavior for any possible vector of parameter values. It predicts purchases for every establishment (purchases for every potential task) as a function of its demographics, individual tastes (unobservables to us) and a parameter vector to be estimated:

$$X_{fj}^{*}(D_f,A_{fj},\theta) = (X_{1fj}^{*}(D_f,A_{fj},\theta),\ldots,X_{Ifj}^{*}(D_f,A_{fj},\theta)) \quad f=1,\ldots,F \text{ and } j=1,\ldots,J_f \tag{14}$$

That is, a vector with the number of units each establishment f uses in task j of each of the I types of PCs. All entries except one are equal to zero in this vector. The non-zero entry can take any non-negative integer value.

If the model is correct, it would perfectly predict every establishment's behavior, had I known D_f, A_f and Θ. The parameter vector is what I want to estimate, hence unknown; the set of random coefficients of firm f, A_f is unknown as well. Hence, our best prediction for each establishment's total[7] purchases of PC type i will be the expectation of X^{*} conditional on the available information, added over j, that is:

$$X_f^{e}(D_f,\theta) = \int_1^{\infty} \Big(\sum_1^{J_f} \Big\{ \int_{-\infty}^{+\infty} X_{fj}^{*}(D_f,A,\theta)\,\mu(dA/D_f,\theta) \Big\} \Big) \,\mathrm{P}(\,dJ(D_f)\,) \tag{15}$$

where $\mu(x)$ is the normal density and P(x) is the Poisson distribution.

This means that the expected behavior of a firm, conditional on the available information (D_f), is given by the sum (over the different tasks) of expected purchases conditional on a specific number of tasks, then averaged by the probability that these are the actual number of tasks. Notice that, if I had data at the task level, I could treat each firm/task as a different unit of observation, that is, conditioning on tasks. As I do not have such detailed data, I have to work at the firm level, aggregating purchases over the different tasks.

Using predicted behavior, I can define the prediction error:

$$\epsilon_f(D_f,\theta) = X_f^*(D_f,\theta) - X_f \tag{16}$$

where X_f is the vector of actual purchases of establishment f. If the model depicts the actual generating process for purchases, then at the true parameter values, Θ_o :

$$E(\epsilon_f / D_f, \theta_0) = 0 \quad for\ f=1,2,...F. \tag{17}$$

Consequently, any function of the available data independent of the unobservables (i.e., contained in the information available to us, D_f) must be uncorrelated with ϵ, when the latter is evaluated at $\Theta=\Theta_0$. That fact, as suggested by Hansen (1982), can be used to generate a method of moments estimator for Θ. The estimation procedure is based on forming the sample analog of those moment restrictions and then, because it converges to the true moment condition, looking for the parameter value Θ_{MM} for which the sample moments are closest to zero.

More precisely let,

$$G(\theta) = E(T(D) \otimes \epsilon(\theta)) \tag{18}$$

where T(.) is any function of the available information.

Then (17) guaranties $G(\Theta_0)=0$.

The sample analog is:

$$G_F(\theta) = \frac{1}{F}\sum_{f=1}^{F} T(D_f) \otimes (X_f^*(\theta) - X_f)$$

Under some mild assumptions, $G_F(\Theta)$ will converge to $G(\Theta)$ uniformly in Θ. Hence, an estimate could be chosen by setting

$$\underset{\theta}{Min}\ |G_F(\theta)| \tag{20}$$

See Hansen (1982) for conditions to assure consistency and asymptotic normality with covariance

$$(\Lambda'\Lambda)^{-1}\Lambda' V\Lambda(\Lambda'\Lambda)^{-1} \tag{21}$$

Where

$$\Lambda = \frac{\partial G(\theta_0)}{\partial \theta} \qquad and \qquad V = \mathrm{E}((T\epsilon)(T\epsilon)') \tag{22}$$

Since the precision of the estimates depends on the weighing functions T, I use Hansen's procedure to find the efficient weighing matrix for a given set of instruments.

Demographic variables were used as instruments. Since I have conditioned the predicted behavior on demographic information, those instruments are valid, that is, orthogonal to the error term, by construction. Among the feasible ones are firm size and labor composition.

The difficulties in integrating over the unobservables led us to substitute the generalized method of moments for the simulated method of moments. The latter accounts for simulating expected behavior rather than computing expectations, i.e., integration. The simulation is performed by choosing an R, the number of random draws to be taken from the different distribution functions describing the uncertainty. The first step of the simulation requires drawing from a Poisson distribution (see Appendix 3 for the way I perform the Poisson simulation) whose mean is a function of D_f in order to get J_{fr} for $f=1,2,\ldots F$ and $r=1,2\ldots,R$. That is, given the functional dependence of the mean of the distribution on the demographics, for every firm in the sample I draw R simulated numbers of potential tasks .

With J_{fr} in hand, I proceed to the second step of the simulation, which is to generate predicted behavior for every firm at each of the R rounds (draws) of the simulation, conditional on the simulated number of potential tasks at every simulation round. Intuitively, the integral (expected value) is nothing other than performing this simulation an infinite number of times. For the simulation method to work, the convergence of the simulated sample moment to the true moment condition is required. That requires, the law of large numbers to work, which in principle requires that both F (sample size) and R (number of simulation draws) approach infinity. The former assures that the simulated sample analog converges to the true sample analog, while the latter assures that the sample analog converges to the true moment condition. McFadden (1989) shows that provided the linearity of the objective function, only F is required to tend to infinity for the convergence of the simulated sample moment to the true moment. Hence, provided a large F, any R suffices for consistent estimates. That is, provided a large enough sample, anyway needed for the generalized method of moments to work, one round of simulation suffices for consistent results. Nevertheless, the asymptotic variance of the estimates does depend on R; that is, the number of simulation draws determines the efficiency but not the consistency of the estimators. Pakes and Pollard's (1989) insightful result shows how to perform the simulated estimation with an affordable number of simulation draws. The main requirement is to keep the uncertainty over the estimation procedure constant.

At this stage, note the importance and restrictiveness of the linearity and additiveness assumed in the model. First, notice that in order to find predicted behavior, I first simulate the number of potential tasks; then, for each task, I draw the random coefficients. The model assumes linearity and additivity in the number of PCs of different brands (X_{ijf}) in performing each particular task. Hence, given an atomless distribution of random coefficients, as assumed, I know that for every task each establishment almost surely chooses one brand, if any. Linearity and additivity are essential to assigning a unique brand per establishment/task. Uniqueness is essential for computational purposes. It makes the purchase decision tractable at the task level, enabling us, given the drawn random coefficients, to compare the payoffs that the J different brands would give to the firm, which facilitates finding predicted behavior. In contrast, if this uniqueness was absent, I would have to compare every combination of brands to every task, making the model non-tractable. The restrictiveness is clear as well. It imposes that for every draw of random coefficients, there is only one brand (outside a set of measure zero) that best suits each task. The limitation is that it does not allow for two brands to perform the same task at the same firm. Were we to observe this in the data, the model would tell us that those are two different tasks.

Second, notice that tasks are a rather abstract[8] concept in our framework. Since the available data is not classified by task, I am unable to define the moment conditions at the establishment/task level. I am compelled to sum predicted purchases over assumed tasks, aggregating purchases at the establishment level. Our moment conditions are then defined as the error term between total purchases of every type of PC ($i = 1, \ldots I$) for each establishment. If tasks are unobservable, can the data say anything about them? How can they help in the estimation? I estimate the relationship between the number of potential tasks and the characteristics of a firm. The former relationship determines the joint distribution of the number of PCs and the number of types of PCs purchased across different firms. Hence, the abstract concept of tasks, not observed in practice, play a role in our estimation approach by determining the profile of purchases conditional on the demographics of firms. For example, imagine that two establishments purchased the same number of PCs but one of them has one task and the other 10 potential tasks, hence according to the model the former would purchase just one type of PCs while the second up to 10 different types, each of them in a lower quantity. That is how tasks determine the joint distribution of units and different types. The joint distribution of types and number of units observed in the data enables estimating the relation between demographic information and tasks. The estimation has two joint levels, the task generating function and the profit function given the tasks.

In Section 2.2, I explained the decision process when a non-integer purchase of computers is feasible. I then presented an example that showed that the process may not accurately describe the case in which purchases are restricted to integer units. The estimation algorithm solves the problem in the following way: find X^*, the quantity of PCs that solves the first order condition, then take the closest two integers and plug them into the profit function in order to get what I defined as latent profits (see Section 2). We know, by the concavity of the profit function, that one of the integers contiguous to X^* is the profit maximand in the set of integers. Then, for every establishment/task I have two times J options to compare, i.e., two for every brand. Among those alternatives the one giving maximum profits is the one chosen according to the model.

To summarize, the estimation procedure is as follows: The behavioral model is a map $M:(D_f,C,P,\Theta)\text{--}>X^e$. The data gives us D_f, C, P and X. I choose Θ_{MM} such that makes X^e resemble X as much as possible. I now present the data sources.

SECTION 4 - THE DATA

What makes the pursuit of this project possible is the availability of two data sets. One is comprised of data on prices and attributes of microcomputers marketed in the US from 1976 to 1989. This data set was collected by E. Berndt and Z. Griliches, see Berndt and Griliches (1993). It provides (C,P), as denoted above, the N-1 objective attributes of the available choices as well as their prices, that is, information on the feasible set of choices. The second and primary one is the Comtec[9] data set, a survey of 8000 representative non-farm establishments in the US economy, containing information on size, sector, weight in the industry, location as well as other demographic information of each business unit. It includes highly detailed data on actual purchases of "information technology" including PCs and software by those establishments. This data set provides the actual behavior of the sampled firms as well as their demographics. The 1984 and 1988 Comtec surveys were generously made available for this work by Manuel Trajtenberg.

The product choice set data (from Berndt and Griliches) was collected from advertisements in specialized magazines (Byte, PC Magazine, PC Week, etc.) from 1978 until 1989. It includes MHZ, RAM, expandable RAM, number of slots, weight, size, ROM, wordsize, as well as prices for different system configurations in various monitors, hard drives and number of floppy disk drives. Appendix 4 shows the values for those PCs included in the estimation.

The micro data on the information processing equipment purchases of establishments is extremely broad, though few of the reported variables proved useful for our purposes. Henceforth the economic unit I will refer to is the establishment rather than the firm -as the survey is performed at the establishment level. The surveyed units are one location of a firm that may have more sites.

This is a stratified sample, i.e., the survey is performed by randomly sampling establishments that belong to a particular size and sector classification. That is, the whole economy is divided by sector and ranges of size; the sampling is performed in each of those cells. Every sampled establishment is assigned a "weight", computed by ascertaining how much of its cell each establishment represents. This weight can be used to extrapolate figures from the sampled establishments to the economy as a whole. Surveyed establishments answered one questionnaire about demographics and another about their information processing equipment which included: telephone systems, facsimile machines, PCs, computer systems and text processing equipment. The demographics include, among other information: sector, employment, number of white collar and desk workers, enterprise size, establishment function, region, time at premises, autonomy in purchasing and activities performed.

Concerning information processing equipment, I only have information relating PCs. The PC questionnaire starts by asking how many PCs the establishment has; what their makes/models are and how many units of each make/model the establishment has. Then, for each make/model, communication as well as software information is reported. The questionnaire ends with requests for information about plans for removal and new purchases for the coming year.

As the whole sample was too big to handle at this stage, and the establishments too heterogeneous, an individual sector was chosen. Banking is an appealing sector as their computer usage is not insignificant and previous studies on its computerization can be used as a test of validity[10]. The following tables present summary statistics for the whole sample and for the banking industry .

Table II presents summary statistics of the unweighted data to give a sense of the sampling procedure (compare these figures to those in Table III using weighted numbers). I observe a negative correlation between establishment size and weight, suggesting that large establishments were oversampled. Hence, they were given a lower weight for the extrapolation from the whole population. Sampled firms have, on average, over 240 employees and 35 PCs, while weighted figures show about 15 employees and 2 PCs for the average establishment in the economy as a whole. Table IV presents those same figures for three of the sectors used for estimation. The sectors chosen in the present research are shown to be better equipped (in terms of general information processing equipment) than the average establishment. See Appendix 5 for the banking sector correlations.

SECTION 5 - RESULTS

5.1. Empirical Results

The results presented here are mainly for banking industry establishments in the Comtec data set although results from other sectors such as education and R&D are presented for comparison. There are 546 establishments in the banking industry; I randomly choose 123 for the estimation.[11] Those 123 establishments made a total of over 1000 PC purchases (I identify purchases by the number of different types of PCs an establishment owns). The banking or FIRE (for financial, insurance and real estate) industry consists of banks, brokers, S&Ls, insurance companies and insurance agencies.

Notation :

emp_f = number of employees in establishment f.

wh_f = number of white collar workers in establishment f.

$soft_f$ = number of different types of software used in establishment f.

$drd_f = 1$ if establishment f belongs to the R&D sector, 0 otherwise.

$ded_f = 1$ if establishment f belongs to the educational sector, 0 otherwise.

$depr_{if} = 1$ if firm f held in stock PCs of brand i the previous year, 0 otherwise.

This last variable is a dummy, introduced as a first step in dealing with the dynamics of purchasing behavior. The model assumes away adjustment costs in PC investment as well as a perfect resale market, in which case, the relevant alternative cost of holding a PC is its rental price. For a new machine, it is simply the price of buying it today and reselling it tomorrow. The alternative cost of keeping an old machine, purchased in the past, for an additional year is its resale price today less the resale price tomorrow. However, I have no data on the prices of used PCs. Hence, I assume that the alternative cost of holding an old PC for an extra year is proportional to that of a new machine of the same type. The proportion in which differ will be captured by the coefficient on the variable $depr_{if}$. A potential validity check on that parameter is to compare it to the gap between new and used prices of the same machine. Estimates show they are in fact very similar.

As I do not have data on used prices, I simply use the Berndt-Griliches result about price indices for PCs. They find that keeping quality constant, PC hedonic prices drop by 25 percent per year. Our working assumption is that users foresee such a price decrease (due to technological improvements and market interactions) in calculating the cost of keeping a machine for one year and then reselling it. Moreover, I do not know the precise added attributes each user purchased. Therefore, I build a fictitious price that reflects the essence of each model, that is, the basic configuration (built-in attributes). The fictitious price is calculated by subtracting all the accessories (such as monitor, hard disk and floppies) from the PC discount price. This fictitious price is used to compute the rental price of the basic configuration.

Distributional assumptions:

The potential number of tasks, $\Gamma(D_f)$, is assume to distribute Poisson with mean:

$$\lambda_f = g_0 + g_1 . emp_f + g_2 . soft_f \tag{23}$$

The quality dimension (for every type i), not observed by us, is assumed to be distributed normally in the population with some mean A_i , to be estimated.

Let us restrict the objective attributes of a PC to be just MHZ, RAM and bits. I allow for a random coefficient for the valuation of MHZ. I estimate the mean valuation and variance in the population in question, assuming that this valuation is distributed normally as well.

Functional forms:

The following are the functional forms selected for $S(D_f)$ and $m(D_f)$. Both terms have already been presented in equations (5) and (6):

$$S(D_f) = s_0 + (s_1 + s_{1rd} \,.\, drd_f) \,.\, emp_f \tag{24}$$

$$m(D_f) = m_0 + m_1 \,.\, wh_f + m_{1rd} \,.\, drd \tag{25}$$

In the choice set, 15 types of PCs are included: IBMs 8088, 8086, 286, 386, Apple 6502, 68000, 68002, Zenith 8088, 386, Compaq 286 and 386, Clones 8088, 286 386, ATT 8086. See Appendix 4 for the attributes of those PCs and a description of the PCs included in the Clone category. The estimated parameters are reported in Tables V-VII. Table V[12] presents the estimated parameters for the banking industry. Table VI presents additional parameters for the R&D and education sectors, when those are added to banking in the estimation procedure. Table VII presents the estimated mean qualities for the different PC brands, that is, their perceived locations in the vertical dimension.

The Fit of the Model

The value of the objective function, $G_F(\Theta)$ (see Table V), equals 405. It is asymptotically distributed Chi-square with 20 degrees of freedom (45 moment restrictions less 25 estimated parameters). The critical value for rejecting equality to 0 at the 0.05 significance level is 30. Thus the model is rejected by the data.[13]

The coefficients g_0 and m_0 were not estimated but kept fixed at the 0.1 and 1.0 level, respectively. The main reason was to reduce the dimensionality while searching for the estimates (given that in previous runs they seemed not to be that different from their assigned values). The model can only improve by removing the constraints.

Parameter Estimates

As expected, α_0, is found to be between 0 and 1.0. This assures the concavity of the objective function, see (5), and therefore provides an interior solution for the number of PCs purchased.

The returns shifter function, $S(D_f)$ in (5), captures the relation between firm demographics, D_f, and the number of PCs purchased. See (11) for optimal purchases as a function of $S(D_f)$. I am assuming $S(D_f)$ is a linear function of the number of employees in firm f, emp_f. s_1 is the coefficient of employees. I allow for this function to differ across sectors. A significant s_1 shows the expected size effect given by the number of employees per establishment, that is, the relationship between firm size and number of PCs purchased. The intercept, s_0, is significant, indicating that PC purchases are lower in small firms but still significant. Table VI presents sector effects in the function $\underline{S}$. In particular, s_{1RD} is significantly larger than zero, showing that the purchases by firms in the R&D sector are more sensitive to the size of the firms. This phenomenon is expected given the presumed complementarity between labor and PCs in that sector.

In (6) we see the role played by the taste for quality parameter, $m(D_f)$. The highly significant m_1

shows the existence of a vertical dimension tied to size, indicating that bigger firms do prefer higher quality, and hence more expensive machines, in a statistically significant way. Table VI reports a non-significant m_{IRD}, rejecting the hypothesis of a vertical dimension tied to sector.

Recall that I assumed that the population of perceptions about the unobservable quality dimension of every brand distributes normally (see Section 3). The same variance was imposed for all the different brands. The variance, $V(A_i)$ is found to be significatively different from zero, showing that firms in fact do have different perceptions of any given computer aside from the objective characteristics (e.g., MHZ, RAM, bits). In other words, a horizontal dimension of differentiation exists.

The coefficient for mhz, A_m, as well as its variance are significant, confirming a random coefficient for mhz across the sample. The coefficient for RAM, A_r, is found to be significant as well. Moreover, an assessment of their relative contribution is achieved. On average, users are indifferent trading 1.6K of Maximum expandable RAM for an additional MHZ. The coefficients for 16 as well as 32 bits are non-significant, indicating that users do not place any value on that attribute or that their effect is already captured by the coefficients of other attributes.

The task generating process was assumed to be distributed Poisson, with a mean being a linear function of the total employment, emp_f, and an index of software diversity in establishment f, $soft_f$. Both coefficients generating the potential number of tasks, g_1 and g_2, are significantly different from zero. A positive g_1 shows that the number of the potential tasks increases with establishment size. Moreover, since the variance equals the mean, for a Poisson random variable, it also means that the number of tasks, J_f, is heteroschedastic. In addition, g_2 shows a significant relationship between variety of software used and the potential number of tasks performed by an establishment.

The coefficient allowing for depreciation, d_0, is significantly different from zero. Moreover, I expect it to be lower than 1.0 (lowering the repurchase price of one's old equipment). The estimated value is 0.79, significantly lower than 1.0 ($1-d_0$ has a t-statistic value of 2.29). Moreover, the estimated

20% gap between the alternative cost of a new machine and that of an old machine of the same type, is very similar to the actual comparison between new and used PC prices.

Looking at the summary statistics in Table IV (Section 4) we can see that R&D establishments have a higher ratio of PC's/emp. The model found a significative s_{IRD}. On the other hand, the average value of those PCs purchased by the R&D sector is not higher than those held by establishments in the banking industry. Consistently, as reported in Table VII, we found a non-significant m_{IRD}. That is, the model was able to capture both effects, the size effect as well as the lack of a vertical dimension.

The model captures the taste for Apple for the educational subsample, as displayed in Table VII. The dummy that interacts Apple computers and educational sector (A_{Apple}) is borderline significant. The size of that dummy variable is equivalent to 4 MHZ. That is, the preference of using an Apple is comparable to having a non-Apple machine with 4 additional MHZ.

The estimated quality location coefficients (all highly significant) are shown in the Table VII. The reported average mean quality is weighted by sales. Brand locations (i.e., average mean quality) seem reasonable, especially for the period in question when IBM's hegemony was undisputed.

Notice that a unique perception (mean quality) was allowed for both Zenith models and both Compaq models (that is, both models of each of these brands were constrained to the same location, while other makes were allowed different perceptions for different models). This decision was quite arbitrary. On the one hand, I cannot allow for a full set of dummies. If I did so, I could not identify the effects of the objective attributes (i.e., MHZ, RAM, bits), at least one dummy must be dropped. On the other hand, is it at the model or at the brand level that companies develop goodwill? There are well-known cases of lemons produced by high quality companies, e.g., the IBM PCJr. The right level of aggregation for the location parameter remains to be determined.

5.2. Implications

The estimated parameters describe the demand structure, in particular demand as a function of both characteristics of the users and the attributes of the products. I can now use the estimated preference structure to evaluate the welfare gains from computerization, to assess the gains from technological advances in peripherals industries and to compute price elasticities of demand in order to compare them to what is implied by theoretical models of pricing behavior.

5.2.1. Welfare Gains from Computerization

I have used the estimated model to compute the welfare implications of the computer revolution. By plugging the estimated parameters into the model, I can estimate the profits level achieved by the sampled firms at their actual or predicted behavior. This calculation gives a number implied by the estimated profit function for each establishment in the sample. I then use the weight of each sampled establishment to extrapolate to the whole banking sector.

The rational behind the calculation is the following. By observing purchases we can infer demand, i.e., preferences. By Hotelling Lemma, the derivative of the profit function with respect to an input price, which we estimated, equals the demand for that input. To compute the gains from computerization, we are taking, the area below the demand curve for PCs, between 1988 PC prices and an infinite price level. This area represent the difference in profits between 1988 and a situation without PCs.

The banking industry, according to extrapolated figures, owned 729,922 PCs in 1988 and extracted a profit of $1.2 billion[14] from the use of PCs during that year. This means, on average, a "user" surplus of $1300 per PC yearly, while the average price of renting the equipment for one year was $1443. To get a better sense of the former result, let's compare that figure with national accounting data (U.S. Dept. of Commerce, 1989). The industry in question (leaving aside some industry definitional problems) employed about 5 million full-time workers and had an output of $500 billion in 1988. The profit level was $41.8 billion; that is, the welfare gains represent 3.5% of profits. An apparently reasonable return and very significant when compared to invested dollar in PCs. The return on investment is 90%. This should be taken as an upper bound as some costs such as labor training are not accounted for here.

As a comparison, let us take Bresnahan's (1986) estimated surplus from mainframe use in the banking industry in 1972. He reports that banks spent $68 million (rental figure in 1989 dollars) on mainframes and extracted a surplus between $225 and $417 million. He found a much higher return on investment although a lower user surplus, but in the same order of magnitude.

Some comments are in place now. First, notice that we have not used profits information but we are concluding about profit levels. PC prices are anchoring our estimates to monetary units. The estimated profit function is cardinal, i.e., cannot be monotonically transformed without affecting behavior (in contrast to utility function that under a budget constraint would still represent equal behavior after a monotonic transformation). Second, the reader may be puzzled how can I estimate total profits if the only input information available is PCs purchases. The answer is that I am not estimating total profits. If you recall our specification, the model presented every firm an outside alternative, of not purchasing PCs at all. Hence, the profit level should be interpreted as profits beyond the outside alternative, which includes minis, mainframes, or no equipment. This latter point also explains where is the substitutability to other equipment (e.g., minis and mainframes) hidden. Finally, notice that we used predicted (by the model) PC purchases rather than actual purchases to assess the gains from computerization. At first glance it seems more convincing to use the latter. But an important piece of information is missing. I do not observe actual tasks; hence, I do not know how to evaluate actual PC holdings in the profit function. The latter, being concave, implies that the larger the number of tasks, the greater the welfare from a given number of PCs.[15] One approach to take is to treat every different type of PC as a different task.

According to the model, this is a lower bound on the number of tasks, as one type of PC may be used for more than one task but no two types are used in the same task. Hence, in theory this approach would lead to a lower bound to the welfare gain. The calculation lead to a welfare gain of $956.25 millions, that is, 25% lower than using predicted purchases. The result is consistent with the model's prediction that actual purchases would lead to a lower bound on the estimated gains.

5.2.2. Price demand elasticities and implied markups

The estimated parameters determine demand elasticities with respect to price. They serve first as a validity check; that is, to test whether the estimated parameters have reasonable implication regarding substitution patterns. Moreover, they can be confronted with different models of oligopolistic behavior both as a validity check and as a hint on the actual conduct on the market. Table VII presents demand price elasticities. Elasticities are computed by taking the numerical derivative of the estimated demand, then multiplying it by the actual price and dividing it by actual purchases.

The table shows all negative elements on the diagonal. It also implies that most of the models are priced in the elastic part of the demand curve, consistent with oligopolistic conduct. Interestingly, the highest elasticity of demand of IBM 286's is with respect to the price of the IBM 386's, and vice-versa. This is expected as those are the most similar compatible models. The demand of the Zenith 8088 is more sensitive to the price of the IBM 8088 than to any other price. The demand for Apple's 68020 is highly responsive to the price of the IBM 286, which is in a similar price and attribute range.

I can use estimated demand combined with conduct assumptions to infer markups. Just suppose firms were playing a static Nash game in prices (overlooking durable good and multi-good firm problems). Then, from demand,[16] I can find the numerical derivative with respect to its own price in order to calculate marginal revenue. From the first order condition of the assumed conduct, marginal revenue should equal marginal cost. As an example, I take the IBM 386; an $8000 machine in 1988.

The marginal revenue according to the estimated demand derivative and actual sales is $6350. The implied markup is $1650, that is, a 25% markup - a reasonable figure for an IBM recently introduced into the market.

5.2.3. Peripherals technological advances assessment

It would be interesting to assess the effect of technological progress in peripherals industries. This would account for decomposing the gains from improvements in different PC attributes; for example, checking the effect of a reduction in microprocessor price to clock speed ratio. Think of an increase in MHZ for a given price. Imagine all available microprocessors running 10% faster at an unchanged price. The 10% increase in performance to price ratio would imply an increase in the end user surplus of 4%. The result should be taken as a simple exercise, assessing the dependence of user welfare on the peripherals industry and not as a truly comparative static, as I did not model the PC suppliers' reactions to such change.

5.2.4 Comparisons to Hedonic Regressions

Hedonic regressions describe the relation between prices and attributes (generally, both in logs) of a differentiated good. In general, they reflect equilibrium prices as a function of attributes. Hence, no interpretation of marginal benefit can be given to the coefficient of the different characteristics of the good, as that slope may be reflecting marginal cost, or a combination of both. They can be compared to my estimates, to check how much they depart from reflecting marginal benefits. Let me note that the model presented here does identify demand parameters, as it uses micro-behavior, rather than aggregated equilibrium prices. For comparison, let us take the exercise mentioned in B.3 of increasing the performance of all microprocessors by 10%. According to the estimates in 5.2.3, the effect is to increase welfare by 4%. Berndt and Griliches (1990) report a price to MHZ elasticity for the 1986-1988 hedonic regression (Table 7 in Berndt and Griliches) of 0.458. That is, an increase of 10% in MHZ increases the price by 4.58%. Hence, the estimated relation between price and MHZ is not far from reflecting marginal benefits.

SECTION 6 - CONCLUSIONS

PCs are differentiated and durable goods in continuous process of technological change. The paper develops a model of PC purchasing behavior designed to deal with these features. In particular, it captures the main feature of PC demand, multiple-discreteness of choice. The proposed model together with new econometric machinery and a new data set on PC holdings permits estimating demand at the micro-level. As data aggregation has proved to be one of the causes of the computerization puzzle, this study uses micro-data in order to provide more reliable estimates than in previous works. New econometric methods, simulated method of moments (Pakes and Pollard (1989) and McFadden (1989)) proved to be extremely useful in enabling the estimation.

The demand structure is the basic ingredient in addressing other interesting issues such as: PC suppliers pricing and location behavior, the construction of precise price indices to account for quality changes as well as welfare assessments of technological advances. In this paper I have focused mainly on the latter, but the instruments developed here can be used to explore the rest in future work.

The estimated parameters together with actual PC holdings provide an assessment of the benefits from computerization. In contrast to the findings regarding the computerization puzzle, this model assesses a surplus of about $1.2 billion in the banking industry in 1988 due to computerization. This figure when compared to investment in PCs leads to a 90% return on investment. Furthermore, estimated price demand elasticities as well as implied markups are within a reasonable range; hence, the estimated structure passes that first validity check.

Princeton University

REFERENCES

BERNDT, E. AND GRILICHES, Z. 1993. "Price Indices for Microcomputers: An Exploratory Study" in FOSS, M., MANSER, M. AND A. YOUNG eds. Price Measurements and their Uses , University of Chicago Press.

BERNDT, E. AND MORRISON, C. 1990. "Assessing the Productivity of Information Technology Equipment in the U.S. Manufacturing Industries." National Bureau of Economic Research Working Paper #3582..

BERRY, S., 1993. "Estimating Discrete Choice Models of Product Differentiation". Rand Journal of Economics, 25, pp. 242-262.

BERRY, S., LEVINSOHN, J. AND A. PAKES, 1992. "Automobile Prices in Market Equilibrium". Mimeo Yale University.

BRESNAHAN, T. 1986. "Measuring the Spillovers from Technological Advance: Mainframe Computers in Financial Services". The American Economic Review, 76, 742-755.

BRESNAHAN, T. 1987. "Competition and Collusion in the American Automobile Oligopoly: The 1955 Price War". Journal of Econometrics, 17, 201-227.

BRESNAHAN, T. 1989. Empirical Studies of Industries with Market Power. In Schmalensee, R. and Willig, R. (eds.) Handbook of Industrial Organization. Amsterdam: North Holland, Volume II, 951-1055.

DUBIN, J. AND MCFADDEN, D. 1984. "An Econometric Analysis of Residential Electric Appliance Holding and Consumption". Econometrica, Vol 52, 345-362.

EICSON, R. AND PAKES, A. 1992. "An Alternative Model of Firm Dynamics". Mimeo Yale University.

GILICHES, Z AND SIEGEL, D. 1992. "Purchased Services, Outsourcing, Computers and Productivity in Manufacturing" in Zvi Griliches eds. Output Measurement in the Service Sector, Chicago University Press.

HANEMANN, M. 1984. "Discrete/Continuous Models of Consumer Demand". Econometrica, Vol. 52, 541-561.

HANSEN, L. 1982. "Large Sample Properties of Generalized Method of Moments Estimators", Econometrica, 50, 1029-1054.

HENDEL, I. 1994. "The Role of PCs in the Manufacturing Industry", Chapter 2, PhD dissertation, Harvard University.

LANCASTER, K. 1979. "Variety, Equity, and Efficiency: Product Variety in an Industrial Society." New York: Columbia University Press.

MANSKI, C. AND SHERMAN, L. 1980. "An Empirical Analysis of Household Choice Among Motor Vehicles". Transportation Research, Vol. 14A, No 5-6, 349-366.

MCFADDEN, D. 1989. "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration". Econometrica, Vol 57, 995-1026.

PAKES, A. AND POLLARD, D. 1989. "Simulation and the Asymptotics of Optimization Estimators". Econometrica, Vol 57. 1027-1057.

SUTTON, J. 1991. "Sunk Cost and Market Structure", MIT Press, Cambridge.

TRAIN, K., MCFADDEN, D. AND BEN-AKIVA, M. 1987. "The demand for local telephone service: a fully discrete model of residential calling patterns and service choices". Rand Journal of Economics, Vol. 18, 109-123.

U.S. DEPARTMENT OF COMMERCE. 1989. National Income and Product Accounts of the United States, Vol II 1959-1988.

Appendix 1

Production Function Approach to the Computerization Puzzle

Variable	Coefficient	Standard Deviation
Materials	0.61	0.011
Employment	0.19	0.015
Adjusted Capital	0.09	0.014
Computing Capability (Mhz)	0.01	0.002
Constant	1.96	0.089

Polynomial Approximation to the Individual Effect Yes

R Squared = 94.58

These are the estimates of a Cobb-Douglas production function. An index of computer capability was added to the usual inputs (capital, employment and material). Olley and Pakes (1992) polynomial approximation approach is used for dealing with the individual effect allowed to change over time.

Appendix 2

Most Preferred Quality With a Continuum of Quality Choices

Equation (12) presented optimal profit level while consuming a brand with quality a and price P. For convenience I reproduce that equation here.

$$\pi^* = (1-\alpha).S^{\frac{1}{1-\alpha}}.\left(\frac{a^m.\alpha}{P}\right)^{\frac{\alpha}{1-\alpha}} \tag{a}$$

Now suppose the consumer faces a continuum of qualities, which prices are given by an increasing and convex function of quality, P(a). Then the consumer maximizes profits given in (6) with respect to a. The first order condition is:

$$\frac{\partial \pi}{\partial a} \propto C.\left(\frac{m.a^{m-1}}{P(a)} - \frac{a^m.P'(a)}{P(a)^2}\right) = 0 \tag{b}$$

where C is a non-zero constant.

The first order condition holds for an a such:

$$m = \frac{\partial P(a^*)}{\partial a}.\frac{P(a^*)}{a^*} \tag{c}$$

Based on the assumptions $P' > 0$ and $P'' > 0$, the elasticity of P with respect to a, increases in a, hence the most preferred quality increases in m.

This way, I introduce the vertical dimension into the model, by tying m to some demographic information like size or sector.

Appendix 3

Simulating Poisson Random Variables

The gamma distribution is given by:

$$g(x/\alpha,\beta)=\frac{\beta^{\alpha}x^{\alpha-1}e^{-\beta x}}{\Gamma(\alpha)} \qquad \text{(a)}$$

$$where \qquad \Gamma(\alpha)=\int_0^{\infty}x^{\alpha-1}e^{-x}dx \qquad \text{(b)}$$

It is easy to check that its moment generating function is:

$$M_x(t)=(\frac{\beta}{\beta-t})^{\alpha} \qquad \text{(c)}$$

Moreover, if $s=\Sigma^{K}_{i=1}x_i$ and $x_i \sim g(x/\alpha_i,\beta)$, then $s\sim g(s/\Sigma^{K}_{i=1}\alpha i,\beta)$. Which follows directly from at the moment generating function.

As an exponential distribution is nothing but a gamma with parameters $\alpha=1$ and $\beta=1$, the sum of K exponential random variables distribute $g(./K,1)$, that is:

$$g(y)=\frac{y^{k-1}e^{-y}}{\Gamma(K)} \qquad \text{(d)}$$

Then:

$$Pr(y<\lambda)=\int_0^{\lambda}\frac{y^{k-1}e^{-y}}{\Gamma(k)}dy=1-\sum_{j=1}^{k-1}\frac{e^{-\lambda}\lambda^{j}}{j!} \qquad \text{(e)}$$

where the second equality follows by integrating by parts K-1 times, and using the fact that $\Gamma(K)=(K-1)\Gamma(K-1)$.

Notice that if a random variable n is distributed Poisson with mean lambda then:

$$Pr(n=j)=\frac{e^{-\lambda}\lambda^j}{j!} \tag{f}$$

Hence, I get

$$Pr(y<\lambda)=\sum_{j=k}^{\infty} Pr(n=j) \tag{g}$$

where y is the sum of k exponentially distributed random variables and n is distributed Poisson with mean lambda.

After some algebra, you can verify

$$Pr(y_k<\lambda<y_{k+1})=\int_0^{\lambda}\int_{\lambda-y_k}^{\infty} e^{-x_{k+1}}dx_{k+1}\frac{y_k^{k-1}e^{-y_k}}{\Gamma(k)}dy_k=\frac{e^{-\lambda}\lambda^k}{k!} \tag{h}$$

Where y_x is the sum of x independent exponentially distributed random variables.

However, the RHS term is none other than

$$Pr(n=k) \quad \textit{for n distributed Poisson with mean } \lambda$$

This shows that to sample from Poisson with mean lambda, one can draw exponentially distributed random variables until the partial sum exceeds lambda. The number of draws required to reach lambda distributes Poisson. Hence, in order to perform the simulation estimation, which requires keeping the randomness constant throughout the minimization (parameter search), one can simply keep the exponential draws constant and let lambda change. Exponential random variables are easily attained by drawing from a uniform distribution (that any statistical software provides) and applying the inverse transformation of the exponential c.d.f..

Appendix 4

MAIN BRAND ATTRIBUTES

	Model [1]	MaxRam	Mhz	MakePrice	Bits[2]	Sales[3]	Shares
IBM	8086	640	8	750	16	1320	6.1
IBM	8088	640	4.77	1069	16	4533	20.1
IBM	286	1024	10	2700	16	8117	37.3
IBM	386	2048	20	4400	32	3673	16.9
Apple	6502	550	1	629	8	165	0.8
Apple	68000	1024	7.83	1800	16	963	4.4
Apple	68020	1024	15.6	2400	16	79	0.4
Zenith	8088	1024	8	1150	16	79	0.4
Zenith	386	1024	16	3219	32	33	0.2
Compaq	286	2048	12	1900	16	821	3.7
Compaq	386	2048	16	3400	32	404	1.9
"clone"[4]	8088	640	7.16	800	16	244	1.1
"clone"	286	830	10	1150	16	904	4.2
"clone"	386	2048	20	2800	32	193	0.9
Att	8086	640	8	700	16	197	0.9

1/Models are basically defined by their processors. Different models with the same microprocessor and similar characteristics were aggregated into one model in order to broaden coverage.

2/Refers to the number of bits used in calculations.

3/Refers to 1986-1988 accumulated sales, representing current non-obsolete holdings (where 3 years was assumed to be the life horizon of PCs).

4/"Clone" includes: American Research, Ast, Kaypro, Leading Edge, Packard Bell, Racore, Radio Shack, Sperry, Wang, Wyse and other small compatible PCs.

Appendix 5

Banking Industry

Correlation Coefficient Matrix of Establishment Data

	Employ	Whitepc	Autonom	Offauto	Totalpc	Pcwp	Value
Employee	1.00						
Whitepc	-0.03	1.00					
Autonomy	0.02	0.04	1.00				
Offauto	0.03	-0.09	0.01	1.00			
Totalpc	0.77	-0.01	0.01	0.17	1.00		
Pcwp	0.52	-0.01	-0.02	0.11	0.85	1.00	
Value	0.02	-0.06	0.07	0.01	0.01	-0.01	1.00
Mainframe	0.46	-0.01	-0.03	0.35	0.42	0.35	-0.01

Employee - number of employees per establishment.

Whitepc - percent of white collar workers.

Autonomy - index of autonomy in purchasing equipment.

Offauto - index of office automatization.

Totalpc - total number of PCs.

Pcwp - number of PCs used for wordprocessing.

Value - average PC value.

Mainframe - total number of minis plus mainframes.

TABLE I

JOINT FREQUENCY OF TOTAL NUMBER OF PCS BY NUMBER OF TYPES OF PCS

PCs\TYPES	1	2	3	4-6	7-9	10-19	20-49	50-99	100+	TOTA
1	1369									1369
2	335	308								643
3-4	271	215	196							682
5-6	123	99	129	29						380
7-10	104	111	152	52	7					426
11-15	39	59	102	56	26					282
16-35	68	84	154	99	65	8	2			480
36-60	24	31	68	45	61	15	5			249
61-100	12	12	51	39	47	16	8			185
101-300	16	17	59	74	100	56	21	3		346
301-600	4	3	18	26	29	23	8	0	1	112
601-1000	2	0	3	8	8	11	6	1	0	39
1001-2000	0	1	1	7	5	7	4	3	1	29
2000+	0	0	2	3	2	2	0	1	2	12
TOTAL	2367	940	935	438	350	138	54	8	4	5234

TABLE II

WHOLE SAMPLE (UNWEIGHTED NUMBERS)

VARIABLE	OBSERVATIONS	MEAN	STD. DEV.	SUM	MIN	MAX
EMPLOYE	7895	241.55	823.85	1907074	0	29000
TOTALPC	7895	35.05	194.47	276740	0	7810

CORRELATION

	EMPLOYEE	TOTALPC	WEIGHT
EMPLOYE	1.00	0.56	-0.16
TOTALPC		1.00	-0.10
WEIGHT			1.00

KEY:

WEIGHT - ESTABLISHMENT'S WEIGHT IN THE ECONOMY.

TOTALPC - TOTAL NUMBER OF PCS PER ESTABLISHMENT.

TABLE III

Whole Sample (Weighted Numbers)

Variable	Observations	Mean	Std. Dev.	Sum	Min	Max
Employee	7895	15.49	373.00	112,000,000	0	290000
Whitepc	7895	73.08	105.00		0	100
Offauto	7895	4.02	11.12		0	29
Totalpc	7895	1.96	103.87	142,00,000	0	7810
Mainfram	7895	0.09	4.45	646,219	0	857
Pcvalue	7895	3.45	3.08		0	20
Csvalue	7895	14.93	1018.26	108,000,000	0	80,550
Pcwp	7895	0.40	20.20	2,881,000	0	1750

<u>KEY:</u>

EMPLOYEE - NUMBER OF EMPLOYEES.

WHITEPC - PERCENT OF WHITE COLLAR WORKERS.

OFFAUTO - INDEX OF GENERAL OFFICE AUTOMATIZATION EQUIPMENT.

MAINFRAM - TOTAL NUMBER OF MINIS AND MAINFRAMES.

PCVALUE - AVERAGE VALUE OF PC HELD, IN THOUSANDS OF DOLLARS.

PCWP - NUMBER OF PCS USED FOR WORDPROCESSING.

CSVALUE - VALUE OF MINIS AND MAINFRAMES, IN THOUSANDS OF DOLLARS.

TABLE IV

WEIGHTED FIGURES BY SELECTED INDUSTRIES

VARIABLE	OBSERVATIONS	MEAN	STD. DEV.	MIN	MAX
		FIRE			
EMPLOYEE	620	13.29	238.99	0	7000
WHITEPC	620	96.81	34.91	4	100
OFFAUTO	620	6.41	11.05	1	27
TOTALPC	620	2.86	94.71	0	3550
MAINFRAM	620	0.14	2.11	0	65
PCVALUE	620	4.02	4.16	1	20
CSVALUE	620	38.30	2084.02	0	80,550
PCWP	620	0.60	38.86	0	1750
		R&D/COMPUTERS			
EMPLOYEE	251	137.90	1052.31	0	14096
WHITEPC	251	84.40	25.67	4	100
OFFAUTO	251	6.90	8.15	2	26
TOTALPC	251	64.20	476.92	0	6516
MAINFRAM	251	0.50	5.83	0	139
PCVALUE	251	2.60	1.97	0	20
CSVALUE	251	84.40	1022.71	0	13300
PCWP	251	8.10	72.46	0	1120
		EDUCATION			
EMPLOYEE	117	18.20	252.73	1	5000
WHITEPC	117	88.90	54.73	3	100
OFFAUTO	117	7.40	11.54	1	29
TOTALPC	117	7.50	109.51	0	1950
MAINFRAM	117	0.70	10.79	0	220
PCVALUE	117	4.00	3.42	1	20
CSVALUE	117	95.90	2101.72	0	30356
PCWP	117	1.00	24.14	0	755

TABLE V

Parameter Estimates

	COEFFICIENT	NUMERICAL STD. DEV.	BOOTSTRAP STD. DEV./C
S_0	0.61	-	0.19
S_1	0.47	0.18	0.06
M_1	6E-4	1.3E-4	1E-5
A_0	0.29	0.03	0.04
$VAR(A_i)$	3.50	0.55	0.92
A_M	0.89	0.24	0.16
$VAR(A_M)$	1.48	0.78	0.23
A_R	0.53	0.11	0.06
A_{B16}	0.78	0.87/B	-
A_{B32}	1.17	2.93/B	-
G_1	2.42	1.34/A	0.32
G_2	0.78	0.18	0.10
D_0	0.79	0.09	0.19
$G_F(\Theta)=405$			

KEY:

S_0- INTERCEPT OF THE FUNCTION $S(D_F)$.

S_1- SLOPE OF $S(D_F)$ WITH RESPECT TO EMP_F.

M_1- SLOPE OF $M(D_F)$ WITH RESPECT TO WH_F.

A_0- PROFIT FUNCTION EXPONENT.

A_i- MEAN VALUATION OF MAKE/MODEL I.

A_M- MEAN VALUATION FOR MHZ.

A_R- MEAN VALUATION FOR RAM.

A_{BX}- DUMMY FOR X BITS.

G_1- SLOPE OF TASK GENERATING PROCESS MEAN W.R.T. EMP_F.

G_2- SLOPE OF TASK GENERATING PROCESS MEAN W.R.T. $SOFT_F$.

D_0- DEPRECIATION PARAMETER FOR OLD EQUIPMENT.

/A -10% SIGNIFICANCE LEVEL.

/B -NON-SIGNIFICANCE COEFFICIENT.

/C - TO PERFORM THE BOOTSTRAP, 50 FIRMS WERE RANDOMLY CHOSEN FROM THE EMPIRICAL DISTRIBUTION. I REPORT THE STANDARD DEVIATION OF THE ESTIMATED PARAMETERS IN THE 50 ROUNDS OF BOOTSTRAP. DEVIATIONS ARE TAKEN AROUND THE MEAN BOOTSTRAP ESTIMATES

TABLE VI:

ADDITIONAL RESULTS/A

	COEFFICIENT	STANDARD DEVIATION
S_{IRD}	0.69	0.11
S_{IED}	1.00	3.33
M_{IRD}	0.70	7.00
A_{APPLE}	4.20	3.02

KEY:

S_{IRD}- EXTRA SLOPE OF FUNCTION $S(D_f)$ WITH RESPECT TO EMP_f, FOR THE R&D SECTOR.

S_{IED}- EXTRA SLOPE OF FUNCTION $S(D_f)$ WITH RESPECT TO EMP_f, FOR THE EDUCATIONAL SECTOR.

M_{IRD}- COEFFICIENT OF THE R&D DUMMY IN $M(D_f)$, I.E., VERTICAL DIMENSION TIED TO SECTOR.

A_{APPLE}- COEFFICIENT OF THE INTERACTION BETWEEN EDUCATIONAL SECTOR AND TASTE FOR APPLE.

/A - THE ESTIMATES REPORTED IN TABLE VI (S_{IRD}, S_{IED}, M_{IRD} AND A_{APPLE}) COME FROM A DIFFERENT ESTIMATION THAN THOSE FROM TABLE V. IN THE FORMER, THE A_s'S WERE OMITTED. A_{APPLE} CAPTURES THE EXTRA VALUATION OF THE EDUCATIONAL SECTOR FOR THE APPLE BRAND.

TABLE VII

ESTIMATED MEAN QUALITIES

MAKE	MODEL	MEAN QUALITY
IBM	8086	7.2
IBM	8088	24.0
IBM	286	64.0
IBM	386	85.5
APPLE	6502	10.2
APPLE	68000	30.8
APPLE	68020	40.2
ZENITH	8088	16.5
ZENITH	386	-
COMPAQ	286	40.9
COMPAQ	386	-
"CLONE"	8088	9.7
"CLONE"	286	16.4
"CLONE"	386	50.0
ATT	8086	2.9

AVERAGE MEAN QUALITY	
IBM	54.1
COMPAQ	40.9
APPLE	28.5
"CLONE"	20.1
ZENITH	16.5
ATT	2.9

KEY:

MEAN QUALITY - IS THE MEAN PERCEPTION IN THE POPULATION OF USERS. WHAT I CALLED A_i.

AVERAGE MEAN QUALITY - IS THE AVERAGE QUALITY OF THE DIFFERENT MODELS OF A BRAND, WEIGHTED BY SALES.

TABLE VIII

DEMAND PRICE ELASTICITIES[17]

	I8086	I8088	I286	I386	A6502	A68000	A68020	Z8088
IBM 8086	-0.93	0.05	0.01	0.00	0.20	0.02	0.13	0.80
IBM 8088	0.18	-2.16	0.06	0.07	0.20	0.63	1.95	5.84
IBM 286	0.29	1.15	-1.34	0.65	0.94	0.01	38.29	1.65
IBM 386	0.01	4.39	0.32	-1.91	0.00	8.76	0.39	0.08
APPLE 6502	0.00	0.00	0.00	0.00	-2.58	0.00	0.00	0.01
APPLE 68000	0.00	0.00	0.00	0.00	0.00	-0.13	0.00	0.00
APPLE 68020	0.40	0.08	0.01	0.00	0.00	0.01	-7.9	0.1
ZENITH 8088	0.02	0.05	0.00	0.00	0.07	0.00	0.04	-6.69
ACTUAL SALES	1319	4532	8105	3675	163	963	79	79

ENDNOTES

[1]See Appendix 1 for a summary of these results.

[2]The term "random utility" is used to describe the standpoint of the econometrician. Although utility is deterministic and all the information is known to the decision maker, it contains some components which are not observed by the researcher. These unobserved variables are therefore treated as random variables, capturing both the idea of variation in tastes among individuals as well as unobservables.

[3]The purchasing process of medium- to-large-sized firms is the following: different divisions in the firm are allocated a budget; based on the budget, they specify their needs in terms of attributes, software and units to the "purchasing division" and then that division transacts the purchase given current available alternatives. This process is consistent (provides an additional interpretation) with our model. Instead of different tasks, there are different operative units demanding a specific PC profile. The purchasing division, in making the purchase, introduces an intra-firm purchasing correlation (i.e., different units of a firm do not behave totally independently).

[4]For those used to the usual vertically differentiated models, it may appear suspicious that m, the parameter capturing the vertical dimension, is an exponent (see (6)) rather than a constant multiplying a_{ijt}. But a closer look at the problem makes it clear that multiplying all the alternatives does not change their relative attractiveness to the consumer; rather, the same choice is always made (the one with the highest W/P ratio). The reason why in the classic model of vertical differentiation (Bresnahan, 1987) a constant multiplying utility level could capture the vertical dimension, is that consumers purchase only one unit.

[5]In "The Role of PCs in the Manufacturing Industry", I find evidence, using Euler equations for PC investment, of non-significant adjustment costs.

[6]There is a market for used PCs. Although a price spread between new and used PCs exists, we do find the rental assumption a reasonable approximation.

[7]By total we mean adding over the different tasks.

[8]We tried to exploit the available data on software by identifying the type of software with a task, and then

predicting the behavior conditional on task. However, this attempt failed because most firms use only a few types of software. Moreover, the estimated parameters for different types were not that different, giving evidence that the approach failed in identifying different tasks.

[9]This data was made available through the Institute for the Study of Business Markets, Penn State University. The data was originally made available to ISBN by Gartner Group, Inc. Interpretations herein are my responsibility and not the sources.

[10]Bresnahan (1986) presents a study of the welfare gains to the banking industry from technological improvement in the mainframe industry. His figures in absolute terms are comparable to those found here although they imply a much higher return on investment.

[11]I take a subsample to reduce computing time, in particular to reduce RAM requirements. Using whole sample even a 32 meg RAM workstation would swap continuously making estimation endless.

[12]The variance matrix involves calculating the expected value of derivatives (of the moment conditions). The essence of the problem precludes us from analytically finding them (the same reason that we used simulation). We computed them numerically. They proved to be step-size sensitive. I report here the average from three different step-sizes, 0.5%, 1% and 5%. Bootstrap estimation was suggested to overcome this problem. Both methods, bootstrap and numerical derivatives, are

reported for comparison purposes.

[13]It should be taken into consideration that it is a stringent test, that generally rejects for large enough samples.

[14]Computing a confidence interval, for instance by using the delta method, could be extremely complex. What I did here is to simulate that interval. That is, I have drawn parameters from their estimated distributions, then computed the welfare gain at every draw of the simulation. The standard deviation of the welfare gain is about 0.2 billions, that is, about 15% of the estimated gains.

[15]As $(1/a)f(ax) < f(x)$ for f increasing and strictly concave and $a < 1$.

[16]I am using Banking Industry PC demand, assuming it is representative of aggregate PC demand. In principle the model can be applied to every industry.

[17]Part of the elasticities reported as zero are nonzero although negligible. Notice that simulated demand is non-differentiable for any finite number of simulation draws. That explains part of the reported zeros.