Econ. Innov. New Techn., 2000, Vol. 9, pp. 353-383
Reprints available directly from the publisher
Photocopying permitted by license only

IDENTIFYING THE DEMAND FOR FEATURES: AN APPLICATION TO MAINFRAME COMPUTERS

KENNETH H. BROWN[a*] and SHANE M. GREENSTEIN[b†]

[a]University of Northern Iowa Department of Economics Cedar Falls, IA 50614–0129 and [b]Northwestern University Leverone Hall 2001 Sheridan Road Evanston, IL 60208–2013

(Received February 25, 1999)

This paper examines the mainframe computer market from 1985–1991 and attempts to identify the types of buyers that demand particular computer features, such as speed and memory. To identify these buyers, demand for computer characteristics is estimated using a demand model based on Rosen (1974). Through these demand estimates we are able to show that the advent of on-line transactions processing was pushing the demand for computer speed and memory to some extent. However, beyond this specialized application, only a few industries seemed to be demanding the newest technology, while the majority of buyers continued to buy small mainframes throughout the sample period.

1 INTRODUCTION

As with many information-technology industries, innovation proceeded rapidly in the 1980s mainframe computer industry. Broadly speaking, there was a large decline in price per unit of performance across a wide set of computer systems and components, and vendors developed many new computing functions and capabilities. The market grew and advanced, led by hardware architecture redesign, software development and system customization. For example, "on-line transactions processing" extended the functions of existing large computing facilities, enabling the development of many new goods and new services.

* E-mail: ken.brown@uni.edu
† E-mail: s-greenstein1@nwu.edu

An open debate (see Dulberger, 1989, and Gordon, 1989) concerns the rate and unevenness of computer users' demand for these changing systems. This paper advances our understanding of this topic by measuring the features of buyers which predict demand for computer system characteristics. This model is adapted from Rosen (1974), Bartik (1987), and Epple (1987). This model's principal strength is that it measures the demand for more memory or more speed in a standard price-theoretic treatment.

We study 21,268 acquisitions of mainframe computers from 1985 to 1991, more than half of all mainframe acquisitions in the U.S. This is as comprehensive a dataset as used in any previous study of large computer buyer demand. We observe characteristics of the purchases being made and the characteristics of the firms making those purchases. Most published work on the computing market only has data on the set of systems available for sale. Our more detailed data allows us to directly measure the demand for product features at the level of the individual user.

By any historical standard for innovation, the rate of change in quality adjusted prices in mainframe computers is quite fast (approximately 25–30 percent per year). The maximum size available to buyers also grew over time. Yet, these data show that most buyers initially bought a "small" mainframe system in the early 1980s and still bought a small system at the end of the decade, even with rapidly declining mainframe prices and large extensions in computing capacity.

We focus on estimating the demand for speed and memory as a function of buyer characteristics. We use two classes of regressors to predict this demand. One class describes features of hardware found at a user location, which we largely view as controls for size, industry effects, and other considerations. The other class describes the features of software used at the location. We show that several types of software from different suppliers predict buyer demand, but that no single factor alone predicts the purchase of more speed or memory. We conclude that a variety of factors induced users to buy something other than a small system. These estimates are consistent with industry norms about differences in the demand for large systems across buyers, and about the importance of demand for on-line transactions processing applications in this time period.

This research is also important for understanding the notion of the "lead user." The diffusion literature has long recognized that some users are first to try new products or new features of established products (Rogers, 1962). These users often try new developments and experiment with new applications even before suppliers have worked out all problems with the technology. Further work on the sources of innovation (von Hipple, 1988) has shown that lead users may generate ideas that spill back to vendors. This interaction is an important addition to the view of Trajtenberg (1990), who placed great emphasis on the welfare benefits of stretching the product space. Thus, the characteristics and motivation of lead users are an important source of economic growth and important to understand. In addition, it is likely that the pattern observed in the computer market of the 1980s will be observed for other technical products in other time periods.

Previous work on computing in the 1960s and 1970s (Greenstein, 1995) argued that an important source of growth in large scale computing arose from the growth of use of computer systems with greater capacity. Careful examination of sales data showed that the use of large capacity had the usual pattern of lead users first exploring ways of using larger capacity and experimenting with new applications. This paper seeks to understand whether a unique set of users can be econometrically identified as lead users and whether these lead users are associated with specific business applications or specific relationships with vendors.

2 MAINFRAME COMPUTERS AND TECHNICAL CHANGE

The following discussion provides a basis for the demand model used in this paper. The discussion focuses on the impact of technical change in the mainframe computing market in the 1980s, arguing that technical change had a large influence on the demand for new capacity.

Prices declined rapidly and pervasively across all systems as measured by prices per CPU speed and memory capacity.* The maximum feasible system capacity, as measured by maximum computing memory and CPU speeds, also expanded rapidly.† This permitted users to address increasingly more complex problems involving more calculations and large databases, and regularly perform tasks that could not previously be

* See Dulberger (1989), Gordon (1990), and Triplett(1989). Similar estimates have been found for peripheral and selective software programs. See Cole *et. al.*, (1986).

† For example, the data in this paper shows an expansion in product space from a maximum of 20 MIPS to 110 MIPS in under 5 years.

accomplished, let alone attempted.* This change had many antecedents in the previous two decades, so it should not have caught buyers by surprise.†

Broadly speaking, by the 1980s users had come to expect lower prices, extensions of capabilities, or entirely new products and planned for it. In response, buyers modified the memory and speed of their CPU, but kept other durable investments in software or peripherals. Or, buyers enhanced particular software programs or peripheral components, but not other parts of their systems. A regular pattern emerged in the mid 1960s and 1970s and continued into the 1980s: peripheral and software upgrading induced bottlenecks in CPUs, which later induced further CPU upgrading, which later induced further peripheral and software enhancements, and so on. This pattern is well-known and widely studied.‡ Thus, for many buyers, demand for greater computing capacity proxies for the demand for new peripherals and improvements in software, reflecting the demand for new goods and services.

Several contrasting forces influenced the benefits buyers received from new computing systems. Many buyers of on-line transactions processing (OLTP) place value on larger computing capacity (embodied in CPU), because it allows more users and faster response times with larger databases. Many researchers of centralized management of computing facilities (*e.g.*, Inmon, 1985) emphasize this notion. In contrast, buyers may not have realized localized economies of scale in their large systems in this period (Friedman and Cornford, 1989). The complaints centered on problems inherent in centralized management (of mainframes) rather than technology alone, and hence, no pure "technology fix" was possible.

Development in the market for minicomputers, workstations, and personal computers also influenced mainframe demand. First, by the mid 1980s minicomputer vendors offered users viable growth paths for their systems if the users needs outgrew large superminis. Small users or divisions within large corporations found them attractive as a means to avoid

* Probably the best documented developments were associated with "on-line transaction processing" (OLTP) applications, *i.e.*, applications that required multiple-users to simultaneously access and update large databases. A wide series of innovations, dating to the late 1970s and early 1980s, to hardware architecture, operating system design, networking technology, and application software enabled and improved OLTP applications (Friedman and Cornford, 1989). Many large users, insurance and banking users, wholesalers, and many large database users employed these developments in new inventory and reservation systems.

† For examination of the diffusion of computing systems and some of its economic determinants during this time period, see Greenstein (1997) or Bresnahan and Greenstein (1997).

‡ See Friedman and Cornford (1989) and Inmon (1985).

centralized mainframe management. Second, by the mid to late 1980s, even smaller platforms, associated with personal computers and workstations, offered a different type of decentralized platform for small computing jobs involving small partitionable databases, word-processing, and spreadsheets.

As a result, not many new buyers were drawn into the mainframe market in the late 1980s. The majority were experienced buyers with their legacy systems. This stable set of mainframe buyers was identified and surveyed each year, which provides us with very detailed data on their behavior.*

In sum, technical change was more than a simple fall in the price level. The willingness to take advantage of new capabilities became associated with a willingness to adopt computing capacity at higher levels. Previous research has not identified significant sources of demand for new capacity. This paper will attempt to identify characteristics of buyers of large systems that will be useful in predicting the demand for new capacity which as stated above proxies for the demand for new goods and services, an important source of economic growth.

3 METHODOLOGY

The computer characteristics that have generally been recognized as important for identifying technical change in computers are memory and speed (Triplett, 1989; Dulberger, 1989). Historically, users have been able to purchase a system with essentially any amount of memory they desired. The hedonic pricing literature on mainframe computers has implicitly made this assumption and some hedonic papers have even used computer system prices composed from unit prices of memory (*e.g.*, Gordon, 1990). In addition, Phister (1979), the primary source of computer data from the earlier era listed prices per unit of memory. Thus, memory is considered a continuous variable.

While it may be assumed that memory is a continuous variable, it would be incorrect to assume that speed is a continuous variable. Computer speed

* There was limited exit from the market by certain types of users. Scientific and engineering users had traditionally been the first to take advantage of faster computing speeds and larger memories, but this was beginning to be less true by the late 1980s. See Bresnahan and Greenstein (1997).

over the time period that this paper investigates was available in many discrete choices. However, over this time period there were many different speed choices available and manufacturers were increasingly providing more configurations of their basic systems. Thus, while not every speed was available for purchase, much choice of speed did exist.

This mixing of continuous and discrete variables into a single model implies an important choice. One approach would be to employ a discrete approach similar to Trajtenberg (1990) in his study of computed tomography scanners. Alternatively, one could employ a hedonic approach following Rosen (1974) which assumes continuity of the product space. Because memory is a continuous variable and because a significant number of speed choices were available, this paper chooses to employ a Rosen (1974) hedonic model to this industry. Future research should return to this data and investigate the robustness of these results using a discrete model approach.

This paper employs an approach for estimating demand for computer capacity based on Rosen (1974). This methodology has the advantage that it directly measures the demand for more speed and memory capacity. This method's econometric strengths and weaknesses are well known, which was extremely helpful in implementing the method.* Finally, it is well suited for measuring the effects of the distribution of demand across the product space, which is an important issue in this market. The next section reviews Rosen (1974) and discusses the issues surrounding his proposed methodology.

3.1 Rosen (1974) Revisited

Rosen (1974) suggested a model and methodology for examining demand and supply in hedonic pricing models. Rosen posited that a characteristic/price surface represents a locus of equilibrium transactions between buyers and sellers. This surface represents an upper envelope of buyers' bid functions and sellers' offer functions. Transactions occur where these bid and offer curves are tangent. As a consequence, the marginal price function represents the locus of intersections between buyers' marginal bid

* For example, we were able to learn from the experiences of Rosen (1974), Brown and Rosen (1982), Diamond and Smith (1985), Bartik (1987), and Epple (1987), and avoid some subtle econometric pitfalls.

and sellers' marginal offer curves. Rosen suggested that the following system of equations described the market:

$$p_i(X) = F^j(x_1, \ldots, x_n; Y_1) \quad \text{Demand} \tag{1}$$

$$p_i(X) = G^j(x_1, \ldots, x_n; Y_2) \quad \text{Supply} \tag{2}$$

Here p_i represents the marginal price for characteristic i, defined as the first derivative of the hedonic price surface at the observed x_i, and x_i, $i = 1, \ldots, n$ represents the observed levels of product characteristics. In our case, these product characteristics are computer speed and memory. Because of the potential nonlinearity of the hedonic price surface, buyers and sellers simultaneously choose both the levels of characteristics and the marginal prices for those characteristics, given by the slope of the hedonic price surface. This implies that the system given by (1) and (2) has $2n$ equations and $2n$ endogenous variables. Y_1 and Y_2 represent exogenous demand and supply shift variables that are characteristics of the individual buyers and sellers in the market.

Rosen suggested that this system of equations represented a "garden variety" econometric problem with a straightforward solution. He suggested a two-step estimation procedure. First, he defined the hedonic price function as

$$P(X) = f(x_1, \ldots, x_n) \tag{3}$$

where P is the product price, $x_1, \ldots, x_n$ are the product characteristics, and f is a function relating the two. The estimation procedure begins by estimating (3) using product price/characteristic combinations observed in the market. Having done this, one differentiates (3) with respect to each x_i and evaluates the derivatives at the observed x_i values. These $p_i(X)$ values represent estimates of the marginal prices for each characteristic i. These estimates are then used on the left-hand side of (1) and (2) in the second step of the estimation procedure where (1) and (2) are estimated using some simultaneous equation estimation procedure.

Rosen points out that a necessary condition for the estimation of (1)-(3) is that $P(X)$ in (3) be nonlinear. If it were not, then the estimates of $p_i(x)$ would be constant and the second-stage estimation of (1) and (2) would be impossible. Although this condition is correct, Brown and Rosen (1982) point out that it is not actually strong enough. Brown and Rosen provided an example similar to the following to illustrate their point.

Assume that $P(X)$ in (3) is second order in the x_i's and that (1) and (2) are linear in the x_i's. Upon differentiating (3) to get the estimates of $\hat{p}_i(X)$ which are obtained by evaluating the derivatives at the observed characteristic levels, it is seen that the $\hat{p}_i(X)$ are linear combinations of the x_i. This implies that in the second-stage estimation, when the $\hat{p}_i(X)$ are regressed against the observed characteristics nothing will be estimated at all. The equations will be identities formed by the initial choice of functional form. In other words, no new information is conveyed from the second stage of the procedure and exact solutions can be obtained with no further estimation. One way to rectify this situation is to require that the orders of (1) and (2) be at least two less than the order of (3).*

In addition to Brown and Rosen's comment on the appropriate orders of (1)-(3), they also offer another suggestion for estimation. Take (1)-(3) as before with no restrictions on order (except nonlinearity of (3)), but now assume that the data includes data from r distinct markets.† Here one would estimate (3) for each of the r markets separately and differentiate each to get the estimates of $\hat{p}_i(X)$ If one imposes the restriction that the demand and supply parameters are identical across markets even though the hedonic functions vary across markets, then identification could be reached assuming sufficient variation across markets.

Diamond and Smith (1985) also discuss issues surrounding Rosen's proposed methodology and, as with Brown and Rosen (1982), argue that only strong assumptions on functional form or data from multiple markets will allow for identification. However, it is required that a minimum of n+1 different markets be observed. Since our data contains eight years of computer acquisitions and since we will only look at a small number of computer characteristics this restriction should not pose problems.

In addition, Diamond and Smith (1985) argue that either side of the market can be estimated without regard to the other side. They argued this was reasonable since the source of simultaneity in this model does not arise between a buyer and a particular product. Rather, the movements by buyers are to new systems rather than along the offer curve of the same sys-

* Brown and Rosen (1982) only consider polynomials for (3) since "any nonlinear function can be represented arbitrarily closely by a polynomial of some order."

† Brown and Rosen (1982) examine spatially distinct markets. One could equivalently examine markets distinguished by time.

tem. The implication is that one only needs data on buyers and can estimate demand by assuming that the supply side of the market is exogenous.

This assumption is crucial for estimating this model since the data in this paper does not include any supply-side information. Since it is possible that certain firms, such as IBM, may have market power, it may seem unreasonable to ignore the supply side of the market.

However, Brown and Greenstein (1994), using the same data, estimated demand twice, once using only IBM and plug-compatible systems and once using all systems. Their analysis discovered no significant differences in the two model estimates. Several things may have produced this result. First, while it may be the case that a few firms have market power, it is not immediately clear that this market power was reflected in hardware prices, which we observe. Instead, this may be reflected in service contracts or software, for example. Second, the change in prices over time is dramatic, essentially driving the main features of the results in this paper. Even if market power matters in any given year, it is overwhelmed by technical change over time. As with the choice of model, better data should be used to investigate the robustness of the results to alternative assumptions about the supply side.

3.2 Estimating Demand for Computer Characteristics

Following the suggestions of Brown and Rosen (1982) and Diamond and Smith (1985) we make these initial assumptions. First, we assume that each buyer purchases only one mainframe computer system in any given time period. This is not implausible for the mainframe market since most "sites" that we observe make only one acquisition in a year, if any. Second, we assume that the demand parameters are the same for all buyers and do not change over time. Each buyer's demand is differentiated only by a set of buyer characteristics used to describe the heterogeneity among buyers. While it may be possible to relax this assumption somewhat in practice, such a relaxation unnecessarily complicates the econometrics and adds little to our main point. Third, we assume that supply is exogenous to buyers so that demand may be estimated without estimating supply. This implies that buyers take the hedonic price function as given and simply locate themselves on it. To estimate demand for computer characteristics, we begin by estimating a hedonic price function for each year that is exogenous to each buyer. These functions take the form

$$\begin{aligned} PRICE_{jt} = {} & \beta_{0t} + \beta_{1t} MIN.MEM_{jt} \\ & + \beta_{2t} MAX.MEM_{jt} + \beta_{3t} MIPS_{jt} \\ & + \beta_{4t} MIN.MEM^2_{jt} + \beta_{5t} MAX.MEM^2_{jt} \\ & + \beta_{6t} MIPS^2_{jt} + \varepsilon_{jt}. \end{aligned} \tag{4}$$

$PRICE_{jt}$ represents the price of system j at time t, and *MIN.MEM, MAX.MEM*, and *MIPS* represent the minimum memory, maximum memory, and MIPS for the particular system. These variables will be defined more carefully in the next section.

Once (4) has been estimated for each year of data, it is differentiated with respect to each of the computer characteristics to obtain three marginal price functions for each year. Next, we compute $\hat{p}_{it}$ where i indexes the three computer characteristics, by evaluating each of the marginal price functions at the observed characteristic levels. This yields three vectors of estimated marginal prices for each year, one for each computer characteristic. Finally, we combine the estimated marginal prices for each characteristic across time to create a single vector of estimated marginal prices for each characteristic. This vector is created by stacking the $\hat{p}_{it}$ $t = 1 \ldots T$, into a single vector containing estimated marginal prices for each computer system across all time periods.

As an example, once (4) is estimated, we differentiate it with respect to *MIN.MEM* which yields

$$\frac{\partial PRICE}{\partial MIN.MEM} = \hat{\beta}_{1t} + 2\hat{\beta}_{4t} MIN.MEN_{jt}. \tag{5}$$

Next, $\hat{p}_{it}$ which is just the left-hand side of (5), is estimated by evaluating the right-hand side of (5) at the observed *MIN.MEM* values. This yields T vectors of marginal prices for *MIN.MEM*, one for each year. These T vectors are then combined into a single vector by stacking the $\hat{p}_{it}$ for each time period on top of each other. This vector, which is denoted $\hat{p}_i$ represents the marginal prices for *MIN.MEM* for all systems purchased during the T time periods. This process is then repeated for *MAX.MEM* and *MIPS* to arrive at three marginal price vectors.

Once the marginal price vectors have been created, we estimated demand for characteristic i by estimating

$$\hat{p}_i = \alpha_0 + \alpha_1 x_i + \alpha_2 Y_{1i} + \eta_i \tag{6}$$

where x_i represents observed values of characteristic i and Y_1 represents individual buyer characteristics as in (1) above. Epple (1987) and Bartik (1987) have shown that x_i is correlated with η_i in (6). Therefore, (6) is estimated using an instrumental variable approach.

The instruments, as suggested by Epple (1987), must be correlated with the observed choice of computer characteristics but uncorrelated with unobserved tastes. In the next section we will discuss the instruments used.

4 DATA

This study's data is at the buyer level and describes the mainframe acquisitions as well as the buyers making the acquisitions.* Previous research on computers has had data on the available systems, their characteristics, and their prices but not on the characteristics of the individual buyers. Because of our superior data, we are able to go beyond estimating the hedonic price function and are able to estimate demand for individual computer characteristics.

The data used in this analysis is a subset of the Computer Installation Data File kept by the Computer Intelligence Corporation (CIC). It contains their complete records for many sites in the United States with a medium to large general purpose computer system from 1984 through 1991. Each year provides data on over 14,000 sites. Each of those 14,000 records includes the name and address of the private company (and parent) at which the system is located, as well as broad information about the company, such as the (four digit) SIC associated with the site, the number of employees, and the amount of revenue.†

Since our interest here is in the demand for new technology, we look only at the acquisitions of new computer systems. This is similar to previous research where researchers typically use the set of systems available for sale in any given time period to perform their analysis. The acquisi-

* The buyers are private firms, educational institutions, and government organizations.

† Other information includes the system name and model, the amount of memory, the amount of peripheral equipment used, the primary language used, the likely market value of the system, the method by which the system was acquired, and at what level such acquisition decisions are made. The file also provides information about the total MIPS and DASD, as well as the number of programmers at the site. Finally, the file provides a multitude of software variable relating to the use of the system at the site. Unfortunately, some of the variables are not reported for the entire time period, making it a difficult task to use all fields.

tions data was generated from a variable for each system indicating whether it was a new acquisition or not.* In order to perform our analysis we require data on the site from the previous year. Therefore, all observations for which there was no site information for the previous year were removed.

To complete the data set, we consulted CIC's Computer Systems Report Users Guide, which contains information on all systems known to them. The Guide contains the system name and a list of characteristics including minimum memory, maximum memory, and MIPS among others. We matched the system names with those in the Guide and merged the characteristics with the list of acquisitions. Due to name discrepancies and the inclusion of some non-mainframe acquisitions in the data set, the size of the data was narrowed to 21,268 acquisitions at unique sites in unique years. This is the data set used in this analysis. In the remainder of this section we give definitions for the different variables used. Descriptive statistics for the variables are given in the Appendix at the end of the paper.

4.1 Price

The system price used in this paper is provided by CIC and is defined as the "estimated value of a `typical' configuration if purchased today." The drawback to this is that all acquisitions of the same system will get the same price associated with them regardless of the true configuration which was purchased. "Typical" is defined by CIC as "an average size system with a normal complement of peripherals and terminals." This is the same type of price that previous work has used. The computer characteristics to be described later are associated with the systems in the same manner, implying that all systems of the same type have the same price and the same characteristics associated with them during a given year. Fortunately, prices for the same system change over time so that there is variation in both the cross-section and time-series. While this is not the most desirable setup, it is consistent with the type of data that has been used in the previous literature.

* To be sure we check two consecutive years of data. Many sites are surveyed on an annual basis, but often in the middle of the year. Hence, any acquisitions late in the year are not recorded in the end-of-year sample if the site has not recently been surveyed. This sampling frame unavoidably produced a smaller number of acquisitions in the data set for 1991.

The price data was adjusted for inflation using the Producer Price Index. This was done so that any shifts in the hedonic function over time would be the results of technological change and not inflation.

4.2 Computer Characteristics

The mainframe characteristics used in this paper are minimum memory (*MIN.MEM*), maximum memory (*MAX.MEM*), and MIPS (*MIPS*). Minimum and maximum memory are the minimum and maximum amounts of main storage supported on the system. This memory is the technical ancestor of the modern notion of RAM, *i.e.*, memory used during CPU processing, found in personal computers today. MIPS is a measure of the speed of the mainframe measured in millions of instructions per second.

Most previous studies have often used both the minimum and maximum memory. This was done either to account for the lack of information as to which size of memory went with the recorded price or to avoid the influence of different pricing schemes for the low-end models when different prices were available for different memory sizes (Phister, 1979; Triplett, 1989; Dulberger, 1989). We use both minimum and maximum memory because we do not possess a price for different memory configurations for each system. The error introduced with this procedure tends to wash out over time with the significant technical change that takes place.

MIPS is the best measure of speed for our purposes, though considerable previous debate influenced this choice (Triplett, 1989; Dulberger, 1989). No less than five measures of speed, including addition time, multiplication time, memory cycle time, MIPS, and KOPS (thousands of instructions per second) have been introduced as independent variables in the specification of the hedonic function. For definitions of these measures and others see Triplett (1989). The most recent studies prefer to use MIPS because it combines the speeds of many instructions and weights each instruction by the relative frequency of that instruction in the job. However, since typical jobs vary across processors, comparability across processors is difficult (Triplett, 1989; Dulberger, 1989). For this reason, Dulberger (1989) chose to exclude all data except IBM and plug-compatible processors for which equivalent MIPS measures were available. However, Triplett (1989) argued that this need not be done since on average the non-plug-compatible systems will be comparable with the plug-compatible systems and that is all that matters for econometric purposes. We follow Triplett's suggestion and maintain all observations that we have available.

4.3 Buyer Characteristics

The buyer characteristics, Y_1 from (6), are the variables that describe the heterogeneity among buyers. We chose a number of variables which were either available or could be generated from the CIC data. Each of the variables is lagged one period similar to the procedure followed by Bresnahan and Greenstein (1997). These variables include dummy variables for various SIC groupings, a dummy variable for whether or not the site owned an IBM system, the estimated purchase value of installed systems at the site, the highest MIPS rating of any system at the site, the total MIPS for all installed systems at the site, and the technical age of the youngest system owned during the previous year. In addition, we included a number of software related variables in order to account for how the systems were being used. All of the buyer characteristics are defined in Table I.

TABLE I Buyer Characteristics Definitions

Variable	*Definition*
MIPS.MAX	The number of MIPS on the system with the largest MIPS rating at the site.
MIPS.TOT	The sum total MIPS of all systems at the site.
AGE.YOUNG	The technical age (year of observation minus the vintage of the system) of the youngest system (most recent vintage) at the site.
SITEVAL	The estimated purchase value of the site in 1000s of dollars. This amount is adjusted by the Producer Price Index.
IBMSITE	A dummy variable which takes the value 1 if the site owned any medium to large IBM system and 0 otherwise.
GDPGRW	The annual GDP growth rate for the site's industry classification during the previous year.
SCITS	Percentage of scientific and technical software used at the site.
STD	Percentage of standard business applications software used at the site.
DB	Percentage of database and applications tools software used at the site.
COMM2	Percentage of communications and networking software used at the site.
PROP	Percentage of software applications written by the hardware vendor.
THIRDPAR	Percentage of software applications written by third-party vendors.

Variable	*Definition*
MPLAT	Percentage of software applications written to run on multiple main- frame platforms.
INHOU	Percentage of software applications written by employees of the site.
*SIC*1	Mining and construction
*SIC*2	Oil and gas extraction
*SIC*3	Manufacturing
*SIC*4	Food products
*SIC*5	Paper and allied products
*SIC*6	Printing and publishing
*SIC*7	Chemicals
*SIC*8	Miscellaneous manufacturing
*SIC*9	Primary metals
*SIC*10	Fabricated metals
*SIC*11	Computers and related
*SIC*12	Electrical apparatus
*SIC*13	Motor vehicles and equipment
*SIC*14	Other transportation equipment
*SIC*15	Instruments
*SIC*16	Transportation services
*SIC*17	Telephone communication
*SIC*18	Other communication services
*SIC*19	Electric service
*SIC*20	Gas and sanitary service
*SIC*21	Wholesale trade
*SIC*22	Miscellaneous trade
*SIC*23	General merchandise stores
*SIC*24	Food stores
*SIC*25	Depository institutions
*SIC*26	Nondepository institutions
*SIC*27	Securing and commodity brokers
*SIC*28	Insurance carriers
*SIC*29	Insurance brokers and real estate brokers
*SIC*30	Hotel, personal services
*SIC*31	Business services
*SIC*32	Health services
*SIC*33	Legal, social, engineering services, museums
*SIC*34	Education services

4.4 Instrumental Variables

Because in general the computer characteristics in the demand estimation are correlated with the error term (Bartik, 1987; Epple, 1987), we need to estimate equation (6) by instrumental variables. The instruments need to be variables that affect the choice of characteristics but do not affect unobserved tastes. The variables we use are time dummy variables, region dummy variables, an SMSA dummy variable and characteristics of the closest systems in the product space as measured by the Mahalanobis distance between systems. This distance is defined as

$$(X_0 - X_i)'\Sigma^{-1}(X_0 - X_i)$$

where X_0 represents the characteristics of the system in question, X_i represents the characteristics of all systems except X_0, and Σ represents the covariance matrix of the variables minimum memory, maximum memory, and MIPS. The idea is that neighboring systems provide information about the shifts in the costs of production without correlating much with buyer demand. A related idea may be found in Berry (1993).

5 RESULTS

Figure 1 presents a boxplot of the variable MIPS over the 1985–1991 sample period. The boxplot shows the distribution of new mainframe purchases each year. The shaded region represents the interquartile range for the distribution, and the bar drawn in the interior of the shaded box is the median. The brackets, or *whiskers*, at the bottom and at the top of the dotted line represent the innermost 99% of the data, and any lines outside the whiskers represent potential outliers.

Figure 1 illustrates a few important points. First, it can be seen from the whiskers that the product space is expanding from a maximum of near 30 MIPS in 1985 to over 100 MIPS by 1991. Second, it can be seen that, although the product space is expanding, the majority of buyers are not moving to the cutting edge technology very quickly. The median MIPS purchase changes little from year to year, going from a low of 2.7 in 1985 to a high of 22.5 in 1991. While the median is moving out much faster than the product space, the median purchase is relatively small over the entire time period.

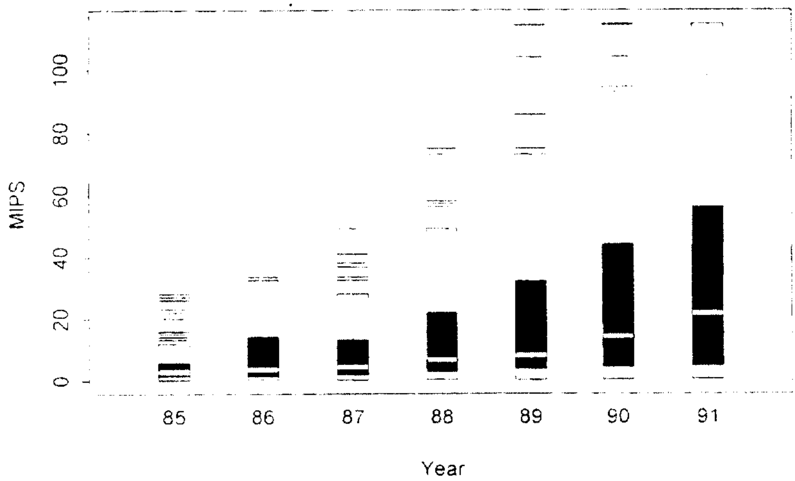

FIGURE 1 Boxplot of MIPS: 1985–1991

This second point is important for our purposes. Recall that our goal is to identify the buyer characteristics that predict demand for new capacity and extension of the product space. Since few sites are moving to the largest MIPS systems, identifying these buyer characteristics with any amount of certainty will be difficult. However, since some sites are pushing the technological frontier, we will be able to identify those site characteristics that increase the likelihood of demanding new capacity. Still, it is unlikely that this data will reveal any specific site types that always demand new technology.

Using the data described in the previous section and the estimation procedure outlined in Section 3, we estimated demand for each of the computer characteristics: *MIN.MEM*, *MAX.MEM*, and *MIPS*. These demand estimates are given in Tables II, III, and IV.

Table II gives the demand estimates for minimum memory. Not surprisingly the variable of interest, *MIN.MEM*, is not significant. One would expect that when a site purchases a system they are most interested in the maximum speed of the computer, as measured by MIPS, and the maximum memory on the system. Thus, while maximum memory may limit a user, minimum memory does not place any real restrictions on the buyer and therefore we would expect the demand for that characteristic to be insignificant.

TABLE II Minimum Memory Demand Estimate

Variable	*Coefficient*	*Std. Error*	*t-value*	*p-value*
CONSTANT	−13.79	46.66	−0.30	0.77
MIN.MEM	0.02	0.09	0.25	0.80
MIPS.MAX	2.18	0.25	8.80	0.00
MIPS.TOT	−0.14	0.10	−1.44	0.15
AGE.YOUNG	2.48	1.00	2.48	0.01
SITEVAL	0.00	0.00	−3.16	0.00
IBMSITE	−53.56	6.25	−8.56	0.00
GDPGRW	−2.26	0.43	−5.29	0.00
SCITS	−113.75	41.47	−2.74	0.01
STD	−105.91	23.44	−4.52	0.00
DB	−59.12	25.60	−2.31	0.02
COMM2	−115.09	22.86	−5.04	0.00
PROP	69.68	43.22	1.61	0.11
THIRDPAR	85.66	42.50	2.02	0.04
MPLAT	64.77	43.11	1.50	0.13
INHOU	−79.66	21.11	−3.77	0.00
SIC1	18.34	36.39	0.50	0.61
SIC2	32.36	27.16	1.19	0.23
SIC3	12.04	20.33	0.59	0.55
SIC4	−1.71	23.06	−0.07	0.94
SIC5	6.50	27.16	0.24	0.81
SIC6	−6.50	21.71	−0.30	0.76
SIC7	42.87	21.62	1.98	0.05
SIC8	−12.92	23.51	−0.55	0.58
SIC9	1.97	28.86	0.07	0.95
SIC10	−0.69	26.78	−0.03	0.98
SIC11	22.61	17.92	1.26	0.21
SIC12	28.27	23.53	1.20	0.23
SIC13	60.85	25.22	2.41	0.02
SIC14	−12.51	24.71	−0.51	0.61
SIC15	−13.91	26.51	−0.52	0.60
SIC16	−13.16	20.56	−0.64	0.52
SIC17	14.89	20.89	0.71	0.48

Variable	*Coefficient*	*Std. Error*	*t-value*	*p-value*
SIC18	67.43	34.50	1.95	0.05
SIC19	8.67	24.40	0.36	0.72
SIC20	16.33	23.78	0.69	0.49
SIC21	9.81	18.83	0.52	0.60
SIC22	2.89	22.74	0.13	0.90
SIC23	−4.89	25.49	−0.19	0.85
SIC24	3.94	28.32	0.14	0.89
SIC25	−0.92	17.91	−0.05	0.96
SIC26	18.87	22.45	0.84	0.40
SIC27	19.37	25.69	0.75	0.45
SIC28	−13.49	17.68	−0.76	0.45
SIC29	2.70	24.34	0.11	0.91
SIC30	−11.19	32.93	−0.34	0.73
SIC31	5.13	16.86	0.30	0.76
SIC32	−4.71	19.16	−0.25	0.81
SIC34	−12.35	18.90	−0.65	0.51

TABLE III Maximum Memory Demand Estimate

Variable	*Coefficient*	*Std. Error*	*t-value*	*p-value*
CONSTANT	55.05	5.54	9.95	0.00
MAX.MEM	−0.01	0.00	−13.18	0.00
MIPS.MAX	−0.34	0.03	−11.41	0.00
MIPS.TOT	0.00	0.01	−0.06	0.95
AGE.YOUNG	−1.00	0.12	−8.38	0.00
SITEVAL	0.00	0.00	8.16	0.00
IBMSITE	8.33	0.74	11.22	0.00
GDPGRW	0.03	0.05	0.61	0.54
SCITS	19.24	4.92	3.91	0.00
STD	26.12	2.78	9.41	0.00
DB	19.76	3.03	6.51	0.00
COMM2	9.36	2.70	3.46	0.00
PROP	−49.32	5.12	−9.63	0.00
THIRDPAR	−46.23	5.03	−9.19	0.00
MPLAT	−56.20	5.11	−11.01	0.00

Variable	*Coefficient*	*Std. Error*	*t-value*	*p-value*
INHOU	3.39	2.50	1.35	0.18
SIC1	−5.60	4.32	−1.30	0.19
SIC2	3.19	3.22	0.99	0.32
SIC3	−2.23	2.41	−0.92	0.36
SIC4	−3.66	2.74	−1.34	0.18
SIC5	−3.10	3.22	−0.96	0.34
SIC6	−3.48	2.58	−1.35	0.18
SIC7	−3.85	2.56	−1.50	0.13
SIC8	−3.39	2.79	−1.21	0.22
SIC9	−1.90	3.42	−0.56	0.58
SIC10	−2.65	3.18	−0.83	0.40
SIC11	−0.35	2.13	−0.17	0.87
SIC12	1.17	2.79	0.42	0.68
SIC13	6.01	2.99	2.01	0.04
SIC14	−0.15	2.93	−0.05	0.96
SIC15	0.33	3.15	0.11	0.92
SIC16	1.13	2.44	0.46	0.64
SIC17	0.68	2.48	0.28	0.78
SIC18	−0.47	4.09	−0.11	0.91
SIC19	2.41	2.89	0.83	0.41
SIC20	5.51	2.82	1.95	0.05
SIC21	−1.13	2.23	−0.51	0.61
SIC22	0.44	2.70	0.16	0.87
SIC23	3.28	3.02	1.09	0.28
SIC24	−2.20	3.36	−0.66	0.51
SIC25	0.35	2.12	0.17	0.87
SIC26	0.75	2.66	0.28	0.78
SIC27	5.62	3.05	1.84	0.07
SIC28	3.98	2.10	1.90	0.06
SIC29	6.71	2.89	2.33	0.02
SIC30	−2.07	3.91	−0.53	0.60
SIC31	−1.46	2.00	−0.73	0.46
SIC32	−2.12	2.27	−0.93	0.35
SIC34	1.48	2.24	0.66	0.51

TABLE IV MIPS Demand Estimate

Variable	*Coefficient*	*Std. Error*	*t-value*	*p-value*
CONSTANT	−165.23	59.96	−2.76	0.01
MIPS	−0.95	0.20	−4.83	0.00
MIPS.MAX	−0.92	0.32	−2.84	0.00
MIPS.TOT	0.31	0.13	2.45	0.01
AGE.YOUNG	4.39	1.29	3.40	0.00
SITEVAL	0.00	0.00	−2.68	0.01
IBMSITE	60.08	8.04	7.47	0.00
GDPGRW	2.17	0.55	3.96	0.00
SCITS	−1.84	53.29	−0.03	0.97
STD	53.76	30.15	1.78	0.07
DB	−1.82	32.92	−0.06	0.96
COMM2	69.94	29.42	2.38	0.02
PROP	57.11	55.51	1.03	0.30
THIRDPAR	−30.96	54.59	−0.57	0.57
MPLAT	84.43	55.36	1.53	0.13
INHOU	127.29	27.13	4.69	0.00
SIC1	11.19	46.77	0.24	0.81
SIC2	−56.57	34.91	−1.62	0.11
SIC3	14.90	26.13	0.57	0.57
SIC4	36.55	29.64	1.23	0.22
SIC5	52.86	34.89	1.52	0.13
SIC6	45.93	27.91	1.65	0.10
SIC7	−1.43	27.79	−0.05	0.96
SIC8	68.61	30.22	2.27	0.02
SIC9	7.60	37.09	0.20	0.84
SIC10	49.65	34.42	1.44	0.15
SIC11	−6.68	23.03	−0.29	0.77
SIC12	−34.51	30.24	−1.14	0.25
SIC13	−97.91	32.41	−3.02	0.00
SIC14	−71.14	31.76	−2.24	0.03
SIC15	56.38	34.08	1.65	0.10
SIC16	−19.15	26.42	−0.72	0.47
SIC17	−18.91	26.85	−0.70	0.48

Variable	*Coefficient*	*Std. Error*	*t-value*	*p-value*
SIC18	−100.86	44.34	−2.27	0.02
SIC19	−43.19	31.35	−1.38	0.17
SIC20	−41.63	30.56	−1.36	0.17
SIC21	17.62	24.20	0.73	0.47
SIC22	3.17	29.23	0.11	0.91
SIC23	0.77	32.76	0.02	0.98
SIC24	45.93	36.39	1.26	0.21
SIC25	5.86	23.02	0.25	0.80
SIC26	16.62	28.85	0.58	0.56
SIC27	−37.19	33.02	−1.13	0.26
SIC28	7.51	22.74	0.33	0.74
SIC29	−6.20	31.28	−0.20	0.84
SIC30	65.35	42.33	1.54	0.12
SIC31	11.97	21.67	0.55	0.58
SIC32	46.30	24.63	1.88	0.06
SIC34	41.50	24.29	1.71	0.09

Tables III and IV give the demand estimates for maximum memory and MIPS. As expected, the demand for maximum memory is downward sloping and significant, as is the demand for MIPS. In addition, the site characteristics related to the existing mainframe base, such as *MIPS.MAX*, *IBMSITE*, and others tend to also be highly significant. For example, *MIPS.MAX* is negative and significant in both demand estimates.

MIPS.MAX represents the MIPS rating for the system at a site with the highest rating. The negative demand coefficient implies that buyers are willing to pay less for additional MIPS the larger is their largest system. Intuitively, if a site has recently purchased a fast system and increased their maximum feasible task to meet their needs and have closed the gap between their existing technology and the frontier technology, it is likely that the site will wait to purchase additional computer capacity.

MIPS.TOT represents the total MIPS rating of the existing systems at a specific site. Not surprisingly, that variable is insignificant in the maximum memory demand estimate and positive and significant in the MIPS demand estimate. The positive coefficient in the MIPS demand estimate implies that those sites that tend to require large amounts of MIPS, as evidenced by their existing mainframe base, tend to demand more MIPS each year.

AGE.YOUNG, another variable included to control for the existing mainframe base, represents the age of the youngest system at a particular site and is intended to control for the age of the existing mainframe base. We would expect that newer systems would embody newer technology so that the older the youngest system is, the more likely a site will need to upgrade to a new system. This would imply a positive coefficient in the demand estimates. In the MIPS demand estimate *AGE.YOUNG* is positive as expected, but the coefficient on *AGE.YOUNG* is negative in the maximum memory demand estimate. We believe this result points out an important dynamic element we are not able to capture with our model. Intuitively, as more time passes without making a new mainframe purchase, the farther the existing mainframe base is from the technological frontier, and the more likely the site will upgrade to a newer technology system. However, there is another effect associated with the frequency of purchase. As the youngest system gets older, the site may be nearing closure. In such a case, investments may be more cautious and they may become less likely to purchase a bigger technology system. Alternatively, as the youngest system gets older, the site may be migrating away from mainframes to workstations or personal computers. We believe the different signs on *AGE.YOUNG* in the two demand estimates are a result of this type of dynamic phenomenon that our model is unable to capture.

The other three variables that control for the existing mainframe base are *SITEVAL*, *IBMSITE*, and *GDPGRW*. *SITEVAL* appears to be unimportant in both demand estimates. *IBMSITE* is positive and significant in both estimates. Recall that *IBMSITE* is a dummy variable indicating whether the site had an IBM system in their existing mainframe base. Finally, *GDPGRW* is positive and significant in both demand estimates. This implies that the better the site's industry is performing, the more likely the site will upgrade to a newer technology system.

While we have controlled for the existing mainframe base when estimating demand, our main area of interest lies in identifying the types of buyers that are pushing out the technological frontier. These sites are best identified by looking at the software and SIC variables in the demand estimates.

In order to better compare the software variables and identify those buyers pushing the technical frontier we create a standardized coefficient. The standardized coefficient is created by multiplying the coefficient in the demand estimate by the standard deviation of the variable. For example, the coefficient on *SCITS* in the maximum memory demand estimate is 19.24 and the standard deviation of *SCITS* is 0.067 giving a standardized coefficient of 1.283. Table V provides standardized coefficients for the software variables computed this way for both maximum memory and MIPS demand. These values represent the change in the willingness to pay for the characteristic per one standard deviation change in the variable. Examining the regression results in this way allows us to gain insight into which variables are more or less important.

TABLE V Standardized Coefficients

Variable	*Standard Deviation*	*Maximum Memory Coefficient*	*Max. Mem. Standardized Coefficient*	*MIPS Coefficient*	*MIPS Standardized Coefficient*
SCITS	0.067	19.24	1.28	−1.84	−0.12
STD	0.166	26.12	4.33	53.76	8.91
DB	0.110	19.76	2.18	−1.82	−0.20
COMM2	0.125	9.36	1.17	69.94	8.74
PROP	0.219	−49.32	−10.79	57.11	12.49
THIRDPAR	0.205	−46.23	−9.46	−30.96	−6.34
MPLAT	0.128	−56.20	−7.17	84.43	10.78
INHOU	0.155	3.39	0.53	127.29	19.76

Interestingly, Table V indicates that the demand for memory and the demand for speed do not have much in common. The signs on the variables tend to be of differing signs in the maximum memory and MIPS estimates and the magnitudes show that some variables are more important for one characteristic than the other. As an example, *INHOU* has the smallest impact on demand for maximum memory with a standardized coefficient of 0.53 while the same variable has the largest impact on the demand for MIPS with a standardized coefficient of 19.76.

Consider the demand for MIPS. The coefficients on *PROP*, *THIRDPAR*, *MPLAT*, and *INHOU* indicate that only *THIRDPAR* and the excluded category (consultants) do not affect the demand for MIPS much. The more deliberately a site invests in one type of supplier of software, whether it is exclusively IBM (*PROP*), or multiple-platform firms (*e.g.*, Oracle), or its own crew (*INHOU*), the more likely the site is large. In our data roughly 41% of the sites get the majority of their software from a proprietary vendor or in-house, while roughly 1% comes from a multiple-platform vendor. Thus, a main driver of demand is specific assets either from a relationship with a big supplier or from one's own development team.

The other software variables send us a different message with regard to demand for MIPS. The small standardized coefficient on *SCITS* probably indicates that scientific sites were leaving mainframes for workstations in this era. The standardized coefficients on *STD* and *COMM*2 indicate that standard database tools and communication software were the biggest drivers in the demand for speed as opposed to tools for database development (*DB*). These would be on-line transactions processing (OLTP) sites. Overall, though, the type of software application is less important than the relationship with the software vendor.

Now consider maximum memory. The coefficients on *PROP, THIRDPAR, MPLAT*, and *INHOU* indicate that *INHOU* and the excluded category (consultants) tend to have little impact on the demand for memory, and the other three indicate that the more heavily one invests in one type of software supplier the less memory one needs. Considering the software applications, we see that, relatively speaking, any software use leads to some demand for memory. Sites specializing in standard database applications will see the highest demand for memory.

Overall, Table V suggests that sites specializing in applications, particularly OLTP, or specializing in software from a particular vendor are the main drivers of demand for extreme features. While OLTP sites appear to be obvious demanders of high speed and memory over this time period, Table V does not suggest any additional types of sites that are pushing out the technological frontier.

In addition to examining the software variables, one could look at the SIC variables to get a handle on which industries are demanding the most advanced technology. Table VI provides a sorting of the SIC coefficients into quartiles in the specific product characteristics. For example, the coefficient on *SIC*34 in the MIPS demand estimate is 41.50 and the maximum memory demand estimate is 1.48. Ordering the SIC coefficients in both demand estimates reveals that these values fall into the top quartile for both estimates. Thus, we place *SIC*34 in the MIPS demand fourth quartile and the maximum memory demand fourth quartile cell in the table. We continued this exercise for all SIC groups to get the results in Table VI.

*SIC*34, education, pushes the technical frontier for both product characteristics. This makes intuitive sense since we typically think of universities being on the cutting edge of research. But more importantly, universities typically have a very large number of simultaneous users demanding computer resources. In order to accommodate the large number of users, universities must equip themselves with fast machines with large amounts of memory, much like the OLTP sites identified in the previous discussion.

TABLE VI SIC Coefficients Grouped by Quartiles Within the Maximum Memory and MIPS Demand Estimates (Each cell contains SIC groups)

	MAX_1	MAX_2	MAX_3	MAX_4
$MIPS_1$		*14, 18*	*12, 16*	*2, 13, 19, 20, 27*
$MIPS_2$	*7*	*11*	*17, 22, 25, 33*	*23, 29*
$MIPS_3$	*1, 3, 4*	*9, 21, 31*	*26*	*28*
$MIPS_4$	*5, 6, 8, 10, 24*	*30, 32*	*15*	*34*

In addition to education, *SIC*15, *SIC*26, and *SIC*28 are all at least in the third quartile in both MIPS and maximum memory demand. *SIC*15 is scientific instruments, *SIC*26 is nondepository institutions, and *SIC*28 is insurance carriers. It is not surprising that these last two SIC groups are high demanders of MIPS and memory. Both maintain large databases of clients and need fast access to account information. However, with *SIC*26 ranking this high, it is a bit surprising that *SIC*25, depository institutions, did not.

The grouping in Table VI indicates that sites tend to demand one characteristic more intensively than the other. This is seen by the large number of SIC groups that fall into the third or fourth quartile in one characteristic and the first or second quartile in the other characteristic. In all, 26 of the 34 SIC groups specified fall into this situation. Few industries are pushing both the speed and memory frontiers simultaneously.

The obvious exercise after examining the software and SIC variables is to go back into the data and to identify the sites, or types of sites, that demand extreme product characteristics and push out the technological frontier. What one discovers when performing this exercise, however, is that no single factor determines whether a site buys a very large system.

Indeed, as already noted in the boxplot, most buyers adopt smaller systems. Hence, the large system buyers tend to be those who have several features, such as a demand for OLTP, the use of specialized assets, and a particular industry background.

6 CONCLUSION

This paper has examined the demand for mainframe computer characteristics in the 1980s and attempted to identify the types of mainframe users that demand features. Using a data set covering mainframe acquisitions from 1984–1991, this paper has applied a Rosen (1974) hedonic model to identify demand for mainframe speed and memory.

In this era of technical change, the demand for large scale computing was quite uneven across users. Among the 21,268 acquisitions we study, most buyers initially bought a "small" mainframe system in the early 1980s and still bought a small system at the end of the decade, even with rapidly declining mainframe prices and large extensions in computing capacity. No single factor alone predicts purchases of more speed and memory. We conclude that a variety of factors induced users to buy something other than a small system. These estimates are consistent with relatively stable demand in this period, with industry norms about differences in the demand for large systems across industries, and with the importance of demand for on-line transactions processing applications in this time period.

Two critical assumptions were made to obtain the results in this paper. First, due to the continuity of the memory characteristic and the large number of speed choices available, we assumed that the product space was continuous. This allowed the estimation of a Rosen (1974) hedonic pricing model. Relaxing this assumption and exploiting the discrete speed variable would imply the application of a discrete choice model. Future research should investigate the robustness of the model choice in this paper by estimating a discrete choice model and comparing the qualitative results obtained with those in this paper.

Second, because the data used in this paper does not include any supply-side information, this paper followed Diamond and Smith's (1985) assumption that the supply side is exogenous. A complete model of this market should include the supply side to determine if market power was an important factor. Should this data become available, this model should be estimated again using this data to determine the impact of firm market power on the results.

Third, our results suggest that only a small number of traits identified lead users here. In this case, a few industries and OLTP software applications were key traits. These findings raise questions about how this pattern has changed as a consequence of the diffusion of client/server systems into large scale computing applications, even those using mainframes. As this technical change has been an important locus of commercial developments in the 1990s, closer scrutiny of lead users of large scale client/server systems deserves study.

Acknowledgements

We would like to thank Steve Berry, Tim Bresnahan, Jan Brueckner, Rob Feenstra, Zvi Griliches, Harumi Ito, Roger Koenker, Ariel Pakes, Pablo Spiller, Manuel Trajtenberg, Jack Triplett, Wesley Wilson, many seminar participants, and an anonymous referee for helpful comments and useful conversations. We would also like to thank Computer Intelligence Corporation for their help and Denise Chachere for her valuable research assistance. This work was supported by NSF under IRI92–09321 and by the Institute for Government and Policy Analysis at the University of Illinois. Greenstein also thanks the hospitality of CEPR and the Computer Industry Project at Stanford University, were he resided when portions of this paper were completed. All errors are our own.

References

Bartik, T. J. (1987). The estimation of demand parameters in hedonic price models. *Journal of Political Economy*, 95(1):81–88.

Berry, S., Levinsohn, J., and Pakes, A. (1993). Automobile prices in market equilibrium: Part I and II. Working Paper No. 4264, National Bureau of Economic Research, Cambridge, MA.

Bresnahan, T. and Greenstein, S. (1997). Technical progress and co-invention in computing and in the use of computers. *Brookings Papers on Economics Activity: Microeconomics*, pages 1–78.

Brown, J. N. and Rosen, H. S. (1982). On the estimation of structural hedonic price models. *Econometrica*, 50(3):765–768.

Brown, K. H. (1994). *Estimating the Economic Benefits of Advances in Computer Technology*. PhD Dissertation, University of Illinois, Champaign, Illinois.

Cole, R., Chen, Y. C., Barquin-Stolleman, Y. C., Dulberger, E., Helvacian, N., and Hodge, J. (1986). Quality adjusted price indexes for computer processors, selected peripheral equipment. *Survey of Current Business*, 66:41–50.

Diamond, Jr, D. B. and Smith, B. A. (1985). Simultaneity in the market for housing characteristics. *Journal of Urban Economics*, 17:280–292.

Dulberger, E. R. (1989). The application of a hedonic model to a quality-adjusted price index for computer processors. In Jorgenson, D. W. and Landau, R., editors, *Technology and Capital Formation*. The MIT Press, Cambridge, MA.

Epple, D. (1987). Hedonic prices and implicit markets: Estimating demand and supply functions for differentiated products. *Journal of Political Economy*, 95(1):59–80.

Friedman, A. L. and Cornford, D. S. (1989). *Computer Systems Development: History, Organization and Implementation*. John Wiley and Sons, New York, NY.

Gordon, R. J. (1990). *The Measurement of Durable Goods Prices*. University of Chicago Press, Chicago, IL.

Greenstein, S. M. (1993). Did installed base give an incumbent any (measurable) advantage in federal computer procurement? *Rand Journal of Economics*, 24(1):19–39.

Greenstein, S. M. (1997). From super-minis to super computers: Estimating surplus in the computing market. In Bresnahan, T. and Gordon, R. J., editors, *The Economics of New Goods*. University of Chicago Press, Chicago, IL.

Inmon, W. H. (1985). *Technomics: The Economics of Technology and the Computer Industry*. Dow Jones-Irwin, Chicago, IL.

Mendelson, H. (1985). Pricing computer services: Queuing effects. *Communications of the ACM*, 28(3):312–321.

Phister, Jr, M. (1979). *Data Processing Technology and Economics*. Santa Monica Publishing Company and Digital Press, Bedford, MA, 2nd edition.

Rogers, E. M. (1962). *Diffusion of Innovations*. Free Press, New York.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82:34–49.

Triplett, J. E. (1989). Price and technological change in a capital good: A survey of research on computers. In Jorgenson, D. W. and Landau, R., editors, *Technology and Capital Formation*. The MIT Press, Cambridge, MA.

von Hippel, E. (1988). *The Sources of Innovation*. Oxford University Press, New York.

APPENDIX

Descriptive Statistics for Model Variables

Variable	*Min*	*Mean*	*Median*	*Max*	*Std. Dev.*
PRICE	33.670	25840.000	10060.000	140000.000	31434.890
MIN.MEM	0.008	28.810	16.380	262.100	38.152
MAX.MEM	0.032	285.100	32.770	4719.000	664.884
MIPS	0.100	16.480	6.500	114.000	22.367
MIPS.MAX	0.000	12.540	4.500	114.400	18.839
MIPS.TOT	0.000	27.450	6.400	772.700	56.778
AGE.YOUNG	0.000	3.344	3.000	24.000	2.608
SITEVAL	0.000	2584.000	744.200	44790.000	4357.368
IBMSITE	0.000	0.790	1.000	1.000	0.408
GDPGRW	−20.000	3.728	3.470	65.400	6.039
SCITS	0.000	0.047	0.040	1.000	0.067
STD	0.000	0.204	0.190	1.000	0.166
DB	0.000	0.210	0.202	1.000	0.110
*COMM*2	0.000	0.254	0.241	1.000	0.125
PROP	0.000	0.500	0.455	1.000	0.219
THIRDPAR	0.000	0.424	0.467	1.000	0.205
MPLAT	0.000	0.086	0.061	1.000	0.128
INHOU	0.000	0.147	0.111	0.938	0.155
*SIC*1	0.000	0.005	0.000	1.000	0.073
*SIC*2	0.000	0.011	0.000	1.000	0.104
*SIC*3	0.000	0.031	0.000	1.000	0.173
*SIC*4	0.000	0.019	0.000	1.000	0.137
*SIC*5	0.000	0.012	0.000	1.000	0.110
*SIC*6	0.000	0.025	0.000	1.000	0.157
*SIC*7	0.000	0.024	0.000	1.000	0.153
*SIC*8	0.000	0.018	0.000	1.000	0.133
*SIC*9	0.000	0.009	0.000	1.000	0.096
*SIC*10	0.000	0.013	0.000	1.000	0.115
*SIC*11	0.000	0.071	0.000	1.000	0.257
*SIC*12	0.000	0.018	0.000	1.000	0.134

Variable	*Min*	*Mean*	*Median*	*Max*	*Std. Dev.*
*SIC*13	0.000	0.016	0.000	1.000	0.126
*SIC*14	0.000	0.017	0.000	1.000	0.129
*SIC*15	0.000	0.012	0.000	1.000	0.110
*SIC*16	0.000	0.036	0.000	1.000	0.186
*SIC*17	0.000	0.031	0.000	1.000	0.173
*SIC*18	0.000	0.007	0.000	1.000	0.081
*SIC*19	0.000	0.016	0.000	1.000	0.125
*SIC*20	0.000	0.017	0.000	1.000	0.130
*SIC*21	0.000	0.050	0.000	1.000	0.219
*SIC*22	0.000	0.021	0.000	1.000	0.143
*SIC*23	0.000	0.016	0.000	1.000	0.124
*SIC*24	0.000	0.012	0.000	1.000	0.109
*SIC*25	0.000	0.084	0.000	1.000	0.277
*SIC*26	0.000	0.022	0.000	1.000	0.146
*SIC*27	0.000	0.014	0.000	1.000	0.118
*SIC*28	0.000	0.083	0.000	1.000	0.276
*SIC*29	0.000	0.018	0.000	1.000	0.134
*SIC*30	0.000	0.007	0.000	1.000	0.084
*SIC*31	0.000	0.141	0.000	1.000	0.348
*SIC*32	0.000	0.048	0.000	1.000	0.214
*SIC*33	0.000	0.022	0.000	1.000	0.148
*SIC*34	0.000	0.051	0.000	1.000	0.220