The Competitive Crash in Large-Scale Commercial Computing

TIMOTHY F. BRESNAHAN AND
SHANE GREENSTEIN

The appearance of a new technology offering lower costs or superior capabilities rarely leads to instant replacement of the old technology. Many important historical examples display this pattern: steam ships versus sailing ships; diesel locomotives versus steam locomotives; equipment for the basic oxygen process for steel versus the open-hearth process; jet engines in commercial aircraft versus propeller engines; numerically controlled machine tools replacing those that were not numerically controlled; and many others (Mansfield, 1968; Rosenberg, 1982; Rogers, 1983; Ray, 1984; Stoneman, 1988). In each case, it is not surprising that the old technology stayed in use; users may be reluctant to retire capital that continues to offer a flow of useful services, even if technical change apparently depreciates the market value of those services. What is surprising is that the old technology continued to sell and to compete viably long after the introduction of the new.

The equilibrium pace of diffusion of a new technology depends not only on developments within that new technology but also on the behavior of customers and older competitors. Buyers may delay their purchase of the new technology until anticipated price/performance improvements materialize. Often buyers need to become informed or to make other investments to take advantage of "enabling" technologies.[1] Sellers of the old

---

We would like to thank participants at many seminars for their comments. Ken Brown, Denise Chachere, and Harumi Ito provided outstanding research assistance. Kathryn Graddy, Tom Hubbard, Scott Stern, Garth Saloner, and the editors provided many useful comments. We also thank the Institute for Government and Policy Analysis at the University of Illinois, the Center for Economic Policy Research, the Sloan Foundation, and the National Science Foundation for funding.

1. See, inter alia, Bresnahan and Trajtenberg (1995) for the widespread importance of this phenomenon in connection with general-purpose technologies.

technology may find their competitive circumstances changed, and react with new pricing or technology strategies. Clearly, the pace of adoption of the new technology, the pace of retirement of the old, and the competition between old and new determine average practice in the economy and, ultimately, the equilibrium pace of creation of social returns.

This historical pattern is reappearing in contemporary information technology. Large complex computer installations are in the process of shifting to a new technological base. For many years, large organizations were forced to rely on expensive mainframe and supermini computers and the proprietary system software and networking technology that accompanied them. More recently, microprocessor-based smaller systems have begun to compete for use in these very large applications. The process of transition has been called many things, but we will call it "downsizing" to "client/server architectures." A transition from old to new has clearly begun. Only its pace and character are still somewhat sketchy.

This transition is more than just a story about the speed of technology diffusion. It also coincides with a major change in information technology firm and industry structure, where the contrasts between old and new structure are hard to miss (Grove, 1990). Most of the old suppliers maintain vertically integrated organizations, with proprietary rights over their technologies. A single firm, the system supplier, influences the development of all hardware and software technologies. It is widely anticipated that the industry structure associated with the new technology will resemble the current structure of the personal computer industry. Competing specialized supplier firms influence different hardware, software, and networking technologies, and no single firm monopolizes the rate and direction of technical change. The anticipation that all of these changes are a serious possibility has already led market observers to devalue the property rights over technologies held by vertically integrated suppliers.[2] This is the "competitive crash" of our title.

The pace of creation of social gains to the new technology has been slow. This is due primarily to slow buyer adoption of the new technology, which contrasts with the rapid advance in the capabilities of that technology (Caldwell, 1994; Ambrosio, 1993). Again, there is (recent) historical precedent for this contrast—it is just an exaggerated version of normal relations in information technology. The information technology industry contains some of the most rapid sustained technical progress in modern economies

2. See "Hardware and Tear" (1992) about destruction of rents at IBM, and Sherman (1991) about DEC. Also, see Hall (1993) for estimates of the decline in the private return to R&D at incumbent large-system vendors in the computer industry.

(consider the integrated circuit) as well as somewhat slower technical progress (consider software) and some very slow progress (consider organizational change and systems development to make full use of computer and data telecommunications technologies). We investigate the competitive crash to understand the forces underlying buyers' slow movements.

The goal of this chapter is to examine the factors underlying buyer demand for large information technology solutions. This goal takes advantage of the natural experiment embodied in the current choice between old and new: recent choice behavior illuminates what demanders really value. Understanding what buyers value not only illuminates the factors underlying the competitive crash, but also the factors underlying the slow realization of the social gains to information technology in large complex applications more generally. We use systematic statistical methods and focus on the early period of diffusion of client/server architectures, through 1991. In this early period, there is very little actual choice of the new technology. Yet it is not competitively irrelevant. Buyers chose, in very substantial numbers, to wait for the new technology to mature. This very substantially lowered demand for the old technology. Demand behavior regarding the old technology is the best available observable information about the early competition between old and new.[3]

Demand for the old technology is well documented in large data sets. Our investigations are based on individual user site data on mainframe hardware and software collected by Computer Intelligence Corporation. We contrast two periods to learn about the competitive crash. The first is in the mid-1980s, late in the period of a mature and stable large-systems market. The other period is the early 1990s, very early in the diffusion of the new client/server technology. Our study provides the first systematic statistical analysis of buyers of large computer systems confronted with the new technological opportunity.

There is controversy about the appropriate theory for understanding the buyer behavior behind the slow diffusion of client/server technology. All reasonable views explain the slow transition as a balance between forces moving buyers forward and other forces holding them back. In the dominant view, the forward-moving forces are the lower costs of the microprocessor-based systems used in client/server architectures. The backward-looking forces are the slow development of client/server software and the sunk investments large users have made with old, proprie-

3. More anecdotal but less consistent and comprehensive information is available from interviews and from the trade press. We take up the relationship between our results and the results of the 1992 and 1993 Bresnahan-Saloner (1994) interview study below.

tary architectures. Yet there are other views as well. Another important hypothesis about the new technologies is that they themselves will alleviate the bottlenecks in information technology commercialization. This view emphasizes the superior features, not lower costs, of microprocessor-based computing. Many buyers would say that the full benefits of client/server architectures, like those of most networking and software technologies, will be difficult to achieve and therefore very slow. We will attempt to clarify the testable implications of these different theories of the competitive crash and then test them.

We do not see this as a backward-looking study of the death of an old technology. We expect a reversal of some of the trends of the late 1970s and 1980s, when small-systems solutions to individual or small-group business problems were the cutting edge and a smaller fraction of total information technology spending went to solving large-business information problems. Networking today, especially over wide areas, is driving a new secular increase in the importance of organization-wide or even inter-organization computing. Understanding the economic process underlying demand for those large-scale computer projects has lasting value.

Investment in Large-Scale Information Technology Solutions

To model the demand for large-scale computing, in either mainframe or client/server form, we begin with the observation that many user organizations have business needs calling for large, complex hardware and software systems. Typically, these systems are not merely purchased from outside the organization, but involve substantial programming at the user's site and even substantial redesign of business practices (Friedman and Cornford, 1989). These projects can be quite large, so that adjustment of the stock of information technology capital is costly. There is a normative literature advising managers how to minimize these adjustment costs, but little quantitative work on their size or origins.[4] In this section, we review the investment process for large projects in general. The next section turns to several specific theories of the adjustment from mainframe to client/server architectures in particular.

We use the Friedman and Cornford "map" (1989, p. 46f) of the position of computer systems in large organizations. It speaks to four distinct complementary assets which are part of adjustment of useful computing capacity. The "computer system core" consists of hardware *and* software acquired from outside the using organization. The "uses of computer applications" are large organization-wide demands for data processing services. These are backbone financial applications such as payroll or accounting, or operations support applications like reservations systems in airlines or accounts processing in banking.[5] The "mediating process" between usage and the computer systems core is undertaken by employees of the using organization (or consultants to it) to make the computer systems core useful. Typically, most of the mediating functions are done by a specialized management information systems (MIS) staff.

They undertake three main kinds of activities. The least frequent and most expensive are whole new applications. End-user departments and MIS jointly work out what broad applications are needed. Then MIS undertakes detailed systems analysis and programming to realize those goals in part. This process is typically denominated in years, not months, and is undertaken by very large teams. More frequently, users and MIS discover problems with existing applications, or request new kinds of reports based on existing data. The maintenance and new-report programming backlog is typically months rather than days. An intermediate category arises when systems usage presses against systems capacity, and MIS manages the tran-

4. Most quantitative literature on the demand for computing uses hedonic measurement in an attempt to quantify the value of computers in use. Triplett (1989) summarizes the literature covering mainframes; more recent evidence in this area comes from Dulberger (1989), Gordon (1989, 1990), and Oliner (1993). Stavins (1996), Berndt and Griliches (1993), and Berndt, Griliches, and Rappaport (1993) have conducted hedonic microcomputer studies. A second branch of the literature on demand for computing focuses on the relationship between computerization and productivity, as in Berndt and Morrison (1991). Loveman (1994) and Brynjolfsson (1993) review this literature. Another branch tries to estimate the aggregate marketwide value of different forms of computerization by demand analysis (Bresnahan, 1986; Flamm, 1987; Brynjolfsson, 1993), sometimes using micro data (Trajtenberg, 1990; Greenstein, 1995). Only a few papers look at the theory of demand, and those are confined to very special groups of demanders (Greenstein, 1991, 1992).

Non-statistical literature on the value of computers in use is largely normative. A positive analysis has been provided by Friedman and Cornford (1989). Scott-Morton (1991) and Allen and Scott-Morton (1994) contain essays that are good examples of the positive and normative literature.

5. This definition excludes personal productivity applications running on personal computers or workstations. The usage category boundaries are hard to define precisely in a technical way. Small systems, for example, replaced many time-sharing usages of mainframes over a decade ago. The same applications that require mainframe power in larger areas can be mini-computer "departmental computing" or even micro-computer "small business computing" in other contexts. So the definition of the category boundary depends on both the size and complexity of the user organization and the business purpose of the application. Our definition is pragmatic, the kinds of applications for which mainframes were deployed in the mid-1980s. Our description of them, and our language, closely follows the standard systems-choice doctrine of that era (Inmon, 1985).

sition to new (frequently compatible but involving work to install) higher-capacity systems. This third category is often caused by the second—better systems get more use, and more reports eat up more computing resources. The third category often merges into the first—increased purchases of hardware and software capacity will often be the occasion for increasing an application's features. These upgrades/improvements also can take significant time to build.[6]

As a result, most important expansions of capacity, whether new systems or major upgrades/improvements, involve changes in hardware, externally acquired software, on-site technical work, and changes in business procedures together. For this reason, we feel confident that using changes in hardware capacity offers a good way to observe large projects. As long as we catch both major upgrades and whole new systems, hardware expansions and new projects should largely overlap.

These expansions and upgrades obviously involve investment costs which are irreversible in part. While mainframe hardware can be leased, and mainframe software typically has annual license fees, the costs of in-house and consultant programming typically are irreversible. From reports on the budgets of a typical MIS staff in our time period, it seems clear that the latter, irreversible budget category is well under a half and probably no more than a third of total investment costs.[7] In earlier work with Harumi Ito, we quantified the fraction of project investment costs which sites appear to treat as irreversible. That led to a much larger estimate, around four-fifths.[8] The discrepancy in the two estimates is probably explained by irreversible investments in changed business practices accompanying projects, suggesting that these internal investments are roughly as large as hardware, acquired software, or local programming.

The analytical literature on investment (Dixit and Pindyck, 1994) and recent theoretical work on competition, standard setting, and the rate of technical progress in information technology industries has highlighted several distinct roles that buyer inertia or caution may play.[9] These are reflected in competing engineering and business theories of buyers' slow response to client/server architectures. In the next section, we attempt to organize these competing theories of the slow switch to client/server. That work emphasizes that the appropriate theory of the irreversible adjustment costs is as important as the size of the irreversible costs themselves.

6. Friedman and Cornford (1989) offer an excellent summary of both anecdotal and quantitative research on these processes.

7. See, for example, data processing budget stories in *Datamation* on April 1, 1986, and May 1, 1993. Friedman and Cornford (1989) also have useful information on this topic.

8. The source of this estimate is in a distinct treatment of increases in capacity versus decreases in capacity. (The present chapter examines only increases in capacity.) The decrease in demand that leads to capacity reduction is approximately four times as large as the increase in demand that leads to capacity expansion. Hence the four-fifths sunk estimate.

9. See David and Greenstein (1990) or Besen and Saloner (1988) for reviews.

Technological and Economic Theories of Slow Diffusion

Each of the currently available competing theories, as we shall see in this section, embodies an important truth about technical forces. Hypotheses about which of these forces are most important, however, are necessarily hypotheses about demand. In this section we go on to illuminate the testable implications of a variety of specific theories of the competitive crash.

The dominant view of the new competition contrasts an old, inferior technology with a new, superior one. Mainframes and other large computer systems, in this view, embody old hardware and software technologies. By contrast, microprocessor-based computer systems are the wave of the future. They are based around technologies that offer lower costs per unit of performance, and that promise more rapid technical progress in the future. In this view, the date of replacement of old systems by new is determined by the timing of technical advance. In particular, two main classes of technical advance were needed. The first was the emergence of a "mainframe on a chip." For some time, microprocessor-based computer systems offered cheaper price/performance, certainly cheaper measured by cost per millions of instructions per second (MIPS) and also by broader performance measures. Now the largest microprocessor-based systems have begun to offer these low costs at levels of performance comparable to large systems. The second advance needed was the emergence of fundamental software technologies such as operating systems, databases, and networks which would permit new systems to perform the traditional tasks of the old. The slow changeover is explained by the difference in technical progress between software and hardware. Throughout the period from 1989 to 1992, the hardware technical progress was typically described as recent, the software technical progress as imminent.[10]

This view is extremely attractive to technologists, in large part because of its compact and compelling description of technical progress. We call this view "competitive MIPS arbitrage." Obviously, it suggests a rosy fu-

10. Compare, for example, Kador (1992) to Keefe (1990) or Radding (1989). All describe the near-term possibilities in much the same terms.

ture for the social gains to information technology once a difficult period of adjustment has passed.

This first view explains the destruction of private rents in the old computer industry as an anticipated increase in replacement of old hardware by cheaper new hardware. That there are potential future substitution opportunities due to different hardware costs is not in serious dispute.[11] When they can actually perform this arbitrage, buyers will destroy the market power of sellers of old technologies—that is, they will flatten the demand curves for mainframe and supermini hardware and software. This is a powerful testable implication. It implies not only that the old system business was unprofitable overall, but also that it was unprofitable in the price-cost margin sense. Since over 80 percent of our sites use IBM mainframe architectures, it is probably appropriate to view our tests of this hypothesis as primarily about IBM mainframe market power.

Another very important technologists' view of recent changes emphasizes the different technical characteristics of traditional large and small systems. Large systems to solve large business problems are very powerful, but very difficult to use. The specialist programmers and others who use these systems, in this view, have also not been organized in a way that makes them very responsive to business end users. Programming backlogs are better measured in quarters than in weeks. This has been an ongoing frustration to computer-using organizations. A change occurred when business people in the organizations saw how quickly and easily easy-to-use microcomputers could solve real (but small) problems. There began to be very substantial demand for business computer systems that were as powerful as traditional mainframes yet as responsive and easy to use as micros. Client/server architectures attempt to accomplish this through the use of linked heterogeneous systems. In the second technologists' view, one should understand the competitive threat to traditional systems as coming from these superior technical features, not just lower costs. This view, too, is broadly held in the technical community.[12] It has even spilled over into the business strategy community. We summarize this view as "client/server best of both worlds."

The "best of both worlds" view is important because it captures something fundamental in the demand for large-systems, and links it to the successes of different technologies in the marketplace before the competitive crash. User organizations are deeply unhappy with the clumsiness of central MIS as an organizational solution.[13] Further, the theory is testable because there is considerable variety in the extent of this unhappiness. Under the best of both worlds theory, the kinds of sites for which professionalized MIS is a particularly unsatisfactory organizational solution should be those most eager to switch to client/server.

There is another theory based on much the same facts and history.[14] This theory agrees that the largest *potential* gains from client/server come in the organizations least satisfied with existing MIS. There is, however, an equilibrium reason for the dissatisfaction. These organizations are those in which the adjustment costs of change to use new information technology for business purposes have been the largest historically. In this story, these sites are simply those for which the problem of coordinated change in business practices and information technology is the most difficult. If the adjustment costs to client/server are very large at these same sites, they may find the switch both more attractive and more difficult than other sites. They could be, counting costs and benefits together, the least rather than the most interested in switching.

The relationship between the best of both worlds and adjustment costs theories is that they are opposites. Both order site organizations according to the degree to which there is dissatisfaction with existing MIS as an organizational solution in the mainframe era. In Figure 1, the horizontal axis captures this. As we move to the right, the existing internal organization of large-scale computing grows more complex and correspondingly less satisfactory. The existing set of information technology solutions is less satisfactory to, or less controlled by, the business organizations using them. Now, as we move to the right, both the benefits (best of both worlds) and difficulties (adjustment costs) of moving to new solutions rise. Under the best of both worlds theory, it is the benefits curve which rises more steeply, so the organizations to the right are the most interested in switching to client/server. Under the adjustment costs theory, we get the reverse. The cost curve rises more steeply than the benefits, and it is those organizations on the left switching to client/server.

11. See interview with John F. Akers, IBM Corp. Chairman, in the July 15, 1991, issue of *Fortune* magazine.

12. See the same articles as in note 10 for journalists' views of this. This view tends to be held more by systems integrators, consultants, and client/server software engineers rather than by technologists from the small-systems world exclusively. An important version of this view links the payoff from information technology to a broader "reengineering of business processes." See, for example, Hammer and Champy (1993, chap. 5).

13. Friedman and Cornford (1989) devote several chapters to the long history of this unhappiness.

14. This view is argued by Bresnahan and Saloner (1995) in connection with their interview study. It is clearly consistent with the theory of adjustment costs advanced by Friedman and Cornford (1989) for an earlier era. By late 1993 or early 1994, the trade press began to pick up these gripes from users. For example, see Caldwell (1994).

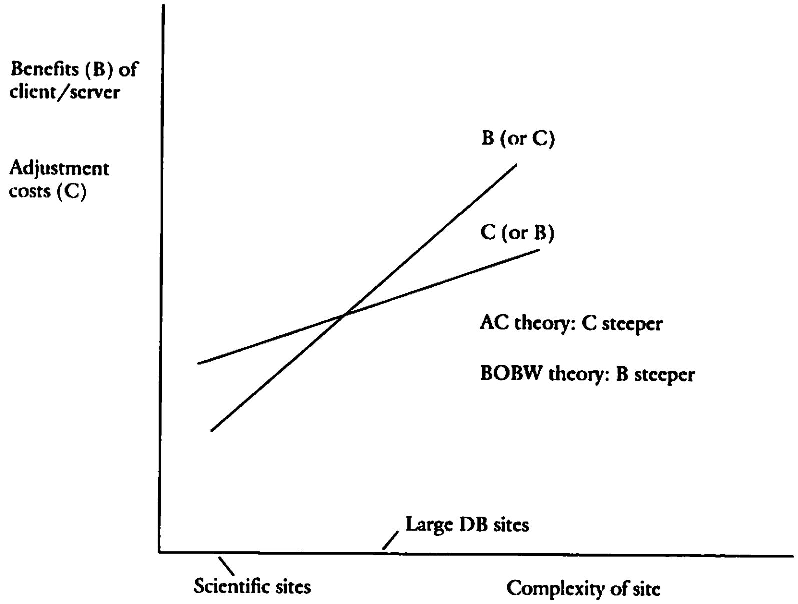

Figure 1. Best of both worlds theory versus adjustment cost theory.

Finally, the diffusion of new technologies may have been slowed by the possible lock-in of proprietary systems vendors at particular sites. The costs of existing ("legacy") applications may be not only irreversible, but irreversibly tied to the systems of a particular vendor. More plausibly, sites may vary in that some of them have very high costs of migrating away from their existing systems vendor, others lower costs. Similarly, the MIS department itself may have locked in a powerful internal political position and be resistant to change.

All of these stories have in common that there are powerful forces pulling demanders forward toward client/server. None of the theories suggests that client/server will not prevail in the long run. The stories differ in whether the client/server attractions are costs or features. More important, the stories differ in the nature of the forces holding back the diffusion of client/server—though clearly every theory must have such a force as well. Some posit a "lock-in" to existing assets: that is, the inertia of already sunk costs is holding back the diffusion. Others posit caution as a source of high forward-looking adjustment costs to new opportunities. Note that the theories do not differ in their predictions for the pace of diffusion in the

early phases. Instead, they differ in the kinds of sites they predict to be faster or slower adopters.

Sample and Data

What kinds of sites change their demand for the old technology? Our strategy focuses on differences between large-system users who continue to add capacity to their installations and those who choose not to do so. We wish to identify which large-system users waited for client/server rather than expand the stocks of their general-purpose mainframes. To accomplish this goal, we use a database of many large-system users in the United States.

We examine individual site locations as measured by Computer Intelligence Corporation in their year-end surveys. We use two "triads" of data, 1984–85–86 and 1989–90–91. While the first triad is the oldest available to us, it also has the virtue that it represents a period of mature mainframe demand. The latter triad represents the beginning of the diffusion of client/server alternatives.[15] Characteristics of a site in a "base" year, 1984 or 1989, predict capacity expansion. We will interpret the kinds of sites with the largest otherwise unexplained downturns in mainframe demand (in a richly specified model) between the two triads as those who are waiting.

Our sample begins with all Computer Intelligence Corporation survey participants with at least one general-purpose mainframe in any of the six years. This is the most complete and richest panel data available on the use of large computing equipment. Roughly 14,000 sites appear in the Computer Intelligence Corporation sample in each year, which comprises somewhere between 70 and 80 percent of all general-purpose mainframe computer users, according to Computer Intelligence Corporation estimates. Each year new sites enter and some old sites exit; turnover is about 10 percent of the sample of sites each year. To be included in our analysis, the site can exit in the third year but not the second of each triad. Also, the site must have general-purpose mainframes and must have filled in the soft-

15. Investigation of periods after this very early one is going to call for more complex models than the simple ones reported here. We have acquired the more recent data for 1992 and 1993 and are in the process of analyzing it. Other issues arise in these periods. For example, sites that decided to wait during our current sample period may later decide not to keep waiting. To many sites, it became clear that client/server applications for their purposes would arrive after 1992 or 1993, not as soon as predicted. Accounting for such dynamically complex behavior calls for more subtle empirical models than the ones we are treating here.

ware as well as the hardware survey. Finally, we must be able to determine the industry of the site.[16] We are left with over 10,000 sites in each triad, over 50 percent of all mainframe users in the United States.

We use Computer Intelligence Corporation's definition of a "site," which corresponds with a unique company address and senior data processing manager. Since Computer Intelligence Corporation designs its database for direct marketing campaigns by value-added peripheral and software vendors, a site corresponds closely to the organization within which decisions are made about acquisition of systems. Thus, it is likely that the same factors influence decisions at the same "site." However, this correspondence may be weaker at the largest sites, such as those devoted to varied research tasks in campus-like settings in private industry. At these sites, Computer Intelligence Corporation's site definition may only partially embed decentralized authority.[17]

We also employ Computer Intelligence Corporation's definition of a *general-purpose mainframe* computer. The advantage of Computer Intelligence Corporation's definition is the accuracy and completeness of Computer Intelligence Corporation's data for large systems. This definition, like any other, is unavoidably arbitrary at the smaller end, where general-purpose mainframes compete against general-purpose superminis. Though we could quibble with some of Computer Intelligence Corporation's choices about what systems to include and exclude as general-purpose mainframes, they tend to follow industry conventions about what is and is not a mainframe. The most important problem arises in limiting the scope of our conclusions. We cannot say, on current evidence, whether proprietary supermini systems have been affected in the same way as have proprietary mainframe systems.

Endogenous Variables

Our dependent variables should capture increases in mainframe capacity, taking into account lumpiness and the time taken to make changes. We construct three different variables with partially overlapping definitions of capacity increases.

We begin with increases in the number of systems in use at the site that persist for at least two years. In each triad, we say that there is an increase

16. We have used the name of the firm or other institution owning the site matched to public sources to increase the coverage and accuracy, especially for government sites.

17. As in many marketing databases, there is some information about the locus of decision making. For the years 1987–91, we know whether large technical decisions are made at the site or at a central authority elsewhere in the company. We have not yet used this information to examine our definition of "site" as decision locus.

in capacity if there are more mainframe systems the second year than there were the first. We say that the increase is persistent if there continue to be more systems in the third year than in the *first*; transitory, if the number of systems falls back to or below the original level. We believe that the persistent increase in the system counts variable, hereafter Systems, measures large increases in the stock of mainframes. Our interpretation is that increases in the ***number*** of mainframes in use represent significant increases in mainframe capacity and reveal large increases in desired capacity. To capture smaller changes in computing capacity, such as those associated with upgrades or system replacements, we turn our attention to the total processing power of a site's mainframes, measured in MIPS. Here, a persistent increase is more MIPS on the site in the second year than the first, and still more MIPS on the site in the third year than in the first.

In Table 1, we present descriptive statistics on these and closely related variables. Note that persistent capacity increases are much less frequent for Systems than for MIPS. In both triads, persistent capacity reductions outnumber persistent increases for Systems but not for MIPS. This reflects the mature state of the mainframe market, where revenue stays high through selling larger systems, in spite of selling fewer of them. Consistent with the description of the difficulty of large capacity projects above, the most frequent outcome in each of our triads is "other," which consists mostly of sites that do not change their stock of mainframe computers.

Another fact in Table 1 also has some implications for the amount of time the investment in large new computer projects takes. There is a dramatic difference in the MIPS and Systems measures. In both triads, half of the increases in Systems counts are transitory—that is, half of the increases are reversed after one year. Only a very small portion of MIPS increases are

TABLE 1

Net Changes in Large Computing Capacity

(*percent*)

	Persistent capacity increase	Transitory capacity increase	Persistent capacity reduction	Site exit	Other (mostly no change)
Counting systems					
First triad	8	8	14	6	64
Second triad	5	5	12	8	70
Counting MIPS					
First triad	33	2	11	6	48
Second triad	25	1	10	8	55

NOTE: Total sample size for the first triad was 10,778 sites; for the second triad, 11,776 sites. Rows may not sum to 100 percent due to rounding error.

reversed in the second year. This is evidence for the quantitative importance of dual systems operation. The investment process for new data processing projects must take a very great deal of time, at least a very substantial fraction of a year, to explain these numbers.

Now let us consider changes over time in demand behavior looking at the raw facts in Table 1. First, consider *reductions* in capacity. There are always some; but there is very little change over time in the fraction of sites that reduce either mainframe MIPS or Systems. If anything, the fraction of sites reducing capacity is slightly smaller later on. On the other hand, far fewer sites *expanded* mainframe capacity in the second triad. Measured by Systems, the rate of capacity expansion fell from 8 percent to 5 percent, by MIPS, from 33 percent to 25 percent. The larger drop in MIPS means that there was a decline in upgrades and replacements above and beyond the decline in whole new systems.

A variety of evidence makes clear that this decline in mainframe expansion is not actual switches to client/server. First, the trade press and the Bresnahan-Saloner interviews (1994) make clear that there is not much downsizing to client/server until 1993, at least not in the sense of switching over real production applications (Caldwell, 1994; Ambrosio, 1993). The switch to massively parallel computers is trivial, despite persistent rumors.[18] About one-fourth of total expected mainframe demand has gone away ([33−25]/33). It is not the case that these are needs met with new technology, but instead unmet needs.

One possible explanation is the recession during our second triad. But this explanation is far from sufficient. First, despite the broader recession, MIS budgets continued rapid growth into our second triad's decision times.[19] Moreover, using our econometric estimates of the impact of demand growth on capacity expansion, we still see a substantial downturn above and beyond the effects of the recession. Finally, we have demanders' frequent statements in the trade press or in interviews that this was a period of "evaluation" or of "wait and see" for downsizing opportunities. Using either the MIS budgets or the econometric estimates, we can calculate the extent of the decline in mainframe-based projects above and beyond recession effects. Both calculations suggest that there are over 1,400 "missing" mainframe projects nationwide, including upgrades as well as new systems.

18. This question is very common in seminars. But the evidence is that there was little replacement, even as late as 1993. Even then, massively parallel systems were typically deployed as complements to, not substitutes for, mainframe systems (Boughten, 1993).

19. MIS budgets continued to grow in 1990 only slightly slower than in the first triad. By 1991, there were clearly decelerations in the growth of MIS budgets. But they continued to have positive nominal growth. For example, see *Datamation*, April 15, 1991.

TABLE 2

"Brand" Switches

	Count of brand switches	As fraction of persistent increases in system counts	As fraction of sites
First triad	93	0.10	0.0086
Second triad	171	0.31	0.0145

NOTE: Definition of "brand switch": If the main system is IBM-compatible, a permanent increase in the count of non–IBM-compatible systems is a switch. Otherwise, a permanent increase in the count of IBM-compatible systems is a "switch." In 1984 the main-system question was not asked, so we use the 1985 reported main system. At one-system sites, we often impute a main-system brand.

Within our sample, which covers about half of the installed base, there are over 700 missing projects. There was very substantial waiting for client/server even though there was little actual adoption of the new technology in this period.

The economics literature on product pre-announcement has for some years posited the importance of this kind of anticipatory demand behavior (Farrell and Saloner, 1986). The strength of the behavior, given that client/server architectures were definitely "vaporware" at this stage, is impressive.

We also report simple statistics on brand switches among vendors of mainframe technology. We consider only two "brands" of mainframes, IBM (and compatibles) and all others. As can be seen in Table 2, switches are very infrequent in our first triad and, while increasing, still rare in our last. Some alternative brand-switch definitions, like changes in the reported main system, would be even rarer. So we do not pursue analysis of brand switches further.

Finally, we add a continuous-valued capacity increase variable, the persistent increase in MIPS at the site. Because of the importance of dual system operation, we define the persistent increase in MIPS as the minimum of the increase from the base year to the first year or to the second year. The simple first difference double-counts the MIPS of the systems in dual system operation, and we know from Table 1 that this double-counting applies to about half of capacity expansions. So that the first and second triad figures will be comparable, we deflate the MIPS figures using a mainframe computer price index from Dulberger (1989).

We will proceed by estimating cross-section models for increases and decreases in capacity, measured by both number and MIPS. These will be probits in the first analyses. Similarly, we will estimate a tobit for the continuous-valued increase in MIPS.

Exogenous Variables in Cross Section

We predict each of these three dependent variables with a long list of regressors. This section defines the regressors. In each triad, the regressors are observed in the "base" year (1984 and 1989). We use them to predict persistent net increases in capacity over the next two years. We begin this section with variables which are included primarily to ensure we capture much cross-section variation in large computing demand. We then describe variables closely linked to our hypothesis.

We use employment data for each industry (two- or three-digit SIC) to proxy for changes in the derived demand for computer systems output. We also include SIC dummies for a more limited set of unusual cases.[20] Employment has several useful properties: though it is an input in production, it is a cyclical indicator of computer systems output and therefore desired computer system investment. Moreover, user institutions in our sample are both public and private, for-profit and not. Thus, employment is probably the best unifying measure of the derived demand for inputs. We would prefer company or institution data rather than industry data, but this is only available for a subset of users.

The maximum and minimum age of the general-purpose mainframe computing systems at a site measure, crudely, the distribution of times since upgrades. As a result, they are related to the gap between the technical frontier embodied in new equipment and the level embodied in the equipment at the site. Of course, these variables are endogenous in a dynamic sense. They are likely determined by (among other things) the site's past history of computing power needs, which could be correlated with current needs. Here and elsewhere, we use lagged technical choices as proxies. We do not make causal inferences about these variables. Their task is to capture much of the cross-section variation in the state of the replacement cycle at the site. If they also pick up persistent heterogeneity in the valuation of computer services, or in "lock-in" to particular systems, we are untroubled by that.

Similarly, we use the MIPS rating of the largest and smallest general-purpose system as an indicator of the maximum and minimum demands on computing capacity. Use of a large-capacity system correlates with a

20. In preliminary research we tried regional dummies interacted with time and a more complete list of SICs than shown in the present results. We found that our results were not qualitatively influenced by dropping or including these variables. Hence, we only show the shorter results below. In work in progress, we have linked many of these sites to microdata sources (Bresnahan, Greenstein, and Ito, 1994).

demand for systems performing a large maximum feasible task (Bresnahan and Greenstein, 1992). Use of a small-capacity general-purpose system ought to correlate with a need to employ mainframes instead of the next smallest alternative, a general-purpose supermini. That is, it may suggest that the buyer anticipates increasing capacity along well-understood mainframe growth paths as user needs grow (instead of the more limited growth paths associated with superminis). So these variables may capture the site's past assessment of the pace of upgrading and replacement.

We include a count of general-purpose systems, with several possible interpretations. First, it may signal that the computing core serves a large end-user community. The coordination problems associated with a large community may slow the pace of change. Second, a large site is likely to realize the economies of scale and scope necessary to try technical solutions with high fixed costs. Therefore, we expect to observe a large portfolio of technical solutions to computing needs.

We also include a dummy variable showing whether the site's "major" system is not from IBM or from an IBM plug-compatible manufacturer. Because of the rarity of vendor switching, this will help us measure differences in the demand facing IBM relative to the other mainframe vendors.

We now describe the variables closely linked to our hypotheses. Using standard descriptive analyses of large computer installations, we identify the kinds of environments associated with organizational dissatisfaction with large systems. To obtain proxies for these environments, we construct a series of variables based on the software in use on mainframes at the site. Computer Intelligence Corporation provides lists of software programs and their provider, categorization of its functionality, and the number of copies in use at a site. This information is rich in detail. Software information captures important activities inside the mediating process at the site. Different software categories point to a more or less costly, complex, localized, or locked-in mediating process.[21]

We categorize software programs into two different sets of dummies. The first uses the software ***author*** to identify the importance of the vendor-user interface for large-system demand. If sites' investments lock them into their hardware vendor, as switching cost theory suggests (Klemperer, 1992), then a site that uses much software written by its general-purpose hardware vendor will be particularly locked in. Switching will require aban-

21. In general, while we use software variables as proxies for the sites' adjustment costs, none of these uses of software variables is a calculation of investment in complementary software, per se.

doning any idiosyncratic investments tied to the software provided by the hardware vendor. A similar argument applies to software that Computer Intelligence Corporation designates "in-house," meaning where the user is also the designer. Such software may incorporate idiosyncratic features of the user and the computing platform, which makes it virtually unportable. However, in-house expertise in software programming may ameliorate some of these lock-in effects. These users may be able to overcome portability difficulties themselves, instead of relying on vendors.

The rest of the software, not written in-house and not from the hardware vendor, is either from consultants or from third-party software firms. We somewhat arbitrarily categorize software as "third-party" if we find more than twenty programs in all the sites in our sample. Under the lock-in theory, users with much third-party software find it less costly to move to new platforms. We further divide third-party software. If the apparent strategy of the software author company was to make its product portable across different brands of mainframe system, we put it in the "multiplatform" category. If the author company appears dedicated to only one type of computer, we put the software into an IBM-specific or other-specific category.

The test of both the vendor lock-in and MIS lock-in theories comes from the behavior of buyers with more specific software. More specific software—that from the proprietary systems vendor or from a third-party software firm writing only for one type of computer—is interpreted as revealing a mediating process with costs more sunk to a relationship with a specific mainframe vendor. Similarly, under the MIS lock-in theory, software that is more local to this site is interpreted as revealing an opportunity for foot-dragging by MIS should it wish to preserve the value of its skill base in the old system. Being tied to a vendor occurs either because vendors force such sunk costs on the buyer who cannot successfully resist, or because managers of information systems prefer their incumbent and have the power to enforce these preferences, even if these conflict with broader organizational goals.[22]

We calculate the fraction of software packages that fall into each author category at each site. The results are in Table 3, along with descriptive statistics of all our other regressors. Note that the fractions are essentially the same in our two triads.

The second set of software variables focuses on the use of software and

22. The MIS lock-in and the vendor lock-in theories are not completely distinct, as this sentence suggests. Outsourcing of the entire MIS function in connection with downsizing is often suggested as a way to solve the two linked problems.

TABLE 3

Site Characteristics in Selected Summary Statistics

Variable label	Definition/categories	First triad Mean	First triad Std. dev.	Second triad Mean	Second triad Std. dev.
Software usage variables (proxies for organizational complexity)					
SCI	% scientific & number crunching s/w	0.037	0.062	0.037	0.085
TS	% technical support required s/w	0.009	0.068	0.008	0.045
STD	% standard business application s/w	0.256	0.209	0.219	0.224
DB	% database & application oriented s/w	0.206	0.149	0.201	0.144
COMM	% communication & network s/w	0.259	0.163	0.243	0.169
MIPCM	% comm * maxmip (defn. below)	0.812	1.502	1.208	2.176
MIPDB	% db * maxmip (defn. below)	0.687	1.361	1.161	2.131
Software author variables (lock-in theories)					
INHOUS	% software written in-house	0.198	0.198	0.189	0.248
PROP	% s/w from a proprietary systems vendor	0.445	0.244	0.451	0.252
PROPDB	% s/w prop and db	0.139	0.132	0.121	0.129
PROPCM	% s/w prop and comm	0.183	0.147	0.172	0.158
TPBLUE	% s/w from a third-party vendor, all IBM	0.285	0.207	0.291	0.224
TPNONB	% s/w from a third-party vendor, one brand	0.002	0.043	0.009	0.047
MPLAT	% s/w from a multiplatform third-party vendor	0.061	0.147	0.062	0.111
Replacement cycle or background variables					
EMPGRW	% employment growth in 1–3 digit SIC	0.024	0.080	0.004	0.033
MINAGE	Age of the newest system	1.724	2.108	2.046	2.137
MAXAGE	Age of the oldest system	3.116	2.538	3.264	2.566
SYSSUM	# of systems	1.733	1.242	1.697	1.337
MMBLUE	Major system is IBM or compatible	0.853	0.353	0.834	0.372
MAXMIP	Maximum MIPS of systems	3.142	5.378	5.545	9.141
MINMIP	Minimum MIPS of systems	1.655	2.540	2.955	5.366
Dependent variable not shown in Table 2					
ΔMIPS	Increase in MIPS (if positive) (deflated)	5.430	11.14	8.749	18.95

the kinds of system it is running on. Here, we make use of Computer Intelligence Corporation's evaluation of the purpose of the software. We group their very detailed categories based on a close reading of the similarities and differences between each market niche. Our reading focused on attempting to predict the horizontal axis in Figure 1 under the best of both worlds and adjustment costs theories.

One category is what we call "scientific computing and other numerically intensive methods." This includes such software as CAD/CAM and standard large spread-sheet applications. Years before client/server, these uses were first to move to workstations because these users tend to possess a high degree of computer sophistication and do not require frequent use of a large centralized database. Another category is what we call "technical support necessary," which includes applications such as manufacturing. These applications are technically demanding—where "technically" means the computing is complementary to technologies other than computer technology—and require frequent interaction between user and vendor. A site with a high percentage of these products will be populated with engineers and will contain needs that are organizationally simple to address. So these first two categories are to the left in Figure 1. Earlier, these users were the first to anticipate leaving large computing platforms and taking advantage of advances in alternative smaller platforms like minicomputers. These users tend to be among the most successfully resistant to centralized management of computing resources, frequently using junior scientists rather than MIS professionals.

A third category of software is what we call "communications and other multiuser tools." This includes many system programs designed to enable mainframe-micro links, and many system programs designed to control communications. A large community of users will exist at sites with a large percentage of these programs. This may signal difficult mediating processes associated with essential computing tasks or costly processes of adjusting applications to new technical alternatives.

Our fourth and fifth categories examine the type of database programs in use. Computer Intelligence Corporation designates these as either "system" or "application" programs. System database programs include software such as file management programs. Database applications include such software as standard financial analysis and large accounting packages. Sites that make use of many application database programs may find it marginally easier to shift, since many of these types of programs are available on different computing platforms. The omitted category includes software that we find on nearly all large computers, like operating systems. These programs should provide little information about a large-system user, since virtually every computing core makes use of similar programs.

Finally, we interact some software variables with other measures to highlight where the mediating process has been problematic. We interact our database application variable with the size of the maximum MIPS system on site. We also treat database software from the systems vendor as a separate category. We do a similar interaction of our communication software variable with the measure of maximum MIPS and treat this software differently if it is proprietary to the system vendor. We think that the interactions with the largest MIPS should capture sites to the right in Figure 1. Under adjustment cost theory, these sites are least likely to move out of mainframes because these users are taking advantage of system size and vendor-specificity in applications using large databases and frequent real-

time communication with computing resources. Under best of both worlds theory, these are many of the users who express the most unhappiness with large-system solutions and are the most likely to move.

These variables, too, can be seen in Table 3. Once again, the figures reported come after a calculation of the fraction of mainframe software packages at the site falling into the category.

Econometric Models

Our econometric models focus on identifying changes in mainframe capacity expansion behavior between our two triads. We have three dependent variables; the persistent capacity increase dummies for MIPS and for systems described above, and continuous-valued increases in MIPS. The capacity expansions are measured in the second two years of each triad (1985/86 or 1990/91). The three dependent variables are treated separately; the first two are estimated by probit, the third by tobit.

The regressors are all measured as of the first year of each triad, 1984 or 1989. We interact all of the X's with a second triad dummy. Call the first-triad coefficients of all the regressors in one of the analyses β_{85}. The second-triad coefficients are $\beta_{85} + \beta_2$. Our specification leaves the β_2, which measure how behavior changes over time, unrestricted.[23] All of the regressors are positive. Thus, negative β_2 identifies the types of sites that tended to expand mainframe capacity less in the second period. Our interpretation of negative β_2 is that it identifies the sites that waited for client/server.

The interpretation is slightly more complicated for the two mutually exclusive sets of software dummies. We include separate intercepts for each year, and we also include the employment variable. Between these two variables, they should capture much of the business cycle effects. Since the software variables within each category sum to one, we must exclude one variable in each category. As a result, they have relative interpretations. A negative β_2 identifies kinds of sites that tended to wait more for client/server; a positive β_2 identifies kinds of sites that tended to wait less.

Specifications Estimated and Results

Results are reported in Tables 4 and 5; the format is that all three estimations are reported together, with the change parameters β_2 in Table 4 and the baseline from the first triad in Table 5.

23. In an obvious notation, we will call $\beta_{90} = \beta_{85} + \beta_2$ below.

TABLE 4

Changes in Behavior Over Time

Variable label	Definition/categories	MIPS tobit	MIPS probit	Systems probit
Software usage variables (proxies for organizational complexity)				
SCI	% scientific & number crunching s/w	−12.4085	−.543227	−.642352
		(4.53918)	(.302894)	(.420265)
TS	% technical support required s/w	−6.71945	−.525692	−.917339
		(6.23987)	(.397538)	(.649192)
STD	% standard business application s/w	−3.47132	−.247190	−.370071
		(2.41683)	(.157595)	(.217556)
DB	% database & application oriented s/w	−3.86638	−.311132	−.601025
		(4.04167)	(.266800)	(.378115)
COMM	% communication & network s/w	−6.71316	−.565845	−1.14444
		(4.79685)	(.316349)	(.468129)
MIPCM	% comm * maxmip	−.392041	−.014073	−.005754
		(.381483)	(.029606)	(.032770)
MIPDB	% db * maxmip	1.55301	.059314	.027342
		(.416548)	(.031260)	(.034424)
Software author variables (lock-in theories)				
INHOUS	% software written in-house	−5.65120	−.553258	1.47796
		(7.39035)	(.475507)	(.918601)
PROP	% s/w from a proprietary systems vendor	−7.42648	−.729804	.746187
		(7.36940)	(.473940)	(.919417)
PROPDB	% s/w prop and db	−4.80366	−.185774	.263288
		(4.52567)	(.297639)	(.427679)
PROPCM	% s/w prop and comm	−4.99698	.621137	1.74229
		(5.24249)	(.344654)	(.510188)
TPBLUE	% s/w from a third-party vendor, all IBM	−6.06387	−.475147	1.20520
		(7.34379)	(.473252)	(.919225)
TPNONB	% s/w from a third-party vendor, one brand	−7.49751	−.571061	1.80417
		(10.1583)	(.657678)	(1.15012)
MPLAT	% s/w from a multiplatform third-party vendor	−2.94401	−.429298	1.23953
		(7.68936)	(.494201)	(.938708)
Replacement cycle or background variables				
C-1990	Constant (change over time)	−21.0692	−.946986	−3.26459
		(4.52015)	(.295153)	(.777093)
EMPGRW	% employment growth in 1–3 digit SIC	5.59733	−.184301	−.126910
		(9.83675)	(.654021)	(1.05055)
MINAGE	Age of the newest system	.743663	−.010190	.008400
		(.213215)	(.014706)	(.022147)
MAXAGE	Age of the oldest system	−.672842	.006289	−.014566
		(.172217)	(.012177)	(.018369)
SYSSUM	# of systems	1.81999	.027313	.022627
		(.301390)	(.023934)	(.031137)
NONIBM	Major system not IBM-compatible	5.16814	.517137	.484600
		(1.18211)	(.078230)	(.108700)
MAXMIP	Maximum MIPS of systems	−.053739	−.003765	.0067614
		(.140150)	(.010458)	(.011645)
MINMIP	Minimum MIPS of systems	−.444075	−.001209	−.046341
		(.107141)	(.007970)	(.009200)

NOTES: These results are drawn from three separate analyses and are a subset of the total parameter vector in each. In particular, each coefficient reported here is a change over time between the first and second triad in the impact of the variable. The rest of the coefficients are in Table 5.

The first column has units d(MIPS)/d(variable). In the second column, multiply each coefficient by .37 to get units d(probability of increasing MIPS)/d(variable). In the third, multiply by .15 to get units d(probability of increasing Systems)/d(variable).

Estimated standard errors are in parentheses.

TABLE 5

Variety in Behavior in First Triad

Variable label	Definition/categories	MIPS tobit	MIPS probit	Systems probit
Software usage variables (proxies for organizational complexity)				
SCI	% scientific & number crunching s/w	.389712	−.004630	.389712
		(.297245)	(.237877)	(.297245)
TS	% technical support required s/w	.182191	−.004600	.182191
		(.278786)	(.243933)	(.278786)
STD	% standard business application s/w	−.082500	−.046636	−.082500
		(.151175)	(.116685)	(.151175)
DB	% database & application oriented s/w	−.036402	−.158785	−.036402
		(.251937)	(.195033)	(.251937)
COMM	% communication & network s/w	.253666	−.367987	.253666
		(.291834)	(.221719)	(.291834)
MIPCM	% comm * maxmip	.006777	.037635	.06777
		(.027698)	(.024987)	(.027698)
MIPDB	% db * maxmip	−.025255	−.039224	−.025255
		(.028623)	(.025703)	(.028623)
Software author variables (lock-in theories)				
INHOUS	% software written in-house	−.168724	.015557	−.168724
		(.491378)	(.374191)	(.491378)
PROP	% s/w from a proprietary systems vendor	.334697	.258297	.334697
		(.484176)	(.369084)	(.484176)
PROPDB	% s/w prop and db	−.041529	.342716	−.041529
		(.272438)	(.210215)	(.272438)
PROPCM	% s/w prop and comm	−.708612	.042241	−.708612
		(.324029)	(.242506)	(.324029)
TPBLUE	% s/w from a third-party vendor, all IBM	.328231	.807962	.328231
		(.484904)	(.371453)	(.484904)
TPNONB	% s/w from a third-party vendor, one brand	−.177615	.757258	−.177615
		(.701146)	(.493164)	(.701146)
MPLAT	% s/w from a multiplatform third-party vendor	.442619	.363190	.442619
		(.499505)	(.382307)	(.499505)
Replacement cycle or background variables				
C	Constant	−23.9410	−.8177	−1.7601
		(5.79986)	(.370131)	(.485662)
EMPGRW	% employment growth in 1–3 digit SIC	4.88205	.420278	.933320
		(2.55363)	(.172980)	(.227801)
MINAGE	Age of the newest system	.201929	.007889	.025299
		(.147617)	(.010322)	(.014457)
MAXAGE	Age of the oldest system	.500196	.037131	.001795
		(.116371)	(.008432)	(.011836)
SYSSUM	# of systems	.090544	.096111	.090544
		(.022566)	(.017483)	(.022566)
NONIBM	Major system not IBM-compatible	−.142170	−.444534	−.142170
		(.078144)	(.058747)	(.078144)
MAXMIP	Maximum MIPS of systems	.342612	.003132	.008892
		(.118147)	(.008730)	(.009519)
MINMIP	Minimum MIPS of systems	.616223	.013679	.057616
		(.094417)	(.007025)	(.007935)
Summary statistics				
	Log (likelihood)	−32777.9	−11382.3	−4714.7
	Observations	18567	18567	18031

NOTES: These results are drawn from three separate analyses and are a subset of the total parameter vector in each. Each coefficient reported here is the estimated first triad impact of the variable. Changes over time are reported in Table 4.

The first column has units d(MIPS)/d(variable). In the second column, multiply each coefficient by .37 to get units d(probability of increasing MIPS)/d(variable). In the third, multiply by .15 to get units d(probability of increasing Systems)/d(variable).

Estimated standard errors are in parentheses.

Before we turn to the hypotheses, we note that these tables reveal quite a bit about how much information there is in the data. In particular, the probits are able to determine the coefficients of the replacement cycle variables reasonably precisely. They are, however, not able to determine the coefficients of very many individual software author or usage variables with much precision at all. We can reject, at extremely high degrees of confidence, the hypothesis that either set of software variables taken as a group has constant coefficients over time, or that the coefficients are zero in the second triad. We cannot, however, say much about individual coefficients. Nor is there much difference—in a statistical sense—between the MIPS and Systems probits. On the other hand, the tobit, with its continuous-valued dependent variable, clearly has information to tie down many of the coefficients. Accordingly, we focus discussion on it, noting the few cases where the probits might lead to a different conclusion.

Which Version of Figure 1 Is Correct?

We begin with changes over time in the coefficients of the software usage variables. These are the first panel in Table 4. We have ordered the coefficients so that going down the page corresponds to movements to the right in Figure 1.

The first coefficients show that intensive users of scientific and numerically intensive software reduced their demand for mainframe hardware in the second triad, relative to other kinds of sites. First read the first row of coefficients, those relating to scientific and other number-crunching software, literally. The −12.4 coefficient in the first column means that a 100 percent increase in the percentage of this kind of software would lead to just over 12 fewer MIPS being bought at the site in the second triad. The standard error of about 4.5 suggests that we can estimate this coefficient reasonably precisely. Now, that is not a within-sample change in the variable—a 100 percent SCI mainframe is rare (recall that the operating system and similar management tools are counted in these percentages). But a 50 percent change in this variable is well within the sample range. It corresponds roughly to the difference between a purely data-processing computer and a mostly dedicated number-crunching computer. So the coefficient means that the number-crunching site would decrease its mainframe acquisitions by about 6.2 MIPS (12.4 × .5) deflated between the two triads, compared to other kinds of sites. That is a huge decrease in demand, corresponding to delaying a very large replacement/upgrade project.[24]

24. In the second triad the mean increase in capacity among expanding sites was only a little over 8 MIPS (deflated).

The next two columns refer to the probability of increasing MIPS (rather than the amount of MIPS increase) and the probability of permanent increases in the number of systems. The −.54 in the "MIPS probit" column means that the same 50 percent increase in SCI would lead to a decrease in this probability of 10 percent (.54 × .37 × .5) for a site in the middle of the sample on all the other variables. (The .37 is the probability derivative from the probit evaluated at the sample mean.) Once again, this is the predicted change in behavior between triads for this kind of site in relation to others. Since about a third of the sites upgrade or expand (increase MIPS), 10 percent is a lot of waiting behavior. We are not, however, able to estimate this coefficient with all that much precision, as the large standard error suggests. Finally, the same logic implies that the 50 percent increase in SCI would lead to a decrease in the probability of permanently increasing the number of systems of almost 5 percent (.64 × .15 × .5). Since the sample average for that probability is about 8 percent, this, too, is a huge change in behavior. Once again, the estimate is statistically imprecise.

The coefficient of TS, the technical and engineering software usage, is similar to that of SCI but less precisely estimated in all analyses.

For the rest of the software usage variables, all three specifications tell much the same story. After SCI, the other reasonably precisely estimated coefficient is that of MIPDB—that is, database and dbms (database management systems) tools software running on very large systems. The rest of the coefficients are, on average, negative and not significantly different from zero. Our choice of omitted category (which is, after all, arbitrary) only hides one statistically significant difference: The coefficient of MIPCM is clearly larger than any of the SCI, TS, STD, or DB. Once again, we only have much in the way of statistical precision with the MIPS dependent variable. Finally, the coefficient of COMM is of the same general size as SCI, but much less precisely estimated.

Relying first on only the statistically significant results, there seem to be two facts here. First, the scientific and number-crunching software sites seem to be waiting for new computer architectures, compared to other sites. Second, sites running very large applications on very large computers, those with large MIPCM or large MIPDB, seem to be waiting less than other sites.[25] In between those two extreme groups, there is little informa-

25. In the first triad, a select number of heavy users of database and communication software show accelerated, not slowed, demand, particularly in the MIPS tobit. This period was well into the diffusion of relational databases and real-time query capabilities, as reflected in the DB and COMM coefficients.

tion in the data to tell the rest of the sites apart.[26] They form a large "middle."

In terms of overall waiting for client/server, the number-crunching kinds of sites do not have too much to contribute. Individual sites' behavior is predicted to change a lot, and at least for the scientific categories we can have a good deal of statistical confidence in the size of that change. There are not, however, many of these sites left in the mainframe world by the 1980s, and their aggregate contribution to the downturn in demand is small. The big contribution comes from the difference between the "middle" category and the non-waiters. There is a smaller but still significant difference in behavior between the large MIPCM and MIPDB sites and the "middle" sites. The "middle" category contains many sites, so the aggregate amount of waiting for client/server that it represents is substantial.

These results argue that the right version of Figure 1 is the one in which the adjustment cost curve is steeper; in other words, the adjustment cost theory rather than the best of both worlds theory is true. The important caveat to remember for this result is that it is based on the early part of the competition between the two technologies. There could be differences in expectations between the different kinds of sites about future standardization or software developments.[27]

Lock-In?

In the second part of Tables 4 and 5, the comparable results for the vertical relations software variables appear. For these, the coefficients are estimated far less precisely and the sign pattern varies between the analyses of MIPS and Systems. In the MIPS tobit specification, where there appears to be the most information in the data, the signs are surprising. The negative coefficient on in-house software means that sites which had written their own applications tended to wait for client/server. This is exactly the opposite of what would happen with defensive and powerful MIS. It is

26. To a large extent, this is caused by the nature of the cross-section distribution of computer usage rather than by behavior in this time period. The scientific-computing sites and the MIPDB or MIPCM sites tend to be quite distinct from other sites. The former are typically doing primarily number crunching (rather than a mix of it and other things). The latter are typically using a "transactions processing" kind of application or something like it. If we remove these two groups of sites, it is very hard to see any clear pattern in the remainder of the software usage in the data. The remaining sites tend to do some of all the remaining categories, and not to vary all that much.

27. All of our results could be turned around by appropriate coincidence theories. The large sites which we say have large adjustment costs might instead be those for whom future standardization and future software developments are the most valuable, for example.

further (and weak statistically) evidence against MIS power in organizations (Lucas, 1984).[28]

Similarly, with one exception the vertical relations variables are insignificant and of the wrong sign given the vendor lock-in theory. Sites that have acquired software from their proprietary systems vendor or from single-platform third-party vendors tended to wait more, not less, for client/server than those buying multiplatform software or using consultants. Once again, these effects are quite weak statistically and the sign pattern changes in the systems probit. Overall, the results offer little support for the vendor relations theory.

An important exception is the positive sign on usage of communications software from the proprietary systems vendor. While the coefficient is not statistically significant, the size of the coefficient is consistent with considerable vendor lock-in for this kind of software. Since this important category of software is numerically dominated by IBM products that differ radically from industry-wide data communications products, it is not surprising that this is one area where we detect vendor lock-in. Recent innovation in this area is important enough that users of these services, even as late as the early 1990s, may still be making long-term commitments to mainframes in order to exploit these innovations.

Overall, however, we must conclude that vendor and MIS lock-in are an unimportant explanation of behavior in this period. It simply is not true that the most backward-looking sites are those with a lot of in-house or systems-vendor proprietary software. We were quite surprised by these results. One possible interpretation is that these sites are indeed locked in but expect that their downsizing to client/server will go forward within the client/server products families that are compatible with the products of their historical mainframe vendor.

IBM

We draw attention to one result from the rest of Table 4 because it is so large. In the probits, the coefficients on the non–IBM-compatible sites are large and positive. If we look at Table 5, we see that the same coefficients are negative in the first triad. What this means is that non–IBM-compatible sites used to purchase mainframes less frequently, but that they catch up in our second triad.

28. In more restrictive specifications we tried, the sign is not reversed but the coefficient is estimated much more precisely. There is thus some fragile statistical evidence that in-house measures MIS capability to undertake large, forward-looking (as opposed to defensive) projects. That this effect does not appear in the systems probit underscores its fragility.

This shift in behavior has many possible interpretations in theory, but only a few plausible ones in practice. Given the choice between interpreting this as either "IBM's fortunes got worse" or "its rivals got better," we are tempted more by the former. Here is why:

There is little to suggest that shifts in the competitive position of IBM's mainframe rivals were responsible. For example, little industry evidence suggests that the non-IBM firms innovated dramatically more.[29] We note, as well, that the non-IBM variable could account for characteristics that we have not successfully measured with either the software variables or the other derived demand variables.[30] The variable shift can potentially stand in for any of a number of changes to the non-IBM or IBM network of suppliers, for changes to the software supported by IBM or non-IBM firms, to the quality of the hardware, and so on.

Though we are not out of theoretical possibilities, they seem less plausible than the simple theory that users anticipated a smaller alternative to mainframes: the increasing reliable and capable open system alternatives associated with microprocessor-based systems. The new open alternatives had developed many standard applications by the late 1980s, and the levels and directions of advance were predictable and understood by professionals. In this view, IBM and non-IBM users alike anticipated a future alternative. Both behaved similarly, resulting in similar demand behavior in the latter triad (in contrast with the earlier triad).

Other Determinants of Demand

Many of the rest of the variables are statistically significant in Table 5 but not in Table 4. The magnitude of coefficient estimates for all the other

29. Control Data's attempts at revival were a well-publicized failure. Unisys's victories were largely measured by the ability to stay out of the red. Honeywell, now part of the Bull group, provided no real competition for IBM in general-purpose mainframes by the late 1980s. Despite being swallowed by AT&T, NCR continued its steady, but unspectacular, advances in niches in which it already specialized. DEC's high-flying days were at a well-publicized end by the early 1990s. The advance in systems using vector processors, which came from several high-profile new firms, had hardly dented the mainframe world by 1990. IBM would only feel such an effect for a few select users of extremely large systems. It is unlikely that the imminent diffusion of vector-processor mainframes would affect behavior at more than several score sites at most. Makers of plug-compatibles will feel this demand shift as much as IBM, since they sell exclusively to sites where we record IBM as the dominant supplier.

30. Even with as much data as we have for these sites, there are many possible interpretations of this coefficient, because IBM is both the largest proprietary software vendor and hardware vendor in the mainframe world. Moreover, IBM has the largest user third-party network—that is, an enormous third-party peripheral and software vendor market, large user group communities, its own magazine, and so on.

variables change very little over time. We conclude that these demand factors continued to be a force in the second triad, even though the frequency of capacity increases dropped off considerably. We briefly summarize the site-specific factors.

While employment growth predicts both systems and MIPS growth, these effects do not change over time. It appears that computer systems expand to meet current needs. The age coefficients also do not change much, and illustrate important features of demand. First, there is obsolescence. A newer system appears more sunk because it has had less time to become obsolete, while an older one is less sunk because it has become technically obsolete. Obsolete systems tend to be retired. This sunkness story explains why the coefficients are irrelevant, as the sunkness of existing investments is irrelevant to further expansion. A second interpretation refers to a slow-moving valuation of the stock of capacity. As for MINAGE, sites that have recently invested have a high desire for capacity. Sites that have not invested for some time have not had any needs for a while. As for MAXAGE, sites with a very old system have not felt the need to expand the capacity of some of their applications for some time. For the second interpretation to be the right one, it requires some further explanation about why the age variables do not matter. The likely source of a story is that the lagged endogenous variables have already been put into the stock. Thus, a low "minage" means that the site has already adjusted to desired capacity.

The size of the smallest system is also statistically and economically important. The larger the smallest system, the most likely the user will increase capacity. Inmon (1985) observes that sites whose smallest system is large have placed themselves on a mainframe growth path and rarely deviate. These users are most likely to resist moving away from these long-term commitments. The size of the *largest* system does not predict behavior nearly as well, which is noteworthy in light of its predictive abilities when interacted with software variables.

How Did New Choices Shift Demand for Old Systems?

The future opportunity to downsize cut the rate of (systems) capacity expansions between our two triads. It is by now standard to interpret this as an increase in competition. Yet sellers of the old technology did not act as if they were now in a more competitive industry. Mainframe price/performance ratios, for example, continued to fall at about the same rate as before (Brown and Greenstein, 1994). The largest vendor, IBM, continued to announce ambitious R&D initiatives closely complementary to its existing proprietary products, and resisted until quite recently portability to open systems for its more important software products.[31] It is by now typical to interpret these actions as evidence that mainframe vendors are stupid, or at least backward.[32] An alternative, economic explanation of the pricing and technology behavior is available in our estimates. This explanation turns on a shift inward but not a flattening of the demand for mainframe systems.

The adjustment-cost results of the last section suggest such a story. Traditional inframarginal mainframe customers (for example, those with large MIPDB) *stayed*, while traditional marginal customers (for example, those with large SCI) *moved* or waited. In this section, we examine the implications of our estimates for shifting mainframe demand more systematically.

We order sites by predicted $X\beta$ in each year. Since all sites face the same prices, this should also be their ordering by (the observable portion of) the value of expanded capacity. High $X\beta$ sites will systematically be inframarginal purchasers, for example.

As a first calculation very close to the data, we ask how general the SCI versus MIPDB anecdote is. Has demand fallen because there are fewer high-value, inframarginal customers? That would be demand-curve flattening. Or has it fallen because low-value, marginal customers have shifted away? That would be demand-curve steepening. In Figures 2 and 3, we use $X_{85}\beta_{85}$ and $X_{90}\beta_{90}$ from the MIPS-capacity increase model reported in Tables 4 and 5. On the vertical axis, we graph $X\beta$; on the horizontal axis, the percentage of sites in the sample that have a higher predicted valuation in each year. As can be seen from examination of the graph, particularly from the marked bars, the shift over time in demand appears to be of the demand curve–steepening variety. There is no tendency for inframarginal customers to be the ones who left the market over time. In percentage terms, the decline in high-value sites is somewhat less than the decline in low-value sites. This is the generalization of the SCI versus MIPDB anecdote, and suggests a decline in quantity demanded but not a flatter demand curve.

In an Appendix, we report calculations that move the analysis closer to a theoretical demand curve. The distributional assumptions behind the

31. On the first point, see, for example, the ongoing importance of the SAA and AD/Cycle initiatives. On the second, it was not until spring 1993, for example, that IBM announced a credible policy of moving key database software tools (like CICS) to open systems.

32. See the extensive discussion on the inadequacy of IBM's organizational form, for example.

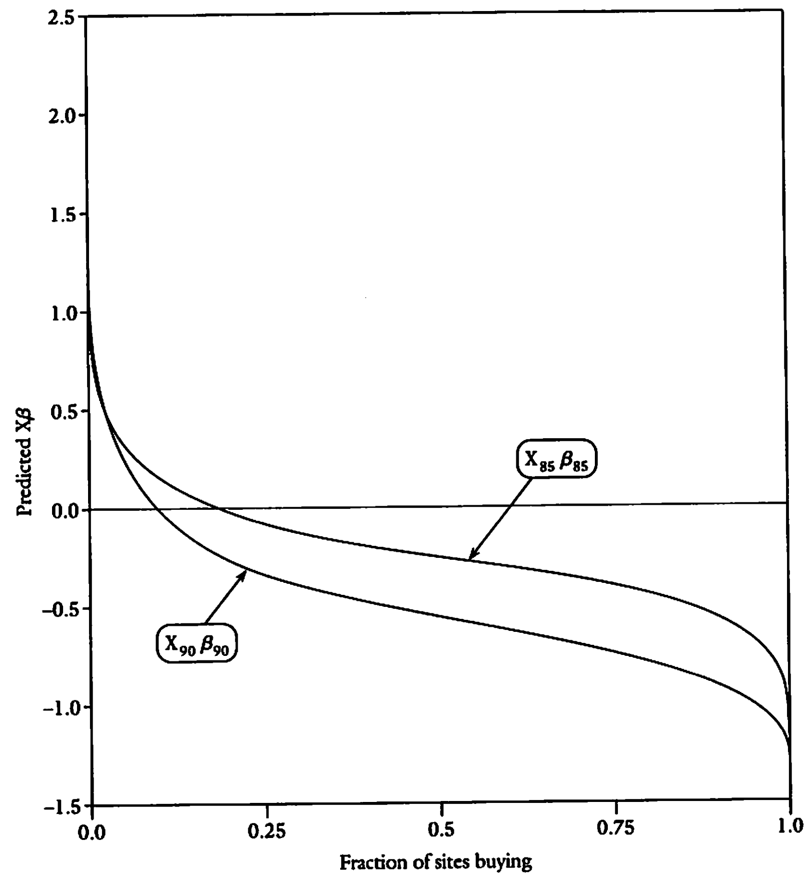

Figure 2. Demand curve implied by probit—fraction of buyers.

probit are relaxed, and the better definitions of predicted quantity demanded and implicit price change are used. The resulting pictures are quite similar to those in Figures 2 and 3.

What changed over time to move the demand curves is closely linked to the increased importance of the outside option, client/server. A simple variance calculation illuminates this. We use the sample distribution of X from the second triad. We take coefficients from the probits and calculate variance statistics with each parameter vector. We find that $X_{90}\beta_{90}$ varies more than $X_{90}\beta_{85}$. The effect of the outside option was not to make the

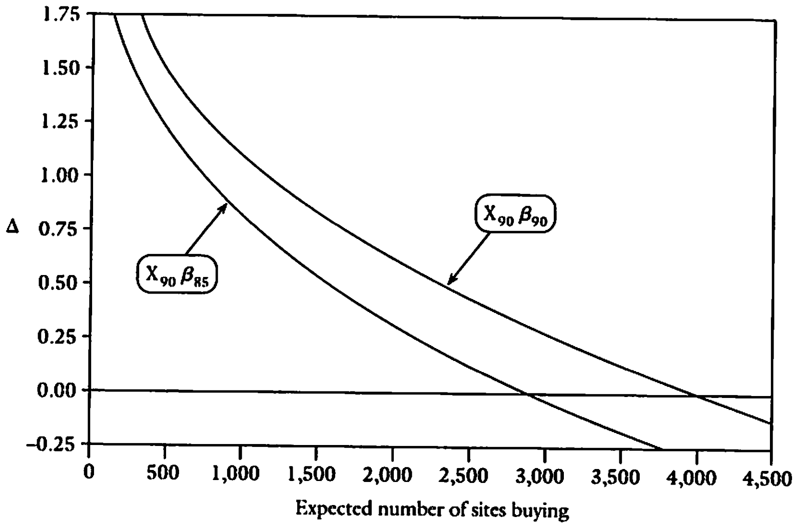

Figure 3. Demand curve implied by probit—expected number of buyers.

sites more alike (reduce variance) as the MIPS arbitrage theory suggests. Instead, the reverse. The demand curves in the figures get steeper because high-value mainframe customers tended not to wait for client/server, while low-value customers waited.

Upshots

While these results are drawn from the early phases of the diffusion of client/server, they resonate with what users think. We propose three interpretations of our results. These relate to the dynamics of investments in large information technology solutions, the commercialization of information technology, and the competitive crash in computing. In each, the technologically active role of the buyer leads to a new interpretation.

The Large-Scale Computing Project as an Investment

We started from the view that expanding capacity for large-scale computing is complex. It calls for new hardware, which is how an expansion project leaves observable tracks in our data set. It calls for new software expenditures. It calls for complementary investments at the site, both within MIS and in the end-user business organization. There is a large body of literature on the management of these investments, but positive

studies of them have been scarce. Our quantitative study of them examines their degree of irreversibility and adjustment costs.

We have found that a large fraction of the investment cost of a large-scale computing project is sunk. Should the need for the project's output disappear (or never appear), reversal of the project will not lead to recovery of these sunk costs. One should expect all the general results about sunk investments, especially the inertia and caution they induce, to hold.

The obvious candidate sunk costs are expenditures on installation and local programming at the site rather than acquired hardware and software. (Hardware can be leased or resold in this market, and software has substantial annual license fees.) Our estimate (in earlier work) of the fraction of investment costs sunk, about four-fifths, is much larger than the fraction of expenditures of a typical MIS department on installation and local programming.[33] Economists frequently draw the distinction between "internal" and "external" adjustment costs. The "external" costs are money spent in the course of making the investment, while the "internal" costs are the disruption to regular business routines that have to be borne while the investment is being made. Since our estimates have the sunk costs too large to be explained in terms of external costs, they suggest internal costs as well.[34] What is interesting about our findings is not that we believe that these costs exist, as that was well established in the descriptive literature. Instead, we emphasize their quantitative importance, roughly as large as the programming expenditure on a large-scale project.

The introduction of a new technological generation, in our case the networked small systems alternative to mainframes, offers an opportunity to study the sources of the adjustment costs. All sites face uncertainty about the future path of technology. When a site shifts from an old technological base to a new one, "legacy" applications matter a good deal. Sites have very different kinds of legacy applications, and as a result can have very different adjustment costs.

We examined two different sets of measures of how legacy applications matter. First, we use software at the site as an indicator of the degree to

33. Surveys of MIS departments reveal that externally acquired hardware and software are well over half the total budget. If we assume (conservatively) that MIS employees and consultants do nothing but big projects, we still get too small a fraction.

34. Some analysts use internal political power language rather than costs language to describe these phenomena. Projects may be difficult to reverse because MIS holds a favored position in the organization after an expensive project is completed, for example. For our purposes, this alternative language is not particularly different. Obviously, the distinction matters a great deal for the practical marketing of downsizing solutions, and so on.

which the site is tied to a particular systems vendor's technology and of the possibility of MIS lock-in. We contrast, for example, sites using much software acquired from their systems vendor with those using third-party software. To our very considerable surprise, the sites more closely tied to the vendor do not appear to be more reluctant to move forward to the new technology. Neither does MIS lock-in appear to be an important problem.[35] In contrast, variation in the application of software does predict failure to adjust quickly. The pattern closely follows that suggested by the organizational adjustment costs model. More complex organizations (like those using big database management system applications) adjust much more slowly than simple ones (like number-crunching sites). We conclude that many of the sources of slow adjustment are in the adjusting organization. User relations problems, not vendor relations problems, appear to be the source of slow adjustment.

While these results refer to mainframe-based computing, we suspect that they apply with little alteration to large projects based on wide-area network or client/server technology. (These are much harder to study in a systematic way at the present time.) To the extent that these newer enabling information technologies gain their value in use by changing business practices, they will be characterized by sunk internal adjustment costs.

The Commercialization of Information Technology

In information technology, as in many other areas, a sustained high rate of technical progress by inventors is not the same as large continuing social gains from use of the technology. The problem of commercialization intervenes. Computer and networking hardware and software are enabling technologies, and the costs of bringing them into use will affect behavior. For information technology, the commercialization problem can be summarized as a very high rate of technical progress in hardware, a reasonably high rate of return in marketed software, and often painfully slow complementary investment in new software and business practices at end-user sites. The last portion has limited economies of scale because of the variety of business practices in a highly decentralized economy and is also characterized by sunkness.

The primary behavioral implications of sunk costs are inertia and caution. We see both in the demand for large-scale computing. All these are rational responses to sunk costs: caution before moving to a new technology, inertia in staying with an old technology, and even caution in making new commitments to an old technology when a new one may be arriving. All of these behaviors are evident in the late period of mainframe usage. The inertia and caution in this case must ultimately break and permit movement to new technologies, at least with regard to hardware. It appears that the transition era is characterized by great technological uncertainty; the theory suggests that this will lead to more caution.

35. This confirms the general finding in the organizational literature since Lucas (1984) that MIS has little internal political power.

A variety of market responses to this problem are in evidence. Consider the recent market successes of system integrators and consultants. Expertise in making the adjustment to new technological opportunities certainly lowers external adjustment costs.[36] In this regard, system integrators and consultants are a mechanism for gaining economies of scale in the on-site portion of information technology investments. In the old industrial organization of information technology, this expertise often could be found in the systems vendor. As information technology moves to a more open-systems arrangement, that source becomes correspondingly less important. This leaves a market opportunity for system integrators and consultants, and quite possibly for sellers of proprietary software.

Yet system integrators, consultants, and the sellers of systems, networking, and database management system software cannot make the internal adjustment costs less sunk, nor can they fundamentally reduce uncertainty about future technical developments. The internal adjustment costs arise from the need to make valuable organizational changes to get the biggest advantages of information technology, a problem that is not going away.[37] This view implies that the current transition era in information technology is not just a time of technical change and the emergence of new standards. Instead, it is a period of definition of new market institutions for commercialization.

Once again, there is every reason to believe that the shift to wide-area network and client/server technologies will increase these forces rather than make them go away. The span of cutting-edge information technology investments is increasing to cover more technologies, more vendor companies, and more markets.

36. It may lower the internal (disruption) costs as well, though this assertion is more controversial.

37. The degree of future technological uncertainty will certainly decline with time as standards for the post–competitive crash era are set. This will reduce the purely technical role of system integrators and consultants, but probably not their adjustment cost–lowering role.

The Competitive Crash

Two technological/economic stories of nascent competition between old and new types of computer systems and between the kinds of companies that sell them have circulated widely among technologists and in the trade press. Both are wrong. What is instead right is not yet completely clear, but the behavior of customers in the early stages of the competitive crash gives many useful clues.

The "MIPS arbitrage" theory correctly identifies an important driver behind the competitive crash, increases in the capabilities of the largest microprocessor-based systems and in networks of microprocessor-based systems. Yet the theory is seriously incomplete in that it ignores a product-differentiation advantage of mainframe software. In our estimates, the size of the market for mainframe systems declines with competition but the degree of market power does not. Most mainframe brands continue to be monopolies, albeit over a smaller body of inframarginal customers.[38] The best of both worlds theory expected that client/server architectures would quickly solve the long-standing user relations problem. To be sure, the user relations problem is more likely to be solved sometime in the future than it was in the past. Yet the view that it was going to be solved quickly by combining the strengths of servers with the strengths of clients was more a fond hope than a technological and organizational reality. At least in the early going, exactly the sites that would benefit least from these advantages were the fastest to switch.[39] Buyers appear to have viewed the advanced claims for client/server architectures with real suspicion.

What instead is actually true? The dynamics of user behavior affected the early competition between the old and new computer systems in a variety of ways. First, the readiness of buyers to wait for new technologies they could not yet use was a huge revenue and public relations shock to old-system suppliers. This was partially offset by their continued ability to command a substantial price premium for their products—the market power alluded to above. A more important offset was the very slow pace of the transition to the new world. This left sellers of the old technology a number of years to come to interpret events and to organize technology and marketing for a competitive response. This "breathing space" may well be important for the future structure of the information technology industry.

Further evidence of a very different kind comes from the supply behavior of vendors. First, the failure of the vertical-relations model as an explanation of preference for specific old vendors is an important part of our story. If we are correct, then old-line vendors should be abandoning the "account management" marketing strategy. That strategy focuses on extracting rents from the existing base of locked-in customers. The switch of most old-style vendors to a somewhat more open-systems approach, while late, suggests that they see the same environment we do. Most current discussion of the old-line vendors discusses the inefficacy and slowness of their decision making. The slow transition to a new technological base is "breathing space" to them and permits these changes of strategies to be visible despite how slowly they have occurred.

Second, the adjustment costs appear to be inherent in the problem of making effective use of the new technologies in large applications. If this is correct, it suggests that old vendors' behavior should change; they should now see the source of their rents in service and in software products that run on large systems or networks. The same argument suggests that new vendors—of database management systems and tools, and (especially) systems integration services—may pursue the same rents. Once again, this is a recognizable description of parts of the technology strategy of old-line vendors, their competitors in open-systems software markets, and systems integrators. Supply behavior as well as demand behavior is consistent with the story.

Our analysis of all three topics is limited by essentially the same problems, and these await further research. We study the very early period in the diffusion of client/server technology. We have little to say about technological expectations, in particular about waiting for software tools and the setting of new standards. Yet we want to finish by emphasizing the element of continuity in behavior we observe, which leads us to believe that the world will not quickly change to make us wrong. A long series of technical initiatives have dramatically increased the potential range of useful information technology applications. Achieving that potential has always been difficult and therefore slow.

38. By late 1993, the trade press had caught on to this, as discussed in a number of citations above. It does not speak particularly well of client/server vendors that they needed to be berated this late in the transition for using MIPS arbitrage arguments for marketing purposes. The falsity of that view was evident in buyers' behavior as early as 1990, and could be clearly heard in the first phase of the Bresnahan-Saloner (1995) interviews with buyers in late 1992.

39. Seeing whether this persists into the 1990s is one very good reason for our current investigation of more recent data. As of the second wave of Bresnahan-Saloner (1995) interviews in spring 1993, there were some interesting exceptions, but this described the overall pattern quite well. The exceptions, for example in the marketing departments of telecommunications companies, related to the value of best of both worlds–style solutions as a reaction to a radical change in competitive circumstances.

APPENDIX *Statistical Models*

Our statistical models for increases in capacity at the site (s) level take the form:

$$V_{s85} = X_{s85}\beta_{85} + \epsilon_{s85}. \quad (1)$$

$$V_{s90} = X_{s90}\beta_{90} + \epsilon_{s90}. \quad (2)$$

So the expected aggregate demand for systems expansion in each triad is calculated by first predicting the probability that each site will expand, then adding them up:

$$Q_{85} = \Sigma_s \text{ Prob (Expand)}_{s85} = \Sigma_s F_{s85} (X_{s85}\beta_{s85}). \quad (3)$$

$$Q_{90} = \Sigma_s \text{ Prob (Expand)}_{s90} = \Sigma_s F_{s90} (X_{s90}\beta_{s90}). \quad (4)$$

It is clear from this definition that the predicted aggregate demand curve shape is determined by two forces. The first is the shape of the distribution function for unobserved heterogeneity, as has been emphasized much in recent modeling work.[40] The second force is the distribution of *observed* heterogeneity in valuation, $X\beta$. Since our data do not contain variation in prices, we can say very little about the first force. Instead, we will try to make inferences that are robust to assumptions about the shape of $F\epsilon$, and to changes in its shape over time. Our estimates do contain a good deal of information about the changing distribution of $X\beta$ over time.

First, assume for a moment that the statistical assumption behind Tables 4 and 5 is correct—that is, that the shape of $F\epsilon$ is unit normal for both 1985 and 1990. Then, we can calculate (3) and (4) to get predicted demand. To hypothetically change prices, we add a fixed "value deviation," Δ, to the valuation of all customers. We recalculate (3) and (4) for a wide variety of Δ, using

$$Q_y(\Delta) = \Sigma_s \phi(X_{sy}\beta_y + \Delta).$$

Now, since we do not estimate a price coefficient, we do not know whether Δ has units of pennies or billions of dollars. Yet it does correspond to a hypothetical price change. We could show the results with Δ on the vertical axis and $Q_y(\Delta)$ on the horizontal axis for y = 1985, 1990. The shape of this is the same as Figure 3, and the tendency to steeper demand over time is preserved.

This analysis relies heavily on the assumption that the errors are known to have a specific shape, the normal, and that the shape is the same over time. Neither assumption can be defended on a priori grounds. To relax them both, we move to semi-nonparametric estimation. In particular, we approximate $F\epsilon_{85}(X_{85}\beta_{85})$ and $F\epsilon_{90}(X_{90}\beta_{90})$ by a fifth-order series approximation around the unit normal. This allows quite different shapes, and we also allow the approximation to vary between the two triads. This body of reported demand curve findings changes very little if we replace the MIPS definition of capacity changes with the coarser systems definition. These figures are available from the authors.

40. See, for example, Berry, Levinsohn, and Pakes (1993) for the importance of this in determining the demand elasticities.

Also, the body of results changes little if we use different conditioning assumptions. Obviously, the distribution of X across sites is an important determinant of $X\beta$. The distribution of X changes quite little between our two triads. The important changes are that the macroeconomic downturn is reflected in smaller employment growth figures and that the sites have more powerful mainframe computers but slightly fewer of them five years later. We have redrawn the figures using the same Xs for both years, and this, too, makes little difference.