week 1
2 of 4

4
Excuses

The Trouble with Computers
Landauer, T. (1995)

Given that the computer business has become a major industry, employing in one way or another well over a million people and causing hundreds of billions of dollars a year to change hands, it's hard to accept the implication that these machines bring no net improvement in economic efficiency. Certainly they must be delivering something of value to the millions of people plunking down hard cash for them. Since computers seem to have been designed and intended as productivity tools, to help work get done more efficiently, productivity would seem to be the first place to look. Thus it is tempting to try to find a way to deny appearances. Here is a bunch of good tries.[1]

Counterarguments

Perhaps the easiest way out is to deny the conclusion and to assert that despite the data, computers do or must contribute substantially to productivity, or at least to net well-being. Three forms of this argument are heard most often; one relies on the undeniable fact of computer power, the second on the undeniable fact of computer popularity, the last on proponents' undeniable personal enthusiasm.

The Productivity of Computer Manufacture

In its simplest form, this argument states the obvious: computers are wonderful machines that can do marvelous things; therefore, they must improve tasks to which they are seriously and vigorously applied, including productivity. Q.E.D.

The most sophisticated form of this argument was put forward by Stanford economist Timothy Bresnehan (1986). He says that when some popular product starts out expensive and later becomes cheap, there is a social welfare spillover; consumers reap the benefits. If because of technical or manufacturing improvements people now pay $1 for something for which someone used to pay $11, the world must have gained $10 in value received for each additional unit purchased at the new price. If we multiply by ten the increased number of purchases at the $1 price, we can calculate a welfare gain from productivity. Applying this reasoning to computers, Bresnehan totes up how much the prices per million calculations per second (CPU power in megaflops, or MIPs) and per million stored data items (memory size in megabytes) decreased between 1958 and 1972, using low and high estimates of 19 percent and 34 percent per year price improvements. Bresnehan used purchase data from the financial services sector from 1958 to 1972, a period when regulatory and business conditions in the industry were relatively stable, as was the basic computer technology and its modes of application. Financial services, the heaviest users of computers in those years, used computers operated by service bureaus to do bookkeeping and record keeping. He multiplies the number of new megaflops and megabytes of computer power delivered to the industry each year by their price declines and comes up with an enormous productivity improvement contributed to the overall economy. By these calculations, the value added by computers was over five times their cost. By extrapolation he concludes that "in current (1986) terms, the downstream benefits of technical progress in mainframe computers since 1958 are conservatively estimated at 1.5 to 2 orders of magnitude [thirty-two to one hundred times] larger than expenditures.

What should we make of this extraordinary conclusion? It certainly paints a different picture from the superficial productivity and business data and from the other econometric analyses as well. It finesses the use of suspect measures that underestimate productivity in financial service industries. However, the analysis has some glaring defects. First, consider overall plausibility. If we update Bresnehan's figures to 1995 and extend to all computer uses, and there is nothing in his logic to discourage us from so doing, we get astronomical numbers. Total cumulative CPU and

memory purchases have mounted to well over $1 trillion, and their unit price has declined by another factor of ten or fifteen. Following Bresnehan's reasoning, the accumulated welfare benefit of computers in the United States comes to at least $300 trillion, their current beneficence to some $40 trillion a year, or over six times GNP. The average person's real standard of living is seven times what it was in 1958 because of the added pleasures and convenience of life supplied gratis by computers!

Bresnehan's method seems silly. Here is an only slightly facetious example of where it could lead. In the late 1950s, all-aluminum skis came on the market. They sold for approximately the same price as the best wooden skis. Sales figures and race results showed that they were not quite as good, and they eventually disappeared, but for a while they sold quite briskly. A purchaser of aluminum skis would have fallen heir to one of history's most remarkable productivity improvements. One hundred and fifty years earlier, the aluminum of which they were made would have cost about $35,000 instead of $3. By Bresnehan's logic, this saving in the manufacturer's potential cost was passed on as a welfare benefit to each buyer. How I regret having bought wooden skis.

Let us apply this analysis more seriously to computers. Suppose my new workstation were actually worth $50,000 in extra output from me over its lifetime, and it cost my employer $10,000 to purchase and $10,000 to support, for a net gain of $30,000. Fifteen years ago I could have bought a maxi-mini and terminal with comparable capacity and almost as good functions for, say, $200,000 purchase and $100,000 support, for a loss of $250,000 relative to my $50,000 increase in output. Obviously a wise employer would have rejected my purchase request in 1980 but granted it with smiles this year. How much benefit has come from the reduction in hardware costs? By Bresnehan's method, the answer is the difference in price, $200,000 − $10,000 = $190,000. But it seems much more sensible to credit the machine with only the $50,000 − $20,000 = $30,000 net increase in output it is responsible for. What happened to that big price decrease Bresnehan wanted to count as a present to society? It never existed. My employer wouldn't have bought me a computer at the higher price; it had negative utility. The "savings" is totally fictional. It's a lot like the $100 my mother-in-law "saves" when she pays $100 for a dress she would never have dreamed of buying at the asserted "suggested retail" of $200.

The same story applies to historical computer purchases in the financial service industries. In 1958, only the largest firms with the greatest volumes and highest manual processing costs would have seen a gain in buying computers (if such there was, as it turned out). Their net utility for the purchase would be the cost savings from doing transactions with their help. It might have been estimated at, say, 10 percent (we don't really know; this is for illustration). For other firms, the net utility, given the total costs of hardware, software, and fixed support expense, was negative, and they sat on their hands. As hardware prices dropped, more firms saw the same 10 percent positive margin and jumped in, so we get the same gradual saturation of the market as the hardware component of prices drops, and processing by computer gets cheaper than processing by hand, but nowhere does the full difference between the old and new price appear in anyone's pocket.

The flaw in Bresnehan's calculation is not its underlying logic. Indeed, applied to food or cotton thread, where the true value to the consumer of the output is more obvious, and—at least in the short run—increases in proportion to the amount of the goods a person can get her hands on, I have no quarrel with it. The problem is in taking megaflops as the index of utility. This is not what purchasers were buying. They were buying computers to help with their businesses. By assuming a direct relation between megaflops and "utility to someone" Bresnehan is able to compute that "someone" must have gained a lot by the decrease in price and increase in demand for megaflops. In a sense, this reasoning is circular. Only if computation measured in megaflops is assumed to be directly valuable to the world will Bresnehan's method tell us how valuable it was.[2]

Holding my five-year-old daughter, I presented our two coach tickets for the cross-country flight. The agent asked, "Are you two traveling together?" Me: "Yes." He: "You've been booked into rows 13 and 37. Would you prefer to sit together?" Me: "Yes indeed; I wonder how we got separate seats?" He: "Face it, computers don't have much compassion."

Despite what advertisements claim, the proper measure of utility of a computer is not its numbers of calculations and stored bytes but the speed, convenience, and effectiveness with which it can serve some useful purpose for the buyer. Flops and bytes are not even a computer's intermediate product—information—much less its economically meaningful product—what can be done with the information. Most of the new capacity of PCs has been devoted to graphics and user interface improvements, and these do improve their usefulness and usability somewhat. The comparison between Gould's early and Card's later text editors, for example, showed an increase in work efficiency of about 25 percent (but none in letter quality) during a period in which byte and flop capacity improved by at least 600 percent. Similar comparisons are sometimes better but never by an enormous margin. Often poor or ill-aimed software uses much more capacity to deliver the same information processing functionality. For example, the search algorithms in most commercial word processors need one hundred times the hardware capacity to deliver the same retrieval speeds as the algorithms used by Bellcore's SuperBook text browser. In addition, new information processing functions provided with new capacity are not necessarily of great value. Many programs use the tremendous capacity lately available to offer computationally expensive, but for the user merely pleasant, graphic decorations, or hundreds of features that are rarely used by anyone. Most computers are never used at full capacity and are idle or lightly used most of the time. Bresnehan's theory implies that computation itself was a valuable end product, and that getting it cheaper was necessarily a boon. Not so. Calculating the utility for price of computers that way is like valuing an automobile by its engine's peak rpm—not even its horsepower—rather than by its utility as a means of transportation and enjoyment.

Popularity and Sales

Computers have sold like few other technologies and have found their way into important functions of almost all of modern society's activities. All the largest economic entities—governments, armies, huge industries—rely on them utterly. They must be good.

Let's take first our main successful example, the telecommunications industry. The nation's telephone companies were among the first to install massive record-keeping computer machinery. It was probably a matter of survival, at least if they were to fulfill their mission of universal service; the growth in customers and business was so phenomenal that it would probably have been impossible to keep track of all the things that needed to be kept track of without computers (although an awfully good job of it was done for an awfully long time). The same is probably true of the large systems used by many other organizations whose information work load has expanded enormously in recent years: banks, insurance companies, airlines, stock and commodity markets, the IRS and the Social Security Administration, for example. Certainly computers must be just as critically useful and productive for them. But there is a danger here of drawing a doubtful, circular conclusion. As Joseph Weizenbaum observed in his influential 1976 book, *Computer Power and Human Reason*, what such facts tell us is merely that these organizations could not function the way they do today without computers—not that they could not have found some other, perhaps equally satisfactory way to perform equally valuable functions.[3] Who is to say that the particular functions firms have evolved based on heavy use of computers are better than others that might have evolved in their absence? As we have seen, the overall economic impact of the chosen course offers no stunning proof of its superiority, so the mere fact that the large, complex organizations and processes that exist today are dependent on computers does not prove that the computers are productive. The huge pyramid building enterprises of ancient Egypt were as utterly dependent on enormous quantities of slave labor as our own enterprises are on computers, yet few would argue that slave labor must have been especially productive or that it improved the lot of humankind.[4]

The argument from sales is equally seductive—and similarly fallacious. It goes thus: Surely large, successful businesses would not have invested billions, and continued to invest billions, if the technology was not a boon to them. The argument begs the question, assumes its own answer. By this line of reasoning, no business decisions, at least no consensual decisions of the whole business community, can ever be wrong, and no industry or national economy can suffer productivity stagnation or major competitive losses.

More pointedly, there are documented cases of widespread adoption of technology that turned out to be unproductive. For example, in the early 1980s (Baily and Chakrabarti 1988) the U.S. electric power industry, for a few years a rival of telecommunications as an exception to the productivity crisis, invested huge sums in two promising new kinds of generating plants, supercritical steam turbines and nuclear. Both turned out to be no more productive than their predecessors—the first because they never operated reliably at predicted capacity and were unexpectedly hard to maintain, the second because unanticipated technical and political problems ballooned their construction and maintenance costs. We cannot assume that a large investment is proof of its own economic wisdom.

Individual Testimonials

Many people are certain that computers are productive. Among them are people who sell computers, ones who buy and use them, and many who do computer research and development. It's probably best to ignore the sales promoters, who may tend to exaggerate. (Following news of the productivity paradox, software companies began advertising their wares as "productivity tools.") The opinions of buyers and users are more interesting, if fairly predictable. Those of professionals responsible for the quality of computers should be the most objective and well articulated.

Insider Testimonials

Early on in this endeavor, I found myself doubting my own pessimistic conclusions. All the negative results on computer productivity did not convince me in the face of their obvious popularity, so I tried to calibrate my attitude against those of colleagues in the computer business. My area within computer science, human-computer interaction, is the fastest growing subspecialty and arguably the one most closely involved with issues of what computers are designed to accomplish, how well they can be used, and with their design and evaluation from the user's point of view. I tested my views on this community in two ways. One was by giving talks on the theme of this book to audiences—numbering by now several thousand researchers and practitioners in the field—at various conferences and meetings. Perhaps more interesting is a second effort. In late 1990, I began an electronic discussion with two special subsets of people in the field: an informal organization whose participants come from about a dozen research laboratories, about half at universities and half in industry, that are among the leaders in this field, and the group of officers and volunteers who run an international professional organization for this field—the Computer-Human Interaction Special Interest Group of the Association for Computing Machinery—and its annual meeting. Participants in both sets keep in touch by electronic mail. I sent out a broadcast message briefly outlining the lack of evidence for productivity gains from computers and raising the issue of what they thought of this—its validity and the reasons for it if true. In follow-up messages I asked for any evidence of major productivity gains from applications of computers.

The lively electronic mail discussion that ensued grew to include many knowledgeable people around the world to whom the query was forwarded. A certain number of my colleagues did not like my attitude; they thought that computers are of such obvious and huge value that my questions must be either frivolous or mischievous. Others took my question in a different sense than I had intended, and described how decisions about computer purchase and support are actually made in large organizations (virtually never on evidence of productivity or profitability enhancement), how to market computer systems and what causes them to be bought, or how to convince upper management in computer or software firms that the added value of usability was worth the expenditure to support people like us. Many offered examples of computer applications that they thought were superb and that they were sure must be great productivity enhancers, although they admitted that they were unable to marshal any hard data to support this opinion. (My own candidate in this genre is electronic mail. The ability to have this wide discussion with experts of all kinds from all over the world in a matter of days, to do it all with little effort, and to have the results in a form that I can easily search and edit seems utterly magical to me. It must be an enormous boon to my productivity, although I have no evidence to prove it.) Another common theme of my respondents was that the compulsion to find

quantitative proof of productivity gains was misplaced. The values of computers are to be assessed some other way, they urged. Academic colleagues stressed esthetic and scientific values and left me feeling rather crass and commercial; business colleagues stressed competitive advantages and left me feeling hopelessly idealistic.

A few people raised examples of cases with apparently irrefutable economic value. In every instance, however, these cases had to be chalked up as victories in automation or use of computers to do work that humans couldn't. No one proffered solid cases with supporting data for augmentation tools, though these are the kinds of computer programs on which my correspondents and their organizations all work and which are currently needed for large productivity gains.

The flavor of the debate is captured in some samples, somewhat paraphrased and interpreted by me, of this correspondence:

Donald Norman, former chair of the Cognitive Science Program, University of California at San Diego, now Apple Fellow, is the author or editor of several books on computer design. After citing an article in *Macweek* (March 13, 1990) that Peat Marwick, an accounting firm, believes its heavy use of computers is worthwhile but cannot document it, Norman goes on to say, "We live in a society obsessed with quantification, but we forget that in order to quantify one must extract the problem and do some measure that can be counted . . . and quite often in doing so, we have thereby lost the important aspect of the behavior. To quantify the savings from computers is also to ignore the changes in morale, in style, in kinds of [things] being performed that otherwise could have not been done." Norman declares that he is not against quantification and assessment; quite the contrary, he thinks they are very important. However, he believes that good things are done that are hard to quantify and that it is important to appreciate and promote qualitative values in the absence of quantitative measures.

I agree. I do not wish my case to be narrowly interpreted as a request for proof only by dollars and cents returns. On the other hand, I do not want to adopt technological solutions solely because proponents claim them to be good. There are plenty of other kinds of objective evidence besides accounting cost savings or profits. There are good ways to find out if the quality of life is improved by the addition of a computer system.

We saw a modestly positive example in the study of telephone company mechanization by Kraut and Dumais. The field needs a whole range of good means to tell whether it is accomplishing its objectives.

Jonathan Grudin, a prolific author of papers on usability engineering and a trenchant critic of systems that don't deliver, has worked at MCC (the Microelectronics and Computer Consortium in San Antonio), Wang research, and the Computer Science Department, University of California at Irvine: "I have never used a secretary. I produce a lot more now. Just typing up a reference section used to be painful enough to make me stop reading anything. A lot of professional writers use word processors, too, probably not just to be chic. Cheer up. It's not all for naught. . . . Another real productivity booster, perhaps, is this ability to transfer documents, with formatting, graphics, scanned-figures, etc. . . . I find it far easier . . . but more important it appears to be more motivating for the recipient to receive a wonderful looking document."

He makes a good point. But how clear is it that he couldn't have gotten equally good results with a typewriter, copy machine, overnight mail service (although these probably owe their existence to some computer technology), and a part-time assistant a little greater cost than his computer facilities and heavily subsidized network?

Ruven Brooks is a usability engineer who works for Schlumberger Laboratory for Computer Science: "Almost all off-shore oil fields are a result of computer processing. These fields have all been developed in the '60's or later, after Texas Instruments pioneered digital recording of seismic data. It is probably possible to track down actual dollar estimates for many of these off-shore fields." A nice score for the "doing something you couldn't without computers" box. The analysis of seismic data is similar in spirit to medical imaging and requires enormous numbers of calculations that couldn't be done by humans. It is indeed the kind of application that has had positive productivity consequences. But note that it was introduced over twenty years ago.

Industry Leader Testimonials

"Everybody's secretary must have a 486 chip in his or her PC because it's much faster. And the question becomes, So what? The metrics for measuring this kind of productivity are not very good" Martin Stein, vice chairman, BankAmerica (National Research Council 1994a)

A committee of the National Research Council, the government and public advising arm of the National Academy of Science, was commissioned to "study the impact of information technology on the performance of service activities." For a major portion of their evidence, the committee conducted a survey of eighty high-placed executives from forty-six companies in seven major service industries. The companies were chosen for size, rate of growth, profitability, innovation, and heavy, reputedly effective use of IT. Just under half of the interviewed executives were chiefs of computing divisions; most of the rest were CEOs. Their advice and example on how to deploy IT is valuable. Unsurprisingly, only a few said they had wasted their company's money on IT. A strong majority claimed their returns were at least comparable to other investments.[5] Here is a sample of their comments—some positive, some not.

J. Raymond Caron, president of CIGNA Systems, told the committee, "When it comes to PC platforms . . . we find it very difficult to develop a useful cost-benefit measure. We've taken the position that we shouldn't waste time trying to do it."

By contrast, James Stewart, executive vice president and chief financial officer of CIGNA Corporation, said, "I'm not yet convinced that dispersal and utilization of PC-based technology have proven to be efficient. . . . I see increasing expenditures for what I perceive occasionally as 'toys in the business world' which don't add up to measurable output or improve our results."

Richard Liebhaber, executive vice president and chief information officer for MCI, said: "It [a system that allows customer service agents to pull up all correspondence on bills] is an opportunity to improve the image, the feel, the touch of MCI. Intuitively, I would tell you it's going to provide market share, customer satisfaction, bottom line. I can't prove it. I'm not going to waste my time trying."

And Robert Elmore, partner and worldwide director of the Business Systems Consulting Group at Arthur Anderson, was quoted thus: "Strategic use of information is often dealing with a 100-to-1 return or at least a 10-to-1 return, and in these cases returns should be obvious. For example our electronic bulletin board was installed about four years ago, and there is no question among our practitioners that it gives us a competitive advantage and that it has made a direct contribution to sales growth."

What to Make of the Opinions

The more enthusiastic of the comments resemble the content of large numbers of articles in computer and business magazines, as well as in a plethora of advice books for managers. Most assert that computers are of terrific value, and many appealing examples are offered, supported by the author's experience and reasoned enthusiasm. I by no means deny the truth of these claims. However, the well-known psychological phenomenon of dissonance reduction—when people have invested heavily in a decision, they are reluctant to judge it a failure—must be reckoned with. So must the many failure stories and negative opinions, which offset the hopeful words and examples. The real problem is separating the wheat from the chaff. If we judge on the basis of the overall economic results of computers, the weighted sum of the effects of all systems put together is near zero, so probably many are good, many are bad, and many are neutral. In all likelihood the proportion of good ideas among those testified to by experts is better than the average of ideas peddled. But, especially in computer technology where hyperbole, fad, and unfulfilled predictions are as dense as pollen in September, there are too many examples of highly touted and widely admired schemes that have turned out to be losers to rest the case on expert testimonials.

"Salomon Brothers expects that employees trained under this new arrangement will complete transactions more quickly" (Gabor 1992).

My survey of case studies of business uses of computers in the popular press, trade magazines, and academic business journals has yielded the following rule: headlines and titles announcing large productivity gains are almost always followed by an article about the gain a firm "expects" or "projects." Those announcing disappointing results are all based on retrospective information about actual performance. Even in the paper telling of "340 percent" productivity gains from centralizing typing, the

bottom line was in the future tense: "Holding the staff steady will more than compensate for the cost of word processing equipment and cost savings will increase in the future" (O'Neal 1976). Of course, few firms install systems without thinking positively, so stories of rosy prospects are not surprising. However, the dearth of rosy histories is somewhat shocking.

We need more and better evidence. It would be unwise to ignore the judgment of so many buyers, so many experienced business people, clever technologists, and wise scholars. There must be a pony in there somewhere. Perhaps all we need is better ways to get better information, better knowledge of what works, what doesn't, and what to do.

Measurement Problems

Another appealing way to make the problem vanish is to reject the validity or importance of measured productivity. The recent National Research Council reports rely heavily on this excuse to dismiss the productivity paradox and turn to other questions. In essence they say, "the productivity statistics are not entirely accurate, so we will discount the hard-to-believe story they tell."

Economists readily admit that productivity measures are sometimes crude, indirect, and insensitive. One criticism points to the way that government agencies collect and analyze the data. The recent decrease in productivity growth is primarily in service industries, and output in services is especially hard to measure because the product does not stay the same over time. Twenty years ago most banks were highly regulated, ultraconservative holders of large savings accounts and reluctant lenders to well-scrubbed borrowers. Now they are purveyors of vast quantities of consumer, business, and real estate credit to the masses and managers of huge pension funds. Are their earnings on deposits and loans paying for the same services as they were of yore? How do you value the output of insurance companies equivalently when they begin selling many more different kinds of policies or of investment brokers when they move into instruments like mutual funds? There are indeed great difficulties, sufficient to cause one of the agencies that puts together these indexes, the Bureau of Labor Statistics, to throw up its hands in measuring some industries, in particular financial services, for which it just equates output with labor input. To pose the matter even more simply, what is the relative value of a cash withdrawal transaction with an ATM and with a traditional teller? Suppose you're one of those who think that the ATM provides more convenience or less waiting time. If a thousand ATM money withdrawal transactions cost half as much in labor time and capital as a thousand teller money withdrawals, is "productivity" only double? Suppose you prefer a human teller. Is the new service worth just half the old?

What, exactly, is the product? The assessment of comparable output is usually easier, more intuitively obvious, in goods manufacturing or construction. The statisticians can count the "constant dollar" production of 200-horsepower four-door passenger automobiles or 2000-square-foot single-family houses. Of course, even that's not completely straightforward. Automobiles added seat belts and fuel injection; houses added double-glazed windows and central air-conditioning. Still, measurement is obviously more slippery in services. Establishing the correct constant dollar—that is, deflating the value of goods and services—is difficult as well. There is no guarantee that the same deflator (e.g., the Labor Department's market basket measure) stays the same relative to different products as the mix of products consumed by the public changes, or that it is equally applicable to goods and services. In some studies, different deflators are used for different industries; the opportunity for error is all too apparent.

Maybe productivity has been improving in services as computers have been brought in to help but hasn't been measured. Economists who have studied and modeled service productivity have taken the matter seriously, and most believe that there could be real unmeasured benefits of IT in services, especially in financial services. However, most also agree that mismeasurement cannot be the principal villain. John Kendrick, professor of economics at George Washington University, says, "Even after allowances for understatement, however, it seems clear that productivity growth in most services groups, other than communications and trade, has been significantly below the U.S. business sector average" (Kendrick 1988). Edward Denison, senior fellow of the Brookings Institution agrees: "I consider it unlikely that such [measurement] errors were large

relative to the observed change" (Denison 1989). The fact is that there has always been a certain amount of slop and underestimate in measuring services, but they used to show healthy productivity growth anyway. As Roach (1988) says, "It seems an extremely remote possibility that the service sector's productivity disappointments of the 1980s would disappear if the numbers could, in fact, be corrected for statistical problems that date back into the 1950s." The measurement problems don't appear to be much different since the advent of computers.

Mr. K needed to know how to get a brochure about his company benefits and had an hour in which to get the answer. He called the human resources office and was 'ransferred to the voice mail of a counselor, for whom he left a call-back message. After twenty minutes he called again and talked to a different recorded voice. And a third. After an hour, he gave up.

Baily and Robert Gordon, of Northwestern University, who examined the mismeasurement issue in detail (Baily and Gordon 1988), found fault with many aspects of the standard measures of productivity, and with the way constant dollars are sometimes computed as part of the process.[6] For example, they note that BLS usually measures output in financial services as equal to labor input; thus there can be no productivity gains. To get a realistic but limited appraisal for this industry, Baily and Gordon studied the number of shares traded, and their total market value, per worker in the security industry. Productivity measured in shares traded grew strongly from 1965 to 1979; while measured in market value, it declined. From 1979 to 1987 both measures showed very strong growth—over 11 percent per year. They also looked closely at insurance and banking and noted that government inflation adjustments probably make insurance productivity figures look worse than they really are, and that factors such as BEA's not counting credit card transactions in the measure of checks handled cheated the banks a bit. They also found problems with the accounting for airline transportation. However, most of these measurement deficiencies existed before 1973, so they don't help much to explain the post-1973 slowdown. (Indeed, for parts of the economy, productivity was more optimistically estimated after 1973 because of changes in government statistical methods.) Putting all the evidence together and adding a dollop of considered judgment, they conclude that errors in measuring output did not contribute more than 0.2 percent to the total growth slowdown.[7]

My wife ordered fifteen items from a mail order clothing company. They arrived in fifteen different shipments spread over two months.

One view of the measurement problem has been a particular favorite of computer professionals: perhaps new ways of doing business create value in the form of greater quality, convenience, or pleasure for consumers without anyone's paying for it or BEA noticing. This conjecture was also prominent in the National Research Council reports that dismissed the productivity problem. Before examining the evidence more closely, let us note what this hypothesis needs to explain. By 1995, the cumulative shortfall in output—the total amount by which measured annual GNP in the United States would have been larger if productivity growth had continued at precomputer rates—is, conservatively, about $1 trillion per year.[8] That's roughly $4,000 per year for each man, woman, and child, about half the average family's disposable income. Whether the increased convenience and variety of products and services attributable to IT is worth that much to the average person is a matter of opinion, but an opinion that might be measured. We could roughly allocate the cost to different products—banking, insurance, airline service—and ask those old enough to remember whether the changes are worth their imputed values. No one has done this. Instead reports like those of the NRC committees are sprinkled with anecdotes and testimonials by the people who supply the services, and they tend heavily to cite quick loan approvals and ignore the demise of price labels and telephone receptionists. The most frequently cited examples of new service applications are airline reservation systems and complex mutual funds. Unquestionably the reservation systems have underwritten vast increases in air traffic. Dynamically changing discount fare "yield management" systems keep planes more fully occupied and total revenues higher. But I doubt that the cattle car crowding, skimpy meals, crap-game fares, complex restrictions, and multi-hop routes that have become part of the unmeasured quality of the

product are viewed by travelers as positive contributions to their pleasure and ease. Unquestionably the number of shares and shareholders in mutual funds has increased enormously. Many people enjoy the opportunity to choose among several hundred different funds, even if their performance is not reliably different from any other or as good as the market average. But I have to wonder whether the touted competitive advantages and increased pace of business and finance also attributed to computers—things like mergers and leveraged takeovers—have not contributed to the decreased stability of the blue-chip companies, the AT&Ts and IBMs, that used to be havens for small investors.

Now let's look more directly at the evidence as to how much IT has actually accelerated changes in service quality. The main changes claimed by the defense are greater variety and convenience. For example, if concentrating everything anyone ever wanted in one giant supermarket, or installing electronic checkouts, on average saves customers time, it might be fair to count that time at some rate of pay. If customized insurance policies are more valuable to buyers but no more expensive, it might be fair to count the extra value as output. Still, this line of reasoning presents several problems. First, almost all the trends that might be counted in this column, such as self-service outlets, a greater variety of items under one roof, and redesigned payment methods, were growing about as rapidly before 1973 as after (Baily and Chakrabarti 1988). For example, the average number of items carried by supermarkets rose from 9,000 in 1972 to 17,500 in 1985, a compounded yearly increase of 5.2 percent; the number had risen an average 6.6 percent per year over the previous twenty-two years.[9] Figure 4.1 shows the data in more detail. There has been exponential growth in supermarket variety ever since World War II but no sign of a change in the *rate* of growth since the computer age. Part of the supermarket variety increases is due to improved preservation and transportation that give us raspberries in the dead of winter and from competitive strategies that proffer almost 200 varieties of breakfast cereal, among them at least a dozen nearly indistinguishable brands of corn flakes, at prices that include only about 10 percent for ingredients.[10] Information technology presumably contributed by making it easier for stores to keep track of it all, thus helping the trend to continue. However, even if we accept greater variety as equivalent to greater quality, the

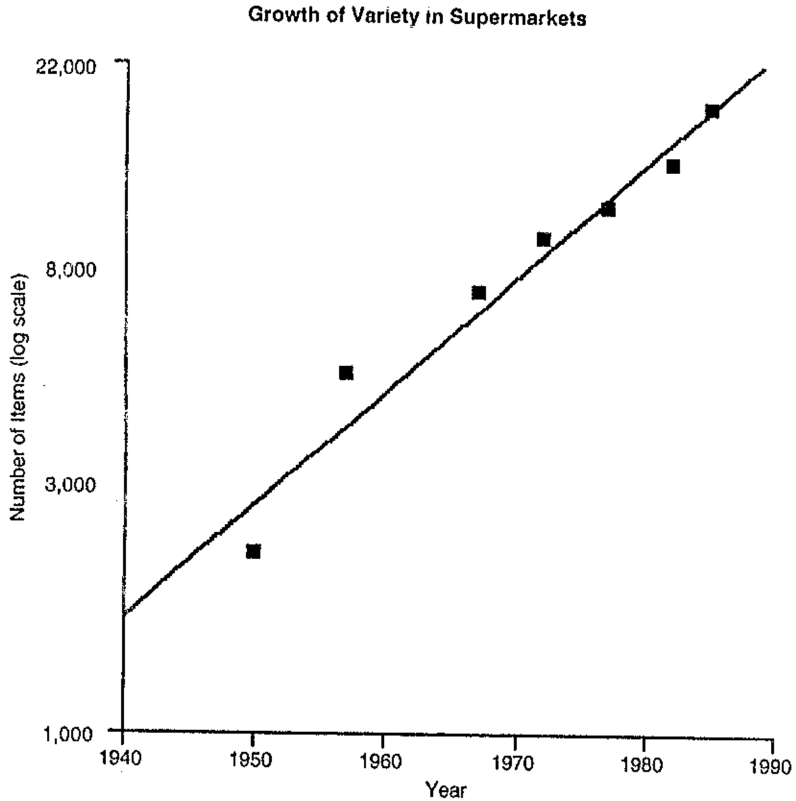

Figure 4.1.
Growth of product variety is frequently cited as an unmeasured quality gain facilitated by computers. A common example is the increased number of items on supermarket shelves (National Research Council 1994). However, as this plot shows, supermarket item variety has been increasing for a long time. The overall average yearly growth of 5.4 percent per year is shown as a solid line. If anything, growth in this quality measure has been slightly slower since computers became widespread.

claim that supermarket quality has grown faster since computers still has no basis.

Second, there is a logical problem for basic economic theory in assuming that businesses would choose to lower their productivity by giving away all the fruits of technology. Indeed, a theoretical modeling effort by Baily and Chakrabarti suggests that the apparent stagnation in services cannot be accounted for by assuming that IT's benefit has all been passed on gratis to consumers. The models indicate that the observed effects could not result if IT was really reducing the labor needed to perform work. I don't trust such disembodied economic theorizing very much, but I also feel queasy about crediting the economy with large unmeasured "convenience" values on the basis of informal impressions. While there are ATMs and car reservation systems that offer improved consumer convenience through computing, there are also those touchtone menu telephone answering systems (if you want information on opening times, touch 1; if you want to listen to Musak forever, touch 2; to enter 3, touch 5 . . .), and impersonal bill payment accounting systems (Dear BigCompany: I've written explaining your error twice. Why is it still on the bill?).

Before explaining away the economic and profitability shortfall with unmeasured customer service improvements, we should at least *try* to measure them. For example, what is to prevent the BEA from asking people how much they like the new financial services they are getting and how much they would be willing to pay to get them back if they were rescinded?[11] The fact that banks and brokers usually supply these conveniences "free," as marketing gimmicks rather than products with a price tag, suggests that the answers would not always be overwhelmingly positive. In any event, the world, if not the producers of these systems, should want service enhancements that increase real revenues or decrease costs rather than just moving constant-value market share around.

It's Too Early to Tell

Significant effects on productivity can take a long time to wend their way through the crooked corridors of business practice, labor resistance, accounting credit, market growth, acceptance, adaptation, diffusion, and perfection. Thus, some large benefits of computerization may not have become visible yet.

Whether IT, and the computer in particular, has had time to show its stuff is open to debate. Some economists, for example, the Nobelist Leontief (Leontief and Duchin 1986), have compared the computer revolution to the industrial revolution, the building of national rail networks, or the arrival of industrial assembly lines, all of which took many decades to produce dramatic improvements in productivity. According to Franke, fifty years passed after the introduction of Watt's steam engine in 1775 before substantial improvements in productivity, and he quotes Leontief and Duchin's conclusion that "the introduction of mechanical power continued to transform western economies and society over a period of some 200 years. The computer revolution became visible only a few years ago, and by the year 2000 it will be no more advanced than the mechanization of European economies had advanced by, let us say, the year 1820" (Leontief and Duchin 1986)[12]

A more recent precedent, the exploitation of electricity, is described by Stanford economist Paul David (1990, 1991). Dating the industrial use of electricity from the first dynamo installed by Edison in New York in 1881, he reports that demonstrable effects on productivity did not appear for over forty years, until the 1920s. But when they finally appeared, they were substantial, contributing, he estimates, almost 2.5 percent per year to a spurt in national productivity growth. What caused the delay was slow diffusion due to ineffective early application and reluctance to replace old plants and machines. He believes that productivity gains are hard to see until about half the potential users have adopted a technology and that this didn't happen with electric motors in manufacturing until around 1920. Counting the beginning of computers from the first VLSI devices about 1970, David likens our current perspective to that of an observer of industry in 1900, and reckons that we shouldn't get impatient with them for another twenty years.

David's time comparison with IT is way off the mark. VLSI processors and memory chips were a major step forward for computers but hardly marked the beginning of their commercial use. Recall that by VLSI-time, 1970, there were already 125,000 computer operators and well over $100 billion worth of IT in the United States. Even the electronic computing machines used by scientists, bomb builders, bankers, brokers, and insurance companies in 1950 were further along the technological learning curve than Edison's primitive direct-current dynamos and employed some 900 operators.[13] By 1980, every bank with more than one vice president had a computer; David estimates that only 5 percent of manufacturers were using electric motors in 1900. A random sample of U.S. firms found 100 percent using computers in 1992 (Kraut and Streeter in press). David's divide of half adoption was well crossed by the mid-1970s in the mass of companies large enough to have payroll or accounting departments, and word processing had nearly reached that level as well. A more realistic comparison would expect IT to have started raising productivity statistics by 1975 if it had followed the electric motor's career path.[14]

But this is quibbling. Different technologies take different amounts of time to mature. There is no reason that computers should take exactly the same time as electric motors. David's main point is merely that superior technologies do not increase productivity immediately, so the lack of an effect now does not prove that there won't be one sometime.

The effects of a superior technology on productivity growth are strongest during the diffusion period, as more and more firms start using it (Kraut and Streeter in press). David's model puts the maximum growth effects somewhere near the 70 percent adoption point. Since computers have long since been almost universally adopted in commerce, the inescapable conclusion is that they are not yet superior.[15] The question before us is when, if ever, will they become so? How long should we wait for business as usual to make computer dreams come true?

None of the historical analogies, even electric motors, is very helpful here because we live in a different world. The rate of change in technology, industry, and patterns of consumption is much faster now than during the historical revolutions of the nineteenth century or the pre–World War II era of the twentieth. We have much greater capital, labor, baseline productivity, and market resources; a more widely educated population (both workers and consumers); much improved means of communication; much cheaper and faster transportation; and dramatically better scientific infrastructure and engineering methods with which to perfect and bring technology to bear. In addition, we have different attitudes: widespread expectation of miracles from technology have replaced nineteenth-century skepticism and turn-of-the century investment caution.

Thus the pace of refinement, adoption, and utilization of new technology is much more rapid than was possible in earlier times. Witness the rapid introduction of jet aircraft and the extraordinary growth rates of television and other consumer electronics. The earlier technologies appealed to in this excuse—railroads, steam engines, assembly lines, electric motors—took much longer to achieve full deployment. Moreover, the investment this time out looks enormously larger. Tens of trillions of (1990) dollars were not pumped into any earlier technology before it paid off, as Leontief's and David's projections would have us tolerate. I doubt that 10 percent of England's GNP was being spent on not-yet productively superior steam engines, or of the United State's on not-yet superior electric motors and their operation, forty-five years after their first commercialization, as was the case with computers.

Perhaps the size of the total investment in energy machines of all kinds, canals and water-powered looms as well as railroads and steam engines, during the first industrial revolution matched the current outlay for information machines. Plausibly this is a better analogy and raises some interesting issues. If we viewed the totality of the first industrial revolution as equivalent to what is happening now, we would want to ask how we are doing now relative to, say, forty or fifty years after the start of the earlier boom. But when did that start? With the invention of the steam engine? The automobile? The birth of commercial computers seems easier to place. However, it is a mistake to view the computer as a single technology. More appropriately, each major application is a new technology harnessing information processing capability, much as the electric motor, the locomotive, and the jet plane all harnessed energy-transforming capability.[16] Suppose we consider all the advances of the energy evolution as one prolonged development. Would any forty or fifty years after introduction of its major technological innovations, say from 1765 to 1810 (steam engine), 1865 to 1910 (steel, railroads, textiles), 1880 to 1925 (electricity generators), 1900 to 1945 automobiles, assembly lines), or 1935 to 1980 (airplanes, everything else), show the near-zero productivity growth in their primary sectors of application, let alone the total economy, that has occurred with information machines and the sum of all their uses to date?

Yes, it is too early to tell. It is too early to conclude that computers will never bring improvements in work productivity, but it is too late to be complacent. No logic compels every innovation, or every immense investment, to improve productivity. Surely computer use is still in a learning stage, still improving. Industry leaders are optimistic and believe that businesses are beginning to use IT effectively. They point to examples of major local gains from coordination and waste reduction. Crude estimates suggest that the efficiency effects of text editing could be improving about two-thirds as fast as the very long-term averages for, say, textiles.[17] But these tantalizing hints and expressions of faith are not enough. We have few repeated winners, few phase two applications with as strong a claim as text editing. And surely we'd rather not wait fifty years to see each application mature—if it does. We don't want to wait as long as from the steam engine to the jet—200 years—for information machines to complete their conquest. Our challenge is to foster faster growth and learning, to make the technology pay off in our young programmers' lifetimes.

Coincidence

Another way to escape the conclusion is to assert that the evidence we have considered is simply an accident or, rather, an elaborate set of accidents. Perhaps there are genuine and large effects on work efficiency, but they are masked by other factors that happen to have reduced productivity in just the same times and places that computers have increased it. For example, the worldwide recession of the mid-1970s was very pronounced in the United States at just the time when productivity growth took a downswing. Recessions cause productivity decline by softening markets, leading to unused but still expensive production capacity. But they also lead to labor force reductions, removing fat in payrolls, thus increasing productivity, a proposition marshaled by Roach (1992b) to sound a note of optimism for the 1990s.

One part of the slowdown period corresponded with the oil crisis. Remarkably, productivity in some transportation industries increased during this period. Banks and some other financial institutions have experienced unusual woes in the last decade as a result of deregulation, oil crises, and unlucky real estate deals. U.S. industry has suffered from the gradual accretion of mountainous government reporting requirements regarding employment, safety, pollution, pensions, and taxes, all of them piling nonpaying loads on management information systems. Indeed, the finger of blame for low output figures has been pointed in many different directions: labor recalcitrance, educational crises, management greed, inflation, taxes, short-term business focus, the takeover craze, service industry complacency, the low personal savings rate, the oil crisis, unfair competition from the Japanese. Perhaps we could explain away each and every time period, industry, and individual firm where computers have disappointed, an exercise that would leave as the only evidence that computers haven't helped a handful of laboratory and field studies showing limp efficiency effects of specific uses. This would be a slender reed to stand against strong winds of favorable opinion. In the late 1970s, some economists gave serious consideration to the multiple-causes explanation of productivity growth slowdowns, which the *Los Angeles Times* labeled "the revised Murphy's Law—Everything that can go wrong will go wrong at the same time." However, the doldrum has persisted for another decade, and this mode of denial now requires so many coincidences as to stretch credulity. As just one example, all the negative factors would have had to be relevant to service industries (except telecommunications)[18] but not to manufacturing, and it doesn't ring true; strikes, environmental impact statements, oil prices, foreign competition, and almost all the others have been at least as burdensome for manufacturing.[19]

More telling, perhaps, are the quantitative analyses, which left a large part, between 20 and 60 percent of recent productivity slowdowns unexplained. After a well-worth-reading examination of all the other factors they consider credible candidates, Baily and Chakrabarti choose as their prime suspect insufficient innovation of productive technology. Even Denison (1983) conjectures that "slow . . . advance in knowledge itself, and . . . the incorporation of knowledge into production" is a primary cause. Since there was, in fact, enormous investment in information technology during this period, and relatively little in other innovation, especially in the service sector, it is nearly impossible to avoid the conclusion

that investment in information technology hasn't paid off in productivity. Moreover, to explain away the studies that correlated investment in IT with low productivity growth or return on investment in individual firms, one would have to find some reason why highly computerized firms were buffeted more strongly by other factors than were their more conservative counterparts, and that in itself would be damning.

Competitive Success

The next two excuses come from well-documented commercial accomplishments of computers. Several applications of computers have yielded dramatic competitive advantages for the people who pioneered them. ATMs provide a good example. They were introduced in large numbers by CitiBank in New York, which as a result increased its market share from 4 percent to 13 percent within a very short period of time. A direct order entry system introduced by Blue Bell, manufacturers of Wrangler jeans, by which their retail outlets could instantly respond to fluctuating consumer style demands, resulted in a large increase in business, amounting to some 13 percent of their total revenues. When Merrill Lynch was the first to introduce a cash management account for its customers, a very elaborate computerized system that required 100 programmer years to construct, it attracted large numbers of accounts from other brokerage houses.

In all of these examples, the other major players in the same markets fairly soon followed suit, and some or all of the initial market advantage dissipated. However, that's not the real issue here. The issue is whether the use of computers as instruments of strategy for gaining market share, that is, shifting sales from one company to another increases productivity. In the case of ATMs, there seem to be contradictory data and opinion as to whether they have brought about any net improvement in productivity for banks. At least for small banks early on, the cost of installing and maintaining these devices was more than the labor saving that they provided (Brand and Duke 1982). When installed beyond bank lobbies, they can entail substantial real estate costs and require expensive support by relatively high-paid technical personnel not needed for human tellers. It is obvious that ATMs have provided something that some customers wanted—some mix of convenience, longer banking hours, and novelty. People apparently make more frequent but smaller withdrawals from ATMs than from tellers (Brand and Duke 1982). My wife chooses the ATM in the bank lobby over the human teller, even when there's a line for the ATM but not for the teller. She says ATMs are more fun than writing checks.

Support of successful competitive strategy alone does not necessarily constitute evidence of productivity. It is possible that ATMs could have increased the productivity of users, if not of the banks. For example, if ATMs free people's time for more productive work, total GNP might have reflected it, but possibly in such a small way as to be invisible. Perhaps ATMs are generating a pure welfare benefit, as economists call it, in which the investment is good for the world but not necessarily for the company or the economy as measured by paid-for output. An example of a welfare investment is industrial air pollution abatement (although this presumably has some long-term payoff in reduced medical bills, lower agricultural costs, and other human benefits). We could postulate that computers and information technology are in large part simply making life better without making it more productive. When a government invests extra money in the routing of a superhighway to make it more scenic, the goal is not higher GNP but greater satisfaction of the citizenry. ATMs may be a similar product. The increased convenience and availability of cash, and the fun of operating a simple device that hands over money, might well be expected to make ATMs popular even if they had no productivity consequences at all—that is, provided that they did not cost banks enough money that a competitive advantage could be gained by doing away with them and reimbursing customers for the loss with higher interest rates or lower fees.

Consider this: There is no reason that ATMs and similar service products that are good, not because they reduce business costs but just because people like them, should not show up as productivity gains. If people like them enough, they should be ~~willing to pay~~ for their use; the service feature should become a new product. If the new product is being produced at low cost relative to what people will pay—for example, in the fees that some banks charge for ATM transactions or in lower interest rates that customers will tolerate—then the industry's ledger sheet should show a productivity gain.

On the other hand, services like this could attract customers from one provider to another with no effect on productivity at all. Advertising is generally considered to be largely such a service. Imagine a hypothetical computer system that produces better decisions on plays for a football team. The team might win more games as a result, but overall productivity in the football industry (the gross per player salary megabuck) will not depend on which football team is the most frequent winner. In a similar manner, leveraged takeovers of companies, by focusing on short-term financial gains for stockholders, can be construed as winning for someone, but they may have negative effects on long-term productivity. Perhaps the game metaphor is more apt than might appear. As an economy gets closer and closer to having all the goods and services that people really need or want, more and more of peoples' proportional effort might come to be devoted to just playing competitively, either for the sheer joy of competition or to see who gets the greater share of the pie. Other countries than the United States may be improving their productivity relatively more by the use of technology, such an argument would go, because they are not as far along on the economic success curve and are thus still devoting more of their energies to greater efficiency and less to the market share game.

Another area of documentable success for computers is in inventory and resource management. Bookstores, for example, have made effective use of International Standard Book Numbers (ISBNs), bar codes, and computers to allow them to stock only popular books and keep on hand only enough of each title to meet day-to-day needs. Ignoring the fact that this use has made every chain book outlet a twin of every other and left customers with a greatly impoverished list of titles from which to choose—a decrease in public welfare presumably—the retail book trade has reputedly increased its productivity. Similar stories are told about manufacturing resource planning (MRP), in which the raw materials, piece parts, and flexibly assigned labor are all kept track of and marshaled in minimal numbers at just the right time so that capital is not unnecessarily tied up in unused resources. Borg-Warner is said to have increased productivity by 20 percent ~~by such~~ means. The consulting company Booz Allen has been quoted as saying that such techniques can yield up to 29 percent gains in profits for many manufacturing companies. Better management of capital should improve multifactor productivity, if not labor productivity directly. The mystery is why these effects are not more prominent. There have been healthy improvements in productivity in manufacturing in the last twenty years, possibly—despite Loveman's (1986) data to the contrary—due in some part to such computer-based methods. However, manufacturing has shrunk to a small part of overall economy, so these gains may not have greatly affected the overall measures at which we looked. Retail trade, however, has continued to experience productivity losses. Perhaps inventory control techniques such as those instituted by bookstores have been insufficiently adopted (although the ubiquity of bar codes and scanners suggests the opposite), have too small an effect on the average, or are being counterbalanced by other forces.

We've Come About as Far as We Can Go

Another possible out is to believe that productivity, at least in the United States, has topped out or come close. As in the case of direct labor costs for manufacturing, the total contribution to the economy has shrunk drastically, so little overall effect can be felt by improvements here. However, jobs in the service industries—financial services, marketing, design, advertising, engineering, maintenance, customer service occupations—which have grown apace as manufacturing has declined, are associated with relatively low productivity and slow productivity growth. Is it possible to believe that these too have topped out on productivity? Normally we'd expect a sector with low baseline productivity to provide much greater opportunity for improvement. But with some rather fanciful arguments, we can hypothesize the opposite in this case. Suppose that most of these service jobs are ones whose dollar value is defined directly as personal labor. For an example, consider babysitting. You hire a sitter by the hour to keep a watchful eye on one or more small children. Excepting some farfetched notions of robotics or telepresence, there is no way

of improving the sitter's productivity; in an hour a sitter can do an hour of sitting; that's all the product there can be. Occupations in which the supplier's personal time and presence are the definitions of value may be growing—professional athletes, servers in elite restaurants, entertainers, personal trainers, masseuses, chiropractors, and nude dancers are other examples—but they are certainly not yet a major factor.

Complacency

Stephen Roach (1992a) offers another, rather imaginative, rationalization. He suggests that the vast computing resources acquired by business in the 1980s and 1990s were grossly underutilized because there was so much regulation, so little competition from abroad, and such easy success that managers just didn't care about productivity. Says he, "The technology paradox is traceable to the service sector's legacy of complacency and inefficiency." Although a great fan of Roach (he has done much to open eyes to the computer paradox), I find this one hard to swallow. Roach's main evidence is a reduction in "back office" white-collar job growth in the first part of 1992, which he attributes to mergers of large banks and subsequent consolidation of back-office processing capacity. One reason why I am so doubtful of his conclusion is that it seems highly unlikely that banks actually had large excess processing capacity. "Sizing" of computer systems for such operations is a fairly accurate engineering art, especially when the processing load grows gradually over time. In telephone office business support systems, for example, capacity is usually stretched to near its limit (Srinivasan, 1992); it is often impossible to add even a modest new software feature without first adding more computer power. That large and growing firms would have enough capacity to spare for consolidation to make a big difference in hardware requirements is not credible. Indeed, studies in the mid-1980s found no cost savings from bank mergers (Srinivasan 1992). As of this writing it is impossible to tell whether the 1992–1994 "restructuring" of the U.S. service business (economists' euphemism for permanent layoffs) will continue, will result in genuine and lasting productivity gains, especially ones that combine increased output with strong employment and high wages as in the past, or will turn out to be a temporary cost-cutting binge that will leave the economy weakened and well-being diminished. Only time will tell.

My second reason for doubting the "complacency" hypothesis is that I see no evidence or basis for the existence of complacency or its complicity in underuse of technology for productivity. Although regulation and weak foreign competition in some parts of the service industry may have had some local moderating effects on the pitch of activity, there has certainly been enough incentive to keep management in hot pursuit of productivity improvements. There has obviously been vigorous pursuit of market share, for example, even in banking; witness ATMs and overdraft accounts. And surely increased productivity—the sword the Germans and Japanese used to maim the American steel business—is in itself a highly effective competitive tool. Wouldn't domestic service firms have used it on each other if they could? Moreover, in the most tightly regulated, least competitive sector of the service industry, the telephone business, productivity improvements continued at a normal clip. Disincentive effects of regulation and weak competition won't do as explanations.

Summarizing the Excuses

We have now reviewed all the plausible ways to convince ourselves that there really is no computer productivity problem. Each is seriously wanting. And even if all of these objections and excuses were to be rather generously credited, they are not sufficient to undo the conclusion that something is wrong with computers. Surely there is plenty and more room for greater efficiency in the white-collar and service sectors of the economy. There must be many ways to make the provision of building maintenance, retail customer services, hospital care, hotels stays, car washes, restaurant meals, doctors' exams, lawyers' advice, and health club workouts more labor and capital efficient. But somehow these jobs seem not to be getting more productive very fast. Why haven't phase two computer applications done something to aid this majority part of our economy? Surely there is a great deal of profit and glory to be accrued by inventors and suppliers of technology to enhance productivity in these domains. For example, if word processing, spreadsheet, and database programs for PCs were as wonderful as their devotees would have us believe, if they increased the efficiency their users by factors of many hundreds as the inventions of the industrial revolution did, they would have been exploited in the businesses just named to more effect. There is enormous room for computer-aided gains in productivity. The fact that it hasn't happened must not be swept under the rug.