THE CHALLENGE OF MEASURING AND MODELING A DYNAMIC ECONOMY

Alan Greenspan,
Chairman, Federal Reserve Board;
before the Washington Economic Policy Conference of the
National Association for Business Economics,
Washington, DC, March 27, 2001

I am pleased to have this opportunity to address an issue of considerable importance to both business economists and policymakers, that is, the challenge of measuring and modeling our dynamic economy. Moreover, I would like to raise the issue of how much of our finite resources should be directed to measuring and how much to modeling.

Business economics endeavors to understand the structure of an economy, how it works and, above all, how to forecast it. It is a bedeviling job because the future, at root, cannot be foretold. The best we can do is to construct probabilistic models that can inform the decisions of business executives and, of course, economic policymakers who, of necessity, will be making their decisions armed with incomplete information.

For a while in the 1960s, we were increasingly mesmerized by the possibilities of econometric models as a crystal ball. However, history was not entirely kind to this endeavor. For one thing, especially against the backdrop of the inflation of the following decade, it soon became apparent that our theories of the macroeconomy were woefully inadequate. For another, even leaving aside the shortcomings of our theory, we soon learned that the economic structure did not hold still long enough to capture its key relationships. Its changing structure frustrated efforts to isolate a reasonably fixed set of coefficients. In turn, the absence of fixed coefficients undermined the usefulness of the model as a basis for projecting the future.

Econometricians recognized many of these difficulties, and so developed a vast and elegant literature in support of this research program, covering a spectrum of topics ranging from maximum-likelihood-estimation techniques to tests for coefficient stability to diagnostics for detecting undesirable properties of the errors of these equations. The creativity that was applied to this effort is all the more impressive because it took place in the context of a computing environment that was, by modern standards, truly primitive.

(Puede reproducirse en Colombia.)

But in time it became increasingly clear that, for all their theoretical advantages, these sophisticated models did not reliably outperform a number of simple and far less costly reduced-form models, from the money supply models that appeared to work well for a while during the 1970s, to astructural vector autoregression models based on a handful of lagged variables that are still employed today.

To be sure, the large econometric models have been refined, incorporating the fruits of later theoretical developments, including perhaps most importantly the insight that monetary policy could not permanently influence the level of the unemployment rate. Moreover, a larger role was given to a range of financial and expectational variables that earlier practitioners steeped in orthodox Keynesian income determination tended to downgrade. Liquidity preference functions, to be sure, were included even in the early versions of these large-scale models as necessary building blocks for determining the equilibrium level of interest rates and income. But, overall, financial sector modeling was primitive. Indeed, only modest progress has been made in this area since the Federal Reserve began to produce our own flow of funds accounts in 1955.

A further, and perhaps more profound, challenge to the underlying validity of this style of modeling is the possibility that, whereas standard models of the real economy determine a unique level of income, the financial system appears to be capable of reaching myriad equilibria. In addition, the fundamental forces that determine which of these equilibria will be selected may themselves be inherently unpredictable.

We have built large-scale models of the United States and global economies at the Federal Reserve. While recognizing their limitations, we do find them useful in research and analysis. But the experience of the last 40 years underscores a fundamental dilemma of business economics. Should we endeavor to continue to refine our techniques of deriving maximum information from an existing body of data? Or should we find ways to augment our data library to gain better insight into how our economy is functioning? Obviously, we should do both, but I suspect greater payoffs will come from more data than from more technique.

Certainly, statistical systems in the United States, both public and private, are world class and, indeed, in many respects set the world standard. But given the rapidly changing economic structure, one could readily argue that more statistical resources need to be applied to understanding the complexities of the newer technologies that confront analysts.

These newer technologies and the structure of output they have created have surfaced a set of definitional problems that (although evident in a world of steel, fabrics, and grains) were never on the cutting edge of analysis. I refer, of course, to the age-old problem of defining what we mean by a unit of output and, by extension, what we mean by price. The dollar value of sales or GDP depends, of course, on the specific accounting rules chosen. And while value in that context is uniquely defined, the split between volume change and price change is always approximate.

In past decades, we struggled about what we meant by output, and hence price, nonetheless an average price of hot rolled steel sheet and a corresponding total tonnage was precise enough for

most analytic needs. By the same token, tons of steel per work hour in a rolling mill yielded rough approximations of underlying productivity for most purposes.

Output per hour in an economy dominated by such goods, or even services, for which the definition of a typical unit of output was reasonably unambiguous, was a meaningful and relatively robust statistic. Our data systems in the early post-World War II years were by and large adequate to the task.

But over time, and particularly during the last decade or two, an ever-increasing share of GDP has reflected the value of ideas more than material substance or manual labor input. This ongoing development is imposing significant stress on our statistical systems. We know, presumably uniquely, the dollar value of a software application. But when comparing software-application values over time, how much of the change is volume and how much is price? The answer, in principle, requires judgments about very fundamental issues in measurement: What are the underlying determinants of consumer value preference, and how does this good or service contribute to that preference, taking account of all the other goods and services being consumed? Problems that were always latent in defining steel prices and quantities but rarely rose to this level of significance are threatening to seriously challenge our measurement systems in an age of the microprocessor, fiber optics, and the laser.

These latent problems have emerged in full view in the pricing of medical services. Perhaps the inherent complexity of this undertaking is most clearly revealed by posing the question, what do we mean by a standardized unit of medical output? Is it the procedure, the treatment, or the outcome? What does the fee charged for the bundle of services associated with cataract or arthroscopic surgery represent? How does one value the benefits to the patient of shorter hospital stays, more comfortable recoveries, and better physical outcomes? Clearly, the unadjusted fee for a single medical procedure does not adequately represent its "price."

The price indexes for medical services used to be constructed by pricing a variety of inputs, for example, a night in the hospital, or an hour of a physician's time. A few years ago, the Bureau of Labor Statistics began moving toward pricing the treatment paths of particular diagnoses, the better to capture changes in the mix of inputs used to treat a given disease. For example, many surgical procedures that used to require an overnight stay in a hospital now can be performed on an outpatient basis, and the producer price index and consumer price index are now better able to measure the price decline associated with that change. Interestingly, when such techniques are applied to individual medical procedures they appear almost without exception to indicate falling prices at least since the mid-1980s. This has raised significant questions as to whether our current measures of overall medical service price inflation are capturing the appropriate degree of productivity advance evident in medicine.

Indeed, the *level* of real gross product per hour for medical services embodied in our overall productivity measures declined between 1990 and 1999 (the last year for which data are available). This is implausible and raises obvious questions of the validity of the price deflators currently employed. Thus, while progress is visible, enormous measurement challenges remain in measuring prices of medical services.

But there are deeper issues as well, associated with the valuation of a consumer's time. Clearly the shorter stay of a cataract patient is of value to that patient. In other words, today's techniques allow the surgeon to deliver more consumer value per hour of operating time, other things being equal.

A full measure of the output of the medical sector would take account of this reduction in recovery time and attribute it to the medical sector.

While many issues of measurement arise in the context of services, even the measurement of some goods prices presents considerable challenges. High-technology goods are a case in point. Academic research in this area dates to the mid-1960s, but its application in the measurement of real output gained prominence with the introduction of hedonic price indexes for computers and peripherals by the Bureau of Economic Analysis (BEA) in 1985. More recently, the efforts undertaken by statistical agencies have intensified, spurred by the accelerated pace of technological innovation, which has yielded an ever-expanding range of new products and product variants, as well as by the rising share of these goods in our economic value added.

Thus, much progress has been made by the BEA, Census, and the BLS, with which I am sure you are familiar. This morning, I should like to alert you to some of the research in these areas coming from the Federal Reserve. Much of our work in this regard has been focused on improving our published statistics on industrial production.

Our staff's multiyear work in this area began in 1998 with the development of new measures of the domestic output of semiconductors. Next, we revised our procedures for estimating the production of computers, and more recently we have introduced new series for an important component of the output of the communications equipment industry, local area networking (LAN) equipment, which provides the infrastructure critical to expanding the productive uses of information technology.

In total, these high-tech goods (semiconductors, computers, and LAN equipment) currently represent less than 8% of total manufacturing output. However, their production, as we measure it, rose at an average annual rate of around 50% in the second half of the 1990s, and, taken together, they contributed two-thirds of the increase in manufacturing output between 1995 and 2000. Indeed, U.S. production of semiconductors in 1996 eclipsed motor vehicle assemblies as the largest four-digit manufacturing industry in nominal value-added terms.

The characteristics of these goods present the range of complexities that one faces in measuring quality-adjusted prices. First, many are wholly new products: For example, switches, the largest single segment of LAN hardware, did not enter the market until 1993. And even "older" high-tech products, such as computers, are now bundled together in ways that offer an enormous variety of combinations of characteristics related to speed, memory, networking capability, and graphics capability, to mention just a few. For all of these goods, product cycles are truncated by rapid innovation. For instance, in 1995, 10 megabit-per-second Ethernet switches dominated that market; last year, the two most popular switches operated at rates of 100 and 1,000 megabits per second. Product lives for semiconductors and computers can be even shorter; some computer models have remained on the market for only a couple of months.

In such an environment, the availability of detailed micro-level data describing the attributes of these goods is crucial. One means of defining the unit of output is to unbundle the characteristics of a high-tech product and to price each of them separately. This so-called "hedonic" technique, now applied by the BEA to items that account for 18% of GDP, is one approach.[1] In our work at the Federal Reserve, we have developed hedonic price indexes for network routers and switches using, in the case of the former, data from product catalogs and, in the case of the latter, privately produced reports evaluating the performance of the products.

However, hedonics are by no means a panacea. Most important, the measured characteristics may be acting only as proxies for the qualities of the services that buyers ultimately value. This, again, raises the difficult issue of the appropriate scope for value measurement and poses the question of whether the correct approach may be to move toward directly pricing the services we obtain from our information processing systems rather than pricing separately the individual hardware components and the software.

The Federal Reserve staff has found that, when detailed data are available on prices and on quantities, we can produce results that are comparable to those based on hedonics, using the conceptually simpler "matched model" approach.[2] Indeed, we have taken this approach in constructing quantity and price indexes for several high-tech items. In the case of semiconductors, we relied on data from three private vendors for information on nearly 100 unique microprocessors, more than 200 types of memory chips, and more than 80 other chips. We also acquired nominal sales and unit value data for about 1,100 distinct computer models.

Neither hedonic nor matched-model techniques are sufficient to deal with the introduction of wholly new products that differ fundamentally in their characteristics from their predecessors. This will continue to be one of our major ongoing challenges.

I am encouraged by the progress that economists and economic statisticians have been making to date in tackling the daunting task of measuring real output and prices in a rapidly changing economy. The challenge that lies ahead is, indeed, large, and to meet it will require the support of the business and academic communities to supply the information and to help develop the tools that our statistical agencies require.

The information revolution, itself, will also surely play an important role. For example, high-tech information systems might some day allow statistical agencies to tap into a great many economic transactions on a basis close to real time. More generally, I am certain that the possibilities for creatively harnessing technology for the improvement of economic measurement are much broader in scope--although, as in many other areas of endeavor, the precise directions those advances will take are difficult to predict. If we had the appropriate database, of course, who knows?

* * * * *

[1] J. Steven Landefeld and Bruce T. Grimm, "A Note on the Impact of Hedonics and Computers on Real GDP," *Survey of Current Business* (December 2000).
[2] Ana Aizcorbe, Carol Corrado, and Mark Doms, "Constructing Price and Quantity Indexes for High Technology Goods," Industrial Output Section, Division of Research and Statistics, Board of Governors of the Federal Reserve System, July 26, 2000.