
Statistical Discrepancies in Our Federal Data Programs

Daniel J.B. Mitchell

The failings have been managerial rather than technical in nature. We need accountability, when administrative problems lead to statistical deficiencies. We need an independent Board of Federal Statistics, with authority to review both statistical programs and top statistical executives.

A recent cover of *Business Week* depicted a blindfolded figure and a headline that read: "The Real Truth About the Economy: How Government Statistics are Misleading Us." Clearly, public confidence in our national statistical base is declining. But why? Is it just a mood of national cynicism? Should the business community be concerned?

In the past, when criticisms of the federal statistical establishment erupted, the usual reason was that someone did not like the statistical message. The critic wanted to kill the messenger, rather than deal with the problem the message had highlighted. Now, it is no longer that simple. The blood-lust to kill the messenger continues to lurk, notably around the Consumer Price Index (CPI) and its use in escalating tax brackets and Social Security benefits.

The CPI has been subject to various criticisms—mainly that it does not adequately take account of quality changes and the substitution effect and, thereby, somewhat overstates inflation. While there is some validity in these criticisms, they have been driven by congressional interest in budget balancing. And elements of the CPI which *understate* inflation

have been neglected in the debate. Moreover, a CPI that is designed around patterns of consumption of elderly Social Security recipients (whose benefits are escalated by the CPI) historically would have risen *faster* than the official index.

But there have also been some major gaffes in statistical provision. And they have made the national statistical apparatus susceptible to alarm that the numbers are being manipulated.

REASON FOR CONCERN

There are three reasons why the business community (along with economists) should be concerned with the growing distrust of official American data:

- Economic policy is made on the basis of the official numbers. For example, the Federal Reserve makes judgments on whether the economy is overheating partly on the basis of the unemployment rate. Interest rates might be raised, if overheating is indicated. If policymakers are unable to see trends clearly, there is the potential for policy errors to occur.

- Business decisions are often based on official data. Those who

DANIEL J.B. MITCHELL is Professor at the Anderson Graduate School of Management and the School of Public Policy and Social Research at U.C.L.A.

make those decisions naturally expect reliable information. Financial markets also react to official data such as inflation rates.

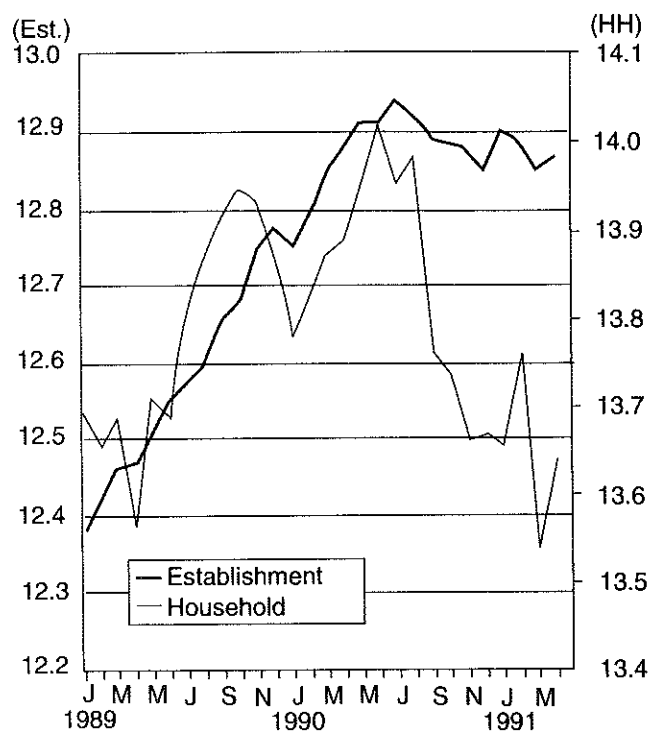
- Managers can readily diagnose the source of recent statistical problems. A commonly blamed villain for statistical inadequacy is the growing “complexity” of the economy. This notion was part of the *Business Week* critique. Not surprisingly, it is often cited by official statisticians themselves. After all, if the problem is complexity, there really is no blame. But complexity is not the root cause. *Managerial disincentives are*. Official statisticians are not as accountable for results as they should be.

Unfortunately, in the arena of federal data-gathering and dissemination, poor performance is not penalized. Neither is good performance rewarded. One need not be a management expert to understand the consequences. Nor does one have to be a high-powered economist or statistician to see the remedy. One recent episode of statistical failure—the inaccurate and contradictory information provided about the California recession of the 1990s—is an example of the problems we face in repairing the national statistical machinery. There is also evidence that the California episode was not isolated, but part of a pattern. Clearly, reform in statistical management is needed.

THE 1990S RECESSION IN CALIFORNIA

While absolute accuracy is not possible for any data series, it ought to be possible to avoid egregious errors. Perhaps the most worst error in recent years involved the depiction of the recent California recession by various employment indicators which were published by the U.S. Bureau of Labor Statistics (BLS). BLS is the main provider of labor-market information and employment trends. They are, of course, key indicators of the health of national, state, and local economies. There are two key sources of BLS employment information. BLS’s establishment survey of nonagricultural payroll employment goes back to the pre-World War I era. Essentially, it involves collection of payroll records from over 390,000 business units. Because of the large numbers of employers included in the survey, and because the establishment survey is annually benchmarked (with a lag) to the universe of employers paying unemployment insurance taxes, establishment employment data have long been regarded as key indicators of the business cycle by forecasters and policymakers. The fact

Figure 1
California Employment:
Establishment vs. Household Survey
(millions)



Source: UCLA Business Forecasting Project as of June 1991.

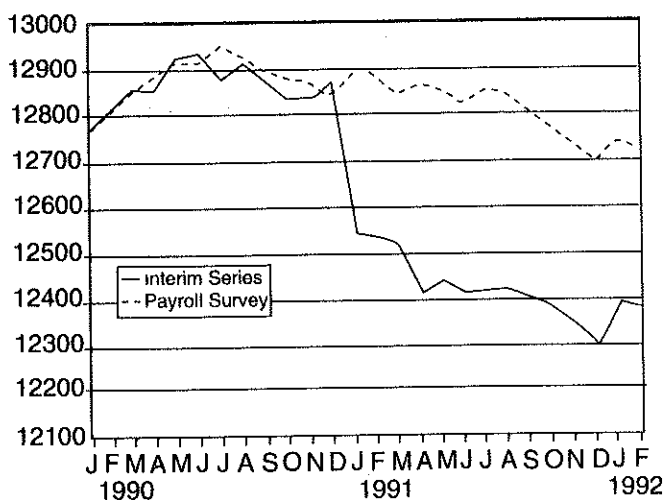
that the data appear monthly, and with only a short lag, enhances their use as a business-cycle indicator. An alternative measure of employment comes from the joint BLS and Bureau of the Census household survey (Current Population Survey). The household survey gathers data from about 60,000 households, rather than from businesses. Its employment series is

Complexity is not the root cause of the problem. Managerial disincentives are. Official statisticians are not as accountable for results as they should be.

limited by the smaller sample size and, therefore, does not provide as much industry or regional detail. Nonetheless, the household survey is the source of the widely cited unemployment rate, other labor-market indicators, and demographic information.

In 1991–92, an unfortunate episode raised questions about the accuracy of the establishment series, and about the care taken by BLS to produce it. Subsequently, problems also developed in the house-

Figure 2 **California Nonfarm Employment:
January 1990–February 1992**
(thousands)



Source: UCLA Business Forecasting Project as of March 1992.

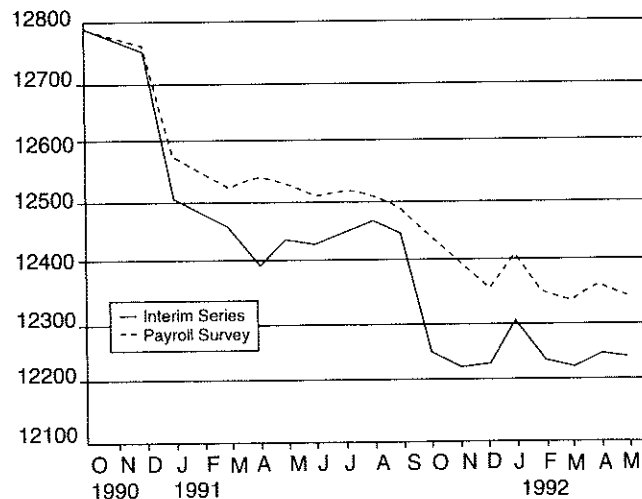
hold series. California—the largest state in the nation—was particularly affected by the recession of that era and by the misreporting of establishment survey employment. *Figure 1* shows that the household survey version of California employment trends and the establishment series began to diverge as early as mid-1990. By mid-1991, the establishment series indicated only a modest decline in California employment. But the household series was indicating a severe recession.

Which was the true story? We now know that, without doubt, California was, in fact, beginning a severe and prolonged recession. But in early 1991, the issue was less clear. Still, other contemporary economic series (such as local help-wanted advertising) suggested that the severe-recession scenario of the household series was closer to the truth than the mild slowdown suggested by the establishment survey. By early 1992, there was no mistaking the fact that California was in the midst of a major economic slide. But the establishment survey continued to paint a picture of relatively mild decline. Forecasters began to take note of another employment series that was maintained by the California Department of Finance (DOF). It was known as the “interim series.” *Figure 2* shows the interim series and the establishment series as of March 1992. Note that the interim series indicated a major decline in employment in January 1991—a judgment with which the establishment series did not concur. Similar results and discrepancies were being reported in other states as well.

The DOF interim series was taken from the unemployment insurance reporting data to which the establishment series is annually benchmarked (with a lag). By early 1992, it was becoming clear that both the California and national establishment data would have to be substantially revised downward at benchmarking time. Indeed, the magnitudes of the downward revision were so large that the BLS issued a special report in June 1992 to explain what had happened (see “Revision of Payroll Survey Employment Estimates to March 1991 Benchmarks” in *For Further Reading*). In fact, BLS’s June 1992 explanation was no explanation at all. It ruled out causes, and did not list the sources of the decline. For example, the report ruled out the hypothesis that the decline was due to a large-firm bias in the prebenchmark series. Nor, according to the same report, was the drop due to a large number of firm failures between December 1990 and January 1991. *Figure 3* shows the DOF interim series and the benchmarked establishment (payroll) series of employment officially in much closer accord—in contrast with the prior prebenchmark establishment data. In particular, the official data showed a marked, but unexplained, decline in employment in January 1991 in this report.

At that point, it simply appeared that there was a large drop in employment (for reasons unknown), which benchmarking had succeeded in capturing. Unfortunately, this was not the end of the story. In January 1993, the BLS announced that the benchmarked results (severe recession) were in error and

Figure 3 **California Nonfarm Employment:
October 1990–May 1992**
(thousands)



Source: UCLA Business Forecasting Project as of June 1992.

that the old prebenchmark results (milder recession) were closer to the truth after all. How could this be? Isn't the benchmark essentially the universe? According to BLS, prior to 1991, it had been discovered that some employers were inadvertently overreporting the number of their employees on their unemployment insurance tax forms. One office of BLS had ordered a correction of this overreporting as of January 1991. But somehow, this order was not known to the BLS office that compiled the establishment data. So the latter office benchmarked its data to a series with an artificial drop in January 1991. It then went on a wild-goose chase looking for large-firm biases and other explanations in blissful ignorance of what another office in the same agency had done (see "Research

Some employers were inadvertently overreporting the number of their employees on their unemployment insurance tax forms. One office of BLS had ordered a correction of this overreporting as of January 1991. But somehow, this order was not known to the BLS office that compiled the establishment data.

Results: March 1991 Benchmark Revisions to Payroll Employment Estimates" in For Further Reading).

A panel of the American Statistical Association was asked to examine this episode. It politely concluded that "there is no evidence to support allegations of deliberate manipulation of the data or of incompetence or carelessness." (See "American Statistical Association Approves BLS Methods in 1991 Payroll Data Revisions" in For Further Reading.) Certainly, there was no deliberate manipulation. After all, the incumbent Bush Administration was hurt in the 1992 elections by the initial revised report of sharp recession. And the incoming Clinton Administration campaign argument that "it's the economy, stupid" tended to be undermined by the retroactive finding of reduced recession severity. But the charges of incompetence or carelessness are not so easily dismissed. Establishment employment is a leading product of the BLS and was used a key indicator of the economy by government and private economists. Could an episode in which one office does not know what another office in the same agency is doing be characterized in any other way than "incompetence or carelessness?"

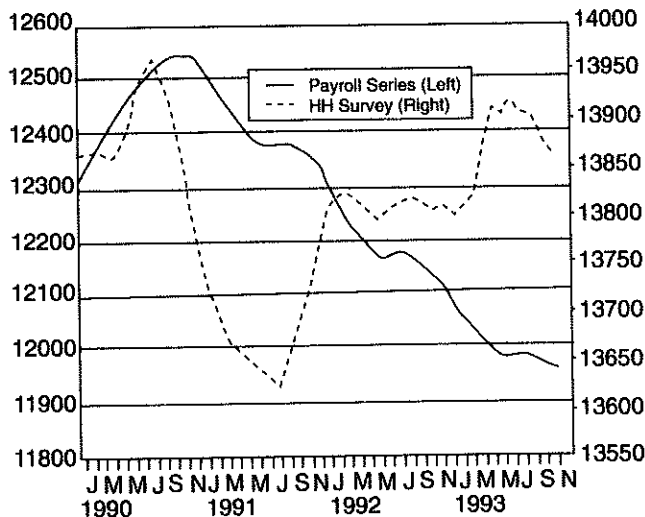
Surely, heads should have rolled for the damaging consequences to the agency and to consumers of data. But it does not appear that any did. What assurance, therefore, has the data-using public that such an occurrence could not be repeated? What incentives are in place in BLS to ensure there will be no repetitions of confusion? Indeed, there was another such incident that involved Current Population Survey only a short time later. Nor is it even clear that the establishment survey problem was fixed. There are new questions about the re-revision of establishment employment data, despite BLS's January 1993 explanation. At least in California, official reports suggest that something more severe than had been indicated by the re-revised establishment series did, in fact, happen in the first quarter of 1991 (see "The Employment Revisions" in For Further Reading). Other data support that conclusion. For example, taxable sales in California—a major state economic indicator—fell sharply in the first quarter of 1991. Thus, there are lingering doubts that the 1993 revisions to the previous benchmarking were accurate in California. And there are reasons to suspect that California's recession was more severe than is now being officially depicted.

For the nation as a whole, the re-revised BLS estimates lowered the 1990 peak in employment by 560,000 jobs to the original benchmarked estimates. Thereby it reduced the subsequent job drop. About 350,000 of these lost peak jobs were in California (see "The New Unemployment Statistics" in For Further Reading). If the California re-revisions are in doubt, then the re-revisions for the United States must also remain questionable. Such lingering doubts about a major series do not build professional support for embattled statistical agencies.

Although the household employment series gave an early (and correct) indication of California's economic troubles, it too was ultimately compromised. Figure 1 has already illustrated a problem of data inconsistency. It shows that, as of mid-1991, the employment data from the household survey and the establishment survey (both series issued by BLS) were showing very different results for the California economy. The establishment series was showing a mild decline after a business-cycle peak. The household survey was showing a major collapse. But this statistical discrepancy had been reversed one-and-a-half years later. Figure 4 (from December 1993) shows the two series, with the establishment series now re-revised as described earlier. Now it is the establishment series that shows a steady decline in California employment after mid-1990 and

Figure 4

**California Employment:
Six-Month Moving Average
January 1990–November 1993**
(thousands)



Source: UCLA Business Forecasting Project as of December 1993.

through late 1993. But it is not as sharp as the initial benchmarking. Indeed, the household series suggests that, while California had a severe recession starting in mid-1990, a significant recovery was underway by mid-1991. If we could believe the household series, California's recession, while initially severe, was no more enduring than that of the rest of the United States.

A year or so later, in early 1995, as *Figure 5* demonstrates, the story had again been changed. Now the household and establishment series appear to coincide due to household series revisions. Both series now tell of a major recession running from mid-1990 to late 1993, and the California recovery beginning in 1994—much later than the rest of the United States. In short, over a period of four years, the establishment and household series have told substantially varying stories of the same events. The revisions and re-revisions of the establishment series have already been described. The problems with it can be traced to BLS. But the Bureau of the Census appears to be involved in the statistical variations in the household series.

Household survey employment is based on a sample. The survey essentially produces *ratios* such as the labor-force participation rate, the unemployment rate, and the employment-to-population ratio, rather than absolute counts. To obtain absolute numbers of employed or unemployed persons, such ratios have to be applied to an underlying population estimate. Population growth is partly determined by relatively predictable birth and death rates. But it is also affected by net migration. At the state level, migration involves

both international movements of people and interstate movements. These elements are both important for California.

It took far too long for the Census Bureau to recognize that people do not migrate to states where jobs are scarce. Census overestimates of population growth in California apparently led to an upward trend bias in the initial household employment estimates. Thus, one agency's error spilled over into another's data series. Again, discrepancies between major data series purporting to measure the same concept both distorted understanding of economic trends and gave credence to the notion that data are arbitrary and subject to manipulation. Had there been more concern about data users and their need to interpret the condition of the labor market, perhaps there could have been a more rapid response to a glaring statistical inconsistency.

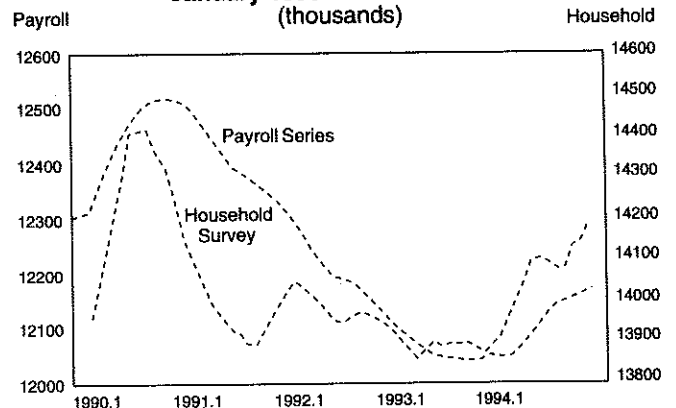
OTHER STATISTICAL SNAFUS

It would be comforting to report that the California recession episode was an isolated event. Sadly, it is one of a series.

- In 1994, BLS introduced a new methodology for collecting the household series. Although it initially promised to provide ways to bridge (some of) the old and new series, in order that trends could be tracked, ultimately the agency admitted a failure to do so (see "BLS Finds Difficulties in Returning to Old Way of Conducting Jobs Survey" in *For Further Reading*). Despite this apparent failure, the involved officials persisted in declaring the change a success: "Leaving aside some difficulties that have arisen in interpreting

Figure 5

**California Employment:
January 1990–December 1994**
(thousands)



Note: Data shown are 6-month averages.
Source: U.S. Bureau of Labor Statistics.

the new data, Census Bureau and BLS officials said they have been pleased with the transition. 'It has been very successful—beyond our wildest hopes and expectations,' said Chet Bowie, Assistant Director of the Census Bureau's Demographic Surveys Division." (See "New Census Household Employment Survey Improves Data-Collection Speed, Accuracy" in For Further Reading.)

Sadly, it is not as easy for *data users* to "leave aside" the difficulties in interpretation as it apparently is for *data administrators*. If statistics are not to be interpreted, then why are they being gathered? The great success to which the above quotation refers is administrative in nature. Part of the new methodology involved the use of laptop computers. They made data collection easier and faster. But the old saying applies here: *A job not worth doing is not worth doing well*. BLS and the Census Bureau should not have instituted the 1994 changes in the household survey until they were sure that they could bridge the old and new series. The above quotation, with its veiled admission of "difficulties," shows that the task was not accomplished.

- While the 1994 press releases of the new household survey data on the labor market at least warned the reader of a break, there have been other, more hidden breaks which may have distorted the interpretation of labor-market trends. In particular, "hedonic" price estimates of computer-manufacturing output—introduced by the Bureau of Economic Analysis (BEA) in the 1980s—have resulted in a significant increase in reported aggregate national output. Output, in turn, is the numerator in the BLS's output per hour (productivity) measure. Hence, the growth rate of officially measured BLS productivity was raised by the BEA's hedonic adjustment.

Hedonic pricing involves the use of regression analysis to determine the value placed by the market on attributes of a product. Where a product is heterogeneous—e.g., housing—or where quality changes over time, the hedonic method allows for an adjustment. The output of computer manufacturers has clearly changed over time. The latest Pentium-based PC has vastly more capacity in terms of speed, memory, and display than the original Apple II. But as econometricians know, regression analysis can be sensitive to specification.

Certainly, in the rapidly changing computer field, the value of computer attributes might be estimated differently by using alternative, but seemingly reasonable specifications. Indeed, the director of BEA has acknowledged that "prominent macroeconomic modeling firms have reservations about the hedonic index-

es for computer equipment and do not use them in certain aspects of their analyses and forecasts." (See "Mid-Decade Strategic Review of BEA's Economic Accounts: Maintaining and Improving Their Performance" in For Further Reading.) The reason for this user skepticism is simple. The hedonic price index for computers falls so rapidly that computer sales, when deflated by the index to estimate real out-

Where a product is heterogeneous—e.g., housing—or where quality changes over time, the hedonic method allows for an adjustment. The output of computer manufacturers has clearly changed over time. The latest Pentium-based PC has vastly more capacity in terms of speed, memory, and display than the original Apple II.

put, show extremely rapid growth rates. Although the computer industry is not a large sector of the economy, if extremely rapid real growth is attributed to the industry, it significantly raises the reported trend in aggregate national output.

Examples are again the best guide to the problem, so consider three widely-believed stories about the 1980s and 1990s: (1) Real wages began substantially lagging productivity growth in the 1980s; (2) the 1990s saw a "jobless recovery" from recession; (3) employers have been making a vast investment in labor-saving equipment, thus causing the seeming jobless recovery. All of these notions depend heavily on the questionable hedonic price adjustments to the computer sector.

Because of the hedonic adjustment, the price index to deflate output differs in trend from the CPI that is used to deflate wages. It is this difference that produces most of the gap between real wage growth and productivity growth. The notion of a jobless recovery depends on the observation that productivity grew unusually fast during the recovery in the early 1990s. The purported result was that less labor input was needed than usual. But the measured rate of productivity growth depends on the hedonic deflator. Alternative deflators do not show above-normal productivity growth and, therefore, do not indicate a jobless recovery.

Finally, although the recovery in the 1990s saw a very high investment ratio of producers' durable equipment to GDP in real terms, the ratio was not so exceptional in nominal terms. Again, the hedonic

deflator seems to be playing an important role in this discrepancy (see "Statistical Methodology May Be Distorting Key Economic Relationships" and "Productivity and Real Wages: Is There a Puzzle?" in For Further Reading).

STATISTICAL LESSONS

The examples we have provided suggest three criteria for official statistical provision which have been violated:

- *Data should be accurate.* The example of the revisions and re-revisions of the establishment employment series indicate what can happen when care is not taken to produce reliable data.

- *Data should be consistent.* If two series are purporting to measure the same thing and they diverge, there is a problem that needs immediate addressing. The conflicts between household employment and establishment employment are examples of the dangers of inconsistent data which give no clear guide for economic policy, interpretation, or forecasts. Inconsistency undermines confidence in the issuing agency.

- *Care must be taken to provide continuity of data.* The 1994 break in the unemployment rate and other labor-market indicators illustrate the danger of changing methodology without concern for time-series users. Similarly, the introduction of hedonic pricing methodology changed long-standing economic relationships in what may be merely a statistical artifact. Thus, some of the most widely held beliefs about recent economic trends may simply reflect a questionable methodological change.

ADMINISTRATIVE LESSONS

As complaints over national statistics have accumulated, there have been calls for an administrative fix. One suggested solution has been to combine all major agencies into a single umbrella organization. Canada, for example, uses such an approach. Perhaps if BLS, Census, and BEA were all combined, they would worry more about the effect of one agency's decisions on another. Perhaps BEA would have weighed more heavily what hedonic pricing of computers would mean for BLS's productivity measures. Perhaps Census would have worried sooner about what its net migration estimates for California were doing to the BLS's household employment estimates for the state. But there are limits to what fiddling with organization charts can accomplish. In the private sector, we like to

think that top managers are held accountable for major blunders. While this ideal may not *always* be achieved in practice in private employment, it seems *never* to occur at federal statistical agencies.

Of course, accountability can be dangerous in the political setting. Recently, there has been some controversy in Congress over whether the CPI inadequately reflects "quality" improvements in the goods and service prices it includes. Since it is used to escalate Social Security payments and income tax brackets, changing its methodology to produce a lower measured rate of inflation would reduce the federal budget deficit. The result was a threat by House Republican leader Newt Gingrich to "zero (BLS) out" of the budget after thirty days if the problem were not fixed, and a counterproposal by Democratic Congressman Richard Gephardt to order the BLS not to fix it. Clearly, we do not want Congress to determine statistical methodology in this fashion.

Privatization is sometimes offered as a solution to making official data-gatherers more user-oriented. But there are many problems which are associated with private data collection. Among them are issues of confidentiality, bias, and possible misuse of information. Although there are some privately gathered data series

There are three criteria for official statistical provision which have been violated: (1) Data should be accurate; (2) data should be consistent; and (3) care must be taken to provide continuity of data.

which are used by economic analysts, care about reliability, sampling, and similar considerations is not guaranteed. Perhaps the most important problem is that *data are costly to gather but cheap to disseminate*. With private data collection, there would be reduced access and less information gathered. Private providers must charge enough to cover the fixed costs of data collection. They also must control distribution to avoid "pirating" of the information.

While basic data collection cannot be contracted out, there are elements of agency operations that can be. Most notable is the research function. It has gradually swelled in importance at BLS and other agencies. Agency researchers have begun to publish in academic journals and government journals have begun to publish academic research articles—despite the tight budget outlook sure to face statistical agencies in the future.

Surely, there are enough universities and think tanks to handle much of the research that is now being done in-house. Unless it is completely esoteric, research is likely to have policy implications and can be controversial. Statistical agencies need to avoid such controversies. They can do so by limiting in-house research and using outside research consultants, as needed.

Statistical agencies currently have advisory boards of outside experts and users of various types. But there is a major difference between an advisory board and a board of control. Generally, advisory boards are composed of friendly appointees who serve with little or no compensation and who are likely to give the benefit of the doubt to agency experts. An alternative model would be a move toward a Board of Federal Statistics, with board members chosen from profes-

Clearly, having a Board of Federal Statistics is a better solution than the current approach—one that shields the statistical establishment from external review and accountability.

sional organizations (e.g., American Economic Association, American Statistical Association, and others) and users (major private-sector data-user groups from business, labor, and academia). Board members could be chosen by the President for fixed lengthy and staggered terms with Senate confirmation—involving the same process now used for mem-

bers of regulatory bodies. Selection guidelines could be specified in statute.

Such board members could be empowered to review not only programs and particular data series, *but also key personnel*. With regard to key personnel, the review powers could include appointments and dismissals, along the lines of a corporate board. Is this a perfect solution? No. There is no perfect solution. After all, in the private sector, corporate boards do not always properly monitor top management. But clearly, having a Board of Federal Statistics is a better solution than the current approach—one that shields the statistical establishment from external review and accountability.

Major statistical failings in recent years have undermined public and data-user confidence in official data. Unreliable information complicates the making of economic policy and business policy and can lead to incorrect decisions. Key criteria for official data are accuracy, consistency, and accountability. Although problems in meeting these criteria are often attributed to growing complexity of the economy, the underlying solution is more managerial than technical. Accountability of statistical executives is the key to reform. Failures must be penalized. Successes must be rewarded. An independent Board of Federal Statistics, based on the corporate board model, would provide the needed accountability.

To order reprints, call 1-800-352-2210;
outside the United States, call 717-632-3535

For Further Reading

- AMERICAN STATISTICAL ASSOCIATION, "American Statistical Association Approves BLS Methods in 1991 Payroll Data Revisions," *Daily Labor Report*, April 13, 1993.
- _____, "BLS Finds Difficulties in Returning to Old Way of Conducting Jobs Survey," *Daily Labor Report*, June 22, 1994.
- _____, "New Census Household Employment Survey Improves Data Collection Speed, Accuracy," *Daily Labor Report*, November 7, 1994.
- BARRY BOSWORTH and GEORGE L. PERRY, "Productivity and Real Wages: Is There a Puzzle?" *Brookings Papers on Economic Activity*, 1:1994.
- CALIFORNIA DEPARTMENT OF FINANCE, "The Employment Revisions," *California Economic Indicators*, May/June 1993.
- _____, "The New Unemployment Statistics," *California Economic Indicators*, March/April 1994.
- CAROL S. CARSON, "Mid-Decade Strategic Review of BEA's Economic Accounts: Maintaining and Improving Their Performance," *Survey of Current Business*, February 1995.
- DANIEL J.B. MITCHELL, "Statistical Methodology May Be Distorting Key Economic Relationships," *The UCLA Business Forecast for the Nation and California*, September 1994.
- U.S. BUREAU OF LABOR STATISTICS, "Revision of Payroll Survey Employment Estimates to March 1991 Benchmarks," U.S. Government Printing Office, June 3, 1992.
- _____, "Research Results: March 1991 Benchmark Revisions to Payroll Employment Estimates," U.S. Government Printing Office, January 8, 1993.

Challenge

THE MAGAZINE OF ECONOMIC AFFAIRS

A GUIDE TO THE DEFICIT

James K. Galbraith
William Darity, Jr.

Worldwide Development or Population Explosion: Our Choice

Gerard Piel

The Minimum Wage Revisited

Jay Mazur

Making the Case for Cities

Peter Dreier

Statistical Discrepancies in Our Federal Data Programs

Daniel J.B. Mitchell

Rising Sun . . . Falling Dollar

William M. Burke

CHA 3877 / 90024UCALO 6 / 1996 1
MANAGEMENT LIBRARY/SERIALS
1400 BRAD SCHMIDT DR
UNIV OF CALIFORNIA
LOS ANGELES CA 90024



JULY-AUGUST 1995/\$7.50
(no vol.)