

Session 5

TIME SERIES REVISION POLICIES

Time Series Revisions: The Effects on Gross Domestic Product

Robert P. Parker and Teresa L. Weadock
U.S. Department of Commerce
Bureau of Economic Analysis

Gross domestic product (GDP) is the most widely used measure of a Nation's overall economic activity. In response to the need for timely estimates, the Bureau of Economic Analysis (BEA) releases an "advance" estimate of quarterly GDP one month after the end of the each quarter. This estimate is based largely on monthly survey data for the first two months of the quarter and BEA judgmental projections for the missing source data. In each of the next two months, revised estimates of GDP that incorporate newly available and revised monthly and quarterly source data are released. Annual and benchmark (comprehensive) revisions of GDP are released on a regular schedule as annual and less frequently collected census-type data become available.

To measure the reliability and accuracy of the quarterly GDP estimates, BEA has conducted a series of studies based primarily on revisions to the successive quarterly estimates. Such studies have been found to be very useful to both users of the GDP estimates and to BEA. For users of the estimates, the studies provide insights into the likely size of future revisions to GDP and its major components and identify components whose reliability they would like to have BEA improve. For BEA, the studies help to identify components with problems in the source data or estimating methodologies. This information helps BEA to work with the agencies who prepare the source data or to devote its own resources to developing improved procedures to reduce revisions. Information from these studies also enables BEA to analyze the impact on the reliability of the GDP estimates of its revision schedule.

In addition, the studies satisfy the requirement of the Office of Management and Budget that agencies producing the principal Federal economic indicators provide periodic evaluations of their performance. This requirement, Statistical Policy Directive Number 3, states that these evaluations will "include an analysis of the accuracy of the series, the effects of revisions, and performance relative to established benchmarks."

This paper, which is based primarily on measures of reliability for the quarterly GDP estimates for 1978-91, consists of four parts: An overview of the preparation of the quarterly GDP estimates and a discussion of the measures used by BEA to measure their reliability;

NOTE: The views expressed in this paper are those of the authors and not necessarily those of the Department of Commerce or the Bureau of Economic Analysis.

highlights of a recently completed study by Allan Young at BEA on the reliability of the first three, or "current" quarterly GDP estimates for 1978-91;¹ an extension of this study to the effects on the quarterly estimates of subsequent annual revisions; and a discussion of revision practices that agencies use for the source data used to prepare the GDP estimates and how these practices affect the accuracy of the quarterly estimates.

The studies reviewed in this paper have implications both for BEA and for the agencies that provide the source data used to prepare GDP. They indicate that BEA should review the need for three current estimates and its policy of not revising prior quarters except at the time of annual and benchmark revisions. They also indicate that other agencies should review their revision practices to provide more timely and accurate revised data.

Part 1. Overview of the Preparation of Quarterly GDP Estimates and Measures of Reliability

Estimating schedule for quarterly GDP estimates

For each quarter, GDP estimates are prepared on a schedule that consists of three successive "current" estimates--"advance," "preliminary," and "final"-- and of subsequent estimates prepared as part of annual and comprehensive NIPA revisions.

The advance estimate is prepared about 1 month after the end of the quarter. For most components, the estimate is based on source data for either 2 or 3 months of the quarter. In most cases, however, the source data are not final and are subject to revision by the issuing agencies. Where source data are not available, the estimate is based primarily on past trends and on BEA analysts' judgment.

One month later, the advance estimate is replaced by the preliminary estimate, which is typically based on source data for all 3 months of the quarter. In most instances, the source data used for the preliminary estimates, particularly the data for the third month of the quarter, are subject to further revision.

One month later, the preliminary estimate is replaced by the final estimate, the last of the current estimates, which incorporates revisions in source data for the third month of the quarter and quarterly source data for some components. For virtually all components, these source data are subject to further revisions by the issuing agencies.

¹ See Allan H. Young, "Reliability and Accuracy of the Quarterly Estimates of GDP" in the October 1993 issue of the Survey of Current Business.

Each quarterly estimate is subject to three successive annual revisions, usually released in July. The first annual revision incorporates further revisions in the monthly or quarterly source data and introduces some annual source data. The second and third annual revisions incorporate a broad range of annual source data. For example, the "final" estimate for the fourth quarter of 1993, which was released last month, will be revised in July 1994 (first annual revision), July 1995 (second annual revision), and July 1996 (third annual revision). Each quarterly estimate is also subject to one or more comprehensive revisions, in which information from the economic and demographic censuses is incorporated in the monthly, quarterly, and annual source data by the issuing agency or by BEA.

Source data

More complete and more accurate information is generally available on an annual basis than on a quarterly or monthly basis. In many cases, annual data are based on larger samples or represent a complete universe count. In addition, annual data often correspond more closely to the desired definitions and therefore require less adjusting, or they may contain more information for making the necessary adjustments. As a result of these factors, quarterly estimates are obtained either by interpolating between annual estimates or by extrapolating from the most recent annual estimate.

Similarly, the annual estimates in many instances represent interpolations or extrapolations of the more complete and accurate information available in economic and demographic censuses, which are conducted every 5 years and 10 years, respectively.

The quarterly and monthly indicators that are used as interpolators and extrapolators are based largely on monthly or quarterly sample surveys conducted by various Federal statistical agencies. Exceptions include budgetary data from the Treasury Department, tabulations of export and import documents filed with the Customs Service, and tabulations of several types on payroll and income tax returns. Another type of exception occurs if no monthly or quarterly data are available--for example, data for some types of consumer purchases of services and of State and local government purchases of goods and services. In such cases, the quarterly estimates are obtained by interpolation and extrapolation based on a BEA analyst's judgment or related information.

An updated summary of the source data used for the NIPA's is included each year in the Survey of Current Business article that presents the annual NIPA revision (see pages 31-42 of the August 1993 Survey). For a list of methodological papers and for additional information about the NIPA's, see "A Look at How BEA Presents the NIPA's" in the February 1994 Survey, pages 31-33.

Sources of error

The GDP estimates contain several kinds of error. The most obvious kind arises in the current estimates either from preliminary or incomplete tabulations of monthly or quarterly source data or, where source data are not yet available, from BEA's judgments. Error also arises in both the current and the latest available estimates because source data do not meet NIPA requirements in terms of timing, valuation, coverage, and definitions. For example, business firms report some types of data on a fiscal year, rather than a calendar year, basis; even though adjustments by BEA reduce the effect of fiscal year reporting, the results differ from those that would have been obtained with calendar year reporting.

Error also arises from the sampling errors and biases in the monthly, quarterly and annual surveys and from biases and other errors in the annual and periodic universe counts. Probably the most troublesome of errors are those due to the delayed recognition of births and deaths of business firms in sample surveys. (These types of errors that affect source data are discussed in part 4 of this paper.)

Seasonal adjustment is another source of error.² Even if the unadjusted source data were free of error, seasonal adjustment would introduce errors. Although some reduction in seasonal adjustment error appears to have been achieved over time in the current estimates through the use of concurrent seasonal adjustment and by combining ARIMA methods with the ratio-to-moving-average method of seasonal adjustment, such errors are still of considerable magnitude.

Measures of reliability

The term "reliability" used in BEA studies refers to the revisions in the estimates, which reflect the following: (1) Replacement of preliminary source data with revised or more comprehensive data, (2) replacement of judgmental projections with source data, (3) changes in definitions or estimating procedures, and (4) in the constant-dollar estimates, updating of the base year.

In its studies of the reliability of the quarterly GDP estimates, BEA uses six summary measures to describe the revisions: Dispersion, bias, relative dispersion, relative bias, upward revisions, and directional misses. (This paper focuses on the dispersion because this measure effectively summarizes the information provided by the other measures.)

² Quarterly and monthly NIPA estimates are seasonally adjusted if necessary. Seasonal adjustment removes from the time series the average impact of variations that normally occur at about the same time and in about the same magnitude each year--for example, weather, holidays, and tax payment dates.

The six measures are calculated as follows. Let P represent the percentage change in the current estimates, L the percentage change in the latest available estimates, and n the number of quarterly changes.

Dispersion is the average of the absolute values of the revisions:

$$\Sigma|P-L|/n$$

Bias is the average of the revisions:

$$\Sigma(P-L)/n$$

Relative dispersion expresses the dispersion as a percentage of the average of the absolute values of the percentage change in the latest available estimates:

$$\frac{\Sigma|P-L|/n}{\Sigma|L|/n}$$

Relative bias expresses the bias as a percentage of the average of the percentage change in the latest available estimates:

$$\frac{\Sigma(P-L)/n}{\Sigma L/n}$$

Upward revisions expresses the number of times that the current estimate of the quarterly change was revised up by the latest available estimate, as a percentage of the number of quarterly changes.

Directional misses expresses the number of times that the sign of the current estimate of the quarterly change differed from that of the latest available estimate, as a percentage of the number of quarterly changes.

In evaluating these measures, they should be viewed in light of two aspects of the estimation process. First, a change in source data or estimating procedures, which one may assume affects the accuracy of the estimates, is not necessarily reflected in the revision of estimates of a given vintage. For example, an improvement in the current estimates results in a permanent decrease in revision size. An improvement in the latest available estimates results in a permanent increase in revision size. Improvement in both the current and latest available estimates results in little change. However, improvement that is introduced retrospectively into the latest available estimates, as is often the case, results in an increase in revision size for a period of years until the improvement is also reflected in the current estimates. Second, the latest available estimates reflect different vintages. The latest estimates up to 1982 at present reflect the incorporation of the benchmark input-output (I-

0) tables, which are based on detailed information from the economic censuses; the latest estimates beginning in 1983 do not yet reflect the incorporation of the recently released benchmark 1987 I-O tables.³ Thus, the size of revisions beginning with 1983 estimates are most likely understated.

Part 2. Reliability of the Current Quarterly Estimates⁴

Summary of Young study, 1978-91

In the most recent BEA study, Young provided an overall evaluation of the reliability of the quarterly GDP estimates by comparing the successive current estimates of real GDP to the latest estimates and asking the following questions:

o Do the current estimates provide a correct indication of the direction of the change in aggregate economic activity?

o Do the current estimates provide a correct indication of whether the change in aggregate economic activity is larger (acceleration) or smaller (deceleration) than in the previous quarter?

Table 1 provides the summary answers to these questions for each of the current estimates. The record for 1978-91 shows that all three estimates correctly indicated direction of change almost 90 percent of the time. They correctly indicate acceleration and deceleration between 75 and 80 percent of the time. (If changes between -1 percent and +1 percent are disregarded, these early estimates correctly indicate direction over 90 percent of the time and acceleration about 85 percent of the time).

Young also found that for the same period, the incorporation of additional or more accurate source data in the second (preliminary) and third (final) quarterly estimates of GDP did not improve the reliability in comparison with the first (advance) estimate. He identified two factors that contributed to this finding. First, the data for second and third months of a quarter play only a small role in determining the change from the previous quarter. Second, the advance estimate is unaffected by certain sources of error in the preliminary and final estimates. In addition, the advance estimates

³ The benchmark 1987 I-O tables were published in the April 1994 issue of the Survey of Current Business. The results of these tables will be incorporated into the GDP estimates in a comprehensive revision presently scheduled for release in late 1995.

⁴ Most of the sections of this part of the paper, as well as several paragraphs of the previous section, were taken verbatim from Young's article in the October 1993 Survey of Current Business.

of GDP and its major components may benefit more from offsetting errors in the detailed components than the later current estimates; that is, the revisions of the advance estimates may be more negatively (or less positively) correlated than those of the preliminary and final estimates.

In the remainder of this part of the paper, Young's findings are presented in more detail.

Reliability of the quarterly estimates

Table 2 shows dispersion for quarterly changes in current- and constant-dollar GDP and its major components for 1978-82 and 1983-91. These measures show that the incorporation of additional or more accurate source data in the preliminary and final current estimates of GDP does not substantially improve the reliability in comparison with the advance estimates. Dispersion declines only slightly over the successive current-dollar estimates of GDP. For 1978-82, it declines from 1.93 percentage points in the advance estimates to 1.82 percentage points in the preliminary and final estimates. For 1983-91, it declines from 1.17 percentage points in the advance estimates to 1.14 percentage points in the preliminary estimates and 1.15 percentage points in the final estimates.

Dispersion actually increases slightly over the successive constant-dollar estimates of GDP. For 1978-82, it increases from 1.64 percentage points in the advance estimates to 1.72 percentage points in the preliminary estimate and to 1.75 percentage points in the final estimate. The corresponding figures for 1983-91 are 1.25, 1.27, and 1.33 percentage points.

A similar picture emerges for the major components of GDP. In many cases, the advance estimates provided a smaller measure of dispersion than did the preliminary or final estimates. In 1978-82, the advance estimates provided the smallest dispersion in 4 of the 11 current-dollar components--PCE nondurables and services, residential investment, and Federal Government purchases--which accounted for almost 60 percent of GDP. In 1983-91, the advance estimates provided the smallest dispersion in 3 components--PCE services, residential investment, and State and local government purchases--which accounted for over 40 percent of GDP. The record for the advance constant-dollar estimates is about the same as that for the current-dollar estimates, though the share of GDP for which the estimates perform the best is smaller for 1983-91. These results raise the question of whether one or both of the two later current estimates might be discontinued.⁵

⁵ Previous studies have also shown that the advance estimates perform well in comparison with the later estimates, but as shown in table 2, the results are not as clear cut in some periods as in others.

Table 2 also permits one to compare the size of the dispersion measure for GDP with that of its major components. In general, dispersion in the components was larger than that in GDP. The components with the smallest dispersion--about the same as that for GDP--were total PCE and PCE services. The components with the largest dispersion--roughly 6 to 8 times as large as that for total GDP--were gross private domestic investment and Federal Government purchases. The unusually large dispersion in these components reflected a change in the treatment of the Commodity Credit Corporation (CCC) that was introduced in the 1991 comprehensive revision, whereby the CCC was shifted from government enterprises to general government. This shift affected the timing and valuation of transactions and resulted in large, essentially offsetting revisions in Federal Government purchases and the change in business inventories. Dispersion was also quite large in current-dollar nonresidential structures in 1978-82 and in constant-dollar imports in 1983-91, reflecting statistical improvements introduced in the 1991 comprehensive revision.

Trends since 1978

Table 3 examines revisions year by year to see if reliability of the GDP estimates appears to have changed in recent years. The table shows annual averages of dispersion and bias in the quarterly revisions between the successive current estimates and between the current estimates and the third annual revision estimates. For the revisions between the current estimates, the measures are shown for 1978-92; for the revisions between the current estimates and the third annual estimates, the measures are shown for 1978-89.

In Young's study, estimates from third annual revisions were used in place of the latest available estimates in order to provide a more nearly comparable standard for the entire period against which to compare the current estimates. Use of third annual estimates abstracts from much of the effect of the economic census and other information that is used in the comprehensive revisions to revise previously prepared third annual estimates. However, it does not remove the effects of definitional changes in the comprehensive revisions, because for most quarters a comprehensive revision intervenes between the current estimates and the third annual estimates. (To more fully study the effects of the annual revisions,

An early study concluded that the advance estimate might be sufficient; see Rosanne Cole, "Errors in Provisional Estimates of Gross National Product," National Bureau of Economic Research Studies in Business Cycles No. 21 (1969). See also Stephen K. McNees, "Estimating GNP, The Trade-off Between Timeliness and Accuracy," New England Economic Review (January/February 1986): 3-10; and Joseph W. Duncan and Andrew C. Gross, Statistics for the 21st Century (The Dun and Bradstreet Corporation, 1993).

a different approach is used in part 3 of this paper.)

The good performance of the advance estimates

The absence of much improvement in the successive current estimates has puzzled both users and estimators for some time. Two seldom recognized factors contribute to the observed result: (1) The small role played by the data for second and third months of a quarter in determining the change from the previous quarter, and (2) certain sources of error in the preliminary and final estimates to which the advance estimates are immune. In addition, advance estimates of GDP and its major components may benefit more from offsetting errors in the detailed components than the later current estimates; that is, the revisions of the advance estimates may be more negatively (or less positively) correlated than those of the preliminary and final estimates.

This section first discusses the two factors and then addresses the problem of quantifying the total error introduced by the second factor, which embodies seasonal adjustment errors and errors related to the estimation process for certain components. The section concludes with a discussion of the implications for the future.

The role played by the data for second and third months of the quarter in determining the change from the previous quarter is small. The change from the second to the third month receives a weight of only one-ninth in the determination of quarterly change. The weight of the second and third months together is only one-third. The weight of the first month is another one-third, and the second and third months of the previous quarter receive the remaining one-third.⁶ Consequently, errors in neither the preliminary source data for the second and third months of a quarter nor in the judgmental projections used in lieu of source data affect the quarterly change as much as one might intuitively expect.

⁶ This may be demonstrated as follows: Let $Q_1 = X_1 + X_2 + X_3$ and $Q_2 = X_4 + X_5 + X_6$, where X_1, X_2, \dots, X_6 are successive months of source data.

Then, if $d_4 = X_4 - X_3$, $d_5 = X_5 - X_4$, and $d_6 = X_6 - X_5$, the months in Q_2 may be stated as $X_4 = X_3 + d_4$, $X_5 = X_3 + d_4 + d_5$, $X_6 = X_3 + d_4 + d_5 + d_6$, and

$$Q_2 = 3X_3 + 3d_4 + 2d_5 + d_6.$$

Therefore, the quarterly change is

$$\begin{aligned} Q_2 - Q_1 &= (3X_3 + 3d_4 + 2d_5 + d_6) - (X_1 + X_2 + X_3) \\ &= [(X_3 - X_2) + (X_3 - X_1)] + [3d_4 + 2d_5 + d_6]. \end{aligned}$$

Introducing the notation for monthly differences, the first bracketed term becomes $[d_3 + (d_2 + d_3)]$, and

$$Q_2 - Q_1 = d_2 + 2d_3 + 3d_4 + 2d_5 + d_6.$$

Normalizing the coefficients on the d 's provides weights of $1/9$, $2/9$, $3/9$, $2/9$, and $1/9$ for the five monthly changes that determine the quarterly change.

The seasonal adjustment of source data for the final current quarterly estimate introduces errors not present in the judgmental projections, which are developed on a seasonally adjusted basis. The seasonal adjustment factors for the current year are derived from the seasonal patterns of recent preceding years. (The concurrent seasonal adjustment method also includes the seasonal pattern of the current year.) The factors are revised as additional data become available, and they eventually reflect the average seasonal pattern of a period of years that extends symmetrically on either side of the given year. The difference between the initial estimate of the seasonal factor and the final estimate prepared some years later is an error that becomes part of the revision in the final current estimate. To the extent that they are based on judgmental projections, the advance and preliminary quarterly estimates do not contain this error.

Future work

The difference between the revisions to the advance estimate of a detailed component and those to the latest available estimate reflects three types of error: (1) The error in the preliminary monthly source data used for the advance estimates that is corrected in the revised monthly source data used for the final current estimate; (2) the error in the judgmental projections used in lieu of source data for the advance estimate; and (3) the error in the source data used for the final current estimate (including seasonal adjustment error) that replace the judgmental projections used for the advance estimate. The total revision in the advance estimate reflects the first two types of error; the total revision in the final current estimate reflects the third type. One should note that the second type of error, like the third, is determined with respect to the data as they stand several years later.

It would be desirable to determine the size of each of the three types of error at the detailed component level. It would also be desirable to determine the extent of correlation among the component revisions. Such analyses presently are not possible, because each vintage of each estimate and the associated source data are not available in a readily usable form. However, the database used by BEA to calculate the alternative measures of real GDP might be extended so as to retain not only the latest available estimates, but all the vintages of estimates at the detailed level at which deflation is carried out.⁷ Over time, this database would be useful in exploring the outcomes of the estimation process and in developing improved procedures. For example, it will facilitate the development of econometric projection techniques and their comparison with judgmental

⁷ For information about the alternative measures, see Allan H. Young, "Alternative Measures of Change in Real Output and Prices, Quarterly Estimates for 1959-91," in the March 1993 Survey of Current Business.

projections. In this respect, it should be noted that a recently completed study found that judgmental projections compared favorably with econometric techniques for certain import and export components.⁸ Thus, such procedures would be difficult to justify if they do not lead to smaller revisions than do judgmental projections.

The question of whether the reliability of the seasonal adjustments on which the current quarterly estimates are based can be improved merits attention. An analysis at a fairly high level of aggregation suggests that revisions in seasonal factors may be large enough to contribute significantly to the observed results. Specifically, in some of the series examined, seasonal-factor revisions are as large as the variation in the irregular component.⁹ Given that the error introduced by a judgmental projection is likely to be smaller than the irregular variation, this result suggests that for some detailed components, seasonal-factor revisions may play a significant role in causing the revision in the final estimate to be as large as that in the advance estimate. In addition, the seasonal adjustments used for source data should be designed from the standpoint of accurately measuring quarterly change. Little attention has been paid to whether the currently used seasonal adjustment procedures are suitable from this standpoint. In addition, BEA should consider whether more use of concurrent seasonal adjustment, with or without ARIMA, would improve reliability.

Finally, because the third month of a quarter receives little weight in the estimate of change for that quarter, there may be instances in which efforts to reduce revisions in the quarterly GDP estimates should focus on improving the final monthly source data rather than the preliminary monthly source data. As shown earlier, for a survey with three successive monthly estimates, two-thirds of the advance quarterly change is based on three monthly final estimates, while only one-ninth is based on the initial monthly estimate for the last month of the quarter.

Part 3. Effects of Annual Revisions on Quarterly GDP Estimates

In his study on the reliability of the quarterly GDP estimates, Young studied the reliability of the three current quarterly GDP estimates. In this part of the paper, the revisions in the quarterly estimates published in the regular annual revisions are studied.

⁸ See Albert A. Hirsch and Michael A. Mann, An Analysis of the Use of Time-Series Models to Improve Estimates of International Transactions, Bureau of Economic Analysis Working Paper 7 (Washington, DC: Bureau of Economic Analysis, April 1993).

⁹ The irregular component is the residual after the systematic components--the seasonal and trend-cycle--are determined by the seasonal adjustment method.

Dispersion

Table 4 shows the dispersion for quarterly changes in current- and constant-dollar GDP and its major components for 1981-90, the years for which comparisons using annual revisions could be made.¹⁰ These measures show that the incorporation of additional or more accurate source data in the first and third annual revisions substantially improves the reliability in comparisons with the third, or "final" current estimates. The dispersion for current-dollar GDP from the final to the first annual revisions declines from 1.34 percentage points to 0.84 percentage point; for constant-dollar GDP, the decline is from 1.45 percentage points to 0.83 percentage point. From the second to the third annual revisions, the dispersion for current-dollar GDP declines from 0.85 percentage point to 0.65 percentage point; for constant-dollar GDP the decline is from 1.08 percentage points to 0.85 percentage point. However, from the first to the second annual revisions for the current-dollar estimates, the dispersion is virtually unchanged, and for the constant-dollar estimates, there is a significant increase in the dispersion from 0.83 percentage point to 1.08 percentage points.

For the major components of GDP, except for nonresidential structures and producers' durable equipment, the first annual revision estimates recorded a smaller measure of dispersion than did the final estimates. Between the first and second annual revisions, the reliability of all major components improved. Between the second and third annual revisions, the reliability of all major components except for durable goods personal consumption expenditures (PCE) and exports of goods and services improved.

Table 4 also permits one to compare the size of the dispersion measure for GDP with that for its major components. The dispersion for the components was larger than that for GDP. The components with the smallest dispersion were PCE nondurable goods, PCE services, and State and local government purchases. The components with the largest dispersion, more than 6 times as large as that for total GDP, were gross private domestic investment, nonresidential structures, and services, and Federal Government purchases. Of these components, the largest dispersion was Federal Government purchases. As noted in the review of Young's study, the size for this component reflected a change in the treatment of the Commodity Credit Corporation that was introduced in the 1991 comprehensive revision.

¹⁰ To conduct this study, it was necessary to reduce the period covered in Young's study because there were no annual revisions in 1980 and 1991, years of benchmark revisions. See footnote 1 of table 4 for additional details.

Findings

Young found that the revisions in the advance current estimates were about the same size as those in the other current estimates. Thus, he questioned the need to continue the preparation of the preliminary and/or final quarterly GDP estimates. The comparisons in table 4 show, as expected, that generally there was continued improvement in the reliability of the estimates in each successive annual revision; therefore, the need for the annual revisions is not called into question. However, the relative size of the improved reliability from the final quarterly estimates to the quarterly estimates from the first annual revisions was larger than expected. The large improvement, which is evident in both the current- and constant-dollar GDP estimates, was unexpected because very few of the annual surveys that are used to prepare the revised GDP estimates become available in time to be incorporated into the first annual revisions. For example, the Census Bureau's annual surveys of retail trade and of manufacturers, which are used for the estimates of PCE goods and of producers' durable equipment, are not available until the second annual revision. If the improved reliability does not result primarily from the incorporation of such new source data, then the improvement might result from two other sources. One source would be the replacement at the time of the annual revisions of seasonal adjustment factors that were derived from the seasonal pattern of preceding years with factors that also reflect the pattern for the most recent year. The second source would be the incorporation of corrections to monthly or quarterly source data series that become available too late to be incorporated into the "final" current estimates.

If research of the detailed components shows that these are major sources of the large revisions in the first annual revisions, then BEA should consider revising previous quarters between annual revisions. For example, when the advance estimate for the fourth quarter of a given year is released in January, newly available corrections and revisions to source data series and updated seasonal factors could be incorporated into revised first-, second-, and third-quarter estimates. This change would allow for a more effective use of concurrent seasonal adjustment -- that is, using these updated factors for all months or quarters of the year. (For many of the source data series for which the issuing agency uses concurrent seasonal adjustment, the new factors are available but are only used for the most recent two or three months.) Changing BEA's revision practice as just described would provide more reliable quarterly estimates earlier than under the present annual revision schedule. If this change were merged with Young's suggestion to eliminate one or more of the current quarterly estimates, BEA might, for example, drop the final current estimates and revise the previous quarters of the year when the advance estimates are released.

Part 4. Revision Practices in GDP Source Data

Accuracy of the GDP estimates

The term "accuracy" refers to the total measurement error. On the assumption that each successive estimate is more accurate than previous ones, revisions can be viewed as measuring part of the total error in earlier estimates. The rest of the error in these estimates, which is unknown, becomes the total error in the latest estimates. The error in the latest estimates results primarily from the following sources: (1) Errors in the most recently available underlying monthly or quarterly, annual, or periodic census source data; (2) errors in the adjustments made by BEA to convert source data to the definitions and conventions used in compiling GDP; (3) errors in BEA's judgmental estimates for components for which there are no source data; and (4) errors because the most recently available source data have not yet been incorporated into the estimates. It is important to note that these types of measurement error have two different effects on the quarterly GDP estimates. The first type of error affects the levels of the estimates of GDP and its components; the second type affects the changes in the estimates. In this paper, the focus is on the second type of errors.

The implication of the presence of measurement error in the latest estimates limits the extent to which the BEA's measures of reliability can be used as measures of accuracy. The questionable relationship between reliability and accuracy is illustrated by the component that has a very high measure of reliability -- that is, very small revisions -- because the source data are never revised by the issuing agency even though the series maybe based on a very small sample. The next section of the paper discusses how three commonly used revision practices adversely affect the key monthly and quarterly source data series used to estimate GDP. For this purpose, revision practices are defined as those that affect only "not seasonally adjusted estimates."

Revision practices that adversely affect GDP

For some surveys, the revision practice consists of a series of regular annual revisions and, if necessary, periodic benchmark revisions. For annual revisions, the monthly or quarterly series are adjusted to reflect annual survey data, which are based on larger samples, or to reflect annual census (universe) data. For periodic benchmark revisions, first the annual survey data and then the monthly or quarterly series are adjusted for all periods since the last benchmark. Examples of GDP source data prepared using this type of revision practice, the issuing agency, and the GDP components affected are as follows: Retail and wholesale trade sales and inventories from the Bureau of the Census (BOC) -- personal consumption expenditures (PCE) and change in business inventories (CBI); farm output and income from the Department of Agriculture--CBI and farm income; manufacturers' shipments and inventories from BOC--producers' durable

equipment and CBI; and establishment employment from the Bureau of Labor Statistics (BLS)--wages and salaries and PCE.

Although the revised monthly and quarterly estimates for source data series prepared using this practice are considered very accurate, they are nevertheless subject to measurement errors that are likely to introduce errors into the latest estimates of quarterly GDP changes. These errors arise because the agencies generally proportionately allocate across months (or quarters) differences (1) between the monthly or quarterly surveys and the annual surveys and (2) between the annual surveys and the census results where proportionate allocation is not appropriate. For example, the most common sources of differences are errors in the initial survey estimates, both monthly and quarterly and annually, due to delayed recognition of births and deaths, to classification errors, or to misreported data. For births and deaths, proportionate allocation of errors introduces new errors into the revised monthly or quarterly series because this type of allocation fails to recognize that changes in the business population are significantly affected by the business cycle and are not likely to have occurred proportionately over the period covered by the revision. For classification and misreporting errors, the errors also are not likely to have occurred proportionately throughout the revision period.

Another common revision practice used for monthly and quarterly series, which can be viewed as a variant of the practice described above, also is likely to introduce errors into the latest estimates of quarterly GDP changes. Under this practice, when periodic benchmark data are introduced, they are used to adjust only the monthly or quarterly and annual estimates for the year for which such data are available and to serve as the basis for the samples to be used for subsequent periods. Data for previous periods are not revised even though the data for these periods may include errors. Examples of GDP source data that are prepared using this practice, the issuing agency, and the GDP components affected are as follows: Foreign direct investment income in the United States from BEA -- corporate profits; State and local government value of construction put-in-place from BOC -- government purchases; and the Consumer Price Index from BLS -- PCE.

A third common revision practice occurs when the agency revises a series to remove selected measurement errors but not necessarily to remove the largest errors, which therefore remain in the revised estimates. The following monthly GDP source data are examples of this type of series: monthly merchandise trade from BOC -- net exports; new residential construction put in place from BOC -- residential fixed investment; and average hourly earnings from BLS -- wages and salaries.

The monthly merchandise trade data are regularly revised to record the export and import transactions based on Customs documents in the correct month but not to record reliable estimates of transactions for which Customs documents are not required to be filed

because of the value of the transaction is below some cutoff or "low value." These unreported low value transactions, which currently account for about 2.5 percent of exports and 4 percent of imports, are estimated using factors based on the amount of such transactions when the exemptions were granted. The adjustments are extrapolated by country, but not by commodity, using changes in reported transactions. This procedure is likely to introduce errors into the monthly changes in exports and imports.

The new residential construction put-in-place series, which is based largely on a sample of housing starts, is regularly revised to reflect additional monthly reports but not to reflect more accurate annual survey data. For the new single-family residential construction component of this series, the value of housing starts are "phased" in over many months based on a fixed monthly pattern of construction activity. This practice introduces errors into the monthly changes in the put-in-place series because the pattern is not updated on a regular basis (the present pattern was estimated for 1976) and there are no data to adjust the pattern for developments such as unusually bad weather. The errors created by this practice can be illustrated using the recent California earthquake. This disaster not only delayed starts, which the series does reflect, but also delayed activity on previously started houses, which the series does not reflect. For new multi-family residential construction, BOC conducts a monthly survey of construction put in place based on a sample of housing starts; this series is not benchmarked.

The final example of a "selective" revision practice is the average hourly earnings series, which is estimated from the BLS monthly establishment survey. Although the employment data collected in that survey are benchmarked annually, the earnings data are corrected only to reflect the revised employment data, which are used to weight the detailed industry earnings estimates to arrive at "all industry" totals. It is likely that a benchmarked hourly earnings series not only would change the levels of the series, but also the monthly changes.

Improving revision practices

This discussion of revision practices identifies some of the types of measurement errors that remain in BEA's latest estimates of changes in quarterly GDP. Although reducing these errors for some series would require the collection of new data, it should be possible for the issuing agencies to reduce certain types of errors with little or no additional resources. For example, errors caused by proportionate allocation of the impact of delayed incorporation of births and deaths could be significantly reduced. Instead of waiting many years until a regular benchmark revision, agencies could continuously track births and deaths and adjust their series annually, even with a one-year lag. This procedure would improve the accuracy of the annual estimates earlier and largely eliminate the proportionate allocation of the errors at the time the agency prepares

their benchmark revision of the survey. (Discussions of such changes are currently underway between BEA and BOC for their annual surveys.)

For series where benchmark revisions are not carried backwards in time, agencies should publish the amount of the sample "drift" since the last benchmark and adjust the historical series. If it is not possible for the agency to make these adjustment, then BEA and other users could make their own adjustments. The latter solution is less desirable because the agencies usually have information with which to make the adjustments that is not available to users, and when different users make the adjustments, they are likely to develop different adjustments.

Table 1.--Reliability of Current Estimates of Quarterly Changes in Real GDP, 1978-91
 [Percentages providing correct indication]

	All quarters			Omitting quarters with changes/differences of 1 percent or less	
	Direction of change (56)	Larger/smaller change than in previous quarter (55)	Change between +1 percent and +4 percent (56)	Direction of change (48)	Larger/smaller change than in previous quarter (43)
Advance.....	88	78	75	92	86
Preliminary.....	89	75	70	94	81
Final.....	89	76	66	94	84

NOTE.--The number of comparisons is shown in parentheses.

Table 2.--Dispersion in Revisions in the Quarterly Changes
in GDP and Its Components
[Percentage points/1/]

	Current dollars		Constant dollars	
	1978-82	1983-91	1978-82	1983-91
Gross domestic product:				
Advance.....	1.93	1.17	1.64	1.25
Preliminary.....	1.82	1.14	1.72	1.27
Final.....	1.82	1.15	1.75	1.33
Personal consumption expenditures:				
Advance.....	1.71	1.40	1.88	1.37
Preliminary.....	1.70	1.41	2.00	1.27
Final.....	1.80	1.35	2.12	1.30
Durable goods:				
Advance.....	5.72	4.20	5.00	3.96
Preliminary.....	5.15	3.88	5.11	3.63
Final.....	5.42	3.97	5.05	3.98
Nondurable goods:				
Advance.....	2.31	1.74	1.75	2.26
Preliminary.....	2.51	1.45	2.37	2.10
Final.....	2.50	1.37	2.39	2.03
Services:				
Advance.....	1.78	1.37	1.38	1.39
Preliminary.....	1.90	1.51	1.50	1.36
Final.....	1.96	1.59	1.56	1.42
Gross private domestic investment:				
Advance.....	13.20	9.38	10.64	9.53
Preliminary.....	12.67	8.62	10.24	9.30
Final.....	12.11	8.68	10.75	9.32
Fixed investment:				
Advance.....	7.01	3.03	5.59	3.74
Preliminary.....	4.96	2.43	4.08	3.29
Final.....	4.45	2.77	3.82	3.64
Nonresidential:				
Advance.....	8.24	3.67	6.36	4.42
Preliminary.....	5.63	3.19	4.15	4.07
Final.....	5.11	3.20	3.62	4.56
Structures:				
Advance.....	13.01	6.39	9.01	5.33
Preliminary.....	9.90	4.54	6.18	4.13
Final.....	9.47	4.92	6.10	4.66

**Table 2.--Dispersion in Revisions in the Quarterly Changes
in GDP and Its Components--Continued**
[Percentage points/1/]

	Current dollars		Constant dollars	
	1978-82	1983-91	1978-82	1983-91
Producers' durable equipment:				
Advance.....	7.09	4.02	6.65	5.21
Preliminary.....	5.17	3.87	4.85	5.39
Final.....	4.20	3.99	4.42	5.77
Residential:				
Advance.....	7.17	4.84	6.91	5.27
Preliminary.....	8.56	4.91	8.67	5.11
Final.....	7.63	4.98	7.89	5.22
Change in business inventories...
Net exports of goods and services:				
Exports:				
Advance.....	8.90	5.49	7.52	5.33
Preliminary.....	8.80	4.72	7.87	4.85
Final.....	8.02	5.19	7.07	5.67
Imports:				
Advance.....	5.48	8.12	7.21	8.92
Preliminary.....	4.98	7.24	5.64	9.29
Final.....	4.71	7.55	5.71	9.61
Government purchases:				
Advance.....	4.25	3.93	3.46	4.83
Preliminary.....	4.37	4.05	3.62	4.79
Final.....	4.34	4.05	3.75	4.89
Federal:				
Advance.....	11.40	9.09	10.36	10.70
Preliminary.....	12.29	9.11	10.48	10.49
Final.....	12.81	8.92	10.99	10.58
State and local:				
Advance.....	2.51	1.53	2.15	1.41
Preliminary.....	2.61	1.63	2.17	1.62
Final.....	2.34	1.65	2.20	1.60

1. Calculated from quarterly percentage changes at seasonally adjusted annual rates.

Table 3.--Annual Averages of Dispersion and Bias in Revisions in the Quarterly Changes in GDP
[Percentage points/1/]

Year	Dispersion				Bias							
	Advance to Preliminary	Preliminary to Final	Advance to Final	Advance to Preliminary	Preliminary to Final	Advance to Final	Advance to Preliminary	Final				
									To third annual revision estimate			
Current-dollar estimates												
1978...	0.5	0.6	1.0	2.1	1.9	1.9	-0.5	-0.2	-0.7	-2.0	-1.6	-1.4
1979...	.5	.3	.3	1.2	1.1	1.0	-.4	.1	-.3	0	.4	.3
1980...	1.0	.5	.8	.7	1.1	1.3	.3	-.1	.2	-.3	-.6	-.5
1981...	2.0	.4	2.3	3.4	2.0	1.9	-2.0	-.2	-2.3	-3.1	-1.0	-.8
1982...	.7	.7	.3	1.8	1.3	1.7	.7	-.5	.1	.7	0	.5
1983...	.5	.2	.7	2.4	2.0	1.8	.3	0	.3	.5	.2	.2
1984...	.7	.4	1.0	1.6	1.8	1.7	-.4	-.3	-.6	.1	.5	.7
1985...	.4	.4	.8	.8	1.2	1.6	.4	.4	.8	-.8	-1.2	-1.6
1986...	.5	.2	.4	1.3	1.5	1.6	0	.2	.2	-.4	-.3	-.5
1987...	.6	.4	.8	1.2	.6	.8	-.6	0	-.6	-1.2	-.6	-.6
1988...	.7	.2	.9	1.4	.7	.5	-.7	-.2	-.9	-1.4	-.7	-.5
1989...	.6	.4	.8	1.3	1.3	1.3	0	.3	.3	.6	.7	.4
1990...	.2	.4	.6	NA	NA	NA	.2	.4	.6	NA	NA	NA
1991...	.4	.3	.4	NA	NA	NA	0	.2	.2	NA	NA	NA
1992...	.9	.2	.9	NA	NA	NA	-.9	-.1	-.9	NA	NA	NA

Table 3.--Annual Averages of Dispersion and Bias in Revisions in the Quarterly Changes in GDP--Continued
[Percentage points/1/]

Year	Dispersion				Bias							
	Advance to Preliminary	Preliminary to Final	Advance to Final	Advance To third annual revision estimate	Advance to Preliminary	Preliminary to Final	Advance to Final	Advance Preliminary	Final			
										To third annual revision estimate		
Constant-dollar estimates												
1978...	.2	.5	.7	2.4	2.3	2.0	-2	-.2	-4	-1.4	-1.2	-1.0
1979...	.8	.2	.6	1.0	.8	.8	-.6	.1	-.5	-.5	.1	0
1980...	.5	.7	.7	.5	.5	.8	.3	-.3	0	-.2	-.5	-.2
1981...	1.0	.3	1.2	2.0	1.3	1.1	-1.0	-.2	-1.2	-1.8	-.8	-.6
1982...	.6	.6	.2	2.4	1.8	2.4	.3	-.5	-.2	1.0	.8	1.2
1983...	.4	.4	.7	1.5	1.4	1.1	0	-.1	-.1	-.3	-.3	-.2
1984...	.7	.4	1.0	1.5	1.8	1.8	-.3	-.1	-.4	-.2	.5	.6
1985...	.7	.5	.7	1.5	1.7	2.1	.2	.5	.6	-1.5	-1.7	-2.1
1986...	.4	.4	.2	2.1	2.1	2.3	0	.1	.1	0	0	-.1
1987...	.3	.3	.3	1.2	1.1	1.3	-.1	-.1	-.2	-1.2	-1.0	-.9
1988...	.5	.1	.6	.9	1.2	1.3	-.5	0	-.5	-.9	-.4	-.4
1989...	.6	.2	.7	1.7	1.6	1.6	0	0	0	1.1	1.1	1.0
1990...	.3	.3	.6	NA	NA	NA	.3	.3	.6	NA	NA	NA
1991...	.5	.3	.5	NA	NA	NA	.1	.3	.4	NA	NA	NA
1992...	.7	.3	.6	NA	NA	NA	-.7	.1	-.6	NA	NA	NA

NA-Not Available
1. Calculated from quarterly percentage changes at seasonally adjusted annual rates.

Table 4.--Dispersion in Revisions in the Quarterly Changes
in GDP and Its Components, 1981-90/1/
[Percentage points/2/]

	Current dollars	Constant dollars
Gross domestic product:		
Advance.....	1.38	1.32
Preliminary.....	1.31	1.36
Final.....	1.34	1.45
First Annual.....	.84	.83
Second Annual.....	.85	1.08
Third Annual.....	.65	.85
Personal consumption expenditures:		
Advance.....	1.48	1.47
Preliminary.....	1.50	1.39
Final.....	1.51	1.49
First Annual.....	1.32	1.19
Second Annual.....	.82	.71
Third Annual.....	.50	.47
Durable goods:		
Advance.....	4.49	4.22
Preliminary.....	4.15	4.16
Final.....	4.46	4.55
First Annual.....	3.27	3.41
Second Annual.....	1.79	1.91
Third Annual.....	1.84	1.94
Nondurable goods:		
Advance.....	1.85	2.15
Preliminary.....	1.61	2.09
Final.....	1.56	2.04
First Annual.....	1.36	1.38
Second Annual.....	1.10	1.03
Third Annual.....	.63	.63
Services:		
Advance.....	1.46	1.34
Preliminary.....	1.62	1.35
Final.....	1.74	1.45
First Annual.....	1.62	1.23
Second Annual.....	1.11	.97
Third Annual.....	.70	.58

Table 4.--Dispersion in Revisions in the Quarterly Changes
in GDP and Its Components, 1981-90/1/--Continued
[Percentage points/2/]

	Current dollars	Constant dollars
Gross private domestic investment:		
Advance.....	11.38	10.48
Preliminary.....	10.37	9.93
Final.....	10.48	10.13
First Annual.....	7.10	7.21
Second Annual.....	5.28	5.12
Third Annual.....	4.88	5.22
Fixed investment:		
Advance.....	3.97	4.22
Preliminary.....	3.28	3.83
Final.....	3.38	3.99
First Annual.....	3.09	3.27
Second Annual.....	2.59	2.93
Third Annual.....	2.09	1.87
Nonresidential:		
Advance.....	4.61	4.84
Preliminary.....	3.99	4.56
Final.....	3.73	4.71
First Annual.....	4.21	4.50
Second Annual.....	3.99	4.65
Third Annual.....	2.37	2.06
Structures:		
Advance.....	7.98	6.19
Preliminary.....	6.17	4.73
Final.....	6.48	5.16
First Annual.....	6.66	5.78
Second Annual.....	6.49	5.87
Third Annual.....	5.07	4.31

Table 4.--Dispersion in Revisions in the Quarterly Changes
in GDP and Its Components, 1981-90/1/--Continued
[Percentage points/2/]

	Current dollars	Constant dollars
Producers' durable equipment:		
Advance.....	4.59	5.56
Preliminary.....	4.10	5.61
Final.....	3.85	5.66
First Annual.....	4.59	6.30
Second Annual.....	3.79	5.31
Third Annual.....	1.40	2.84
Residential:		
Advance.....	6.30	6.61
Preliminary.....	6.48	6.58
Final.....	6.35	6.46
First Annual.....	4.69	5.72
Second Annual.....	3.81	4.22
Third Annual.....	3.54	3.62
Change in business inventories.....
Net exports of goods and services:		
Exports:		
Advance.....	6.13	5.80
Preliminary.....	5.57	5.29
Final.....	5.63	5.61
First Annual.....	4.41	4.25
Second Annual.....	3.09	3.52
Third Annual.....	3.73	4.20
Imports:		
Advance.....	7.99	8.74
Preliminary.....	7.85	9.38
Final.....	7.95	9.46
First Annual.....	4.76	5.50
Second Annual.....	3.44	3.94
Third Annual.....	3.25	4.80

Table 4.--Dispersion in Revisions in the Quarterly Changes
in GDP and Its Components, 1981-90/1/--Continued
[Percentage points/2/]

	Current dollars	Constant dollars
Government purchases:		
Advance.....	4.22	4.74
Preliminary.....	4.46	4.87
Final.....	4.51	5.01
First Annual.....	4.19	4.49
Second Annual.....	3.52	3.75
Third Annual.....	2.91	3.21
Federal:		
Advance.....	10.01	11.30
Preliminary.....	10.54	11.48
Final.....	10.61	11.78
First Annual.....	10.45	10.75
Second Annual.....	8.75	9.55
Third Annual.....	7.14	7.55
State and local:		
Advance.....	1.66	1.56
Preliminary.....	1.81	1.76
Final.....	1.83	1.71
First Annual.....	1.49	1.43
Second Annual.....	1.18	1.17
Third Annual.....	.85	.71

1. As previously indicated in the text, for the comparisons shown in this table, it was necessary to use the period 1981-90. In addition, because the annual revisions in 1985 and 1991 were replaced by comprehensive revisions, the comparisons exclude years with no annual revision. The first annual revision comparisons exclude 1984 and 1990, the second annual comparisons exclude 1983 and 1989, and the third annual comparisons exclude 1982 and 1988.

2. Calculated from quarterly percentage changes at seasonally adjusted annual rates.

RAISING THE NATION'S UNEMPLOYMENT RATE

John E. Bregger
U.S. Bureau of Labor Statistics

I. Introduction

Data released on February 4, 1994, reflected major revisions in the questionnaire and collection methodology that were introduced into the Current Population Survey (CPS), following a planning and developmental process extending over the previous 8-year period. Looking back over these 8 years, was the process worth the effort and cost? Did the data improvements exceed the losses engendered by the breaks in time series? Were public understanding and appreciation of the data negatively affected by the changes? These and many other questions will hopefully be answered in this paper.

II. What happened?

In a nutshell, the nation's overall unemployment rate was found to be somewhat higher in early 1994 than it was in late 1993, resulting directly from a wholesale, stem to stern, series of changes to the survey questionnaire and the total conversion to computer-assisted interviewing. In addition to the rate of unemployment, a number of other important data series were affected by the CPS revisions, including the estimation of discouragement and of persons working part time involuntarily.

Based on tests of the new system, the effect on the overall annual average unemployment rate for 1993 was estimated to be about half a percentage point. That is, utilizing data gathered from a totally separate, parallel, survey, the newly redesigned questions, asked by interviewers using laptop computers or calling from a centralized interviewing facility, identified more people unemployed than under the then current procedures. And if that wasn't dramatic enough, changes in definition of discouragement caused the figure to tumble some 60 percent, and measurement refinements lowered economic part-time employment by 20-25 percent. In other words, of the three most important measures of labor market slack, the one most people point to and talk about -- the unemployment rate -- has now been raised (on a statistically significant basis), while the other two have been lowered markedly. Early results (January-April 1994) from the implementation of the new survey and procedures confirmed the direction of the expected changes.

III. History

How (or why) did this come about? The answer to this question comes in several parts, involving secular changes in the economy, a Presidential commission, the advent of new technologies and surveying knowledges, and, of course, careful planning. These will all be briefly described in a whirlwind tour of the historical backdrop for changing the Current Population Survey.

First, the secular changes. The last time that the questionnaire had been changed to any degree was in 1967, resulting from a period of research in the aftermath of the President's Committee on Employment and Unemployment Statistics (the Gordon Committee). In the subsequent years, many societal changes have taken place, including the more prominent role of women, especially mothers, in the labor force; the continuing shift from a goods- to a service-producing economy; changes in the way business operates, such as opening and closing hours; and, somewhat related to the other factors just cited, shifts in the nature of employment, including more part-time work and less permanent attachment of employees to their employers.

The next Presidential commission to study the statistics, the National Commission on Employment and Unemployment Statistics (the Levitan Commission), issued its report, including a number of recommendations in the labor force area, on Labor Day 1979. For our purposes, its most significant recommendation was for major conceptual changes in the way we measure labor market discouragement. And, while this particular recommendation was accepted for implementation by Secretary of Labor Donovan two years later, it was not implemented, owing to a lack of available funding in the early 1980s for a parallel test panel of households for testing potential questionnaire changes.

At the same time that these developments were going on, there have been many innovations in the way data are collected, innovations that could be expected to improve the quality of data. Foremost among these have been the recognition of the relevance of the theories and methods of cognitive psychology in designing survey instruments and the use of the computer in the interviewing process. With respect to cognitive psychology, under the auspices of the Committee on National Statistics, National Research Council, psychologists, other behavioral scientists, and survey methodologists had come together in 1984 to discuss the contributions that each discipline could make to survey design and, in so doing, helped to launch the cognitive aspects of survey methodology movement. One of the legacies of that advanced seminar is a four-component cognitive model of the question-response process -- comprehension,

retrieval, judgment, and response -- that has provided a very useful framework for designing and evaluating survey questions.

Use of the computer for data collection has been around for some time, but perhaps not for such a large undertaking as a monthly sample survey of 60,000 households, and certainly not for use in personal decentralized interviewing. Testing had suggested that, not only was a large-scale application doable, but, more importantly, it offered incredible gains in a variety of ways. Among these were fewer constraints on the number or variations of questions that could be included in the instrument; greater accuracy of data collection, in that interviewers were more likely to ask questions as worded (some had been anticipating what question they would be asking next); and accuracy of data transcription and transmission. When coupled with the desire to change and add questions to improve overall accuracy of identifying labor force status, the potential for improvement was ever so much greater, because the computer could permit intricate skipping and the storage of earlier information for later use that no interviewer could carry out in a pencil and paper environment.

The planning process for carrying out all of this commenced with a series of conferences involving the senior staffs of the Bureau of Labor Statistics and the Census. The two agencies held a series of meetings beginning in April 1986 and two years later had a detailed plan to redesign the CPS, essentially in its entirety, with the questionnaire-related changes being the centerpiece. Budget submissions, with extensive, year by year, spending plans, were sent forward to OMB in time for the 1990 budget cycle. And, with favorable indications of approval forthcoming, work actually began in late 1988 toward a comprehensive survey redesign, with 1994 being the principal target date.

With respect to revisions to the questionnaire, a number of BLS-Census work groups were set up to develop a new questionnaire. The questionnaire was to be designed under the following guidelines: 1) It would not be constrained by the one-page limitation then in place; 2) it would take advantage of all aspects of automated data collection; 3) it would build upon recommendations of the 1978-79 Presidential Commission (and, to a lesser degree, the 1961-62 Commission); and 4) it would utilize to the maximum extent possible the knowledges available from cognitive science. Behavioral science laboratories were established in both agencies that brought in volunteers from the outside to react to various questions and question sequences. Questions explored included: What are the meaning of terms such as "work," "last week," "layoff," and "private company?" What method of collecting information on

the "actual number of hours worked last week" produces the most accurate data? How can response options be revised to simplify reporting and improve the categorization process -- and, in doing so, reduce measurement error? How could sensitive questions, such as on earnings, be revised to minimize nonresponse and improve reporting accuracy? How might the process of verifying information from a prior month's interview, rather than asking for the same information every month affect the quality and accuracy of the data? By the time we were through, we had managed to come up with satisfactory answers, for ourselves at least, to most of these as well as many other questions.

As a new questionnaire began to take shape, field testing became the next order of business. By this time (1990), two alternative versions of a potential new set of questions were in hand, along with the CPS questions then in use. The goal was to determine the best overall question wording from the three. To do this, the two Bureaus conducted a computer-assisted telephone interview (CATI), random digit dialing (RDD) test at the Census Bureau's centralized interviewing center in Hagerstown, Maryland. The first phase of this CATI/RDD test extended from July 1990 to January 1991, involving approximately 72,000 persons. Its purpose was to compare the then current version of the CPS questionnaire with the two test versions.

The principal product of the first phase was the selection of a single alternative questionnaire, close to the official version now in place, with appropriate additions and improvements that were deemed necessary due to the results of the testing. A second test phase was conducted between July and October of 1991 with approximately 30,000 persons, again via CATI/RDD; with very limited changes, this became the final version to be used in an 18-month parallel survey. During both the phase-one and phase-two testing, as well as the parallel survey phase, researchers employed a variety of methodologies to evaluate alternative question formats. These included respondent debriefings (via follow-up probe questions and vignettes), interviewer debriefings (via focus groups and debriefing questionnaires), response-distribution analysis, item nonresponse analysis, and behavior coding.

IV. The parallel survey

As researchers have long understood and as was once again verified in the CATI/RDD testing, if one or more important questions are changed (even slightly) in a continuing survey setting, we can expect different results. In other words, changing several questions in the CPS could be expected to have an effect on major measures, such as the rate of unemployment. Since the total number of potential questions increased from about 45 to 128 (no one, of course,

ever is asked more than a few of these!) and the wording of almost every question was changed, there was a virtual guarantee that we should expect differences on most of the statistical measures emanating from the survey. Thus, it was more than prudent to plan for a parallel, or overlap, survey for an adequate period of time in order to get some handle on the pact of these changes.

Ideally, we would have liked to have had a parallel sample extending for at least 2-1/2 years, with the same number of households as the ongoing CPS. This would have guaranteed a fully seated set of sample data for a full year, in terms of the 4-8-4 rotation group pattern. But, because our funding was limited, we had to settle for a 12,000 household sample covering the 18-month period, July 1992 through December 1993. Termed the CATI-CAPI Overlap Survey (CCO) internally and the Parallel Survey (PS) externally, this survey introduced the laptop computer (the CAPI portion of the CCO) into large-scale data collection.

One of our initial concerns was how well interviewers would adapt to using laptop computers and whether respondents would react favorably as well. We need not have worried: Both groups seemed to be happier. Interviewers, while concerned that questions did not pop up on their screens fast enough, appreciated the accuracy of the computer and thought that using it made them appear more professional. Respondents who were interviewed in person appeared to be more interested in the survey -- some, for example, invited interviewers to "plug in" -- and paid closer attention to the questions.

Ideally, with changes to the ultimate CPS coming from two directions -- the questionnaire itself and computer-assisted interviewing -- it would have been desirable to isolate the data effects on differences of these changes (questionnaire and interview mode). Alas, this was not possible. As a consequence, the significant differences between the on-going CPS and the PS that were identified can only be ascribed to the overall change in the survey and not specifically to the questionnaire or collection mode. Thus, we have been unable to discern, for example, what the specific effects have been on, say, the overall unemployment rate from changes in the questionnaire wording and question sequencing.

As soon as early PS figures started becoming available to BLS and Census researchers, it became obvious that we were indeed seeing marked changes in important statistical measures. The overall unemployment rate was higher, particularly among women and older workers but essentially across all worker groups. The employed differences were especially interesting, because more women but fewer men were found to have jobs.

There were also other dramatic changes. As expected, the new, more restrictive, measurement of discouraged workers resulted in some 60 percent fewer persons being counted in that category. To be classified as discouraged under the revised scheme, persons who wanted a job but had not looked for work in the prior 4 weeks had to have searched for work during the prior 12 months and not currently looking for work because of discouragement over the job market, while also being available to take a job during the reference week. Similarly, as a result of better question specificity, there was a 20-25 percent reduction in the number of persons working part time for economic reasons (that is, working less than 35 hours during the reference week because of poor business conditions or because of an inability to find full-time work). To be so classified, a person who usually works part time must now indicate that s/he wants a full-time job and was available to take one during the reference week.

V. Communications with official Washington

As soon as the researchers were able to verify that they had accurate data from the PS and thus could estimate the differences that the new questionnaire and collection methodology were yielding, taking appropriate measures of statistical significance into account, it was time to start communicating "up the line." We were, quite naturally, concerned about what kind of reactions there would be to a significantly higher rate of unemployment. There was, after all, a new administration in office that perhaps did not need to be saddled with yet another major issue. It had enough on its plate already.

The researchers had put together a formidable package of tables and analysis, with explanations for the many diverse changes observed over the comparison periods. Initially, 6-month comparisons of the PS with official CPS, covering the period September 1992-February 1993, were utilized, and these were the first figures to be viewed at higher levels, first of all with the heads of the two agencies. Soon thereafter, the first annual average data became available, representing the period September 1992-August 1993. Using these figures, a memorandum detailing the changes that we were expecting to introduce in January 1994 and the expected data effects was sent to Secretary Reich in late October, and this memorandum was forwarded on by him to the President. An hour-long meeting was held with Secretary Reich and top Department of Labor staff on November 1, and this was followed in short order by other high-level briefing sessions with other members of the Administration and the Federal Reserve Board.

Reaction was reasonably swift. All of the changes that were contemplated for implementation in January 1994 were fully acceptable, despite a concern that, with a higher rate of unemployment, the public might fail to recognize that the economy was still gaining steam. Indeed, it was this concern that led directly to a request to sustain the collection of parallel survey data using the paper and pencil methodology beyond the year and a half that had been planned and funded. Monies were found and commitments made to sustain the parallel survey beginning in January 1994, this time with the old questionnaire and procedures. That way, after an initial period where respondents and interviewers might be affected by the previous test, we would have a continuing measurement of the differences that were identified for 1993 as the data on the new basis became available.

VI. Communications with the outside world

A detailed planning document that had been produced and constantly updated had long identified November 16, 1993, as the date of the first public announcement of the plans for introducing changes into the Current Population Survey and detailing what the expected data effects were. Armed with briefing packets and a plethora of other useful information, Commission Katharine Abraham and members of the BLS and Census Bureau staffs presented an extensive array of information to the national economic media. Articles appearing throughout the country the next morning, as well as the more immediate wire service stories, suggested that the press well understood what was transpiring. In particular, the notion of "gender bias," which had emerged from data findings, was significantly played up. With few exceptions, they got it right. Not all did, however, as suggested by the headline, "U.S. won't ask women if they cleaned house."

BLS didn't stop with a one-day media session in Washington, D.C. There was a full-day session with technical users on the next day (November 17), also in Washington, which some 150 persons attended. Interest was running high. In December and January, combined data user-media sessions were held in 13 other large cities throughout the entire country, including New York, Chicago, and Los Angeles. Many people turned out to learn what was expected to happen and how their local unemployment rates might be affected.

By this time, we had unemployment rate comparisons for the Census regions and divisions, as well as some data for seven large states. Our uncertainty was quite high as to the reliability of our sub-national comparisons, and this was carefully communicated. Fortunately, our concerns did not fall on deaf ears, and most people, including the local

media, did not play up some of the wide differentials, such as rates that were slightly more than a point higher in the Middle Atlantic Division (covering New York, New Jersey, and Pennsylvania). This was all the more fortunate when, in actual fact, this particular region did not show large jumps in early 1994. In other words, our concerns about the reliability of sub-national data were well-founded, especially since the PS, unlike the CPS, was not a State-based design and the sample size was so much smaller.

VII. The final outcome and lessons learned

By the time the data for January 1994 were released on February 4th, it seemed that everyone -- government officials at all levels, the media, financial analysts, and the public at large -- was well aware that big changes were to be expected. And thus there were seemingly no surprises. The overall rate of unemployment for January was 0.3 percentage point higher than the December 1993 figure, quite reasonable, given expectations that we could expect as much as an 0.6 increase on an annual average basis (all other things remaining equal), 0.5 from the questionnaire and methodological changes and 0.1 from the introduction of the 1990 census-based population figures (adjusted for the estimated undercount) into the estimation procedures.

Did these data results from January, which were followed by an 0.2 percentage point drop in February, imply that our expectations based on parallel survey results -- the population effect was "guaranteed" -- were too large? Or, did January and then February really show large improvements in the underlying rate of unemployment? Even with two more months in, I think that we are still waiting to answer these two questions with more data. (Isn't that always the case?) It would appear, however, that our expectations for 1994 results for the official figures are essentially accurate, that is, the new questions and methodology suggests that the old questions did a good job of measuring mainstream labor market behavior, but not as well for more marginal types of activities, such as might be typical for certain women, youth, and older persons, for whom more jobseeking and more jobholding were found. Now, these missed activities tend to be of a seasonal nature and thus more likely to occur in certain months of the year. January, February, and March are months for which this sort of seasonality is fairly low; it can be expected to be much higher in months like May, June, and July. This implies that we can therefore anticipate higher levels of activity, particularly jobseeking activity, in the spring and summer months. And our seasonal adjustments, which are for the moment necessarily based on experience under the former procedures through the end of 1993, are somewhat "off." So, the answer to the second question regarding the January and February (plus March and April) results would appear to be

that, yes, we were seeing some improvement in the economy, but perhaps not quite as good as implied by these figures.

These early results also suggest that it will be quite some time before we have a full, clear realization as to all of the data effects that have been brought about with the new CPS. Ideally, for example, we should have had a longer lead-time than a year and a half with the parallel panel, so that it could have settled in better and given us more direct comparisons with the official CPS figures. Budget exigencies rarely resolve the hindsight "shouldas." New seasonal patterns are not fully discerned for at least 5 years, and we therefore may have to wait that long to expect to attain a degree of accuracy in month-to-month movements in employment and unemployment that we are fully comfortable with. The seasonal adjustment process will improve over time as data based on the new procedures are gradually taken into account. It is also possible that other improvements could be made to expedite the process.

The seasonality issue just discussed and the potential breaks in series for a number of measures, particularly those of labor market slack, inevitably raise the question as to the whether the process should have been embarked on at all. From my own viewpoint, the answer is clear: Breaks in series and comparatively short losses in time-series comparisons, while never desirable with any degree of frequency, are vitally necessary to ensure that we are accurately measuring what is occurring in our economy. We must recognize that there is always a cost to bring about improvements in data collection of economic phenomena. If we focus only on data consistency and therefore take our eyes off the prize of data improvement in a constantly changing society, we will never even attempt to undertake improvements in the measurement and collection of statistical surveys in the first place. Once undertaken, it is imperative that we go all the way, that is, make all of the improvements that are discernable and viable and then carefully measure their impact through a separate parallel survey. That is precisely what we have done with the CPS, and I firmly believe that the payoff was well worth the short-term losses that we are experiencing.

Perhaps the most significant lesson we learned from all of this was one that was a major winner: full, extensive communication. By careful interaction with, firstly, our internal customers -- i.e., the Administration and the Congress -- and then our external customers -- the media and the public -- there were few, if any, surprises. Friday, February 4th, turned out to be a business as usual, ho-hum day. Everyone knew or thought they knew what was going to happen, their expectations were more or less met, and thus not a lot of news was good news. My recommendation for any statistical agency undertaking major changes in surveys or

data series, therefore, is to err heavily on the side of both extensive and continuing communication with every possible group -- not just the media, not just here in Washington, but with everyone everywhere.

Carrying out the questionnaire-related redesign of the Current Population Survey cost the taxpayers an estimated \$30 million. Was it a worthwhile expense? Coming from a highly biased person, one can take my answer with a grain or two of salt, which is a resounding yes! In return, the Nation is getting better, more accurate figures on the labor force activities of the population. It is getting new kinds of important statistics, such as monthly data on labor market discouragement (on a totally revised conceptual basis) and on multiple jobholding. And it is getting the assurance that the measurements of the labor force, employment, unemployment, and those not in the labor force have been carefully studied and researched. And that, I would argue, is an incredibly great bang for our bucks.

A potential lesson that I hope we will not forget is that the total job is not as yet completed. The data comparisons for 1993, based on parallel survey and official statistics, need to be studied much closer than we have been able to thus far. The new figures for 1994 and beyond will require careful analysis. Continuing research on bridging data estimation both back in time and forward in time should continue, with the intent of assisting time series users in their analytical endeavors. We should have learned well the benefits that behavioral science has given us in terms of all future data collection; thus, for example, survey supplements, whether ongoing ones such as income and work experience or new ones such as the upcoming inquiry into contingent work, should be subjected to careful cognitive testing. Lastly, we should carry forward what we have learned into other CPS-related areas -- such as instituting improvements into the "control card," in which demographic characteristics are identified, or in improving the coverage of minority groups in the data collection process.

Finally, the last Presidential commission to examine labor force statistics issued its report 15 years ago. It took that long to implement a couple of its important recommendations. It would not be too radical to suggest that another commission ought to be established in the not too distant future to assess the viability and adequacy of the 1994 changes and then to determine appropriate future directions as we embark on the 21st century.

References

Chester E. Bowie, Lawrence S. Cahoon, and Elizabeth A. Martin, Overhauling the Current Population Survey: Evaluating changes in the estimates, Monthly Labor Review, September 1993.

John E. Bregger, The Current Population Survey: a historical view and BLS' role, Monthly Labor Review, June 1984.

John E. Bregger and Cathryn S. Dipbo, Overhauling the Current Population Survey: Why is it necessary to change?, Monthly Labor Review, September 1993.

Thomas B. Jabine, Miron L. Straf, Judith M. Tanur, and Roger Tourangeau, Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines, 1984 (National Academy Press, Washington, D.C.).

Elizabeth A. Martin, "Surveys as Social Indicators: Problems in Maintaining Trends, Chapter 16 to Handbook of Survey Research, 1983 (Academia Press, Inc.).

National Commission on Employment and Unemployment Statistics, Counting the Labor Force, Labor Day, 1979 (U.S. Government Printing Office, Washington, D.C.)

Anne E. Polivka, Comparisons of Labor Force Estimates from the Parallel Survey and the CPS During 1993: Major Labor Force Estimates, CPS Overlap Analysis Team Technical Report 1, March 18, 1994.

Anne E. Polivka and Jennifer M. Rothgeb, Overhauling the Current Population Survey: Redesigning the Questionnaire, Monthly Labor Review, September 1993.

President's Committee to Appraise Employment and Unemployment Statistics, Measuring Employment and Unemployment, 1962 (U.S. Government Printing Office, Washington, D.C.)

Jennifer M. Rothgeb, Revisions to the CPS Questionnaire: 1994 Effects on Data Quality, CPS Overlap Analysis Team Technical Report 2, April 6, 1994.

Jenny Thompson, Mode Effects Analysis of Labor Force Estimates, CPS Overlap Analysis Team Technical Report 3, April 14, 1994.

COMMENTS ON PARKER AND WEADOCK, TIME SERIES REVISIONS:
THE EFFECTS ON GROSS DOMESTIC PRODUCT

Murray F. Foss
American Enterprise Institute

I. Introduction

This paper by Robert Parker and Teresa Weadock is an interesting study that extends the series of evaluations of quarterly GDP estimates undertaken in BEA by Allan Young in 1974 (and most recently in the October 1993 Survey of Current Business) and initiated by George Jaszi a decade earlier.

In my comments I will discuss some differences between the current and constant dollar figures. I suggest other things the authors might have looked at, some of which would be quite easy. I then raise the question about which figures are the best ones for gauging "the true quarterly change," a point implied by the authors in their criticism of source data. And then I ask what lessons we should learn from all of this.

II. Lack of symmetry between current and constant dollar measures

The Parker-Weadock (henceforth PW) measures of reliability are typically presented in terms of current dollars and constant dollars. We are interested in both but the two sets of figures are really not symmetrical, and I wish the authors had discussed this asymmetry. While the monthly PPI data are revised once going back 4 months as a result of the incorporation of late returns and the correction of errors by respondents and by BLS in the initial reports, this is not true of the current monthly CPI. Yet if we look at the succession of dispersion measures in Table 2 from advance to preliminary to final, for GDP as a whole as well as for personal consumption the dispersion gets worse in the constant dollar series.

Aside from making very few revisions on a monthly basis it is not the practice of BLS to conduct a bigger survey after the calendar year is over--what might be an Annual Survey of Prices, analogous to, say, the Annual Survey of Manufactures. Every ten years BLS changes its market basket for the CPI to take account of changes in consumption patterns. New and different products are appearing on the market constantly, and these BLS treats in a variety of ways, depending on continuing probability sampling to pick up new products and types of outlets.

It would be good to know whether this deterioration in the reliability of successive constant dollar estimates is statistically significant. Is it simply a reflection of newer seasonal factors, which are revised by BLS each year going back several years, and to what extent does it reflect a benchmark (10-year) change?

III. Some other perspectives of reliability

PW, like Allan Young, examine the reliability of the quarterly estimates of GDP from an historical point of view. That is useful because it permits one to say something about possible long-term trends in reliability. But outlined below are other ways that I would like to see examined.

A. Business cycle perspective

The GDP statistics are the single most important indicator about what is going on in the national economy. But as we all know economists are often in the dark about whether an expansion has begun or whether the economy has slipped into a recession. So I would like to see how these reliability measures--both dispersion and bias--behaved around turning points. Looking at the historical record we can ask if there are any patterns, for example, in the four quarters up to and including the business cycle peak and in the first four quarters of downturns (which average not quite a year in length). Do these patterns differ from one another? Would they differ from the pattern in the first four quarters of an upturn? It would not be hard to find a rationale for any differences that might turn up; for example something concerning the quality of statistics within the firm over the business cycle but any patterns would be of interest in themselves.

B. Inflation

It would be interesting to examine the data for possible differences when the rate of inflation differed. Is there a difference between 1973-80, when inflation was very high and 1983-90, when inflation was much lower? It is more difficult to capture a change in real output when inflation is high than when it is low. When buyers resort to new sources of supply or when sellers change their discounts from list prices, the Producer Price Indexes may be slow to adapt even though the current dollar figures on sales reflect these changes immediately.

C. The current data

The first three estimates of a given quarter--the advance, preliminary and final--carry a lot of weight because these are the figures that affect decisions by business and government in the short run. So one could use this criterion: given the advance and the final, how often did the preliminary move in the direction of the final (third) change? For example, if the advance change is 1.9 and the final is 1.5, we can ask if the preliminary moved down from 1.9 or exceeded it. Small misses in direction would be ignored, following the authors' approach.

D. Calendar quarters separately

It would also be a simple matter to collate the measures of dispersion and bias by calendar quarters, to see if the fourth quarter differs from the other three quarters. The reasoning behind this is that in organizing their work accountants put most emphasis on the annual report, which comes out a few months after the end of the year. (I realize that fiscal years pose some problem). Accountants do things at the end of the year that they don't do during the year. For example, they may take a physical count of inventories at year-end but use shortcut methods to estimate inventories for months and quarters. Earlier errors that show up at the end of the year are corrected in the final quarter. If the errors are in one direction the fourth quarter correction will tend to be reduced if not eliminated by seasonal adjustments. After year-end, accountants may go back and revise earlier quarterly figures; this may be a regulatory requirement but I am not certain. If accountants in fact do a lot of estimating during the year such a practice could give rise to revisions between advance-preliminary, on the one hand, and final-first annual, on the other.

E. Final sales and inventory change

Estimating the quarterly change in business inventories is an inherently difficult task and remains so even with the many improvements made by the Census Bureau and BEA over the years. The change in inventory change is ordinarily a significant part of the average quarterly change in GDP. The inventory estimates are not shown explicitly by the authors because of the particular measures they employ for GDP and all other components. It would seem from table 4 that revisions in inventories are a significant source of total revisions. It would be a good idea to examine a common measure published by BEA, namely, total final sales, which excludes inventory change. Obviously the shifting of farm inventories between the Federal Government and private business creates a problem but it would not seem to be too difficult to make allowance for this.

As a matter of fact, the change in business inventories ought to be shown explicitly with its own reliability measures because it is so difficult and involves much judgment not only by BEA but also, I would guess, at the firm level. This suggests an additional reason why the authors should show successive revisions of nonfarm CBI: the monthly CBI's are subject to far greater variation than any of the flow components. Maybe exports and imports as now calculated would be close runnersup. PW and Allan Young point out that the final month in the quarter has a weight of only one-ninth and the second and first months weights of two-ninths and three-ninths, respectively. If the expenditure components were random numbers, the fact that they have such "small" weight would not be so important. The fact is that this month's seasonally adjusted retail sales must be very close to last month's. A one percent

difference is a big seasonally adjusted change. But that is not true of CBI. One month of inventory change can be positive, the next month, negative. Such a pattern is possible because sales can be higher or lower than expected by the firm and, with production plans based on expected sales, inventories will be correspondingly lower or higher. The same is true of incoming supplies to the purchasing firm--a consequence of capacity limitations, strikes, natural disasters, etc. This is not to deny that during the expansion firms tend to build stocks and during the contraction they cut them. But I urge the authors to do the dispersion measures of nonfarm CBI (and the GDP) in constant dollars. This would be a good test.

IV. What is the "true" quarterly change?

Given the way quarterly data are revised to make them compatible with subsequent annual figures and benchmark annual totals from the quinquennial censuses, how can we be sure that the very final quarterly pattern that emerges is superior to all previously published quarterly data for a given year? Parker and Weadock criticize the Census Bureau for making proportional adjustments in originally published monthly and quarterly data. This is an old problem. For example, Morris Cohen raised the same issue at an Income and Wealth Conference 15 years ago. He said that the data were being oversmoothed and that cyclical fluctuations were being damped if not eliminated. The late Otto Eckstein agreed with this point of view but it remains a minority opinion. The dominant revision philosophy is above all to get the long-term trend correct. If that is so, cyclical fluctuations must be fitted into the trend (that is, benchmark) values for a given year.

The answer, of course, is to get more and better within-the-year data. There is no substitute for this. That was said 15 years ago and, I am sure, many times before that. People who make decisions in business and government have a big stake in the currently available quarterly numbers and after some 50 years deserve more improvements than the agencies have made. Economists studying the business cycle have an important interest in getting the record straight.

V. What lessons should we learn from these studies?

Parker and Weadock, like Allan Young, raise the possibility of dispensing with the second and third quarterly estimates for a given quarter. A single current estimate, namely, the advance, might conceivably save some money. I am not sure that the nation would be better off. As Allan Young points out, the detailed estimates might suffer. As for the total, there are so many people and firms engaged in this business today that several estimates would make their appearance to fill the void left by BEA. Unfortunately these estimates would differ from one another. I would guess that large organizations like the Federal Reserve would

make their own estimates. Estimates of GDP made by outsiders are not likely to be as good as those made by BEA. So I am inclined to stick to the present system. It is less bad than what might supplant it.

Parker and Weadock are impressed by the improvement in reliability from final current to the first July estimate. They should, of course, find out how much of the improvement is due to better seasonals. But they should not be surprised that firms send in better data after the year is over.

I think that both the source agencies and BEA should do more field work to find out why numbers submitted to the government change. The agencies ought to do more to find out how firms obtain their monthly and quarterly numbers, the extent to which they rely on within-the-firm estimates and end-of-year adjustments. Sales probably are not much of a problem. But other statistics like inventories and profits may be. Irving Rottenberg and I found that firms using LIFO accounting had great trouble estimating inventories on a monthly basis. The fact is that firms use shortcuts of all sorts; these are doubtless useful to the firm but may be subject to biases that are functions of the stage of the business cycle or the rate of inflation. I don't have in mind huge field surveys checking data quality. One can get many insights from small field trips. I believe that BEA at least does too little in this regard. And, to repeat, Census and BLS ought to do better in obtaining within the year figures to accompany more comprehensive annual and benchmark surveys.

**Reports Available in the
Statistical Policy
Working Paper Series**

1. Report on Statistics for Allocation of Funds (Available through NTIS Document Sales, PB86-211521/AS)
2. Report on Statistical Disclosure and Disclosure-Avoidance Techniques (NTIS Document sales, PB86-211539/AS)
3. An Error Profile: Employment as Measures by the Current Population Survey (NTIS Document Sales PB86-214269/AS)
4. Glossary of Nonsampling Error Terms: An Illustration of a Semantic Problem in Statistics (NTIS Document Sales, PB86-211547/AS)
5. Report on Exact and Statistical Matching Techniques (NTIS Document Sales, PB86-215829/AS)
6. Report on Statistical Uses of Administrative Records (NTIS Document Sales, PB86-214285/AS)
7. An Interagency review of Time-Series Revision Policies (NTIS Document Sales, PB86-232451/AS)
8. Statistical Interagency Agreements (NTIS Documents Sales, PB86-230570/AS)
9. Contracting for Surveys (NTIS Documents Sales, PB83-233148)
10. Approaches to Developing Questionnaires (NTIS Document Sales, PB84-105055/AS)
11. A Review of Industry Coding Systems (NTIS Document Sales, PB84-135276)
12. The Role of Telephone Data Collection in Federal Statistics (NTIS Document Sales, PB85-105971)
13. Federal Longitudinal Surveys (NTIS Documents Sales, PB86-139730)
14. Workshop on Statistical Uses of Microcomputers in Federal Agencies (NTIS Document Sales, PB87-166393)
15. Quality on Establishment Surveys (NTIS Document Sales, PB88-232921)
16. A Comparative Study of Reporting Units in Selected Employer Data Systems (NTIS Document Sales, PB90-205238)
17. Survey Coverage (NTIS Document Sales, PB90-205246)
18. Data Editing in Federal Statistical Agencies (NTIS Document Sales, PB90-205253)
19. Computer Assisted Survey Information Collection (NTIS Document Sales, PB90-205261)
20. Seminar on the Quality of Federal Data (NTIS Document Sales, PB91-142414)
21. Indirect Estimators in Federal Programs (NTIS Document Sales, PB93-209294)
22. Report on Statistical Disclosure Limitation Methodology (NTIS Document Sales, PB94-165305)
23. Seminar on New Directions in Statistical Methodology (NTIS Document Sales, PB95-182978)

Copies of these working papers may be ordered from NTIS Document Sales, 5285 Port Royal Road, Springfield, VA 22161 (703)487-4650