REVISION OF THE CURRENT EMPLOYMENT STATISTICS SURVEY

Thomas J. Plewes
U.S. Bureau of Labor Statistics

Over the next several years, the Bureau of Labor Statistics will mount a fairly intensive review and revision of the survey of employment, hours and earnings of wage and salary workers in nonagricultural industries. It is a project whose time has come. It will be no simple undertaking.

In sheer magnitude and impact, the survey has few rivals. Known alternatively as the Current Employment Statistics Survey (after the formal title of the program) or the BLS-790 Survey (after the form number), the establishment survey of employment hours and earnings is the largest monthly sampling operation in the field of economic statistics.

Each month, the Bureau of Labor Statistics and the cooperating State agencies collect, tabulate, and publish data from a sample of establishments representing all nonagricultural industries and government. The current sample includes approximately 165,000 reporting units every month. Over 2,600 separate series are available at the national level, and another 8,700 series covering 3,400 industries are published for the 50 States, the District of Columbia and 220 areas. The data include series on total employment, production or nonsupervisory worker employment, women employed, average hourly earnings, average weekly hours, and average weekly overtime hours (in manufacturing). For many series, seasonally adjusted data also are published.

Description of the survey: Employment, hours, and earnings are measured for the pay period including the 12th of the month, which is standard for all Federal agencies collecting employment data on an establishment basis. They are among the most timely of economic statistics with initial monthly estimates for the Nation, published generally on the first Friday of the month following the reference period--on a target schedule which puts data in the hands of users just 2-3 weeks after the reference period in a press release, which contains preliminary national estimates of nonagricultural employment, weekly hours, and gross average weekly and hourly earnings in the preceding month, for major industry categories. The release also includes seasonally adjusted data on employment, average weekly hours, and average overtime hours. The preliminary estimates are based on tabulations of data for less than the full sample to permit early release of figures. The press release also includes a brief analysis of current trends in employment, hours and earnings, pointing up current developments as compared with those for the previous month and the same month in the preceding years.

The primary collection of the current sample data is conducted by State agencies which have cooperative agreements with the Department of Labor. The agencies implement the sample design, drawn from a universe of establishments which is stratified first by industry, and within each industry, by size of establishment in terms of employment. The primary sample "frame," or universe, is the listing of establishments covered under State unemployment insurance laws (now numbering over 4,000,000 establishments), supplemented by locally developed employer name and address files for industries and size classes not covered under State unemployment insurance laws.

Sample-derived estimates, by nature, differ a certain amount from the data that would be derived from a complete census or universe count of all establishments. To remove the effects of these small sampling errors from the estimates, as well as to reflect changes in the industrial classification of firms and the formation of new establishments, the Bureau annually "benchmarks" the survey data against the latest available census of universe information. The source of the census--or benchmark review--is the periodic tabulation of employment data by industry and size of establishment compiled by State agencies from reports of establishments covered by State unemployment insurance laws (the ES-202 program). In the course of the annual benchmark adjustment, the sample-based total employment estimates are revised by tapering the differences between the census and the sample. Coincident with this adjustment, estimates of labor turnover, employment of women, hours and earnings, and the seasonally adjusted series are revised, because they are derived from, or weighted by the total employment estimates. The employment estimates, on a monthly basis, are prepared for over 800 estimating cells, or groups of establishments in an industry defined by a 3- or 4-digit Standard Industrial Classification (SIC) code. The estimates for these estimating cells are then aggregated to provide estimates for higher-order industrial groupings. Three basic steps are followed in the computation of estimates for each cell:

(1) A total employment figure (benchmark) for the estimating cell, as of March of each year, is obtained in the annual benchmark process described above.
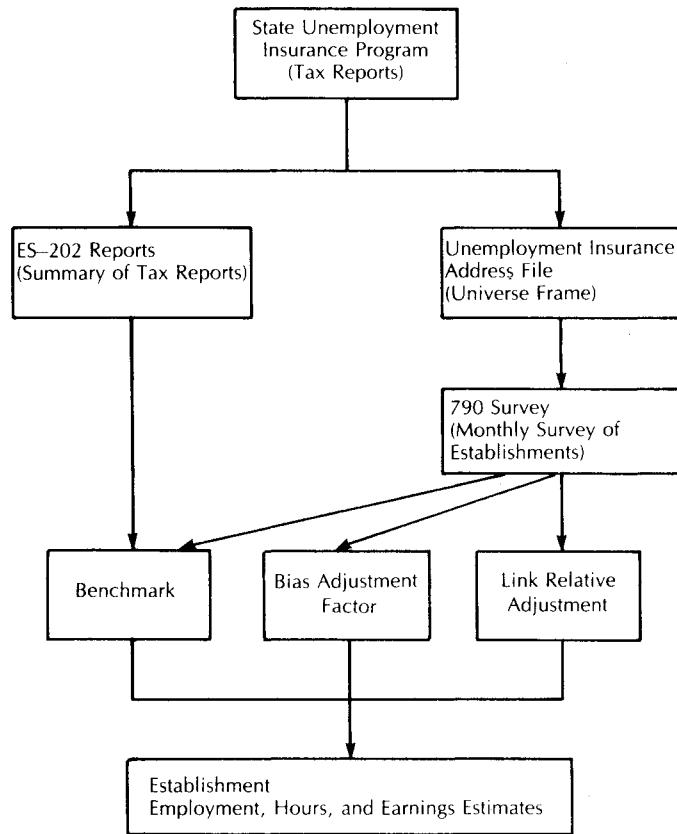
(2) For each cell, the ratio of employment in one month to that in the preceding month (i.e. a link relative) is computed for sample establishments reporting in both months.

(3) Beginning with the benchmark month, the estimates for each month are obtained by multiplying the estimate for the previous month by the link relative for the current month.

During the year, the sample-based estimates are adjusted on a monthly basis to correct for historical downward bias in the employment estimates caused by the late introduction of new firms into the sample by use of a bias adjustment factor, computed coincidently with the annual benchmark review. 1/

The estimation procedure is depicted in the following flow chart, which is extracted from the report of the National Commission on Employment and Unemployment Statistics, Counting the Labor Force, 1979.

```
┌─────────────────────┐
│  State Unemployment  │
│   Insurance Program  │
│     (Tax Reports)    │
└─────────────────────┘
```

```
┌─────────────────────┐              ┌─────────────────────────┐
│    ES–202 Reports    │              │  Unemployment Insurance  │
│ (Summary of Tax Reports)│           │       Address File       │
│                      │              │     (Universe Frame)     │
└─────────────────────┘              └─────────────────────────┘
```

```
                                      ┌─────────────────────┐
                                      │     790 Survey       │
                                      │   (Monthly Survey of │
                                      │    Establishments)   │
                                      └─────────────────────┘
```

```
┌─────────────┐      ┌─────────────────┐      ┌─────────────────┐
│  Benchmark  │      │ Bias Adjustment  │      │  Link Relative  │
│             │      │      Factor       │      │   Adjustment    │
└─────────────┘      └─────────────────┘      └─────────────────┘
```

```
┌─────────────────────────────────────────────────┐
│  Establishment                                    │
│  Employment, Hours, and Earnings Estimates        │
└─────────────────────────────────────────────────┘
```

Program Scope: The employment, hours and earnings data system provide vital data for the appraisal of current labor market trends and the assessment of overall economic conditions. At the national level, the data are used as major components of national product and income estimates, as leading economic indicators, and as a basis for current appraisal of labor market trends. The earnings data are used extensively in escalator clauses of procurement contracts at both the national and subnational levels. At the State and local level, the data are a primary basis for the assessment of current economic conditions, and in analysis of industrial structure and trends. The data are key elements in the computation of State and local unemployment rates.

The program of Federal/State cooperative relationships that has produced these data is one of comparatively long history. The first monthly studies of employment and payrolls began in 1915. The program has evolved to its current scope and coverage over the years, and has existed in its present configuration since 1949, when all States joined with the Bureau of Labor Statistics into a fully integrated program which provides comparable employment, hour and earnings information on a national, State, and area basis in considerable

industrial detail with monthly frequency. Currently, cooperative arrangements are in effect with employment security agencies in 47 States and the District of Columbia, and with 3 State labor departments.

The nature of the program, as well as the basic administrative configuration, has been relatively unchanged for the past 30 years. Although innovations in sampling procedures, expansions in geographic and industry coverage, and enhancements in data processing capacity and techniques have been accomplished, the program has undergone no major revision through its recent history.

Need For A Review: Why now, at this juncture, is the Bureau of Labor Statistics initiating a major review of this program, which would lead, if warranted, to a program revision? Or, to put this central question another way, what evidence is there that the program is "broke"--following the dictum: "If it ain't broke, don't fix it."

Over the years, the program has been the subject of a series of external and internal evaluations. Each of the reviews has identified program shortfalls, and some have recommended revisions in the program. We are confronted with an accumulation of evidence that the data have fairly significant shortfalls.

In 1962, the President's Committee to Appraise Employment and Unemployment Statistics, known as the Gordon Committee, recommended the development of probability sampling, improvements in methods of estimation, measures to expand coverage and speed up the reporting, as well as improvements in benchmarking and quality control. 2/ Over the 18 year period since this comprehensive study was completed, a number of the recommendations, especially those pertaining to coverage, timeliness and techniques for benchmarking and estimation have been implemented. Still, the sample was never put on a true probability basis, nor were the recommended quality control procedures put in place. The result is that the Bureau has serious questions about the adequacy of the data, does not have a capacity for identification of error (in the usual sense that those familiar with probability sampling techniques would feel comfortable) on an ongoing basis. On the bottom line, the survey is still not what might be termed a truly "statistical" survey, with fairly sophisticated procedures for computation of the validity of the estimates that are available, for example, in the Current Population Survey.

The need for an integrated, ongoing review capacity for the survey was again addressed in the early 1970's by a task force, composed of representatives of the Bureau, the Employment and Training Administration, cooperating State agencies, OMB, and other parts of the Department of Labor. The task force examined program content, data sources, benchmark estimates, sample selection, schedule design and data processing and estimation. In addition to emphasizing the need for integration of systems of statistics covering the Nation, States and local areas and the desirability of Federal/State cooperative programs, the task force recommended continuing reviews of: (a) the benchmark adjustments and their frequency, (b) sample adjustment procedures, (c) a regular response analysis program, (d) upgrading the SIC coding, and (e) the greater use of personal visits and/or telephone contacts with respondents. 3/

More recently, an intensive internal review of procedures in the cooperating State agencies in 1978, conducted by BLS staff, provided a snapshot of State program operations. The final report of this review indicated the need for improving and standardizing the State agencies' approach to sample design and solicitation, benchmarking, estimating, publications, and automation. The report indicated the need for additional research, continuing review and validation, training, and the upgrading of automation capacity.

Data users have also weighed in with their reviews and have suggested program and conceptual improvements. One such group, the Advisory Committee on Gross National Product Data Improvement (1977), also known as the Creamer Committee, recommended improvements in the employment, hours and earnings data that are used as input into the gross national product accounts. 4/ The Creamer Committee recommended expanding the scope of the survey to obtain hours and earnings data for supervisory workers, and "a broadly-based research and development program to strengthen the sampling, estimating, and reporting aspects of the monthly payroll survey."

Among the specific areas for strengthening were: (a) the response rates for the sampled firms and the procedures for introducing new firms into the sample; (b) collection of total wage payments for the calendar month; (c) the need for a periodic drawing of a complete new sample and implementing a full probability sample; (d) ways of improving processing of the data; (e) the State and local government component of the estimates; and (f) seasonal adjustment procedures.

Another user-oriented group—the Panel to Review Productivity Statistics of the National Academy of Sciences—has recommended the collection of hours worked data to improve estimates of industrial productivity. 5/

The most recent comprehensive review of the establishment survey of employment, hours and earnings was conducted by the National Commission on Employment and Unemployment Statistics (the Levitan Commission). The Commission's final report, issued in September 1979, assessed the adequacy of the 790 program and made recommendations to upgrade the scope and coverage, as well as methodology of the survey. 6/ The Commission recommended: (a) increasing the sample size to provide better estimates for underrepresented industries; (b) expanding the geographical coverage to permit aggregate employment estimates for all SMSAs and balances of States; (c) research into improving sample design; (d) better current documentation of the survey; (e) initiation of a quality control program; (f) strengthening the hours and earnings estimates; and (g) testing the feasibility and cost-effectiveness of data for nonproduction workers, full- and part-time work, and hours worked. The Commission held that the BLS should devote substantial resources to the enhancement of the program, and stressed the urgent need to upgrade the design and implementation of the program, and to document what is being done. The Commission recommended that "efforts in this direction begin at once."

Approach to the Review and Revision: The accumulated evidence that the establishment survey data have shortcomings can be considered in light of the primary function of the survey. The limitations and opportunities for improvement may be discussed in terms of the statistical methodology functions—sample design, data collection, and estimation—and in terms of the program-related support functions—computer systems, procedures, training and publications and analysis of the data. The remainder of the paper will concentrate on the Bureau's approach to developing a program of research, testing and development leading to improvements in the statistical methodology underpinning the survey. The principle statistical methodology issues relate to the sample design and its implementation, and the estimation procedures.

1. Sample Design. The fundamental issue that must be addressed in considering the design of the 790 survey sample is, quite simply, that there is some difficulty in defining the statistical character of the present sample design. The objectives of the sample design are fairly straightforward; that is, to (a) provide for the preparation of reliable monthly estimates of employment, hours of work, and weekly and hourly earnings which can be published promptly and regularly; (b) through a single general system,

yield considerable industry detail for metropolitan areas, States, and the Nation: (c) be appropriate for the existing framework of operating procedures, administrative practices, resource availability, and other institutional characteristics of the program; and (d) because this is a Federal/State cooperative program, to provide a technical framework which meets the objectives of the national program and within which State and area sample designs can be determined. In practice, this latter objective is usually attained as a fall-out of the national design. Since estimates for States and areas are generally not prepared at the same level of detail as the national estimates, the national design usually provides a sufficient basis for the subnational estimates, with some minor supplementation.

The current sample design in large measure meets these objectives, but how well if fulfills the higher order statistical objective of producing reliable estimates has never been adequately assessed. That is because the current sample design is more empirical than statistical.

In developing the sample design, the universe of establishments is first stratified by industry and within each industry by size of establishment in terms of employment, using six standard size classes. Within each industry, an optimum allocation design is obtained by sampling with probability proportionate to the average size of establishment within each of the strata. Within each strata, the sample members are selected at random.

In effect, we have not one sample, but a combination of samples ranging from a fairly high degree of coverage in some industries to less satisfactory coverage in other industries. Within the industry samples, large establishments tend to fall into sample with certainty—most establishments with 250 or more employees, for example, are included with certainty.

While the selection of sample units is guided by common sense rules, aimed at good representation of the various industries and types of establishments, it is not made under the same well-defined rules that underpin the household survey, and that would permit the calculation of the amount of variation to be expected as a result of the sampling process. Over the years, the procedures have been criticized for not being grounded in classical inference theory. Some have suggested implementation of a full-scale probability design, since the present design does have many of the attributes of a probability design.

Is the establishment survey of employment, hours and earnings susceptible to a straight probability sampling structure? The character of the survey suggests that it would be difficult to design an optimal probability design. The survey is voluntary, and will probably remain so, and is conducted in an exceedingly tight time frame. We perceive that the problem of non-response is very real and will likely stay with us, even if our efforts reduce it below current levels. A standard probability design does not handle nonresponse very well, though some encouraging work in the area of adjusting for nonresponse is now underway. 7/ The Levitan Commission considered endorsing a random probability design, but stopped short of making that

recommendation, finding that:

(a) A radical change in the design would be expensive. Many firms have a structure already set up within the firm to complete the forms efficiently on a regular basis. It would be foolish, as well as costly, to reject reports from good reporters in favor of attempting to bring on establishments selected in a probability design scheme.

(b) Many reports not needed for the national estimates are needed for estimates at the State and local level. This gives a useful purpose to oversampling, which has no convenient place in probability sampling schemes.

(c) Although probability sampling would provide ongoing estimates of error, whether the estimates would be improved is open to question. The benchmark revisions, for most industries, are within acceptable ranges, suggesting that even in the absence of standard errors and in the light of potential biases, the results are somewhat reliable. 8/

Still, the current design is, from a statistical point of view, unacceptable. In the review and revision process, alternatives that build on the present strengths of the survey without radically changing its character should be explored. Once such promising option, well worth exploring, would be to take advantage of recent model-based developments in finite population sampling theory. 9/ Early in the redesign effort, the Bureau will investigate possible advantages of incorporating prediction theory as well as probability techniques to produce an improved strategy for redesign.

2. Estimation. The sample selection and estimation processes are inexorably joined in the survey design. Any changes in the design of the sample must filter through into changes in the estimation process. Thus, if for example, a model-dependent sample design is selected, a model-dependent estimation procedure should follow. These issues will be jointly considered in our revision work.

The need for modernizing the current estimating procedure is widely recognized. Currently, estimates consist of an annual series of measures provided by the benchmark data source (the ES-202 counts) and an ongoing system of estimates derived from payroll data provided by employers on a month-by-month basis. The monthly employer data, obtained on the shuttle schedules, permit extrapolation of the series from the most recent benchmark datum. The continuity of employer reporting on the shuttle schedule form permits derivation of estimates for the more than 800 estimating cells by use of the link relative technique, which was described earlier.

This fairly neat and understandable procedure is complicated, however, by the timing aspect of the survey. The requirement to publish initial estimates some three weeks after the reference period has evolved, over the years, into a pratice of computing three current estimates, or closings. The first estimate is the one on which we must concentrate our attention, since these data are released in the Bureau's monthly employment situation press release, are the most closely watched, and form the basis for the Commissioner's analysis of current trends in her monthly testimony before the Joint Economic

Committee. The public perception of employment trends is largely governed by first closing data, yet, historically, the first closing data are the weakest because of delays in receipt of employer reports. In recent times, despite heroic efforts on the part of the Bureau and the cooperating State agencies, the initial estimates have been based on returns from 30-45 percent of the employer respondents. The second and third closings follow the first closing at one month intervals, and have yielded 85 percent and 98 percent of sample returns on average. By basing the preliminary estimates on somewhat less than the full sample, an additional source of variation is introduced.

Just as is the case with design of the sample, there are certain characteristics of the current estimating procedure which imply that it would be fruitful to explore a model-based approach. The current estimation procedure is based on the simple assumption that those who do not respond have the same characteristics as those who do respond, at the estimating cell level. This is, in essence, a crude model. It seems obvious that a first step toward developing a statistically-grounded survey would be to test the validity of this model for employment on the universe file. This line of investigation could well provide a methodology for the computation of valid variance estimates.

Plans for the Revision: A revision effort offers the opportunity for considering the very basis of a survey operation. This will be the initial emphasis of the project, for which we have done considerable planning and upon which, depending on the availability of resources, we intend to embark.

The Bureau starts with an open slate. That is, we make only two assumptions about the future program:

(a) The survey will continue to be conducted as a Federal-State cooperative program.

(b) The survey will continue to be a monthly survey of establishments.

The Bureau will be reviewing all aspects of the program. In addition to the basic research on sample design and estimation, discussed above, we plan to engage in a number of review and modernization activities, including:

(a) Conducting an in-depth review of employer records. The information that employers maintain on their payroll records largely delimits the data types that can be collected in an establishment survey. The last systematic survey of employer records was conducted in 1956, upon which our current perceptions of the type of data that could be collected with ease and reliability was based.

(b) Conducting a survey of employer perceptions of the survey program. It is a constant source of wonder to me why over 160,000 establishments each month complete a voluntary survey form and return it to the government. Despite talk of a growing anti-government environment and dislike with the response burden, the real strength of the survey is the faithful cadre of respondents, especially in the manufacturing sector of the economy. But if we are to improve coverage, the survey must be extended in the smaller size classes and, likewise, in the construction, trade and services industries. In these size classes and industries, we experience our largest refusal rates.

(c) Improving solicitation methods. Recent research by cooperating State agencies has pointed the way to improving employer response by better, if more costly, solicitation methods. The Bureau's experience in a series of pilot surveys of job openings indicates that response rates of 80-95 percent of establishments selected in a probability design can be attained with sufficient solicitation efforts. The challenge here is to balance improved response rates with increased costs of solicitation, and we are just now on the threshold of coming to an understanding of this relationship.

(d) Improving the sample frame. With expanded coverage of the unemployment insurance program, the ES-202 more closely approximates a true universe of establishments. There are significant problems, however, with identification of multi-establishment firms and adequate industrial classification of establishments covered by the U.I. laws.

(e) Assessing user needs. The Bureau plans to conduct an intensive survey of user needs for employment, hours and earnings information. This will help guide future program developments.

These and a myriad of other activities (some 500 sub-projects in all that have been identified in our plans) will consume our attention and skills over the next several years. We have, and will continue to benefit from the work that has been accomplished and reported on today by our colleagues in the U.S. Bureau of the Census and in Statistics Canada, and we fully expect that the work we are about to undertake will be of benefit in the design and conduct of other establishment-based surveys.

1/ Madow, Lillian H., "An Error Profile: Employment as Measured by the Current Employment Statistics Program," Proceedings of the Social Statistics Section, ASA, 1977, p.33-44; Bureau of Labor Statistics, Handbook of Methods, Chapter 3, GPO: Washington, D.C., 1978

2/ President's Committee to Appraise Employment and Unemployment Statistics, Measuring Employment and Unemployment, GPO: Washington, D.C., 1962

3/ Bureau of Labor Statistics, Report of Task Force on Establishment Employment and Related Statistics, February 1971, Unpublished

4/ U.S. Department of Commerce, Office of Federal Statistical Policy and Standards, Advisory Committee on Gross National Product Data Improvement, Gross National Product Data Improvement Project Report, GPO: Washington, D.C., 1977

5/ National Research Council, Measurement and Interpretation of Productivity, National Academy of Sciences: Washington, D.C., 1979

6/ National Commission on Employment and Unemployment Statistics, Counting the Labor Force, GPO: Washington, D.C., 1979

7/ National Research Council, Symposium on Incomplete Data: Preliminary Proceedings, Social Security Administration: Washington, D.C., 1979

8/ National Commission on Employment and Unemployment Statistics, op. cit., p. 158

9/ Hansen, Morris H., William G. Madow and Benjamin J. Tepping, "On Inference and Estimation from Sample Surveys," Proceedings of the Survey Methods Section, American Statistical Association: Washington, D.C., 1978; and Royall, Richard M. and William G. Cumberland, "An Emperical Study of the Ratio Estimator and Estimations of its Variance," Journal of the American Statistical Association, forthcoming.