PROBLEMS OF ANALYSIS OF URBAN EMPLOYMENT SURVEY DATA

Howard V. Stambler, U.S. Department of Labor

It would not be far from the truth to say that all new or one-time surveys create unique problems of analysis and interpretation. But in the case of the Urban Employment Survey, which is the focus of our discussion, there may well be a larger than usual number of difficulties. These analytical problems exist even though much of the information being collected in the survey conforms closely to the traditional and familiar Current Population Survey labor force concepts and questions which have gained nearly universal acceptance.

My purpose today is to describe to you some of the major difficulties the BLS is encountering in our analysis of the results and to discuss with you some of our attempts at their resolution. Although only a limited amount of the survey data has been published as yet, there has been an early recognition of the problems that the complete analysis will hold. And, as the BLS begins its examination of the first year's results later this month, these considerations will be well etched in our minds.

To provide a framework for my remarks, let me briefly describe again the purpose and scope of the survey, and how even these create problems in analysis. First of all, the Urban Employment Survey, or UES for short, was designed to examine the employment problems and barriers to meaningful employment of persons in Concentrated Employment Program (CEP) areas--target areas in which the Department of Labor has combined under one administrative structure all its manpower programs in order to concentrate their impact. Its purpose is to provide information for use in understanding the barriers to employment in these and, hopefully, other poverty neighborhoods, so that programs best suited to meet the needs of their residents can be developed. As a result, and this gives rise to our first analytical problem, the UES is less a measurement of the economic and social characteristics of poor people than it is of the situation of all people (only some of whom are poor) who live in specific geographic areas recognized through 1960 Census data as being poverty neighborhoods. And, as we might have expected, these are diverse neighborhoods which include a significant number of persons and families not living in poverty or who do not have serious employment problems.

What this heterogeneity means, is that analyses of the target areas really must be made in three stages or levels. The first level encompasses an investigation of the overall employment situation in the particular poverty area. The next step is the analysis of differences in the situation for the poor and the nonpoor in these areas. Finally, we need to develop as much insight as possible into the particular problems of poor people in these areas.

It is important to note also that these neighborhoods are not necessarily the worst or even the only bad areas in the UES cities. The CEP areas were chosen by the Department of Labor based in large part on the extent of unemployment and poverty in the areas as shown by the 1960 Census. Thus, only part of the total slum area in each city is being surveyed and current employment conditions in these areas may or may not be significantly different from those in other poor neighborhoods in the same cities. In one of the cities--New York--this phenomenon was particularly striking, and tracts outside the CEP area were added to the UES sample.

It is worth noting, too, as was touched on by Mr. Gerson, that the survey is presently being conducted also in the balances of 2 cities, Atlanta and Detroit. Information on these areas will permit comparisons of the employment situation of persons living in the target areas of the two cities with that of the general population outside the target areas. However, this in itself causes problems because just as the six target areas do not have a completely homogeneous population, the balance of the two cities included in the survey are also diverse areas which include other slums, as well as middle- and upper-class neighborhoods. These facts must certainly be kept in mind in making any comparisons between the two sample areas in Atlanta or in Detroit.

Another geographic consideration is the continual population movement into and out of these areas. Slums are increasingly being recognized as "staging areas" for many people who leave them soon after arrival. In addition, as a slum dweller improves his economic status, probably one of his first acts is to move out of the slum. This means that in each succeeding year of the survey we will be measuring changes in the situation in that area, but not necessarily of the people who lived there the previous year. Therefore, any changes in these neighborhoods over time will not necessarily reflect a measurement of the improvement or worsening of the situation for a particular group of people, but rather of the area itself. It will therefore be difficult to detect a change in the condition of particular people by observing changes in the area characteristics.

To compensate for this, we are attempting a longitudinal study in the second year of the survey. Beginning this summer, we are identifying and then following up a subsample of persons who were interviewed during the first year of the survey and have subsequently moved to a new address. This addition to the survey will enable us to find out what happens to both movers and non-movers over time.

In this initial stage of my remarks, I have tried to give you an idea of the geographic composition of the UES areas and of the problems

this has created. Let me now turn briefly to one aspect of the sample design that has major analytical implications, with particular reference to statistical reliability.

The weekly sample take in each of the eight survey areas (the six slum areas plus the balance of the cities of Atlanta and Detroit) is approximately 70 households. Thus, over the full year of the survey, about 3,500 households will be interviewed in each area. Since the population size in each CEP area differs considerably, the sampling ratio and thus the estimates of reliability vary somewhat from area to area. For example, in the Atlanta and Los Angeles target areas, approximately one out of every eight households are scheduled for interview each year. At the other extreme, an annual sample of about 1 in 110 persons will be interviewed in the balance of the city of Detroit. Since the same amount of detailed information will be tabulated for each of the areas, analysts and researchers will have to use extreme caution in the interpretation of comparative findings for these areas. Findings which may be statistically significant in one area may be washed out in another area due to the wider range of sampling error. The estimates of reliability are quite different in the UES than in the Current Population Survey, the source of most of our comparative data.

There is a similar problem which arises in relation to the comparison of UES data with information from the CPS--this is the matter of rotation group bias in the CPS.

In the CPS, the monthly estimates are essentially compilations of information obtained from respondents in eight rotation groups. Persons interviewed thus may be in the CPS sample anywhere from the first to the eighth time, since each household is in the sample for 4 consecutive months, drops out for 8 months, and returns to the sample for 4 more months. Long experience with the CPS has shown, however, that the respondents in the first and fifth months of the rotation pattern are more likely to be classified as unemployed than they are in the other months of the survey. By way of contrast, the UES is a weekly survey in which each person is interviewed only once, and the data are accumulated over a given period to provide the necessary reliability. Thus, it is conceivable that all the data from the UES could very well show the same kind of biases as do the first and fifth month households in the CPS.

There are, however, other differences between these 2 surveys which may mitigate such a situation. The length and detail of the UES questionnaire may have some effect on this phenomenon, as could the type of respondent being interviewed in these target areas. Another difference is that the CPS questions for all members of the household are usually answered by one respondent, usually the housewife, whereas in the UES most of the questions are answered by the individual himself (more about this later). In any case, the possible existence of rotation group bias has important implications for analysts, researchers and users of the data who may be making UES-CPS comparisons.

It is worth noting that the UES sampling design will change somewhat during the second year of the survey. Approximately one-half of the UES respondents originally interviewed in one year will be in the sample a second time the following year. The other half of the sample will be comprised of new sample households that will also be interviewed in two consecutive years. Thus, the second year's sample will be comprised of both first and second time respondents with a resulting different rotation group bias effect. This has implications too for UES year-to-year comparisons.

Another source of possible difficulty in the analysis of UES data is the comprehensive and detailed race and ethnic origin information which is being tabulated. Where the sample is adequate and in those cities in which the designation is warranted, data on UES residents will be provided separately for the following groupings--white; Spanish American, both white and nonwhite separately; Mexican American (in Los Angeles and Houston); Puerto Rican (in New York City); and American Negroes. Since the race/ethnic origin composition of the survey areas is quite diverse, it is necessary to be extremely careful in drawing overall conclusions about the data for an entire survey area. Of course, one reason that the particular areas were selected was to enable researchers, policy planners and program administrators to determine to what degree the problems of slum dwellers vary according to their color or national origin.

Although the expanded race-ethnic break will be of great value in examining the situation of different groups in a particular area, it is important to realize that overall data for these particular areas will not be strictly comparable with other area data. For example, a race/ethnic comparison of the situation for whites in the survey area and in the country as a whole would be misleading, since persons of Spanish American birth or parentage make up a very small proportion of the Nation's white population, but about half of the white population in some of the UES areas. Thus, for the analyst and researcher, disaggregation of the data as much as possible will be essential.

Still another area of difficulty which we expect to encounter is in the interpretation and analysis of questions which delve into new and/or sensitive areas. The UES will be attempting to determine if the traditional concepts of "work" and "employment" have the same meaning for poverty residents as they do for most people. It will also attempt to obtain information on illegal or sporadic activity that brings in money but is not generally thought of as work. To do this, the UES will not only use the traditional CPS labor force questions but will also probe further about remunerative activities. For this, we are asking the following questions of all men:

"During the past 12 months did you engage in any kind of activity for which you received money but which you would not normally consider work?"; if "yes", "what was this activity?"; and "Do you have any income from sources other than those previously mentioned, so that you don't need to work?", followed by "what are the other sources of income?", "How long have you been receiving this income?", "How long do you expect it to continue?", "Is it enough for your needs?", and "Do you expect to go to work at a regular job?". There is, of course, the possibility that the questions will not be properly understood, but even when they are understood, what degree of reluctance and evasiveness will we encounter from respondents? What degree of credibility can we attach to the responses? Despite our wariness about these questions, we feel that this probing may provide some valuable insights into those areas.

Other, more traditional types of questions which may be viewed as "sensitive" ones are those on earnings and income. The UES is collecting information on weekly earnings for those employed in the week prior to the survey and on total personal and family income over the past year. Based on past survey experience, we know that there is often a reluctance on the part of respondents to divulge sources and amount of income. In addition, difficulties in recollection over long periods cast doubt on the accuracy of income data for an entire year. But even more important, how will these particular target groups respond to these questions? Will poor people be more reluctant or less reluctant to discuss levels and types of earnings and income, especially welfare payments or illicit income. These facts also have to be taken account of in our analysis.

There is a separate aspect of the survey which may pose some difficulty to analysts and researchers, since it represents quite a departure from the usual objective and factual socioeconomic data that the Bureau of Labor Statistics has traditionally been concerned with. The new survey delves into the very subjective area of attitudes about jobs and about life in general, motivations, aspirations, and opinions. In addition to being a departure from the usual work of the Bureau, further difficulty in analysis will be encountered due to the fact that there are little comparable data for persons not in poverty neighborhoods by which to gauge survey results.

Two different types of information will be available from these subjective questions. The first is on what might broadly be called "job satisfaction." A series of questions will probe the attitude of poverty residents toward their jobs, determining what they like or dislike about the job, their commitment to work, and whether the job plays a meaningful role in their lives. These will be correlated with many characteristics, such as occupation, education, and income to develop some insights into whether attitudes of poverty residents are significantly affected by these characteristics.

The second general area relates to discrimination as perceived by poverty residents. This too will be analyzed in relation to many of the job characteristics and attitudes of poverty residents. Although responsibility for this area of work will be in the hands of social scientists with the proper background and training, we nonetheless expect to encounter many problems in the proper analysis and interpretation of the new data.

I have already discussed problems of comparability in relation to some of the unique characteristics and concepts of the Urban Employment Survey. But the UES data will have other problems of comparability, even where it uses traditional labor force concepts and measures. The major difficulties in the comparison of seemingly similar measures, such as from the UES and CPS, arise mainly from the different time periods covered. As I indicated, the UES is an accumulated sample, and the information will be analyzed and published after a full year's collection. This means that annual data on the current employment status of UES residents will actually be an accumulation of the situation for 52 different weeks, as opposed to the CPS, in which an annual average is an arithmetic average of the 12 monthly observations covering the same week of each month (the week containing the 12th).

Similarly, other data for the annual periods covered in the UES will be quite different than in the CPS. Data on annual income and annual work experience in the CPS are collected in February and March each year, and always refer to the calendar year prior to the survey. In the UES, however, data for the previous year's work experience of income will, in essence, relate to a sliding reference period covering the 12 months immediately prior to the week in which the interview was held.

For income and other questions, this also has important recall implications. It is probably easier to recall income for a calendar year (particularly when the questions are asked around income tax time as in the CPS) than it is to recall income for a June-to-June or a September-to-September period. Similarly, there is some doubt as to whether the income response covering the past 52 weeks may not in reality cover merely the previous calendar or income tax year. This sliding reference period also affects comparability with other CPS data usually obtained from supplementary questions to the monthly survey (such as educational attainment information) which always relate to a single month. In the UES, they will be accumulations of weekly responses covering a particular period. Differences like these in periods covered in the two surveys also raise questions of both seasonality and differences in economic climate at varying periods of time, items which must be considered in the analysis of UES data.

There are still other problems and difficulties mentioned by the previous speaker, which I will note only in passing. These are the serious problems of underenumeration, the lack of

population controls for the slum areas, and non-response. But there is one last problem in the UES that I would like to mention in closing, one posed by the mode of response. As I indicated earlier, in the UES the household respondent answers only current employment status questions for other members of the household. However, the bulk of the questions, such as on last job held for persons not in the labor force and on work experience, are answered only by the individual himself. In addition, if the person indicates his current employment status to be something different than that reported by the household respondent, answers are changed to reflect the individual's responses. In the CPS, on the other hand, the housewife or whoever is at home at the time of the survey is the person who usually answers the questions on labor force status and work experience. All of these differences must be taken into account by users of the data.

Let me now conclude with one general comment which is undoubtedly needed to place these analytical problems in perspective. I certainly do not want to leave the impression that the data will be unusable or that the survey results will not be meaningful; quite the contrary. Despite the acknowledged limitations of the data, we feel that the information from this survey will be extremely valuable to researchers, program planners, policy makers and others. It can not only provide new insights into barriers to employment in poverty areas but can also help to uncover problems about which more should be known. Nevertheless, the interpretation and analysis of the UES findings will present major challenges for both the Bureau of Labor Statistics and other users; caution and restraint will have to be essential tools for all those who use the data.