

## APPENDIX B

# Note on Data Sources

### Current Population Survey

The Current Population Survey (CPS) is a monthly survey of about 60,000 U.S. households conducted by the Bureau of the Census. It is currently the official source of data on income, poverty and labor force characteristics including unemployment in the U.S.

Data for hourly median and average wages are calculated from each year's Current Population Survey Outgoing Rotation Group (CPS ORG) files. The sample of individuals utilized for each year's analysis includes those between 18 and 64 years of age who were employed in the public or private sector (excluding the self-employed) who have valid hourly wage or weekly earnings data. The analysis excludes individuals who earned less than \$.50 and more than \$100 per hour (in 1989 CPI-U adjusted dollars).

Although the CPS ORG datasets have the advantage of being large enough to generate reliable estimates for different subgroups within the population, there are some well-known problems. In addition to the fact that individuals underreport various components of income, income at the high end is top-coded. That is, for individuals above a certain level of income the actual amount of income is suppressed.

Reporting omissions in the high-income range should not affect median wage data reported here. The problem of top coding was handled by the Economic Policy Institute as explained below.

### Census 2000 Supplementary Survey

The Census 2000 Supplementary Survey (c2kss), is a new survey that the Census Bureau conducted for the first time in the year 2000 as an experiment. If funded, it will become an annual survey — American Community Survey (ACS) — starting in 2003. The survey will include 700,000 households.

The c2kss test has some shortcomings. The numbers are based on contacts with houses in just one-third of the 3,141 U.S. counties. And the nearly 3% of the population in the army and institutions was left out entirely this time. The omissions don't affect comparisons of household characteristics, but could skew other numbers. Leaving out the largely less-educated prison population, for example, could artificially boost the education results.

Methodological differences between the “long forms” of the decennial censuses of the past and the annual ACS would make comparisons of results from the ACS and from the long forms of the past difficult.

### Insured Employment Series (ES-202)

The employment and wage data reported by the New York State ES-202 program are commonly referred to as “insured” or “covered,” because they include those employees who are covered under the State Unemployment Insurance (UI) program. The UI program derives its data from quarterly reports submitted to State Employment Security Agencies by employers subject to State UI laws and from Federal agencies subject to the Unemployment Compensation for Federal Employees (UCFE) program.

ES-202 employment and wage data are available on a monthly, quarterly and annual basis. The reports represent a full universe count of employment, covering nonagricultural industries, along with partial information on agricultural industries and employees in private households. ES-202 wages reported include bonuses, some stock options, retroactive pay, tips and the cash value of meals, lodging, or other payments in kind.

The analysis on growing and declining industries in this report was done largely at the level of

the more detailed three-digit industrial classification code to achieve greater descriptive detail. In cases where three-digit industries were similar, the broader 2-digit category was used. Use of the three-digit level may distort growth in cases where reclassifications of employment in industries create the illusion of employment growth or decline. Reclassifications that this report corrected involved the following industries (preceded by the 3-digit SIC code): 451 Air transportation, 421 Trucking, 781 Motion picture production, and 822 Colleges and universities. In the Mohawk Valley, corrections were also made for 833 Job training and related services, 836 Residential care and 832 Individual and family services.

### **Current Employment Statistics**

The Current Employment Statistics (CES) Survey is a monthly survey of business establishments that provides estimates of employment, hours and earnings data by industry for the nation as a whole, all states and most major metropolitan areas. The survey is administered by state employment security departments in cooperation with the Bureau of Labor Statistics under a Federal-State cooperative program.

In New York State, a sample of 20,000 employers voluntarily submit monthly employment information. Data from individual firms are grouped by industry and used as a sample for estimating industry-wide statistics, and are benchmarked each year using data from the State UI program. The employment estimates are for jobs in an area, regardless of the place of residence of the workers holding those jobs.

The CES employment count differs from ES-202 employment in two ways. First, the CES survey excludes all persons employed in farming. Second, the CES survey counts some persons who are not covered by unemployment insurance, such as employees of railroads, and employees of churches and other nonprofit organizations. Like the ES-202, CES does not track the unincorporated self-employed, unpaid family workers, or domestic workers in private households.

### **Analysis by the Economic Policy Institute**

Using the CPS ORG files, Economic Policy Institute imputed hourly wages for individuals who did not report an hourly wage, but who reported weekly earnings, and for individuals whose income was top coded.

The hourly wage was calculated using weekly earnings divided by usual weekly hours. The hours of those who reported varying hours worked are estimated based on the usual hours worked of persons with similar characteristics. The imputation of wages for individuals who were top coded assumes that the distribution of wages is a Pareto distribution.

To make comparison over time possible, two adjustments were made. First, figures were translated into constant dollars using the consumer price index CPI-U. In addition, periodically the definitions of variables (e.g., education) and methods used (top coding) in the CPS ORG files were changed. EPI made adjustments to account for these changes as well.

Because individuals may report “round” instead of actual figures for their wages, the median wages that the Economic Policy Institute reports are “smoothed” medians.

For a more detailed discussion of EPI’s handling of the CPS ORG, see Appendix A in Mishel, Bernstein and Schmitt, *The State of Working America, 2000/2001*, pp. 407-410.

In order to adjust family income data in Tables 1.2 and 1.3 for changes in consumer prices, EPI uses the Consumer Price Index Research Series Using Current Methods (CPI-U-RS), a new series released in 1999 by the Bureau of Labor Statistics (BLS). The CPI-U-RS indexes national consumer price changes from 1978 to the present and incorporates most of the improvements in methodology made in the CPI-U series over that time span into the entire series. Among other improvements, the CPI-U-RS makes quality adjustments for the aging of housing units and for the prices of used cars, personal computers and televisions, and it employs a geometric mean formula to account for consumer substitution within CPI item categories. Currently, the Census

## APPENDIX B

Bureau uses the CPI-U-RS as the inflation adjuster for historical income statistics. The BLS considers the CPI-U the official consumer price index.

### **Analysis by the Fiscal Policy Institute**

The CPS ORG files that FPI used for its own analysis were supplied by EPI and include the adjustments that the EPI had made. Median wages reported by FPI differ slightly from those reported

by EPI because the medians reported by FPI are the actual, not the smoothed, medians.

Throughout this report, FPI used the CPI-U index, the official BLS national consumer price index, to adjust wage and income data for changes in consumer prices. Two tables in Chapter 1 generated by EPI use the CPI-U-RS series to deflate income and are the exception to this rule (see above).