Chapter 4.
Measurement of Unemployment in States and Local Areas

Unemployment estimates for States and local areas are key indicators of local economic conditions. These estimates, which are produced by State employment security agencies, are used by State and local governments for planning and budgetary purposes and as determinants of the need for local employment and training services and programs. Local area unemployment estimates are also used to determine the eligibility of an area for benefits in various Federal assistance programs.

Under the Federal-State cooperative program, the Bureau of Labor Statistics develops the concepts, definitions, and technical procedures which are used by State agencies for the preparation of labor force and unemployment estimates. Currently, monthly estimates of employment and unemployment are prepared in State agencies for some 6,700 geographic areas, which include all States, labor market areas (LMA's), counties, cities with a population of 25,000 or more, and all cities and towns in New England regardless of population.

IN THIS CHAPTER

Background

Unemployment estimates have been developed for subnational areas for approximately 50 years. The program began during World War II under the War Manpower Commission to identify areas where labor market imbalance had been created as a result of an inadequate labor supply, material shortages, and transportation difficulties. After the war, emphasis was placed on identifying areas of labor surplus, and the program of classifying areas in accordance with severity of unemployment was established.

In 1950, the Department of Labor's Bureau of Employment Security (now Employment and Training Administration) published a handbook, *Techniques for Estimating Unemployment,* so that comparable estimates of the unemployment rate could be produced for all States. This led, during the late 1950s, to the formulation of the "Handbook method," a series of computational steps designed to produce local employment and unemployment estimates. This method relied heavily on data derived from the unemployment insurance (UI) system. (See section on "Estimates for sub-State areas—the Handbook method.")

In 1972, the Bureau of Labor Statistics assumed technical responsibility for the program and began to refine the concepts and methods to be used by States to estimate labor force, employment, and unemployment. In 1973, a new system for developing labor force estimates was introduced which combined the Handbook method with the concepts, definitions, and estimation controls from the Current Population Survey (CPS), the Bureau of the Census survey sponsored by BLS and used to measure the labor force status of individuals. The CPS estimates are based on data from a sample of households. The sample is designed to provide reliable monthly unemployment estimates for the Nation and reliable annual average estimates for the 50 States and the District of Columbia. (See chapter 1.)

36

Since 1976, the CPS has been increased in size several times and the design has been modified to improve the quality of State labor force estimates. As a criterion for using the monthly CPS data directly for official publication of labor force estimates, BLS established a maximum expected coefficient of variation (CV) of 10 percent for unemployment assuming an unemployment rate of 6 percent. (The coefficient of variation of an estimate can be defined as the standard error of the estimate divided by the estimate itself.) Based on this criterion, monthly CPS data were used, beginning in 1978, for official statewide labor force estimates for 10 large States—California, Florida, Illinois, Massachusetts, Michigan, New Jersey, New York, Ohio, Pennsylvania, and Texas—and for 2 sub-State areas—the Los Angeles–Long Beach metropolitan area and New York City.

In 1985, a State-based design for the CPS was fully implemented to incorporate 1980 census information and to provide for improved reliability for each of the 50 States and the District of Columbia. North Carolina was added as another "direct use" State, and the CV requirement for monthly unemployment was reduced to 8 percent for these 11 large States. For each of the other 39 (non-direct use) States and the District of Columbia, the reliability requirement was established at an 8-percent CV for annual average unemployment, assuming a 6-percent unemployment rate.

Until 1989, official monthly estimates for the non-direct use States were based on the Handbook method adjusted to CPS controls using a 6-month moving-average ratio adjustment. Since 1989, estimates for these 39 States and the District of Columbia have been based on time series models developed by BLS and tested by State employment security agencies, using standardized procedures. Estimates for sub-State areas are based on the Handbook method, adjusted to State control totals.

In 1996, the number of households in the CPS sample was reduced in response to budget constraints, resulting in the elimination of direct use of the CPS for monthly estimation in the 11 large States, the Los Angeles-Long Beach metropolitan area, and New York City. Beginning with January 1996, labor force estimates for these sub-national areas have been based on the time series modeling approach used since 1989 in the other 39 States and the District of Columbia.

The State UI system provides a count of the insured unemployed, an integral input to State and area estimation. Over the years, major improvements have been made to the UI database. The UI database project, conducted in 1976-78, standardized all UI claims data used in State and area labor force estimates, so that these data would be more consistent with the conceptual underpinnings of unemployment used in the CPS. The result of this project was the regular development, from computer files, of data on UI claimants certifying to unemployment during the week including the 12th day of the month (the CPS reference week). These data are based on the claimants' State/county/city of residence and exclude those who had earnings from employment in the certification week.

Estimation Methodology

Estimates for States

Monthly labor force data for all States, the District of Columbia, the Los Angeles-Long Beach metropolitan area, New York City, and the balances of California and New York are based on the time series approach to sample survey data (Scott and Smith, 1974; Bell and Hillmer, 1990). The purpose of this approach is to reduce the high variability in monthly CPS estimates for these geographic areas due to small sample sizes. The actual monthly CPS sample estimates are represented in signal plus noise form as the sum of a stochastic true labor force series (signal) and error (noise) generated by sampling only a portion of the total population.

where:
$$y_t = \theta_t + e_t$$

y_t = CPS estimate
θ_t = true labor force value
e_t = sampling error

The signal is represented by a time series model that incorporates historical relationships in the monthly CPS estimates along with auxiliary data from the Unemployment Insurance and Current Employment Statistics (CES) programs. This time series model is combined with a noise model that reflects key characteristics of the sampling error (SE) to produce estimates of the true labor force values. This estimator is optimal under the model assumptions and has been shown to be design-consistent under general conditions by Bell and Hillmer (1990).

Two models—one for the employment-to-population ratio and one for the unemployment rate—are developed for each State using over 15 years of data. The signals for both models are based on a core model of the following form:

$$\theta_t = \beta_t X_t + T_t + S_t + I_t$$

where X_t is a single explanatory variable with coefficient β_t and T_t, S_t, and I_t are the trend, seasonal and irregular components. The variable used in the employment model is the statewide monthly estimate of workers on payrolls in non-farm industries from the CES program divided by the intercensal estimate of the State's population of working age. The unemployment model uses the ratio of the number of State workers claiming unemployment insurance benefits to the payroll employment estimate. The regression coefficient is allowed to change over time to adapt to changing relationships between the CPS and the explanatory variable.

The trend and seasonal components change smoothly over time to control for systematic variation in the CPS not accounted for by the explanatory variable. The irregular component accounts for transitory residual variation not captured by other components of the model.

The degree to which the regression coefficient and the time series components vary over time is determined empirically for each State. Occasionally, the trend is a constant, acting as a fixed intercept. In some cases, the seasonal component is estimated to have a fixed pattern from year-to-year. For most models, the irregular component is zero.

Occasionally, there are sudden changes, either temporary or permanent, in the CPS that are not predictable from past history. These effects, manifested as aberrant observations or outliers, are handled by intervention analysis techniques which introduce dummy variables into the model components. Shifts in level are incorporated into the trend component and transitory changes into the irregular component.

The second major component of the signal plus noise model deals with CPS standard errors. Because of this survey's complex design, the behavior of the observed sample estimates differ in important ways from the true values. Sampled households are rotated in and out of the CPS over a period of 16 months, such that 75 percent of the sample from month-to-month consists of the same households and 50 percent from year-to-year. (See chapter 1.) Also, redesigns and major fluctuations in the size of the labor force cause major changes in the variance of the standard errors. These two features of the CPS, an overlapping sample design and changes in reliability, induce strong positive autocorrelation and heteroscedasticity in the standard error. These characteristics can seriously contaminate estimates of the true labor force if the standard error is ignored in the estimation process. For this reason, it is important to specify a model of the standard error process and combine it with the model of the signal to estimate the unobserved components of the CPS. The standard error model is specified as follows:

$$e_t = \gamma_t e_t^*$$

with e_t^* reflecting the autocovariance structure, assumed to follow an ARMA process and γ_t representing a changing variance over time. The parameters of the ARMA model are derived from standard error autocorrelations developed independently of the time series model from design based information. The standard error variances (equivalent to the square of the standard error described in chapter 1) are estimated using the method of generalized variance functions (Zimmerman and Robison, 1996).

The unknown hyperparameters of the signal are estimated by maximum likelihood using the Kalman filter algorithm. Given these estimated hyperparameters, the Kalman filter is used to decompose the observed CPS into its signal and noise components. This algorithm efficiently updates the model estimates as new data become available each month. For the latest month, the Kalman filter calculates estimates based on all available data, but does not revise estimates for the previous months with the latest data. Previous estimates are updated by a Kalman Filter "smoother" which revises an estimate at time t using all available data before and following time t. Smoothing is performed at the end of each year.

Benchmarking. This process is a general statistical procedure used to adjust estimates to a control total. Each year, historical model estimates are benchmarked to the annual average CPS State estimates of employment and unemployment. The goal of benchmarking is twofold: (1) To insure that the annual average of the final benchmarked series equals the CPS annual average, and (2) to preserve the pattern of the model series as much as possible. In practice, these two goals are conflicting, and some changes to the pattern of the time series are made to meet the first goal. The Denton benchmarking method has been used since the introduction of model-based estimates in 1989. It avoids discontinuities between December and January estimates and, through a constrained quadratic minimization approach, minimizes the distortion to the original time-series estimates. The benchmarked series are seasonally adjusted with X-11 ARIMA.

Estimates for sub–State areas—the Handbook method

Until 1973, the Handbook method was the only means used to develop State and local area labor force and unemployment estimates. With the exception of the Los Angeles-Long Beach metropolitan area and New York City, it continues to be the method used for sub-State estimation. It is an effort to estimate unemployment for an area using available information, comparable to what would be produced by a random sample of households in the area, without the expense of a large labor force survey like the CPS. The Handbook presents a series of estimating "building blocks," where categories of unemployed workers are classified by their previous status. Two broad categories of unemployed persons are: (1) Those who were last employed in industries covered by State UI laws, and (2) those who either entered the labor force for the first time or reentered after a period of separation. Until 1996, estimates also were made for those who were last employed in industries not covered by State UI laws, including private household workers, the self-employed, unpaid family workers, and segments of agriculture, nonprofit industries, and State and local government. These industries account for a very small percentage of employment. In a few States, where the agriculture component is significant and where UI coverage of agricul-

tural employment does not go beyond the Federal minimum, an exception procedure was established to allow direct estimation of agricultural unemployment.

Unemployment. In the current month, the estimate of unemployment is an aggregate of the estimates for each of the two building–block categories. The "covered" category consists of: (1) Those who are currently collecting UI benefits, and (2) those who have exhausted their benefits. Only the insured unemployed are obtained directly from an actual count of current UI claimants for the reference week under the State UI, Federal, and Railroad programs. The estimates of persons who have exhausted their benefits are based upon the number actually counted in the current period, plus an estimate of those expected still to be unemployed from previous periods.

The "new entrants and reentrants into the labor force" category cannot be estimated directly from UI statistics because unemployment for these persons is not immediately preceded by the period of employment required to receive UI benefits. Instead, total entrants into the labor force are estimated on the basis of the national historical relationship of entrants to the experienced unemployed and the experienced labor force. The Handbook estimate of entrants into the labor force is a function of: (1) The particular month of the year, (2) the level of the experienced unemployed, (3) the level of the experienced labor force, and (4) the youth proportion of the working-age population. The estimate of total entrants for a given month is derived from the following equation:

$$ENT = A(X+E) + BX$$

where:

ENT = total entrant unemployment
E = total employment
X = total experienced unemployment
A, B = synthetic factors incorporating both seasonal variations and the assumed relationship between the proportion of youth in the working-age population and the historical relationship of entrants to either the experienced unemployed (B factor) or the experienced labor force (A factor).

Employment. The total employment estimate for a particular area may be based on data from several sources. The principal sources include the Federal-State CES survey for the area, a State-designed survey of establishment, or extrapolated ES-202 employment data. These are designed to produce estimates of the total number of employees on payrolls in nonfarm industries for the particular area.

The "place-of-work" employment estimates must be adjusted to a place-of-residence basis, as in the CPS. Estimated adjustment factors for several categories of employ-

ment have been developed on the basis of employment relationships which existed at the time of the most recent decennial census. These factors are appropriately weighted and combined into a single factor which is then applied to the place-of-work employment estimates for the current period to obtain adjusted wage and salary employment estimates. Synthetically developed estimates for employment not represented in the establishment series—agricultural workers, nonfarm self-employed and unpaid family workers, and private household workers—are derived by extrapolation from the decennial census. These components plus the wage and salary component represent total Handbook employment.

Sub-State adjustment for consistency and additivity. Each month, Handbook estimates are prepared for labor market areas that exhaust the each State geographically. To obtain an estimate for a given area, a "Handbook share" is computed for that area; this is defined as the ratio of the area's Handbook estimate to the sum of the Handbook estimates for all LMA's in the State. This ratio is then multiplied by the current statewide estimate for the State to produce the final adjusted LMA estimate:

$$U_a(t) = U_s(t) * UHB_a(t) / UHB_a(t)$$

where:
U = total unemployment
UHB = Handbook unemployment
a = area
s = State
t = time

A comparable adjustment is performed for employment.

Benchmark correction. At the end of each year, sub-State estimates are revised to incorporate any changes in the inputs, such as revisions in the place-of-work–based employment estimates or claims data and updated historical relationships. These revised area Handbook estimates are then readjusted to sum to the revised (benchmarked) State estimates of employment and unemployment.

Producing estimates for parts of LMA's

Current labor force estimates at the sub-LMA level are required by several Federal allocation programs. However, for areas this small, the data required to compute independent Handbook estimates are generally not available, nor was the Handbook procedure designed to produce sub-LMA estimates. Based on data availability, two alternative methods are used to disaggregate the LMA estimates to the sub-area level.

The population-claims method is the preferred technique for unemployment. If residence–based UI claims data are available for the sub-areas within the labor market area, the

ratio of the sub-area to the total number of claims within the LMA is used to disaggregate the estimate of experienced unemployed to the sub-area level. The estimate of unemployed entrants is allocated based on the latest available decennial census distribution of adult and teenage population groups. When the population-claims method is used for unemployment, employment is disaggregated using current population distributions prepared by the Bureau of the Census and weighted by each area's decennial census relative share of employment to population. This preferred combination of techniques is used to derive estimates for all counties in multi-county LMA's and for cities in some States.

If the necessary UI claims data are not available at the city level, the census–share method is used. This method uses each subarea's decennial census share of county employment and unemployment, respectively, in order to disaggregate the monthly estimates of employment and unemployment.

Uses and limitations

Estimates of unemployment and the unemployment rate are used by Federal agencies to determine the eligibility of an area for benefits in various Federal programs. These include the Job Training Partnership Act (JTPA), the Emergency Food and Shelter Program (EFSP), the Public Works Program (PWP), the Temporary Emergency Food Assistance Program (TEFAP), and Labor Surplus Area designation. Under most programs, unemployment data are used to determine the distribution of funds to be allocated to each eligible area. In the case of the Labor Surplus Area designations, the data are used in the determination of area eligibility for benefits.

The annual CPS estimates used to benchmark statewide labor force estimates are based on probability samples of households and are subject to both sampling and nonsampling errors. (See Limitations in chapter 1 for a discussion of errors in the CPS.)

Approximate error measures for annual average estimates of major labor force characteristics for State and sub-State areas are estimated and published in *Geographic Profile of Employment and Unemployment.* Standard error estimates are developed from generalized variance functions which use CPS data for the areas and independently developed parameters. As with all sampling errors produced for CPS State and area data, a number of assumptions are required;

as a result, the sampling errors calculated using generalized variance functions should be used with caution.

Estimates not directly derived from sample surveys are subject to errors resulting from the estimation processes used as well as the limitations of the data sources used. The error structure associated with these estimates is complex, and information on the magnitude of the overall errors is not available.

Data products. Data from the LAUS program are made available to users in a variety of ways. Labor force and unemployment data are published monthly for all States and selected metropolitan areas in a news release entitled *State and Metropolitan Area Employment and Unemployment* and in the Bureau's periodical *Employment and Earnings.* Estimates of labor force, employment, and unemployment for all States, metropolitan areas, labor market areas, counties, cities with a population of 25,000 or more, and other areas used in the administration of various Federal economic assistance programs are provided in *Unemployment in States and Local Areas*, which is available monthly in microfiche form by subscription from the U.S. Government Printing Office. The order number for a subscription is 029-001-81003-2.

Annual average employment status data are provided each year in a press release entitled *State and Regional Unemployment, Annual Averages,* which is typically issued in the spring. It presents data on the population, civilian labor force, employed, unemployed, and unemployment rate for regions, divisions, and States. Annual average information for States and selected metropolitan areas is also published each spring in *Employment and Earnings.*

The annual publication, *Geographic Profile of Employment and Unemployment*, provides annual average data for census regions and divisions, the 50 States and the District of Columbia, 50 large metropolitan areas, and 17 central cities. Data are provided on the employed and unemployed by selected demographic and economic characteristics.

Current and historical data from the LAUS program are also available on-line via LABSTAT, the Bureau's public database. (Access via anonymous FTP or Gopher is *stats.bls.gov*; via World Wide Web, *stats.bls.gov/blshome.html*) Additional information about the LAUS program, including frequently asked questions, contacts, and technical references are on-line via the LAUS homepage (*stats.bls.gov:80/lauhome.htm*).

Technical References

Bell, W.R. and Hillmer, S.C. "The Time Series Approach to Estimation for Repeated Surveys," *Survey Methodology,* 1990, pp. 195-215.

Denton, F.T. "Adjustment of Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization," *Journal of the American Statistical Association,* 1971, Volume 66, pp. 99-102.

Evans, T.D., Tiller, R.B., and Zimmerman, T.S., "Time series Models for State Labor Force Estimates," in *Proceedings of the Survey Research Methods Section,* American Statistical Association, 1993, pp. 358-363.

Scott, A.J.; Smith, T.M.F.; and Jones, R.G. "Analysis of Repeated Surveys Using Time Series Methods," *Journal of the American Statistical Association,* 1974, pp. 674-678.

Tiller, Richard B. "A Kalman Filter Approach to Labor Force Estimation Using Survey Data," *Proceedings of the Survey Research Methods Section,* American Statistical Association, 1989, pp. 16-25.

Tiller, Richard B. "An Application of Time Series Methods to Labor Force Estimation Using CPS Data," *Proceedings of the Survey Research Methods Section,* American Statistical Association, 1990.

Tiller, Richard B. "Time Series Modeling of Sample Survey Data from the U.S. Current Population Survey," *Journal of Official Statistics*, 1992, pp. 149-166.

U.S. Department of Labor, Bureau of Employment Security. *Handbook on Estimating Unemployment,* Employment Security Research Methods, Handbook Series (BLS Reprint No. R–185), 1960.

U.S. Department of Labor, Bureau of Labor Statistics. *Manual for Developing Local Area Unemployment Statistics,* July 1979.

U.S. Department of Labor, Bureau of Labor Statistics. *Questions and Answers on the New Method for Developing State Employment and Unemployment Estimates,* January 1989.

Zimmerman, T.S., Evans, T.D., and Tiller, R.B., "State Unemployment Rate Time Series," in *Proceedings of the Survey Research Methods Section,* American Statistical Association, 1994.