Chapter 2.
Employment, Hours, and Earnings from the Establishment Survey

The Bureau of Labor Statistics (BLS) conducts the Current Employment Statistics (CES) survey, collecting data each month on employment, hours, and earnings from a sample of nonagricultural establishments. The sample includes about 140,000 businesses and government agencies, which cover approximately 440,000 individual worksites drawn from a sampling frame of roughly 9.0 million Unemployment Insurance tax accounts. The active CES sample includes approximately one-third of all nonfarm payroll employees. From these data, BLS, along with State labor market information agency partners, prepares and publishes a large number of employment, hours, and earnings series in considerable industry and geographic detail. CES data are available at **http://www.bls.gov/ces/home.htm**.

IN THIS CHAPTER

Background

The first monthly studies of employment and payrolls by BLS began in 1915 and covered four manufacturing industries. With increasing interest in employment data during the Great Depression, BLS increased its output; and by 1933, employment, average hourly earnings, and average weekly hours were published for total manufacturing, 90 manufacturing industries, and 14 nonmanufacturing categories. Early estimates of hours and earnings were made for production and nonsupervisory employees, who represented about 80 percent of all employees in the private sector. In 2010, BLS published official hours and earnings for all private-sector employees for the first time.

Interest in employment statistics for States and areas also grew. Even before BLS entered the field, in 1915, three States—Massachusetts, New York, and New Jersey—were preparing employment statistics. In 1915, New York and Wisconsin entered into cooperative agreements with BLS, whereby sample data collected from employers by a State agency would be used jointly with BLS to prepare State and national series. By 1940, estimates of total nonfarm employment for all 48 States and the District of Columbia were available. Since 1949, the CES program has been a Federal-State program that provides employment, hours, and earnings information by industry on a national, State, and metropolitan area basis. By 1980, cooperative arrangements were in effect with all 50 States, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands. (National estimates exclude data for Puerto Rico and the U.S. Virgin Islands.)

Concepts

Establishment

An *establishment* is an economic unit, such as a factory, mine, store, or office that produces goods or services. It generally is at a single location and is engaged predominantly in one type of economic activity. Where a single location encompasses two or more distinct activities, these are treated as separate establishments, if separate payroll records are available, and the various activities are classified under different industry codes.

Employment

Employment data refer to persons on establishment payrolls who received pay for any part of the pay period that includes the 12th day of the month.

Data exclude proprietors, the unincorporated self-employed, unpaid volunteer or family workers, farm workers, and domestic workers. Salaried officers of corporations are included. Government employment covers only civilian employees; military personnel are excluded. Employees of the Central Intelligence Agency, the National Security Agency, the National Imagery and Mapping Agency, and the Defense Intelligence Agency also are excluded.

Persons on establishment payrolls who are on paid sick leave (for cases in which pay is received directly from the firm), on paid holiday, or on paid vacation, or who work during a part of the pay period—even though they are unemployed or on strike during the rest of the period—are counted as employed. Not counted as employed are persons who are on layoff, on leave without pay, or on strike for the entire period, or who were hired but have not yet reported during the period.

In addition to employment data for *all employees*, the total number of *women employees* is collected. In private industries, CES collects data on production and related employees in manufacturing and mining and logging, construction employees in construction, and nonsupervisory employees in private service-providing industries; collectively, all these workers are often referred to as production employees.

Production and related employees include working supervisors and all nonsupervisory employees (including group leaders and trainees) engaged in fabricating, processing, assembling, inspecting, receiving, storing, handling, packing, warehousing, shipping, trucking, hauling, maintenance, repair, janitorial, guard services, product development, auxiliary production for plant's own use (for example, power plant), recordkeeping, and other services closely associated with the above production operations.

Construction employees in the construction sector include: Working supervisors, qualified craft workers, mechanics, apprentices, helpers, laborers, and so forth, engaged in new work, alterations, demolition, repair, maintenance, and the like, whether working at the site of construction or in shops or yards at jobs (such as precutting and preassembling) ordinarily performed by members of the construction trades.

Nonsupervisory employees include those individuals in private, service-providing industries who are not above the working-supervisor level. This group includes individuals such as office and clerical workers, repairers, salespersons, operators, drivers, physicians, lawyers, accountants, nurses, social workers, research aides, teachers, drafters, photographers, beauticians, musicians, restaurant workers, custodial workers, attendants, line installers and repairers, laborers, janitors, guards, and other employees at similar occupational levels whose services are closely associated with those of the employees listed.

An *employment benchmark* is a complete count of employment used to adjust estimates derived from a sample. CES sample-based estimates are benchmarked annually. The basic source of benchmark data for the CES survey is the Quarterly Census of Employment and Wages (QCEW) program, which collects employment and wage data from States' unemployment insurance (UI) tax records. The QCEW represents a virtual census of employment in the United States, covering about 97 percent of all jobs on civilian payrolls. (The benchmark process is explained in detail in later sections of this chapter.)

Indexes of diffusion of employment change measure the dispersion of employment change in industries over a specified time span. The overall indexes are calculated from seasonally adjusted employment series for 4-digit NAICS industries and cover all nonfarm payroll employment in the private sector. Diffusion indexes are also calculated for manufacturing using employment in 4-digit NAICS industries.

To derive the indexes, each component industry is assigned a value of 0, 50, or 100 percent, depending on whether its employment showed a decrease, no change, or an increase, respectively, over the time span. The average value (mean) is then calculated, and this percent is the diffusion index number.

The reference point for diffusion analysis is 50 percent, the value indicating that the same number of component industries had increased as had decreased. Index numbers above 50 show that more industries had increasing employment and values below 50 indicate that more had decreasing employment. The margin between the percent that increased and the percent that decreased is equal to the difference between the index and its complement—that is, 100 minus the index. For example, an index of 65 percent means that 30 percent more industries had increasing employment than had decreasing employment [65-(100-65) = 30]. However, for dispersion analysis, the distance of the index number from the 50-percent reference point is the most significant observation.

Although diffusion indexes commonly are interpreted as showing the percent of components that increased over the time span, the index reflects half of the unchanged components, as well. (This is the effect of assigning a value of 50 percent to the unchanged components when computing the index.)

Hours and earnings

The CES hours and earnings series are derived from reports of payrolls and the corresponding paid hours for all employees and also for the various types of production employees. Hours and earnings are for private-sector employees.

Payroll refers to the payroll for full- and part-time workers who received pay for any part of the pay period that includes the 12th day of the month. The payroll is reported before deductions of any kind, such as those for old-age and unemployment insurance, group insurance, withholding tax, bonds, or union dues; also included is pay for overtime,

holidays and vacation, sick leave paid directly by the firm, and commissions paid at least monthly. Bonuses (unless earned and paid regularly each pay period); other pay not earned in the pay period reported (such as retroactive pay); and the value of free rent, fuel, meals, or other payment in kind are excluded. Employee benefits (such as health and other types of insurance, contributions to retirement, and so forth, paid by the employer) also are excluded.

Total hours during the pay period include all hours worked (including overtime hours), hours paid for standby or reporting time, and equivalent hours for which employees received pay directly from the employer for sick leave, holidays, vacations, and other leave. Overtime and other premium pay hours are not converted to straight-time equivalent hours. The concept of total hours differs from those of scheduled hours and hours worked.

Average weekly hours relate to the average hours per worker for which pay was received and is different from standard or scheduled hours. Factors such as unpaid absenteeism, labor turnover, part-time work, and stoppages cause average weekly hours to be lower than scheduled hours of work for an establishment. Group averages further reflect changes in the workweek of component industries. Average weekly hours are the total weekly hours divided by the employees paid for those hours.

Overtime hours represent that portion of average weekly hours that exceeded regular hours and for which overtime premiums were paid. If an employee were to work on a paid holiday at regular rates, receiving as total compensation his holiday pay plus straight-time pay for hours worked that day, no overtime hours would be reported. Overtime hours data are collected only from manufacturing establishments.

Because overtime hours are premium hours by definition, weekly hours and overtime hours do not necessarily move in the same direction from month to month. Such factors as work stoppages, absenteeism, and labor turnover may not have the same influence on overtime hours as on average hours. Diverse trends at the industry group level also may be caused by a marked change in hours for a component industry in which little or no overtime was worked in both the previous and current months.

Average hourly earnings are on a "gross" basis. They reflect not only changes in basic hourly and incentive wage rates, but also such variable factors as premium pay for overtime and late-shift work and changes in output of workers paid on an incentive plan. They also reflect shifts in the number of employees between relatively high-paid and low-paid work and changes in workers' earnings in individual establishments. Averages for groups and divisions further reflect changes in average hourly earnings for individual industries.

Averages of hourly earnings differ from wage rates. Earnings are the actual return to the worker for a stated period; rates are the amount stipulated for a given unit of work or time. The earnings series do not measure the level of total labor costs on the part of the employer because the following are excluded: benefits, irregular bonuses, retroactive items, payroll taxes paid by employers.

Average hourly earnings, excluding overtime-premium pay, are computed by dividing the total worker payroll for the industry group by the sum of total worker hours and one-half of total overtime hours. No adjustments are made for other premium payment provisions, such as holiday pay, late-shift premiums, and overtime rates other than time and one-half. Average hourly earnings excluding overtime are calculated only for manufacturing industries.

Average weekly earnings are derived by multiplying average weekly hours estimates by average hourly earnings estimates. Therefore, weekly earnings are affected not only by changes in average hourly earnings but also by changes in the length of the workweek. Monthly variations in such factors as the proportion of part-time workers, stoppages for varying reasons, labor turnover during the survey period, and absenteeism for which employees are not paid may cause the average workweek to fluctuate.

Long-term trends of average weekly earnings can be affected by structural changes in the makeup of the workforce. For example, persistent long-term increases in the proportion of part-time workers in retail trade and many of the services industries have reduced average workweeks in these industries and have affected the average weekly earnings series.

Real earnings data (those expressed in constant 1982-84 dollars) result from the adjustment of average hourly and weekly earnings by the BLS Consumer Price Indexes. Real earnings for production and nonsupervisory employees are deflated by the Consumer Price Index for Urban Wage Earners and Clerical Workers (CPI-W), while real earnings for all employees are deflated by the Consumer Price Index for All Urban Consumers (CPI-U). Real earnings indicate the purchasing power of money earnings after adjustment for changes over time in the prices of consumer goods and services. These data cannot be used to measure changes in living standards as a whole, which are affected by other factors such as total family income, the extension and incidence of various social services and benefits, and the duration and extent of employment and unemployment. The long-term trends of these earnings data also are affected by changing mixes of full-time and part-time workers, high-paid and low-paid workers, and so on.

Indexes of aggregate weekly hours and payrolls. Indexes of aggregate weekly hours are calculated by dividing the current month's aggregate hours by the average of the 12 monthly figures, for the base year. Indexes are based on 2007 averages for all employees and on 2002 averages for production and nonsupervisory employees. For basic industries, the hours aggregates are the product of average weekly hours and employment of workers to which the hours apply (all employees or production and nonsupervisory employees). At all higher levels of industry aggregation, hours aggregates are the sum of the component aggregates.

Indexes of aggregate weekly payrolls are calculated by dividing the current month's aggregate by the average of the 12 monthly figures for the base year. Indexes are averages for production and nonsupervisory employees. For basic industries, the payroll aggregates are the product of average hourly earnings and aggregate weekly hours. At all higher levels of industry aggregation, payroll aggregates are the sum of the component aggregates.

Industrial classification

All data on employment, hours, and earnings for the Nation and for States and areas are classified in accordance with the 2007 North American Industry Classification System (NAICS), specified by the U.S. Office of Management and Budget. The United States, Canada, and Mexico share this classification system, and, thus, it allows a direct comparison of economic data across the three countries.

Establishments are classified into industries on the basis of their primary activity. Those that use comparable capital equipment, labor, and raw material inputs are classified together. This information is collected on a supplement to the quarterly unemployment insurance tax reports filed by employers. For an establishment engaging in more than one activity, the entire employment of the establishment is included under the industry indicated by the principal activity.

Sample Design

The CES sample design is a stratified, simple random sample of worksites, clustered by UI account number. The sample strata, or subpopulations, are defined by State, industry, and employment size, yielding a State-based design. Sampling rates for each stratum are determined through optimum allocation, which distributes a fixed number of sample units across a set of strata to minimize the overall variance or sampling error on the primary estimate of interest, the statewide total nonfarm employment level.

The sampling frame, and the CES sample itself, are updated twice a year with new quarters of UI-based universe data. This helps keep the sample up-to-date by adding firm births and deleting business deaths. In addition, the design specifies an annual update process, which includes sample frame maintenance and the redrawing of the entire sample for the first quarter of each year. Frame maintenance provides for the updating of industry, employment size class, and metropolitan area designations and for the merging of semiannual birth samples into the overall frame.

Data Sources and Collection Methods

Sample data

Each month, BLS collects data on employment, payroll, and paid hours from a sample of establishments. To encourage participation in this voluntary survey, BLS uses a variety of collection techniques, tailored to individual firm preferences. Data collection centers (DCCs) perform initial enrollment of each firm via telephone, collect the data for several months via Computer Assisted Telephone Interviewing (CATI) and where possible transfer respondents to a self-reporting mode such as touch-tone data entry (TDE), FAX or Internet collection. Very large, multi-establishment firms are often enrolled via personal visit, and ongoing reporting is established via electronic data interchange (EDI). These firms provide electronic files to BLS that include data from all their worksites.

For the few establishments that do not use the above methods, data are collected using mail, and transcript.

Sample enrollment. BLS has a comprehensive program of new sample unit solicitation in the DCCs. Approximately 55,000 new sample units are enrolled in the CES survey each year to account for the births of new firms, to realign the sample distribution with the universe distribution, and to rotate a portion of the sample. All firms with 1,000 or more employees are asked to participate in the survey, as is a sample of firms across all employment sizes. When firms are rotated into the sample, they are retained for 2 years or more. When a respondent is rotated out of the sample, BLS will not ask the firm to participate for at least 3 years.

Data reporting. Each month, respondents extract the employment, hours, and earnings data from their payroll records and submit it to BLS. Data are collected for the pay period that includes the 12th of each month.

A CES reporting form (BLS form 790 series) is provided to all CES respondents except those that report via electronic file. The form provides a convenient means to record payroll data each month. Six variations of the basic CES form are used, and each variation is tailored for the data items, concepts, and definitions of major industry sectors. Separate forms are used for mining and logging, construction, manufacturing, service-providing industries, public administration, and educational services.

CES data collection forms are available on the BLS Web site at **http://www.bls.gov/ces/idcfcesforms.htm**.

The design of the CES form is important for maintaining continuity and consistency in reporting from month to month. The use of a single form for a 6-month period allows the respondent to compare the latest data with data submitted in prior months.

All reported data, regardless of method of collection, are edited by BLS to ensure the information is correctly reported and is consistent with the data reported by the establishment in earlier months. The data are further edited to detect processing and reporting errors that might have been missed during collection. When questionable reports are discovered at any stage of the editing process, BLS contacts the respondents for clarification or correction. The staff of the BLS Washington office prepares national estimates of employment, hours, and earnings using the edited data. The State agencies also cooperate with BLS to develop State and metropolitan area estimates.

Estimating Methods

Benchmark data

For the establishment survey, annual benchmarks are constructed to realign the sample-based employment totals for March of each year with the UI-based population counts for March. These population counts are less timely than sample-based estimates and are used to provide an annual point-in-time census for employment. For National series, only the March sample-based estimates are replaced with UI counts. For State and metropolitan area series, all available months of UI data are used to replace sample-based estimates. State and area series are based on smaller samples and are, therefore, more vulnerable to both sampling and non-sampling errors than National estimates.

Population counts are derived from the administrative file of employees covered by UI. All employers covered by UI laws are required to report employment and wage information to the appropriate State workforce agency four times a year. Approximately 97 percent of total nonfarm employment within the scope of the establishment survey is covered by UI. A benchmark for the remaining 3 percent is constructed from alternate sources, primarily records from the Railroad Retirement Board and County Business Patterns. The full benchmark developed for March replaces the March sample-based estimate for each basic cell. The monthly sample-based estimates for the year preceding and the year following the benchmark are also then subject to revision.

Monthly estimates for the year preceding the March benchmark are readjusted using a "wedge back" procedure. The difference between the final benchmark level and the previously published March sample estimate is calculated and spread back across the previous 11 months. The wedge is linear; eleven-twelfths of the March difference is added to the February estimate, ten-twelfths to the January estimate, and so on, back to the previous April estimate, which receives one-twelfth of the March difference. This assumes that the total estimation error since the last benchmark accumulated at a steady rate throughout the current benchmark year.

Estimates for the 7 months following the March benchmark also are recalculated each year. These post-benchmark estimates reflect the application of sample-based monthly changes to new benchmark levels for March and the computation of new business birth/death factors for each month.

Following the revision of basic employment estimates, estimates for women employees and production and nonsupervisory employees are recomputed using the revised all-employee estimates and the previously computed sample ratios of these workers to all employees. All basic series of employment, hours, and earnings are re-aggregated to obtain estimates for each sector and higher level of detail. Other derivative series (such as real earnings and payroll indexes) also are recalculated. New seasonal adjustment factors are calculated and all data series for the previous 5 years are re-seasonally adjusted before full publication of all revised data in February of each year.

Monthly Estimation

Stratification. The CES sample is stratified into basic estimation cells for purposes of computing national employment, hours, and earnings estimates. Basic cells are defined primarily by detailed industry at the 3-, 4-, 5-, or 6-digit NAICS level. Aggregation results in additional summary cells.

In addition to the basic and summary estimation cells, a small number of independently estimated cells exist and do not aggregate to the summary cell levels.

The Current Employment Statistics (CES) or establishment survey estimates of employment are generated through an annual benchmark and monthly sample link procedure. Annual universe counts or benchmark levels are generated primarily from administrative records on employees covered by unemployment insurance (UI) tax laws. These annual benchmarks, established for March of each year, are projected forward for each subsequent month based on the trend of the sample employment and an adjustment for the net of business births and deaths employment. Benchmarks and monthly estimates are computed for each basic estimating cell and summed to create aggregate-level employment estimates.

Matched sample. CES uses a matched sample concept to produce estimates. A matched sample is defined to be all sample members that have reported data for the reference month and the month prior. Excluded from the matched

sample is any sample unit that reports that it is out-of-business. This aspect of the estimation methodology is more fully described in the section on estimation of business births and deaths employment below.

Estimates of all employees require that only the number of total employees be reported for the current estimated month and the prior month. The matched sample for estimates of production/nonsupervisory employees includes reporting units that reported both all employees and production/nonsupervisory employees for both the previous and current months. The matched sample for estimates of women employees includes reporting units that reported both all employees and women employees for both the previous and current months.

The matched sample for average weekly hours and average hourly earnings of all employees includes responding units that have reported all employees and the corresponding worker hours and payrolls for both the previous and current months. The matched sample for average weekly hours and average hourly earnings of production/nonsupervisory employees, includes reporting units that have reported production/nonsupervisory employees and their corresponding worker hours and payrolls, as well as all employees, for both the previous and current months.

For average weekly overtime estimates of all employees, which are calculated for manufacturing industries only, the matched sample includes reporting units that have reported all employees and the corresponding work hours, payrolls, and overtime hours for both the previous and current months. The matched sample for average weekly overtime hours of production/nonsupervisory employees includes production/nonsupervisory employees and their corresponding worker hours, payrolls, and overtime hours, as well as all employees, for both the previous and current months.

Variables for national estimates. The weighted-link-relative formula is used to calculate estimates of all employees, while the difference-link-and-taper formula is used to calculate all other data types. Both formulas use data reported that meets the matched sample criteria. The difference-link-and-taper formula also uses estimates for the month prior to the reference month or derivatives of estimates, such as ratios. See the table of variable definitions for the link-relative and difference-link-and-taper formulas on page 7.

National employment. The weighted link-relative estimator for the all employee series uses the sample trend in the cell to move the previous level to the current-month estimated level. A model-based component is applied to account for the net employment resulting from business births and deaths not captured by the sample.

The weighted link-relative technique is efficient in that it takes advantage of a reliable, complete count of employment and of the high correlation between levels of employment in successive months in identical establishments.

Current-month estimate of all employees is defined as

$$\widehat{\mathrm{AE}}_c = \left(\widehat{\mathrm{AE}}_p - \sum_j \mathrm{ae}^*_{p,j}\right) \times \frac{\sum_i (w_i \times \mathrm{ae}_{c,i}) - \sum_j (w_j \times \mathrm{ae}^*_{c,j})}{\sum_i (w_i \times \mathrm{ae}_{p,i}) - \sum_j (w_j \times \mathrm{ae}^*_{p,j})} + \sum_j \mathrm{ae}^*_{c,j} + \mathrm{b}_c$$

$$\text{for all } i \in I \text{ and } j \in J$$

Business birth and death estimation. In a dynamic economy, firms are continually opening and closing. These two occurrences offset each other to some extent. That is, firms that are born replace firms that die. CES uses this fact to account for a large proportion of the employment associated with business births. This is accomplished by excluding business death units from the matched sample definition. Effectively, business deaths are not included in the sample-based link portion of the estimate, and the implicit imputation of their previous month's employment is assumed to offset a portion of the employment associated with births.

There is an operational advantage associated with this approach, as well. Most firms will not report that they have gone out of business; rather, they simply cease reporting and are excluded from the link, as are all other nonrespondents. As a result, extensive follow-up with monthly nonrespondents to determine whether a company is out-of-business or simply did not respond is not required.

Employment associated with business births will not exactly equal that associated with business deaths. The amount by which it differs varies by month and by industry. As a result, the residual component of the birth/death offset must be accounted for by using a model-based approach.

Birth-death residual = Population - Sample-based estimate + Error

During the net birth/death modeling process, simulated monthly probability estimates over a 5-year period are created and compared with population employment levels. Moving from a simulated benchmark, differences between the series across time represent a cumulative birth/death component. Those residuals are converted to month-to-month differences and used as input series to the modeling process.

Models are fit using X-12 autoregressive integrated moving average (ARIMA). Outliers, level shifts, and temporary ramps are automatically identified.

Variable definitions

- All estimated values are shown in upper case.
- All sample measures (shown in lower case) are based on a matched sample.
- The estimator for women employees takes the same form as the estimator for production/nonsupervisory employees, where PE and PER are the estimates for women employees and women-to-all employee ratio, respectively, and matched sample totals pe are the matched sample totals for women.
- The estimator for average weekly hours for production/nonsupervisory employees takes the same form as average weekly hours for all employees, where AE and AWH represent estimates of production/nonsupervisory employees and average weekly hours of production/nonsupervisory employees, respectively, and the matched sample totals ae and wh represent matched sample totals for production employees and wh for production/nonsupervisory employees, respectively.
- The estimator for average hourly earnings for production/nonsupervisory employees takes the same form as average hourly earnings for all employees, where AHE and WH represent estimates of production/nonsupervisory employees and their work hours, and the matched sample totals pr and wh represent matched sample totals of payroll and work hours for production/nonsupervisory employees
- The estimators for average weekly overtime take the same form as average weekly hours, where AWH represents the estimates of average weekly overtime hours and wh represents the matched sample for total overtime hours reported. Overtime estimates are calculated for manufacturing industries only.

Variable	Description	All employees	Production/ nonsupervisory (or women employees)	Average weekly hours (or average weekly overtime hours)	Average hourly earnings
*	Atypical sample data.	X	X	X	X
α	0.9		X	X	X
β	0.1		X	X	X
c	Current month.	X	X	X	X
p	Previous month.	X	X	X	X
$\widehat{AE}$	Estimated employment for all employees (or for production/ nonsupervisory employees when estimating their respective hours).	X	X	X	
ae	Reported all employees (or production/ nonsupervisory employees when estimating their respective hours).	X	X	X	
$\widehat{\text{AHE}}$	Estimated average hourly earnings for all employees (or for production/ nonsupervisory employees when estimating their respective earnings).				X

Variable	Description	All employees	Production/ nonsupervisory (or women employees)	Average weekly hours (or average weekly overtime hours)	Average hourly earnings
$\widehat{AWH}$	Estimated average weekly hours for all employees (or for production/ nonsupervisory employees when estimating their respective hours).			X	
b	net birth/death factor.	X			
i	a matched CES report for sample data variables shown in lower case.	X	X	X	X
j	a matched CES report where the current month is atypical.	X	X	X	X
$\widehat{PE}$	Estimated production/nonsupervisory (or women) employees.		X		
pe	Reported production/nonsupervisory (or women) employees.		X		
$\widehat{PER}$	Estimated ratio of production/ nonsupervisory (or women) employees to all employees.		X		
pr	Reported weekly payroll for all employees (or for production/ nonsupervisory employees when estimating their respective earnings).				X
w	Weight associated with a CES report.	X	X	X	X
wh	Reported weekly hours for all employees (or for production/ nonsupervisory employees when estimating their respective hours and earnings).			X	X
$\widehat{WH}$	Estimated weekly hours for all employees (or production/ nonsupervisory employees), derived from estimates of average weekly hours and employment.				X

The net birth/death model component figures are unique to each month and exhibit seasonal patterns that can result in negative adjustments in some months.

Weighted-link-and-taper estimator is used for all data types except for "all employees." The estimator accounts for the over-the-month change in the sampled units, but also includes a tapering feature used to keep the estimates close to the overall sample average over time. The taper is considered to be a level correction. This estimator uses matched sample data; it tapers the estimate toward the sample average for the previous month of the current matched sample before applying the current month's change; and it promotes continuity by heavily favoring the estimate for the previous month when applying the numerical factors.

Production and nonsupervisory employees. To obtain estimates of production (or construction or nonsupervisory) worker employment, the ratio of weighted production employees to the weighted all employees in the sample is assumed to equal the same ratio in the universe. The current month's production worker ratio, thus, is estimated and then multiplied by the all-employee estimate. The weighted-difference-link-and-taper formula, described in the section on hours and earnings, is used to estimate the current month's production worker ratio. This formula adds the change in the matched sample's production worker ratio (the weighted-difference link) to the prior month's estimate, which has been slightly modified to reflect changes in the sample composition (the taper). (See page 10.) An analogous method is used to estimate the number of women employees.

Women employees. Estimation of the series for women employees is identical to that described for production employees, with the appropriate substitution of women employees values for the production worker values in the previous formulas.

Estimates for each type of series (all employees, production employees, and women employees) for individual basic estimating cells are summed to obtain corresponding totals for broader industry sectors.

Hours and earnings

Average weekly hours and average hourly earnings. Independent benchmarks are not available for the hours and earnings series; consequently, the levels are derived directly from the CES weighted-sample averages. (See pages 10 and 11.) Before hours and earnings sample averages or estimates are calculated, all employees, production employees and aggregate hours and payrolls must be multiplied by sample weights both for the month for which estimates are being made and for the prior month. To establish average weekly hours for a basic estimating cell, the sum of reported worker hours for the establishments classified in the cell is divided by the total number of all employees or production employees reported for the same establishments. To establish average hourly earnings, the reported payroll is divided by the reported worker hours for the same establishments.

Average weekly hours and average hourly earnings for industries and groups above the basic estimating cell level are weighted averages of the figures for component cells. Average weekly hours for each basic estimating cell are multiplied by the corresponding estimate of the number of all employees to derive aggregate worker hours. Payroll aggregates are the product of the aggregate worker hours and average hourly earnings. Payroll and worker-hour aggregates for industry groups and divisions are the sums of the aggregates for the component industries.

Average weekly hours for industry groups are obtained by dividing the worker-hour aggregates by the corresponding all-employee estimates. Average hourly earnings for industry groups are computed by dividing payroll aggregates by worker-hour aggregates. This method is equivalent to weighting average weekly hours by the estimated number of all employees in the universe and weighting average hourly earnings by the estimated worker hours for the universe.

For all levels, from basic estimating cells to sector level and higher aggregates, average weekly earnings are computed by multiplying average hourly earnings by average weekly hours.

Current month estimate of overtime hours

Estimation of overtime hours is identical to that described for average weekly hours, with the appropriate substitution of overtime hours values for the weekly hours values in the previous formula.

Estimation formulas for *hours and earnings for production employees* are essentially the same as hours and earnings for all employee series, whereby all 'AE' estimates and 'ae' sample terms are replaced by 'PE' estimates and 'pe' sample, respectively.

Current-month estimate of production or nonsupervisory employees is defined as

$$\widehat{\mathrm{PE}}_c = \left(\left(\widehat{\mathrm{AE}}_c - \sum_j \mathrm{ae}^*_{c,j} \right) \times \widehat{\mathrm{PER}}_c \right) + \sum_j \mathrm{pe}^*_{c,j}$$

, where

$$\widehat{\mathrm{PER}}_c = \left(\alpha \times \widehat{\mathrm{PER}}_p\right) + \left(\beta \times \frac{\sum_i \left(w_i \times \mathrm{pe}_{p,i}\right) - \sum_j \left(w_j \times \mathrm{pe}^*_{p,j}\right)}{\sum_i \left(w_i \times \mathrm{ae}_{p,i}\right) - \sum_j \left(w_j \times \mathrm{ae}^*_{p,j}\right)} \right) +$$

$$\frac{\sum_i \left(w_i \times \mathrm{pe}_{c,i}\right) - \sum_j \left(w_j \times \mathrm{pe}^*_{c,j}\right)}{\sum_i \left(w_i \times \mathrm{ae}_{c,i}\right) - \sum_j \left(w_j \times \mathrm{ae}^*_{c,j}\right)} - \frac{\sum_i \left(w_i \times \mathrm{pe}_{p,i}\right) - \sum_j \left(w_j \times \mathrm{pe}^*_{p,j}\right)}{\sum_i \left(w_i \times \mathrm{ae}_{p,i}\right) - \sum_j \left(w_j \times \mathrm{ae}^*_{p,j}\right)}$$

for all $i \in I$ and $j \in J$

Current-month estimate of average weekly hours for all employees

$$\widehat{\mathrm{AWH}}_c = \alpha \times \widehat{\mathrm{AWH}}_p + \beta \times \left(\frac{\left(\frac{\sum_i \left(w_i \times \mathrm{wh}_{p,i}\right) - \sum_j \left(w_j \times \mathrm{wh}^*_{p,j}\right)}{\sum_i \left(w_i \times \mathrm{ae}_{p,i}\right) - \sum_j \left(w_j \times \mathrm{ae}^*_{p,j}\right)} \right) \times \left(\widehat{\mathrm{AE}}_p - \sum_j \mathrm{ae}^*_{p,j} \right) + \sum_j \mathrm{wh}^*_{p,j}}{\widehat{\mathrm{AE}}_p} \right)$$

$$+ \left(\frac{\left(\frac{\sum_i \left(w_i \times \mathrm{wh}_{c,i}\right) - \sum_j \left(w_j \times \mathrm{wh}^*_{c,j}\right)}{\sum_i \left(w_i \times \mathrm{ae}_{c,i}\right) - \sum_j \left(w_j \times \mathrm{ae}^*_{c,j}\right)} \right) \times \left(\widehat{\mathrm{AE}}_c - \sum_j \mathrm{ae}^*_{c,j} \right) + \sum_j \mathrm{wh}^*_{c,j}}{\widehat{\mathrm{AE}}_c} \right)$$

$$- \left(\frac{\left(\frac{\sum_i \left(w_i \times \mathrm{wh}_{p,i}\right) - \sum_j \left(w_j \times \mathrm{wh}^*_{p,j}\right)}{\sum_i \left(w_i \times \mathrm{ae}_{p,i}\right) - \sum_j \left(w_j \times \mathrm{ae}^*_{p,j}\right)} \right) \times \left(\widehat{\mathrm{AE}}_p - \sum_j \mathrm{ae}^*_{p,j} \right) + \sum_j \mathrm{wh}^*_{p,j}}{\widehat{\mathrm{AE}}_p} \right)$$

for all $i \in I$ and $j \in J$

Current month estimate of average hourly earnings for all employees

$$\widehat{\mathrm{AHE}}_c = \alpha \times \widehat{\mathrm{AHE}}_p + \beta \times \left(\frac{\left(\frac{\sum_i \left(w_i \times \mathrm{pr}_{p,i} \right) - \sum_j \left(w_j \times \mathrm{pr}^*_{p,j} \right)}{\sum_i \left(w_i \times \mathrm{wh}_{p,i} \right) - \sum_j \left(w_j \times \mathrm{wh}^*_{p,j} \right)} \right) \times \left(\widehat{WH}_p - \sum_j \mathrm{wh}^*_{p,j} \right) + \sum_j \mathrm{pr}^*_{p,j}}{\widehat{\mathrm{WH}}_p} \right)$$

$$+ \left(\frac{\left(\frac{\sum_i \left(w_i \times \mathrm{pr}_{c,i} \right) - \sum_j \left(w_j \times \mathrm{pr}^*_{c,j} \right)}{\sum_i \left(w_i \times \mathrm{wh}_{c,i} \right) - \sum_j \left(w_j \times \mathrm{wh}^*_{c,j} \right)} \right) \times \left(\widehat{\mathrm{WH}}_c - \sum_j \mathrm{wh}^*_{c,j} \right) + \sum_j \mathrm{pr}^*_{c,j}}{\widehat{\mathrm{WH}}_c} \right)$$

$$- \left(\frac{\left(\frac{\sum_i \left(w_i \times \mathrm{pr}_{p,i} \right) - \sum_j \left(w_j \times \mathrm{pr}^*_{p,j} \right)}{\sum_i \left(w_i \times \mathrm{wh}_{p,i} \right) - \sum_j \left(w_j \times \mathrm{wh}^*_{p,j} \right)} \right) \times \left(\widehat{\mathrm{WH}}_p - \sum_j \mathrm{wh}^*_{p,j} \right) + \sum_j \mathrm{pr}^*_{p,j}}{\widehat{\mathrm{WH}}_p} \right)$$

for all $i \in I$ and $j \in J$

Robust Estimation Procedure

The matched sample sometimes contains a small number of observations that may have a large and adverse effect on the estimate of the relative change. The influence of such observations may be due to large survey weights, an unusual level or change in reported employment, or the combined effect of these factors. If left untreated, influential observations may cause unreasonable jumps in the monthly estimates, especially at detailed publication levels.

The Robust Estimation procedure is designed to reduce the effect of the influential observations on the estimate of the relative over-the-month change. At the same time, it is recognized that the unusual and influential sample movements may represent similar behavior in the target population and a heavy intervention to the regular estimation procedure may lead to biased estimates. This is especially true if the sample is large. Therefore, the estimator is designed to reduce the volatility of the estimates due to extreme outlying reports while controlling the intervention to protect against the incurred bias.

Definition of influential reports. The CES weighted link relative estimate is based on the ratio of the two survey weighted sums. A scatterplot of the weighted employment reported in two consecutive months provides an insight on what units influence the estimate. Two examples are shown in Figure 1: the survey weighted employment reported for the month (t) is plotted against the weighted employment reported for the previous month (t-1). The line shows the survey weighted link relative trend of matched sample data.

An influential report would have a relatively large survey weight and/or a large change in its reported employment. Numerically, the influence of a report on the sample link relative estimate can be expressed in the form of weighted residuals:

$$d_{i,t} = w_i \left(y_{i,t} - R_t \, y_{i,t-1} \right) \qquad (1)$$

where,

$y_{i,t}$ and $y_{i,t-1}$ denote, respectively, a unit's current and previous months reported employment; w_i is the selection weight; R_t is the sample link relative estimate in a given basic estimation cell. The formula is identified by (1).

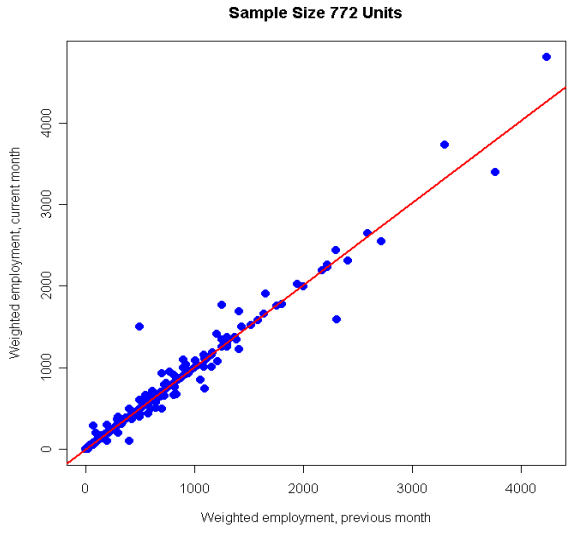
The influential reports are those having large positive or negative values of the weighted residuals compared to the other sample units. The extreme residuals are reduced to specific cut-off values. The cut-off values depend on the distribution of the weighted residuals in a given series and are determined independently for each month and industry series. Pushing the extreme residuals to the cut-off values is accomplished by using an appropriate weight adjustment factor.

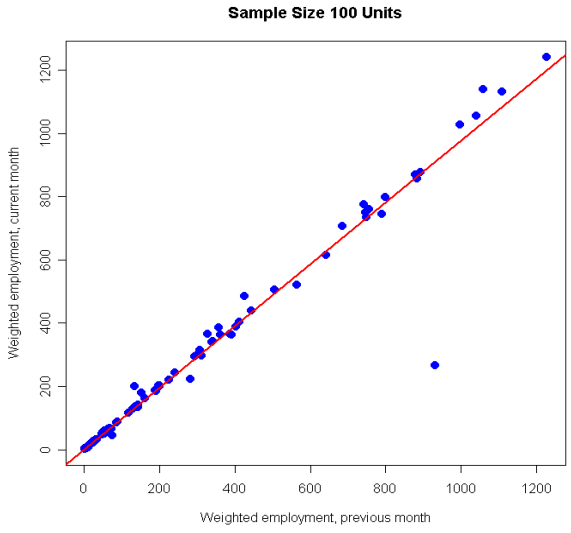
The procedure used for the CES robust estimation is a particular variation of a general method of weight reduction known as "winsorization." See Kokic and Bell (1994).

The actual cut-off values are determined by examining the relative distances of units with extreme weighted residuals to the nearest but less extreme values in the same cell and month.

Determining Weight Adjustment Factors. The first step consists of calculating the weighted residuals using formula (1). The weighted residuals from individual establishments are aggregated to the UI account level within each estimation cell. This is done because reports within a UI account may have similar reported change in employment. Since they are similar, it is possible that none of the individual reports will be identified as an outlying unit. At the same time, the UI level residual aggregated from all the responding establishments may be extreme and very different from other responding UI accounts in the cell.

Figure 1





The residuals may legitimately have very different values depending on the employment size class of the sampled UI accounts. To remove the effect of the size class, the residuals are "centered" within each size class, i.e., the average of the residuals within size classes is subtracted from each original residual.

The cut-off values are determined separately for the extreme positive and extreme negative residuals. The procedures are similar for the positive and negative residuals and are described here only for the positive residuals.

First, sort the residuals in each cell in descending order. Let d_i denote the *i*-th largest centered positive residual:

$$d_1 \geq d_2 \geq \ldots \geq d_n.$$

Set the value $F_1 = d_1$ and $F_2 = 2d_2 - d_1$.

The general formula for F_k is:

$$F_k = (k + 1)\, d_k - (d_1 + d_2 + \ldots + d_k). \qquad (2)$$

Proceed with the computations of the sequence of $F_1,\ldots,F_k$ until, at some step *k*, $F_k \leq 0$. Typically, this point is reached after only a few steps. Next, compute the cut-off value as the point between residuals d_k and d_{k-1} using the formula:

$$L = a \cdot d_k + (1 - a) \cdot d_{k-1}, \qquad (3)$$

where

$$a = 0.8F_{k-1} / (0.8F_{k-1} - 0.2F_k).$$

The initial adjustments for units whose residuals are greater than L are:

$$init.adj_i = L/d_i.$$

The cut-off values defined using the described procedure are always placed between some neighboring ordered residuals, so that all the residuals on the right from a cut-off value are greater than the cut-off value, although they may be very close to it. Based on the historical CES estimates, it has been found that the following rules for determining the final adjustments work the best.

For certainty units, if the initial adjustment is less than 0.5, then a unit is declared to be atypical, representing only itself. The atypical data is removed from the matched sample set and is not used in the estimation of the sample link relative. If the initial adjustment for a certainty unit is greater or equal to 0.5, no intervention is required and the adjustment is reset to 1.

For non-certainty units, if the initial adjustment is less than or equal to 0.3, then a unit is declared atypical and it is not used in the sample link relative estimation. If the initial adjustment is greater than 0.4, then the final adjustment is reset to 1. Adjustments between 0.3 and 0.4 are applied to the sample weight. The final adjusted weight is required to be equal or greater than 1. For example, if the resulting adjusted weight falls below 1, the final weight is reset to 1.

An intervention, such as a weight adjustment, into the regular estimation procedure would reduce the variance of the estimate but it may introduce a bias. Therefore, the intervention, especially in samples of moderate to large sizes, should be done with caution. For example, it is possible that there exist units in the non-sampled part of the population that are similar to the influential observations in the sample. Moderating the effect of the sample's influential units may lead to a reduction in the representativeness of the sample. Since the non-sampled part of the population is not available, it is difficult to judge the amount and the need of intervention based only on the observed sample. One way to protect against unwarranted intervention is to verify its necessity using historical CES estimates. If the estimate falls within the historically observed bounds, then the intervention is deemed unnecessary and the weight adjustments are discarded.

At the very first step of the procedure, the sample link relative estimate is used when defining the residuals. This estimate may itself be affected by the extreme influential observations. Therefore, the whole procedure is performed a second time. The atypical units determined during the initial run are not used in calculating the adjustment factors during the second run.

The reports identified by the robust estimation techniques are treated as atypical in the link-relative technique, while all other matched sample responses are treated as typical in the link-relative technique.

Special Estimation Situations

Small domain model. Relatively small sample sizes in some industries limit the reliability of the weighted-link-relative estimates of all employees. For a few industries (identified in the annual benchmark article, **http://www.bls.gov/web/cesbmart.htm**), BLS uses the CES small domain model (SDM). In addition, BLS and some cooperating State partners use the CES SDM for those State and metropolitan area employment series that have small samples. Estimation of nonsupervisory employees, average weekly hours, and average weekly and hourly earnings uses the standard weighted link-and-taper methodology.

The CES SDM is a weighted least squares (WLS) model with two employment inputs: (1) an estimate based on available CES sample for that series, and (2) an ARIMA projection based on trend from 10 years of historical QCEW data.

Estimator Based on Fay-Herriot Model. To estimate employment for State supersector cells with smaller sample sizes, the CES program uses an estimator based on the Fay-Herriot model. See Fay and Harriot (1979). In the smaller cells, a direct sample-based estimate of the over-the-month change in employment is often unreliable due to the large variance, although the direct estimator is assumed to be approximately unbiased.

In order to make more stable estimates, additional information is used. The model is formulated for a set of States in a given supersector *b* and at a given month *t*. (Since the indexes *b* and *t* in the following description are the same for all States involved in the model, they are suppressed to simplify the notation.)

Let R_a denote a true value of the relative over-the-month employment change in State *a*, $R_a^{(1)}$ is a direct sample-based estimate, and $R_a^{(2)}$ is an ARIMA forecast of the relative change. The model used in CES is formulated as follows:

$$R_a^{(1)} = R_a + e_a \qquad (1)$$

and

$$R_a = \beta R_a^{(2)} + u_a, \quad (2)$$

for a set of States $a = 1,\ldots,K$ (within a supersector *b* at time *t*). The error terms e_a and u_a are assumed to be independent and normally distributed, with mean zero. The variance of the direct sample estimate $R_a^{(1)}$ (and thus the variance of e_a) is $V_a^{(1)}$. It is estimated based on the sample and is smoothed using several years of data, to add stability. The values for the variance of the random effects of u_a, denoted A, and the parameter β are estimated from the model using the method described in the Fay and Herriot (1979) paper.

The resulting model estimator $R_a^{(FH)}$ can be presented as a weighted average of the sample-based estimate and an adjusted ARIMA forecast, as follows:

$$R_a^{(FH)} = \gamma_a^{(1)} R_a^{(1)} + \gamma_a^{(2)}(2)\,\beta\, R_a^{(2)} \qquad (3)$$

The weights $\gamma_a^{(1)}$ *and* $\gamma_a^{(2)}$ are:

$$\gamma_a^{(1)} = \frac{A}{A + V_a^{(1)}} \text{ and } \gamma_a^{(2)} = 1 - \gamma_a^{(1)}$$

The $\beta R_a^{(2)}$ component of the weighted average *a* described in equation (3) is called a synthetic part of the estimator. The variance, A, depends on the strength of the relationship between R_a and the forecasts $R_a^{(2)}$, across the States in a given supersector, as specified by equation (2). The relative magnitudes of A and $V_a^{(1)}$ indicate how much weight should be given to the synthetic part relative to how much weight should be given to the sample-based estimate. If the linear relationship described in (2) holds without error, A = 0, all weight would go to the synthetic part $\beta R_a^{(2)}$ however, if $R_a^{(2)}$ is a poor predictor, then A is large and more weight would be given to the direct sample estimate. The strength is "borrowed" across States within a supersector to estimate β, thus "correcting" the time series forecast using the "adjusted" $\beta R_a^{(2)}$ value, and to obtain A, the estimated "strength" of the prediction.

The model estimate of the employment level is obtained by applying the model-based estimate of the relative change to the preceding month's level of employment. For a month *t*, the estimator of the employment level is a State *a* supersector *b* is:

$$\gamma_{ab,t}^{(FH)} = \gamma_{ab,t-1} R_{ab,t}^{(FH)} \qquad (4)$$

Education and Religious Organizations. Due to the small sample in religious organizations (NAICS 8131), and definitional exclusions in the collection of data for educational services (NAICS 611), certain ratios for these series are recalculated with each benchmark to allow for the creation of aggregate totals. Production worker and women worker ratios, average hourly earnings, and average weekly hours for these series are calculated based on the weighted average of the previous year's professional and technical services, education and health services, leisure and hospitality, and other services' annual averages. BLS sets the March benchmark values based on the prior calendar year's annual averages.

The education services series uses the nonsupervisory employee ratio, average hourly earnings, and average weekly hours calculated from the weighted average. The religious organizations series uses the production employee ratio, women employee ratio, average hourly earnings, and average weekly hours calculated from the weighted average. In both cases, the ratios, average hourly earnings, and average weekly hours are held constant through the next benchmark.

Railroad estimates. BLS obtains monthly employment counts for class 1 railroads, which are not included in the QCEW universe. The Department of Transportation Surface Transportation Board (STB) publishes a mid-month employment count for the survey week of the previous month. (**http://www.stb.dot.gov/econdata.nsf/**) BLS uses this data to estimate employment for railroads. The data from STB are also used to set the benchmark employment levels and hours and earnings.

Residential and Nonresidential Specialty Trade Contractors estimates. Residential and nonresidential employment estimates in Specialty Trade Contractors (NAICS 238) are produced as breakouts under the standard NAICS coding structure. Benchmarks for these series are developed from the QCEW data and independent estimates for these series are made on a monthly basis and raked to the estimates produced under the standard structure to ensure that the sum of the residential specialty trade contractors and nonresidential specialty trade contractors series is consistent with the published total for specialty trade contractors at the 3-digit NAICS level.

The raking adjustment follows the following methodology:

Estimates are derived independently for the residential and nonresidential groups at the 4-digit NAICS level for each region. Regional estimates are rounded and summed to the 4-digit NAICS level for both the residential and nonresidential groups. Within each 4-digit NAICS series, ratios of residential-to-total employment and nonresidential-to-total employment are calculated.

At the 4-digit NAICS level, the sum of the residential/nonresidential series is subtracted from the official industry-region cell structure total to determine the amount that must

be raked. The total amount that must be raked is multiplied by the ratios to determine what percentage of the raked amount should be applied to the residential group and what percentage should be applied to the nonresidential group.

Once the residential and nonresidential groups receive their proportional amount of raked employment, the two groups are aggregated again to the 4-digit NAICS level. At this point, employment is equal to the 4-digit NAICS total derived from the official industry-region cell structure. This raking process also forces additivity at the 3-digit NAICS level.

No estimates of women employees, construction employees, or hours and earnings are made for the residential and nonresidential series.

Real earnings data are expressed in constant 1982–84 dollars. Real earnings are computed by dividing average hourly earnings and average weekly earnings by the BLS Consumer Price Indexes. Real earnings for all employees are deflated by the Consumer Price Index for All Urban Consumers (CPI-U), while real earnings for production and nonsupervisory employees are deflated by the Consumer Price Index for Urban Wage Earners and Clerical Workers (CPI-W).

Average hourly earnings, excluding overtime-premium pay, are computed by dividing the total worker payroll for the industry group by the sum of total worker hours and one-half of total overtime hours. No adjustments are made for other premium payment provisions, such as holiday pay, late-shift premiums, and overtime rates other than time and one-half. Average hourly earnings excluding overtime are calculated only for manufacturing industries.

Seasonally adjusted series

Many economic statistics reflect a regularly recurring seasonal movement that can be measured from past experience. By eliminating that part of the change attributable to the normal seasonal variation, it is possible to observe the cyclical and other non-seasonal movements in a series. Seasonally adjusted series are published regularly for selected employment, hours, and earnings series.

X-12 ARIMA software, developed by the U.S. Census Bureau, is used to seasonally adjust CES data on a concurrent basis, meaning the software incorporates estimates up through and including the current month's data to achieve the best possible series. Using special features of X-12 ARIMA, adjustments are made to remove the effect of the variable number of weeks between surveys from month to month (about 1 month in 3 has a 5-week instead of a 4-week interval) and to remove the effect of the variable number of work days in the reference month, to adjust for moving holidays, and to adjust for the variations in the number of election poll workers in November from year to year. CES processes concurrent seasonal adjustment on a monthly basis using the latest estimates of employment, hours, and earnings. Seasonally adjusted employment series for broader industry groups are obtained by summing the seasonally adjusted data for the component industries. Seasonally adjusted hours and earnings averages for broader level industry groups are weighted averages of the seasonally adjusted component series. For more information on seasonal adjustment of CES series, see **http://www.bls.gov/ces/cesseasadj.htm**.

Data Presentation

The national series on employment, hours, and earnings are available on the Internet and appear in several BLS publications. The summary data are first published each month in *The Employment Situation* news release (**http://www.bls.gov/ces/#news**), which contains preliminary national estimates of nonfarm employment, average weekly hours, and average hourly and weekly earnings in the preceding month for industry sectors. Preliminary estimates are based on tabulations of data for less than the full sample to permit early release of these widely used economic indicators. This release is normally issued on Friday, 3 weeks after the reference week. The news release also includes a brief analysis of current trends in employment, hours, and earnings. The *Real Earnings* news release is published concurrent with the *Consumer Price Index* news release.

Detailed employment, hours, and earnings data also are available on the Internet on the morning of The *Employment Situation* news release. Data can be accessed through the CES database section on the CES homepage (**http://www.bls.gov/ces/#data**). Where sample adequacy and response rates allow, estimates at the NAICS 4-, 5-, and 6-digit detail are published on the Internet on a 1-month lag. Final (prebenchmark) figures are issued 1 month later. In addition, special articles describe technical developments in the program. The *Monthly Labor Review* also presents CES data in articles analyzing industry employment, hours, and earnings trends.

National data also are disseminated in the publications or online databases of other Federal agencies, such as the U.S. Department of Commerce, the Board of Governors of the Federal Reserve System, and the Council of Economic Advisers. Data also are regularly republished in summary form or for specific industries in many trade association journals, the labor press, and in general reference works.

In addition to national estimates, monthly employment estimates for all 50 States, the District of Columbia, Puerto Rico, the U.S. Virgin Islands, and selected metropolitan areas are published in the online *Employment and Earnings*.[1]

Detailed State and metropolitan area industry data also are available monthly in releases published by the State employment security agencies that cooperate with BLS in producing the State and area estimates. State and area data also are available from the State and area current employment statistics homepage **http://www.bls.gov/sae/**, which contains extensive information related to the CES State and area program, including contacts, news releases, and data.

[1] Employment, hours, and earnings for the Nation as a whole exclude Puerto Rico and the U.S. Virgin Islands.

Comparison with the Current Population Survey

The Bureau of Labor Statistics (BLS) has two monthly surveys that measure employment levels and trends: the Current Employment Statistics (CES) survey, also known as the payroll or establishment survey, and the Current Population Survey (CPS), also known as the household survey. Employment estimates from both surveys are published in *The Employment Situation* news release each month. The estimates differ because the surveys have distinct definitions of employment and distinct survey and estimation methods.

The Current Employment Statistics survey, also known as the payroll survey, excludes unpaid family workers, domestic workers in private homes, agricultural workers, proprietors, and other self-employed persons, all of whom are covered by the CPS. Moreover, the payroll survey counts a person who is employed by two or more establishments at each place of employment, but the household survey counts a person only once, and classifies the individual according to the major activity. Certain persons on unpaid leave for the entire reference period are counted as employed under the household survey but are not included in the employment count derived from the payroll survey.

The household survey emphasizes the employment status of individuals and provides information on the demographic characteristics (sex, age, and race) of the labor force. The survey is not well suited to furnishing detailed information on the industrial and geographic distribution of employment. The establishment survey provides limited information on personal characteristics of workers; however, it is an excellent source for detailed industrial and geographic data. In addition, it provides hours and earnings information that relates directly to the employment figures. The payroll and household surveys thus complement each other.

To better understand differences in the surveys' employment measures, as well as divergences that sometimes occur in their trends, see **http://www.bls.gov/web/ces_cps_trends.pdf**. Additional information on the methodologies of the two surveys can be found in the *Quick Guide to Methods and Measurement Issues* on the BLS Web site at **http://www.bls.gov/bls/empsitquickguide.htm**.

Uses

Data from the CES program, along with CPS data, are the first major economic indicators released each month. As such, they are used in the formulation of fiscal and economic policy. CES employment estimates are a primary component of the Index of Coincident Economic Indicators and have proved to be an extremely reliable measure of current economic activity. The manufacturing average weekly hours series is used in the Index of Leading Economic Indicators (LEI), which forecasts changes in the business cycle.

Aggregate earnings data are the major component of the preliminary personal income estimates in the National Income and Product Accounts. Productivity measures (chapters 10 and 11) and the Industrial Production Index are based on the aggregate hours data. Employment series are a basic input for employment projections by BLS (chapter 13) and State labor market information agencies.

The series also are used in the private sector by business firms, labor unions, universities, trade associations, and private research organizations to study economic conditions and to develop plans for the future. Business firms, for example, use the employment, hours, and earnings data for guidance in plant location, sales, and purchases.

Reliability of Estimates

The establishment survey, like other sample surveys, is subject to two types of error, sampling and non sampling error. The magnitude of sampling error, or variance, is directly related to the size of the sample and the percentage of universe coverage achieved by the sample. The establishment survey sample covers over one-third of total universe employment; this yields a very small variance on the total nonfarm estimates.

Most sample surveys publish sampling error as their only measure of error; however, the CES can derive an annual approximation of total error, on a lagged basis, because of the availability of the independently derived universe data. While the benchmark error is used as a measure of total error for the CES survey estimate, it actually represents the difference between two independent estimates derived from separate survey processes (specifically, the CES sample process and the UI administrative process) and thus reflects the errors present in each program. Historically, benchmark revisions have been very small for total nonfarm employment, ranging from -0.7 to +0.6 percent.

The estimation of sample variance for the CES survey is accomplished through use of the method of Balanced Half Samples (BHS). This replication technique uses half samples of the original sample and calculates estimates using those subsamples. The CES survey uses a modification to the basic BHS method known as Fay's method. Rather than using only half of the sample in deriving each replicate estimate, this method uses adjustments to the original sample weights applied to both halves of the sample, thus allowing use of all sample units for each replicate estimate. The sample variance is calculated by measuring the variability of the replicate estimates. The sample units in each sampling strata are divided into two random groups. Columns of the Hadamard matrix (which is a special 0-1 matrix) of appropriate order are mapped to the strata. Each row of the Hadamard matrix defines a replicate subsample: the random group indicators are matched to the 0-1 entries of each row of the Hadamard matrix thus defining a set of units for each replicate. Weights for units that belong to a replicate half-sample are multiplied by a factor of $1+y$, where weights for units in the other half of the sample are multiplied by a factor of $1-y$. Replicate estimates are calculated using the same estimation formula as used for the full-sample estimate.

The formula used to calculate CES variances

$$v_{\kappa}^{+}\left(\hat{\theta}\right)=\frac{1}{\gamma^{2}\kappa}\sum_{\alpha=1}^{\kappa}\left(\hat{\theta}_{\alpha}^{+}-\hat{\theta}\right)^{2}$$

where

$\hat{\theta}_{\alpha}^{+}=\theta\left(\hat{Y}_{\alpha}^{+},\hat{X}_{\alpha}^{+},.....\right)$ is the α th replicate estimate,

$\gamma=1/2$,

k is the number of replicates, and, $\hat{\theta}$ is the original full–sample estimates.

Variances statistics are useful for comparison purposes, but they do have some limitations. Variances reflect the error component of the estimates that is due to surveying only a subset of the population, rather than conducting a complete count of the entire population. However, they do not reflect the non-sampling error, such as response errors and bias due to non response. The overall performance of the CES employment estimates is best measured in terms of the benchmark revisions. The variances of the over-the-month change estimates are very useful in determining when changes are significant at some level of confidence.

Technical References

Employment from the BLS household and payroll surveys: summary of recent trends. Describes differences in the employment measures from the CES and CPS surveys, as well as divergences that sometimes occur in their trends. (See **http://www.bls.gov/web/ces_cps_trends.pdf**.)

Seasonal Adjustment in the Current Employment Statistics Program. Describes in detail the seasonal adjustment methodology and software employed by the Current Employment Statistics Program. It is important to note that this describes seasonal adjustment only as it relates to the CES program's implementation. (See **http://www.bls.gov/ces/cessa_oview.pdf**.)

Technical Notes To Establishment Data Published in Employment and Earnings. An up-to-date, concise description of concepts and methods used to develop establishment-based employment, hours, and earnings data from the Current Employment Statistics program. Provides tables that present measures of the reliability of the data and the magnitude of revisions due to benchmark adjustments. (See **http://www.bls.gov/ces/#technical**.)

Technical Information: Estimation Methods for Business Births and Deaths. (See **http://www.bls.gov/ces/cesbdtech.htm**.)

Mueller, Kirk, "Impact of business births and death in the payroll survey," May 2006, *Monthly Labor Review*.

Kokic, P. N., and Bell, P. A. (1994), "Optimal Winsorizing Cutoffs for a Stratified Finite Population Estimator," *Journal of Official Statistics*, 10, 419-435.

Fay, R.E. and Herriot, (1979). Estimates of Income for Small Places: an Application of James-Stein Procedure to Census Data, *Journal of the American Statistical Association*, 74, 269-277.