Monthly Labor Review
November 2017

A quarterly benchmarking procedure for the Current Employment Statistics program

We propose an improved procedure for benchmarking employment estimates from the Current Employment Statistics (CES) survey. The procedure involves more frequent updating of data, whereby seasonally adjusted CES estimates are benchmarked to seasonally adjusted Quarterly Census of Employment and Wages estimates on a quarterly, rather than annual, basis. We provide simulation results illustrating the advantages of the method, which can be used for benchmarking both national estimates and state and area estimates.

Mark A. Loewenstein
loewenstein.mark@bls.gov

Mark A. Loewenstein is a senior research economist in the Office of Employment and Unemployment Statistics, U.S. Bureau of Labor Statistics.

Matthew Dey
dey.matthew@bls.gov

Matthew Dey is a research economist in the Office of Employment and Unemployment Statistics, U.S. Bureau of Labor Statistics.

The U.S. Bureau of Labor Statistics (BLS) Current Employment Statistics (CES) survey is a quick-response business survey that provides information on employment, hours, and earnings each month.[1] The CES program benchmarks its employee series in order to re-anchor sample-based employment estimates to full population counts. This process is designed to improve the accuracy of the CES all-employee series by replacing estimates with full population counts that are not subject to the sampling or modeling errors inherent in the CES monthly estimates. These population counts, derived from administrative records, are much less timely than the sample-based estimates, but provide a near census of establishment employment. The CES program is examining possible improvements in its benchmarking procedures.

The major 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. All businesses subject to UI laws are required to report, on a quarterly basis, employment and wage information to the appropriate state workforce agency. UI records cover about 97 percent of nonfarm wage and salary jobs on civilian payrolls. Benchmark data for the remaining 3 percent of CES employment are constructed from alternative sources, primarily the County Business Patterns series and the Annual Survey of Public Employment and Payroll publications of the U.S. Census Bureau. Benchmark employment, also called population employment, is the sum of the QCEW employment count and the noncovered employment estimate from these other sources. In the rest of this article, the term QCEW denotes QCEW plus noncovered employment.

The size of the benchmark, or March, revision is equal to the difference between the QCEW and CES estimates of March employment. It is widely regarded as a measure of the accuracy of CES estimates. For the national total nonfarm series, absolute annual benchmark revisions averaged about 0.3 percent over the past decade. For national series, only March sample-based estimates are replaced with population data. In contrast, for state and metropolitan area series, all available months of population data are used to replace sample-based estimates.

BLS is exploring ways to improve the benchmarking procedures. The discussion in this article is an outgrowth of the work of the CES benchmarking team.[2] Two improvements to the benchmarking procedures are being considered. First, the CES program is looking into the possibility of benchmarking the CES quarterly instead of annually. More frequent benchmarking will provide more timely revisions and generally reduce the size of the March revisions. Second, the program is exploring the possibility of adopting the same benchmarking procedure for state and metropolitan area estimates as that used for national estimates. This will achieve greater consistency between national estimates and state and metropolitan area estimates. Although the empirical work and simulation exercise presented in this article are confined to the national all-employee series, the analysis has broader applicability. The theoretical results and the results of the simulation exercise apply equally to any series, be it total or industry, or national or local.

Benchmarking the nonseasonally adjusted CES series to the nonseasonally adjusted QCEW series throughout the year is problematic because of the substantial difference in the seasonal patterns of those series.[3] The QCEW has always shown larger seasonal movements than the CES, both before and after the CES conversion to a probability sample design in 2003. Consequently, the benchmarking team is proposing to benchmark *seasonally adjusted* CES estimates to *seasonally adjusted* QCEW estimates.

This article is organized as follows. In the next section, we describe the benchmarking procedure currently used for the national series. We then discuss how more frequent benchmarking can affect the March revision. With this information as background, we present the proposed quarterly benchmarking procedure and examine results from its implementation for the total private CES employment estimates.[4] We conclude with a simulation exercise to evaluate our proposed procedure when there are potential errors in both the CES and the QCEW and when seasonal factors are measured with error.

Current benchmarking procedure

Newly benchmarked CES national estimates are released with the January Employment Situation in early February of each year. Nonseasonally adjusted data are revised for 21 months, and seasonally adjusted data are revised for 5 years. For example, with the March 2012 benchmark release, nonseasonally adjusted data were revised from April 2011 through December 2012, and seasonally adjusted data were revised from January 2008 through December 2012.

Like all sample surveys, the CES survey is susceptible to two sources of error: sampling error and nonsampling error. Sampling error is present any time a sample is used to make inferences about a population. The magnitude of the sampling error, or the variance, relates directly to sample size and the percentage of the universe covered by the sample. The CES survey captures slightly under one-third of the universe employment each month, which is exceptionally high by usual sampling standards. This coverage ensures a relatively small sampling error at the total nonfarm employment level for the statewide and national series. Both the universe counts and the CES estimates are subject to nonsampling errors common to all censuses and surveys—coverage, response, nonresponse, and processing errors. The error structures for both the CES monthly survey and the UI universe are complex. Still, the two programs generally produce consistent total employment figures. Over the last decade, annual benchmark revisions at the national total nonfarm level have averaged 0.3 percent (in absolute terms), with an absolute range of 0.1 percent to 0.7 percent over the past decade.

While the benchmark revision is often regarded as a proxy for total survey error, this interpretation does not consider error in the benchmark source data. The employment counts obtained from quarterly UI tax forms are administrative data that reflect employer recordkeeping practices and differing state laws and procedures. The benchmark revision can be more precisely interpreted as the difference between two independently derived employment counts, each subject to its own error sources. Overall, however, the universe employment counts are subject to less error than the CES sample-based estimates and therefore serve as a valuable input data to improve the accuracy of the CES through benchmarking.

At the time of annual benchmarking, the monthly sample-based estimates for the 11 months preceding and the 9 months following the March benchmark are also subject to revision. Each annual benchmark revision affects 21 months of data for nonseasonally adjusted series.[5]

Monthly estimates for the 11 months preceding the March benchmark are recalculated with the use of a “wedge-back” procedure. In this procedure, the difference between the final benchmark level and the previously published March sample estimate is calculated and distributed 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 method assumes that the total estimation error (in levels) since the last benchmark accumulated at a steady rate throughout the benchmark reference year.

Estimates for the 9 months following the March benchmark also are recalculated each year.[6] These postbenchmark estimates reflect the application of final sample-based monthly changes to new benchmark levels for March. The sample changes are the ones calculated and used for the previously published final sample-based estimates.

Recall that we have no knowledge about the individual monthly errors, but we do have information about the total error from the size of the March revision. The natural choice of an error-correction model would be to assume that the original rates are all mismeasured by different amounts, but that the implied errors in monthly employment sum to the March discrepancy. In a typical year, one would expect monthly errors to be roughly independent of each other and to have roughly constant variance. In years with a large March revision (typically at turning points in the cycle), the errors are likely to be correlated, possibly resulting from errors in the estimate of the birth–death factor.[7]

If one uses only the March QCEW, there is no alternative to wedging back for an entire year. However, if one benchmarks more frequently, using the QCEW at other times, then one will need to wedge back for a shorter period of time, mitigating the issue.

The effect of quarterly benchmarking on March estimates and March revisions

Although the March employment revision does not provide information about individual monthly errors, the March estimate does provide a natural diagnostic tool for evaluating the effect of more frequent updating. To facilitate the discussion, it is helpful to introduce some notation. Letting $E^{\text{QCEW}}_{\text{March},t}$ denote the March employment count from the QCEW in year t and letting $\hat{E}^{\text{CES}}_{\text{March},t}$ denote the initial CES estimate of March employment in year t, we can express the annual March revision as

(1) $$R_{\text{March},t} = E^{\text{QCEW}}_{\text{March},t} - \hat{E}^{\text{CES}}_{\text{March},t}.$$

Letting $\hat{r}_m$ denote the CES estimate of the employment growth rate in month m and letting r_m be the true employment growth rate in month m,[8] we can rewrite equation (1) as

(2) $$R_{\text{March},t} = E^{\text{QCEW}}_{\text{March},t} - E^{\text{QCEW}}_{\text{March},t-1} \times \prod_{m=\text{April},t-1}^{\text{March},t} \hat{r}_m$$

$$= E^{\text{QCEW}}_{\text{March},t} - E^{\text{QCEW}}_{\text{March},t-1} \times \prod_{m=\text{April},t-1}^{\text{March},t} r_m \times \prod_{m=\text{April},t-1}^{\text{March},t} \left(\frac{\hat{r}_m}{r_m}\right)$$

$$= E^{\text{QCEW}}_{\text{March},t}\left(1 - \prod_{m=\text{April},t-1}^{\text{March},t}\left(\frac{\hat{r}_m}{r_m}\right)\right).$$

In interpreting equation (2), note that if $\frac{\hat{r}_m}{r_m} = 1$, the estimated CES growth rate in month m contains no error. If $\frac{\hat{r}_m}{r_m} > 1$, the CES overstates employment growth in month m, and if $\frac{\hat{r}_m}{r_m} < 1$, the CES underestimates that growth. Equation (2) makes explicit the relationship between the monthly errors and the March revision—namely, the March revision depends on the *cumulative* (over-the-year) monthly errors in the CES rates, and the closer the cumulative error is to zero, the smaller the March revision. The March revision tells us very little about the individual monthly errors in the CES rates. Indeed, one would hope that monthly errors largely average out, leaving us with a relatively small revision. A large March revision likely means that monthly errors in a given year are highly correlated. An example of this was the start of the Great Recession (2007–09), when CES underestimated the sizes of the early employment losses. Throughout this article, we focus on quarterly updating.[9]

At present, there is a 10-month lag between the QCEW reference period and the time that the QCEW employment figures are available for benchmarking. Consequently, CES estimates are currently revised every January. With quarterly benchmarking, CES estimates would also be revised about 10 months after each quarter, using the June, September, and December QCEW numbers.[10]

For convenience, in the analysis that follows, we work with the proportional error ε_m in the CES estimated rate of change in month m employment. This error may be expressed as $\exp(\varepsilon_m) = \frac{\hat{r}_m}{r_m}$ and substituted in equation (2):[11]

(2') $$R_{\text{March},t} = E^{\text{QCEW}}_{\text{March},t-1} \times \left(1 - \prod_{m=\text{April},t-1}^{\text{March},t} \exp(\varepsilon_m)\right).$$

The initial CES estimate of March employment in year *t* consists of two parts: the March QCEW employment level in year $t-1$ and the product of CES monthly growth-rate estimates from April of year $t-1$ to March of year t. Using a similar derivation as that used earlier to obtain equation (2), one can show that this estimate is equal to

$$(3) \qquad \hat{E}^{\text{CES}}_{\text{March},t} = E^{\text{QCEW}}_{\text{March},t} \times \exp\left(\varepsilon_{2,t-1} + \varepsilon_{3,t-1} + \varepsilon_{4,t-1} + \varepsilon_{1,t}\right),$$

where $\varepsilon_{i,t}$ is the proportional error in the CES estimate in quarter i of year t. The estimate is again composed of two parts: the March QCEW employment level in year t and the product of the quarterly errors in the CES rates from the second quarter of year $t-1$ to the first quarter of year t.

Let us assume for the moment that the June QCEW employment level in year $t-1$ is the true population value.[12] Incorporating QCEW information for the second quarter (June) of year $t-1$ into the CES estimate when it first becomes available (in April of year t) results in a revised March estimate that can be written as

$$(4) \qquad \hat{E}^{\text{CES}(2)}_{\text{March},t} = E^{\text{QCEW}}_{\text{March},t} \times \exp\left(\varepsilon_{3,t-1} + \varepsilon_{4,t-1} + \varepsilon_{1,t}\right).$$

Essentially, including the June QCEW information in year $t-1$ into the March estimate in year t eliminates the errors in the CES monthly rates from April to June of year $t-1$. The variance of the revised estimate, $\hat{E}^{\text{CES}(2)}_{\text{March},t}$, around the true population value is therefore lower than the variance of the initial estimate, $\hat{E}^{\text{CES}}_{\text{March},t}$, around the true population value.[13] The next revision in July will yield a still better estimate and, of course, the same will be true of the October revision.

Using the March estimate as a convenient benchmark, we have shown that more frequent updating yields improved employment estimates. But what can we say about the revisions themselves? Let $R^{(1)}_{\text{March},t}$ denote the first March revision in April, and let $R^{(2)}_{\text{March},t}$, $R^{(3)}_{\text{March},t}$, and $R^{(4)}_{\text{March},t}$ denote, respectively, the second, third, and fourth March revisions. It is straightforward to show that

$$(5) \qquad R^{(1)}_{\text{March},t} + R^{(2)}_{\text{March},t} + R^{(3)}_{\text{March},t} + R^{(4)}_{\text{March},t} = R_{\text{March},t}.$$

Equation (5) reflects that the four quarterly benchmark revisions sum to the benchmark revision when we benchmark only annually. One can also show that

$$(6) \qquad \left|R^{(1)}_{\text{March},t}\right| + \left|R^{(2)}_{\text{March},t}\right| + \left|R^{(3)}_{\text{March},t}\right| + \left|R^{(4)}_{\text{March},t}\right| \geq \left|R_{\text{March},t}\right|,$$

with equality holding if and only if the quarterly benchmark revisions are all positive or all negative. When CES errors are opposite signed in different quarters, the revision in one quarter will at least partially offset that in another quarter. This leads to a seeming puzzle. Earlier, we saw that each revision yielded a better estimate of March employment than the previous one. If this is so, why should the revisions be offsetting, with a positive revision in one quarter being followed by a negative revision in a subsequent quarter or vice versa?

A positive revision in, say, April of year t means that the CES initially underestimated employment growth from April to June in year $t-1$. Conversely, a negative revision in, say, July of year t means that the CES initially overerestimated employment growth from July to September in year $t-1$. Occurrences like these should be common if errors in the CES are uncorrelated over time. When the initial errors in different quarters are in different directions, the revisions (which will be opposite signed from the errors) will also be. Actually, having quarterly revisions that are not all of the same sign should be reassuring, because this means that the CES program is not making systematic errors. When the CES errors are all in the same direction, so will be the revisions. Quarterly benchmarking is especially valuable in such cases, because it allows us to begin correcting the systematic errors more quickly than annual benchmarking would.[14]

In sum, offsetting quarterly revisions should not be a reason for concern. They suggest that the CES program is not making systematic errors. In those hopefully rare occasions when errors are correlated (which tend to occur at turning points in the business cycle), quarterly benchmarking provides an important safeguard against systematic errors being baked into the estimates for an unnecessarily long period.

A benchmarking methodology using seasonally adjusted CES and QCEW data

It is well documented that the QCEW and CES series have different seasonal patterns.[15] These seasonal patterns are, of course, not an issue if one benchmarks annually. However, they must be accounted for if one benchmarks more frequently. BLS has examined several methods for dealing with the different seasonal patterns in the QCEW. It has determined that it is best to explicitly model and estimate the seasonal patterns in CES and QCEW data, because methods that implicitly adjust for seasonality have been found lacking.

The proposed procedure is simple. We assume that the ratio of the seasonally adjusted quarterly growth rate from the QCEW to the seasonally adjusted quarterly growth rate from the CES is an unbiased, albeit noisy, signal of the over-the-quarter error term. When a new quarter of QCEW data becomes available, the proposed method adjusts monthly rates in that quarter by a constant determined by the ratio of seasonally adjusted quarterly growth rates. Mathematically, the adjusted monthly growth rate can be expressed as

(7) $$\tilde{r}_{m(q)} = \hat{r}_{m(q)} \times \exp\left(\frac{1}{3}\ln\left(\frac{\hat{r}_q^{\text{SA-QCEW}}}{\hat{r}_q^{\text{SA-CES}}}\right)\right),$$

where $\hat{r}_{m(q)}$ denotes the original CES estimate of the rate of change in employment in month m of quarter q, $\tilde{r}_{m(q)}$ denotes the benchmarked CES estimate of the rate of change in employment in month m of quarter q, $\hat{r}_q^{\text{SA-QCEW}}$ denotes the seasonally adjusted rate of change in QCEW employment in quarter q, and $\hat{r}_q^{\text{SA-CES}}$ denotes the seasonally adjusted CES estimate of the rate of change in employment in quarter q.[16]

As an example, suppose that the estimated CES employment growth rate is 0.997, 1.0025, and 1.01 in April 2014, May 2014, and June 2014, respectively. In addition, suppose that the seasonally adjusted CES growth rate from April to June 2014 is 1.01 and that the seasonally adjusted QCEW growth rate over the same period is 1.007. Because the seasonally adjusted CES growth rate from April to June 2014 exceeds the seasonally adjusted QCEW growth rate, our benchmark procedure will lower the estimated employment growth throughout the quarter.

Specifically, since $\exp\left(\frac{1}{3}\ln\left(\frac{\hat{r}_q^{\text{SA-QCEW}}}{\hat{r}_q^{\text{SA-CES}}}\right)\right) = .999$, the benchmarked growth rate for April 2014 is given by .997 × .999 = .996. Similarly, the benchmarked growth rates for May and June 2014 are, respectively, 1.0025 × .999 = 1.0015 and 1.01 × .999 = 1.009. Therefore, the benchmarked growth rate over the quarter is given by .996 × 1.0015 × 1.009 = 1.0065.

Let r_q denote the true quarterly growth rate, and let γ_q denote the proportional error in the adjustment. This error has two possible sources, namely, error in the estimation of the seasonal factors or error in the QCEW itself.

Multiplying the monthly rates in equation (7) over the quarter, one can show that the benchmarked quarterly growth rate, $\tilde{r}_q$, is given by

(8) $$\tilde{r}_q = r_q \times \exp(\gamma_q).$$

Having calculated the adjusted rates for the months in quarter q, one can readily obtain revised employment growth estimates for quarter q. These estimates yield a revised estimate for end-of-quarter employment, which, in turn, leads to a revised employment level in subsequent months. For example, in May 2015, we would obtain corrected values for April 2014–March 2015 employment growth. The corrected values from April 2014 to June 2014 are obtained by applying the adjusted rates for the period April 2014–June 2014. The corrected values from July 2014 to March 2015 are obtained by applying the new June 2014 employment base, but using the initial CES growth rates. For example, the revised estimate for March 2015 is given by

(9) $$\tilde{E}_{\mathrm{March,2015}} = E^{\mathrm{QCEW}}_{\mathrm{March,2014}} \times \exp(\tilde{r}_{q=2,t=2014}) \prod_{m=\mathrm{July,2014}}^{\mathrm{March,2015}} \hat{r}_m.$$

Similarly, revised estimates could be published when QCEW estimates in September and December become available.

Continuing with our example, suppose that the QCEW employment estimate in March 2014 is 120 million. Also, suppose that the accumulated growth rate from July 2014 through March 2015 is 1.02 (that is, suppose that $\prod_{m=\mathrm{July,2014}}^{\mathrm{March,2015}} \hat{r}_m = 1.02$). Then the revised estimate in May 2015 for March 2015 employment is given by 120,000,000 × 1.0065 × 1.02 = 123,200,000.

In the February revision, as in the three preceding revisions, adjusted growth rates for the months of January, February, and March of the previous year are calculated according to equation (7). These adjusted growth rates in turn yield revised estimates for employment growth for the January–March period. However, the published revision in January also includes a second component arising from the fact that the adjusted quarterly growth rates will have some error. This component, given by the difference between actual March QCEW employment and the employment level predicted from the adjusted quarterly growth rates, reflects errors in the estimation of the seasonal factors.

Table 1 shows the revisions that result from applying the previously outlined method to revising estimates of total private employment for the period 2007–13.[17] Although the March revision does not tell us much about the accuracy of the individual monthly errors, it is nevertheless an informative statistic. Recall from our earlier discussion that a large revision based on annual benchmarking suggests that the monthly estimates are correlated over the past year and therefore do not average out to zero.

Table 1. Analysis of proposed employment estimates and revisions, March 2007–March 2013

Date	QCEW value	Original CES estimate	Revised estimate as of April	Revised estimate as of July	Revised estimate as of October	Revised estimate as of January (never published)
Employment estimates						
March 2007	108,250,800	109,124,676	109,179,508	108,977,936	108,893,811	108,365,917
March 2008	105,535,500	106,097,960	105,951,628	105,626,170	105,679,140	105,570,462

See footnotes at end of table.

Table 1. Analysis of proposed employment estimates and revisions, March 2007–March 2013

Date	QCEW value	Original CES estimate	Revised estimate as of April	Revised estimate as of July	Revised estimate as of October	Revised estimate as of January (never published)
March 2009	107,565,200	107,194,505	107,355,770	107,415,776	107,415,776	107,467,022
March 2010	110,009,400	109,777,090	109,604,541	110,096,767	109,795,585	110,155,060
March 2011	112,699,400	112,152,430	112,265,409	112,076,221	111,923,700	112,427,355
March 2012	114,907,500	114,940,754	115,001,394	114,969,354	114,991,067	114,855,962
March 2013	117,717,000	117,925,267	118,033,062	117,982,639	118,048,168	117,688,749
Signed difference between estimate and QCEW						
March 2007	—	873,876	928,708	727,136	643,011	115,117
March 2008	—	562,460	416,128	90,670	143,640	34,962
March 2009	—	-370,695	-209,430	-149,424	-149,424	-98,178
March 2010	—	-232,310	-404,859	87,367	-213,815	145,660
March 2011	—	-546,970	-433,991	-623,179	-775,700	-272,045
March 2012	—	33,254	93,894	61,854	83,567	-51,538
March 2013	—	208,267	316,062	265,639	331,168	-28,251

Note: QCEW = Quarterly Census of Employment and Wages; CES = Current Employment Statistics.

Source: U.S. Bureau of Labor Statistics.

Looking at the estimates in the bottom panel of table 1, we see that the estimate for March employment generally improves as it is revised throughout the year. The improvement is quite noticeable for some years, especially those in which the initial CES estimate differs substantially from the QCEW value. The estimates in the last column simply reflect the errors in the estimation of the seasonal factors. These estimates indicate how well the adjustment model is estimating March employment. In every year but one, the absolute error stemming from the estimation of the seasonal factors is smaller than the error in the initial CES estimate (compare the last and second columns of the table). At the very least, this smaller error suggests that the ratio of the seasonally adjusted QCEW quarterly growth rate to the seasonally adjusted CES quarterly growth rate contains some useful information regarding the nature of monthly errors in CES growth rates. This information is particularly useful when the monthly errors are correlated, as they seem to be from April 2008 to March 2009 and from April 2009 to March 2010. The large negative revision indicates that the CES may be overestimating the growth rate in those years, and it is noteworthy that the quarterly rate adjustments are always negative. This suggests that capitalizing on the information contained in non-March QCEW data may give an earlier (earlier than the next March) indication that the CES is systematically overestimating or underestimating the growth rates.

A simulation exercise

The results presented so far strongly indicate that the proposed benchmark procedure results in an improved March estimate. Although the revised quarterly estimates likely improve on the initial CES estimates, there is insufficient information to show this definitively (after all, if we knew the true quarterly estimates, we would not need the benchmarking procedure). To handle this issue, we now perform a simulation exercise.

The simulation exercise also addresses another important question. The success of the proposed benchmarking procedure depends crucially on how accurately we estimate the seasonal factors. As one obtains finer CES estimates by industry or area, the estimates will have a larger error component. This, in and of itself, makes the proposed benchmarking procedure more advantageous. However, there is an offsetting effect: the greater the error in the CES estimate, the greater the errors in the resulting estimates of the seasonal factors. At some point, will the errors in the estimates of the seasonal factors be sufficiently great that the quarterly benchmarking procedure actually results in errors that exceed those in the unadjusted CES estimates?

In laying out our simulation model, we let the "true" employment growth that we would like to measure with the CES be given by $r_{q,t}^{\mathrm{true-CES}} = r_{q,t} \times \exp(\delta_q^{\mathrm{CES}})$, where $\exp(\delta_q^{\mathrm{CES}})$ represents a seasonal factor. The CES estimate of employment growth in quarter q of year t is then given by

(10) $$\hat{r}_{q,t}^{\mathrm{CES}} = r_{q,t} \times \exp(\delta_q^{\mathrm{CES}}) \times \exp(\varepsilon_{q,t}^{\mathrm{CES}}),$$

where $\exp(\varepsilon_{q,t}^{\mathrm{CES}})$ is the error in the CES estimate of the growth rate.

We are seemingly implicitly assuming that the seasonal variation in the CES series represents true seasonal variation in the underlying employment. In reality, the seasonal factors could also reflect systematic seasonal errors in the CES. However, without additional information, one cannot distinguish empirically between true seasonal variation in the underlying employment series and seasonal variation that reflects systematically seasonal measurement error.

Similarly to the way we treat CES employment, we let the QCEW estimate of employment growth be given by

(11) $$\hat{r}_{q,t}^{\mathrm{QCEW}} = r_{q,t} \times \exp(\delta_q^{\mathrm{QCEW}}) \times \exp(\varepsilon_{q,t}^{\mathrm{QCEW}}).$$

Note that we are no longer assuming that the QCEW growth rate is measured without error. Rather, like the CES, the QCEW has an error with a random component $\varepsilon_{q,t}^{\mathrm{QCEW}}$. In addition, there is a QCEW seasonal factor that may partly reflect QCEW measurement error that has a systematic seasonal component.

In our simulations, we normalize $r_{q,t}$ to 1 (i.e., no employment change) for all four quarters q and all years t (t runs from 1 to 5). We set the seasonal factors to be the following:

(12) $$\exp(\delta_1^{\mathrm{CES}}) = 0.995,\ \exp(\delta_2^{\mathrm{CES}}) = 1.003,\ \exp(\delta_3^{\mathrm{CES}}) = 0.998,\ \exp(\delta_4^{\mathrm{CES}}) = 1.004$$

(13) $$\exp(\delta_1^{\mathrm{QCEW}}) = 0.990,\ \exp(\delta_2^{\mathrm{QCEW}}) = 1.004,\ \exp(\delta_3^{\mathrm{QCEW}}) = 0.997,\ \exp(\delta_4^{\mathrm{QCEW}}) = 1.009$$

Note from equations (12) and (13) that we have set the QCEW to be more seasonal than the CES.[18]

Finally, we assume that the random errors in the CES and the QCEW are normally distributed, with respective variances σ_{CES}^2 and σ_{QCEW}^2. Let p denote the ratio of these variances:

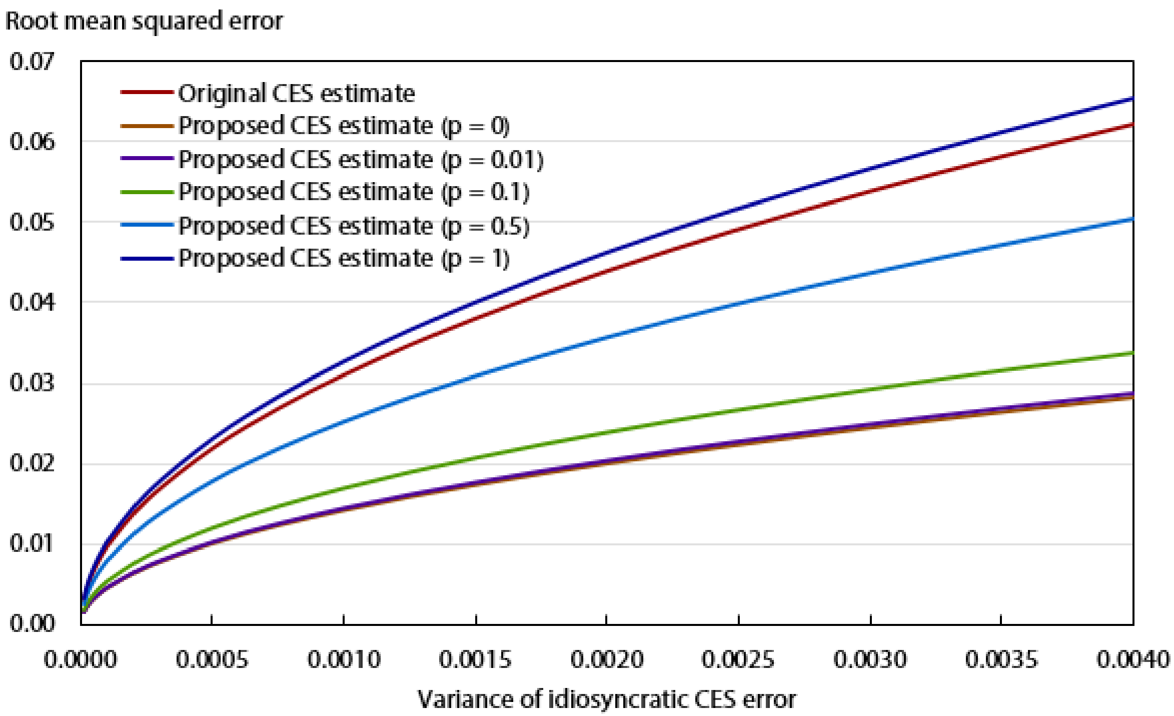
(14) $p = \sigma^2_{\mathrm{QCEW}} / \sigma^2_{\mathrm{CES}}.$

In the simulations to follow, we allow both p and σ^2_{CES} to vary. For a given value of σ^2_{CES}, a lower p is equivalent to a decrease in σ^2_{QCEW}. Thus, the lower is p, the more precise is the QCEW relative to the CES. We should expect the increased precision of the QCEW to lead to improved performance of the benchmark estimator relative to the initial CES estimator. For a given p, a higher value of σ^2_{CES} also means an increase in σ^2_{QCEW}. As a consequence, both the initial CES estimate and the benchmarked estimate will be less precise. There is another likely effect of an increase in σ^2_{CES} and σ^2_{QCEW}. The greater these variances are, the less precise the estimates of the seasonal factors are likely to be. The result is a less precise benchmarked estimate. The net effect on the performance of the benchmark estimator relative to the initial CES estimator is an open question. It is not obvious a priori whether, for a given value of p, a greater σ^2_{CES} is associated with improved or worsened performance of the benchmark estimator relative to the initial CES estimator.

For any given combination of p and σ^2_{CES}, we ran 1,000 simulations of both $\varepsilon^{\mathrm{CES}}_{q,t}$ and $\varepsilon^{\mathrm{QCEW}}_{q,t}$. These simulations yielded 1,000 estimates of the original CES estimate, $\hat{r}^{\mathrm{CES}}_{q,t}$, and the proposed CES estimate, $\tilde{r}^{\mathrm{CES}}_{q,t}$, for all quarters for 5 years. The X12 procedure in SAS was used to estimate seasonal factors.

Figures 1, 2, and 3 summarize our simulation results. First consider the curve labeled "Original CES estimate" in figure 1. Noting that the root mean squared error (RMSE) is graphed on the vertical axis and σ^2_{CES} on the horizontal, we see that, as expected, the RMSE of the CES estimate increases with σ^2_{CES}.

Figure 1. Performance measures of original and proposed CES estimators



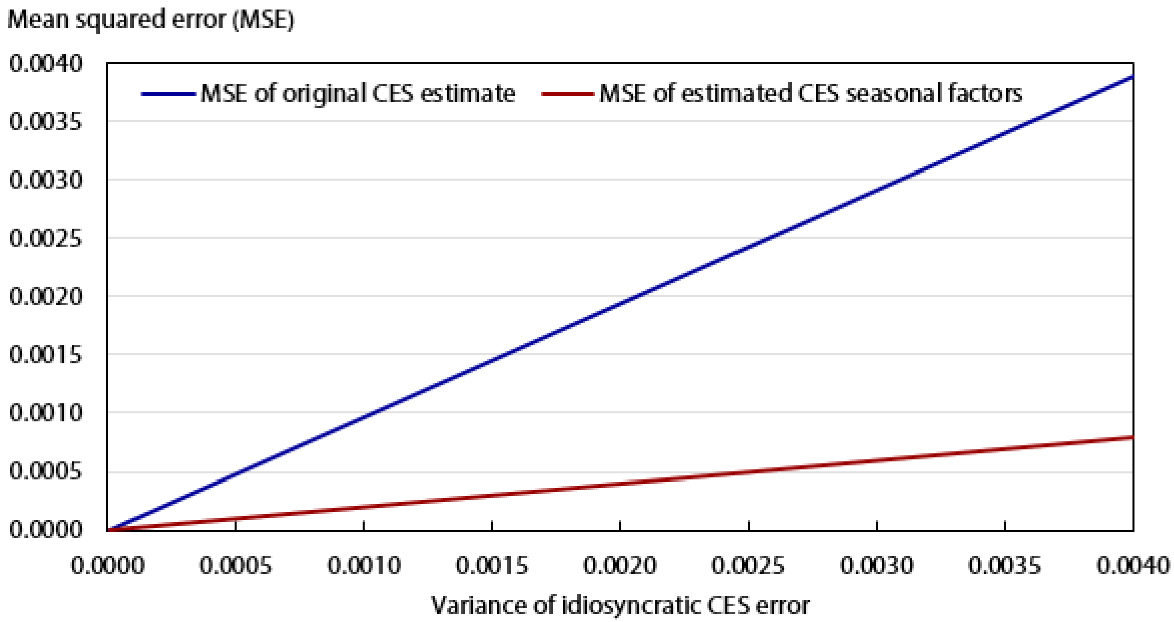
Note: CES = Current Employment Statistics.
Source: U.S. Bureau of Labor Statistics.

The remaining curves in figure 1 illustrate the performance of the proposed benchmark estimator. Each of these curves is associated with a different value of p. Rightward movement along any of the curves is associated with a higher RMSE. This is, of course, expected, because, for a given p, a higher σ^2_{CES} also means a higher σ^2_{QCEW}. Now consider the effect of variations in p, recalling that, for a given σ^2_{CES}, a lower p is associated with a lower value of σ^2_{QCEW}. The lowest curve in the figure is that corresponding to a value of p equal to 0. This curve lies well below that corresponding to the original CES estimate: when p (and therefore σ^2_{QCEW}) is 0, the RMSE of the proposed benchmark estimate is always smaller than that of the original CES estimate for any given value of σ^2_{CES}.

As p increases, the associated RMSE curve moves up, indicating a degradation in the performance of the proposed benchmark estimator. However, for plausible values of p, the proposed estimator still yields a substantial performance gain. This gain disappears only when p is close to 1. As indicated by the uppermost curve in the figure, when $p = 1$ (so that $\sigma^2_{\text{QCEW}} = \sigma^2_{\text{CES}}$), the RMSE of the benchmark estimate exceeds that of the original CES estimate. This reflects the fact that when $\sigma^2_{\text{QCEW}} = \sigma^2_{\text{CES}}$, the error in the benchmark estimate stemming from the imperfect estimation of the seasonal factors is not offset by an informational advantage of the QCEW estimate over the CES.

Figure 2 shows the effect of an increase in σ^2_{CES} on the mean squared error (MSE) of the CES estimate and the MSE of the estimated seasonal factors. As expected, the MSE of the CES estimate increases one for one with the increase in σ^2_{CES}. The MSE of the seasonal factors also increases with the increase in σ^2_{CES}, but at a much slower rate (note that the slope of the red curve is well below 1). The same is true for the relationship (not pictured) between the MSE of the QCEW seasonal factors and σ^2_{QCEW}. Recall from figure 1 that, for a plausible value of p, the RMSE of the proposed estimator is smaller than that of the simple CES estimator for any value of σ^2_{CES}. This is indicated by the fact that the RMSE curve for the proposed estimator lies entirely below that for the CES estimator. This result is reasonable in light of our finding in figure 2 that increases in σ^2_{CES} and σ^2_{QCEW} cause the MSEs in the estimated seasonal factors to increase, but at a slower rate.

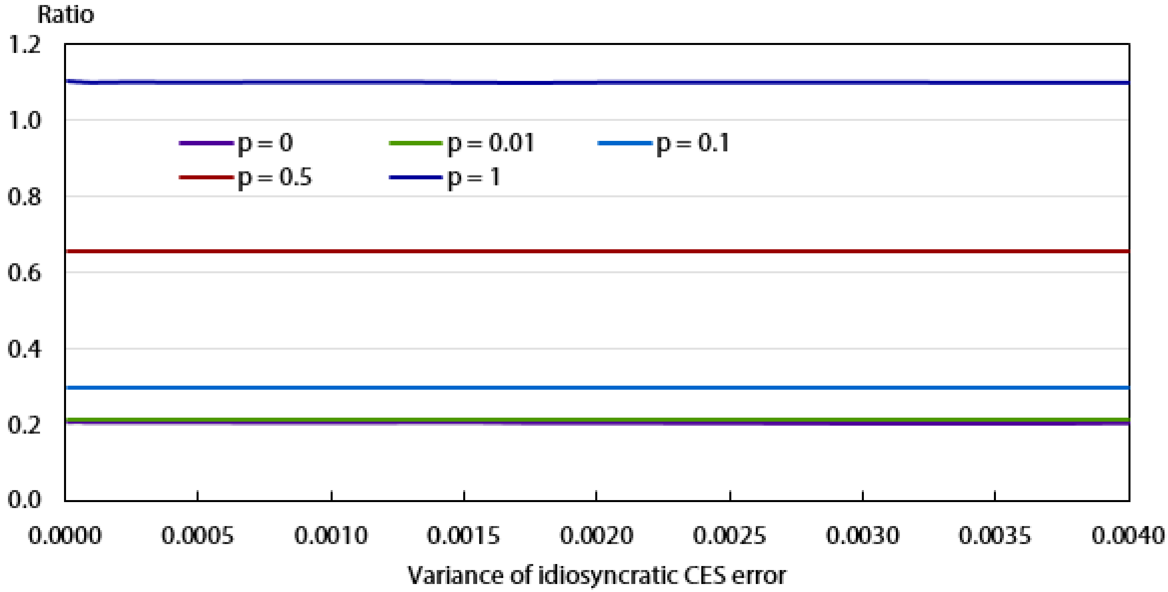
Figure 2. Performance of CES seasonal-factor estimates and original CES estimates



Note: CES = Current Employment Statistics.
Source: U.S. Bureau of Labor Statistics.

We get an even stronger result in figure 3, which shows the performance of the proposed estimator relative to the performance of the CES estimator. As in figures 1 and 2, we plot σ^2_{CES} on the horizontal axis. On the vertical axis, we now plot $\mathrm{MSE}\left(\tilde{r}^{\mathrm{CES}}_{q,t}\right)/\mathrm{MSE}\left(\hat{r}^{\mathrm{CES}}_{q,t}\right)$, the ratio of the MSE of the proposed estimator to the MSE of the CES. Each curve in the figure corresponds to a different value of p. As expected, the lower the value of p, the lower the corresponding curve in the figure. This relationship reflects that, for a given value of σ^2_{CES}, the performance gain of the benchmark estimator increases as σ^2_{QCEW} falls. Less predictable is the fact that the curves in figure 3 are all (nearly) straight lines with a slope equal to 0. This tells us that the ratio of the MSE of the proposed estimator to the MSE of the CES depends only on p and not on σ^2_{CES}. An increase in σ^2_{CES}, accompanied by an increase in σ^2_{QCEW} (so as to maintain a constant p), causes the MSEs of the proposed estimator and the CES estimator to increase by the same proportion. This means that the arithmetic difference between the two MSEs increases (as shown by the distance between the curves in figure 1).

Figure 3. Ratio of MSE of proposed estimate (varies by *p*) to MSE of original estimate

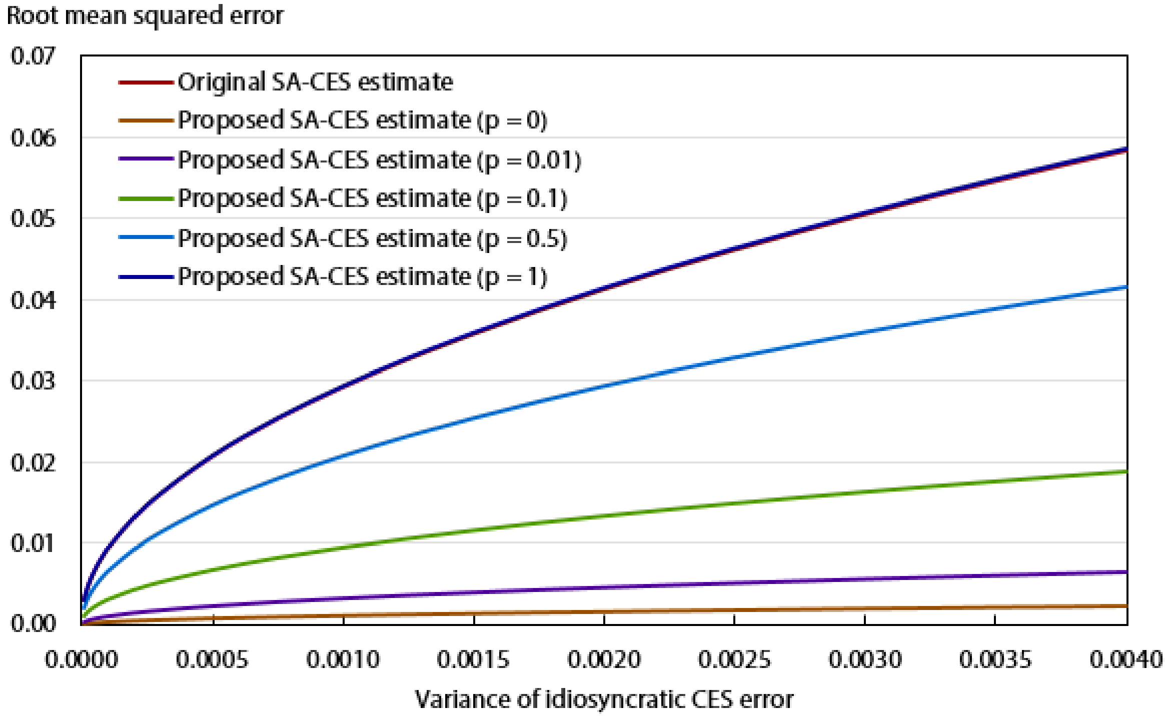


Note: MSE = Mean squared error; CES = Current Employment Statistics.
Source: U.S. Bureau of Labor Statistics.

It is worth explicitly tracing the effect of increasing σ^2_{CES} while holding σ^2_{QCEW} constant. When σ^2_{QCEW} is held constant, an increase in σ^2_{CES} implies a fall in p. In both figures 1 and 3, this would be represented by both a drop to a lower curve and a movement to the right. We therefore find that, for a given precision of the QCEW and other things the same, the less precise the CES estimator, the greater both the relative and absolute gains offered by the proposed benchmark estimator.

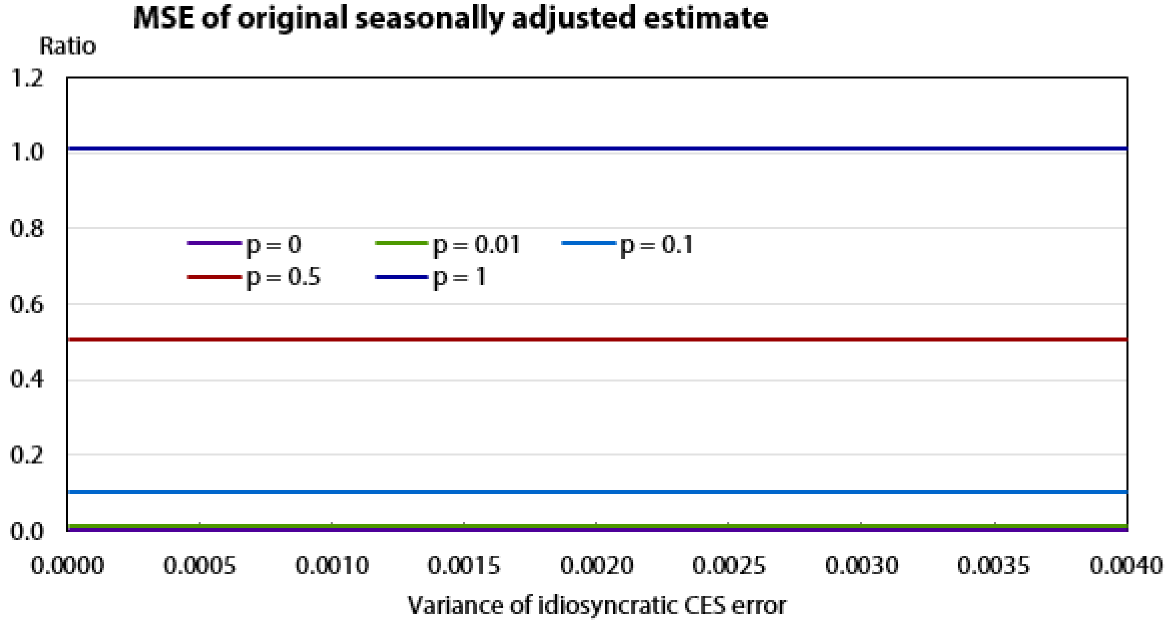
Seasonally adjusted estimates are often of greater analytical interest than nonseasonally adjusted estimates. Figures 4 and 5 summarize the performance of the proposed estimator when it is seasonally adjusted. The curves in figure 4 have the same general shape as those in figure 1, and the curves in figure 5 have the same shape as those in figure 3.

Figure 4. Performance measures of original and proposed SA-CES estimators



Note: SA-CES = Seasonally adjusted Current Employment Statistics.
Source: U.S. Bureau of Labor Statistics.

Figure 5. Ratio of MSE of proposed seasonally adjusted estimate (varies by *p*) to MSE of original seasonally adjusted estimate



Note: MSE = Mean squared error; CES = Current Employment Statistics.
Source: U.S. Bureau of Labor Statistics.

Conclusion

We propose replacing the annual benchmarking procedure currently in place for national estimates with one based on quarterly benchmarking of seasonally adjusted CES estimates to the seasonally adjusted QCEW. The proposed estimator performs well when applied to the national all-employee series. The gain from more frequent updating is especially large when monthly CES errors are positively correlated, as was the case at the beginning of the Great Recession.

The results of our simulation exercise apply equally to *any series*, be it total or industry, or national or local. We used the simulation exercise to compare the performance of the proposed quarterly benchmarking estimate with the initial CES estimate. The results demonstrate that, even when we control for a loss of precision in the estimation of seasonal factors, *the greater the variance of the CES estimate, the greater both the relative and absolute gains provided by the proposed quarterly benchmarking procedure*. The CES industry and area estimates have a greater variance than the national all-employee series. Therefore, one can reasonably argue that there is an even stronger case for applying the proposed quarterly benchmarking procedure to the industry and area estimates.

The discussion in this article has focused mostly on the national CES estimates. However, the methodology can also be applied to the state and area estimates. As noted in the introductory discussion, state and metropolitan area estimates are currently benchmarked annually by replacing sample-based estimates with all available months of population data. An undesirable feature of the resulting hybrid series is the confounding of QCEW and CES seasonality. Difficulties remain in seasonally adjusting the hybrid series.

As originally noted by Franklin D. Berger and Keith R. Phillips, it is best to seasonally adjust the QCEW and CES components of the hybrid series separately.[19] However, as discussed by others, a problem arises at the seam where the QCEW data end and the CES data begin, because differences in the seasonal factors at the seam will affect the growth rate of the hybrid series at the seam point.[20] An advantageous feature of our proposed methodology is that it produces a series that has CES seasonality throughout and can therefore be seasonally adjusted without undue complication.[21]

References

Berger, Franklin D., and Keith R. Phillips. “Reassessing Texas employment growth.” *The Southwest Economy* (Federal Reserve Bank of Dallas, July/August 1993), pp. 1–3.

Berger, Franklin D., and Keith R. Phillips. “Solving the mystery of the disappearing January blip in state employment data.” *Economic Review* (Federal Reserve Bank of Dallas, second quarter 1994).

Dey, Matthew, and Mark A. Loewenstein. “Quarterly benchmarking for the Current Employment Survey.” Working Paper 496 (U.S. Bureau of Labor Statistics, June 2017).

Groen, Jeffrey A. “Seasonal differences in employment between survey and administrative data.” Working Paper 443 (U.S. Bureau of Labor Statistics, 2011).

Manning, Christopher D., and John R. Stewart. “Benchmarking the Current Employment Statistics national estimates.” *Monthly Labor Review*, October 2017, https://doi.org/10.21916/mlr.2017.25.

Mueller, Kirk J. "Benchmarking the Current Employment Statistics state and area estimates." *Monthly Labor Review*, November 2017, https://doi.org/10.21916/mlr.2017.26.

Phillips, Keith R., and Jiango Wang. "Seasonal adjustment of state and metro CES jobs data." Working Paper 1505 (Federal Reserve Bank of Dallas, 2015).

Robertson, Kenneth. "Benchmarking the Current Employment Statistics survey: perspectives on current research." *Monthly Labor Review*, November 2017, https://doi.org/10.21916/mlr.2017.27.

Scott, Stuart, George Stamas, Thomas J. Sullivan, and Paul Chester. "Seasonal adjustment of hybrid economic time series." *Proceedings of the Section on Survey Research Methods* (American Statistical Association, 1994), https://www.bls.gov/osmr/research-papers/1994/pdf/st940350.pdf.

Walstrum, Thomas. "'Early benchmarking' the state payroll employment survey." *Employment, Midwest, Seventh District* (Federal Reserve Bank of Chicago, June 2015).

ACKNOWLEDGMENT: We thank Steve Mance, Steve Miller, Kirk Mueller, and Ken Robertson for providing helpful comments in the drafting of this article.

SUGGESTED CITATION

Mark A. Loewenstein and Matthew Dey, "A quarterly benchmarking procedure for the Current Employment Statistics program," *Monthly Labor Review,* U.S. Bureau of Labor Statistics, November 2017, https://doi.org/10.21916/mlr.2017.28.

NOTES

1 For more background on the CES, see Kenneth Robertson, "Benchmarking the Current Employment Statistics survey: perspectives on current research," *Monthly Labor Review*, November 2017, https://doi.org/10.21916/mlr.2017.27.

2 In addition to the authors of this article, the BLS benchmarking team includes Greg Erkens, Larry Huff, Christopher Manning, Kirk Mueller, Steven Mance, Kenneth Robertson, and David Talan.

3 For further discussion of this issue, see Robertson, "Benchmarking the Current Employment Statistics survey."

4 With the aim of making this article accessible to a larger audience, we have omitted a number of technical details and nearly all mathematical derivations. These can be found in our more technical working paper, which provides the basis for the present article. See Matthew Dey and Mark A. Loewenstein, "Quarterly benchmarking for the Current Employment Survey," Working Paper 496 (U.S. Bureau of Labor Statistics, June 2017).

5 For a more detailed description, see Christopher D. Manning and John R. Stewart, "Benchmarking the Current Employment Statistics national estimates," *Monthly Labor Review*, October 2017, https://doi.org/10.21916/mlr.2017.25.

6 New birth–death factors are also calculated with the use of the more recent data. For more information on the birth–death model, see "CES net birth/death model," *Current Employment Statistics—CES (national)* (U.S. Bureau of Labor Statistics), https://www.bls.gov/web/empsit/cesbd.htm.

7 The birth–death model corrects for business births and deaths that are not captured in the CES survey. For an explanation of the model, see "CES net birth/death model," *Current Employment Statistics—CES (national)* (U.S. Bureau of Labor Statistics), https://www.bls.gov/web/empsit/cesbd.htm.

8 More precisely, $\hat{r}_m$ is the ratio of the CES employment estimate in month m to the CES employment estimate in month $m-1$. When employment growth is positive, $\hat{r}_m$ is greater than 1. When employment growth is negative, $\hat{r}_m$ is less than 1. The definition of $\hat{r}_m$ is similar to that of all other growth rates presented in our analysis.

9 Not only does QCEW employment information become available on a quarterly basis, but the information in the last month of a quarter is more reliable than the information during the first 2 months of a quarter.

10 A preliminary QCEW employment estimate is available 3 months earlier. It is quite possible that little accuracy is lost in using the preliminary QCEW estimates rather than the final QCEW estimates. We plan to explore this possibility in future work.

11 Note that $\varepsilon_m > 0$ if the CES employment growth estimate is too high, and $\varepsilon_m < 0$ if the CES employment growth estimate is too low.

12 The assumption that true employment levels can be observed quarterly helps simplify the exposition in this section, allowing us to focus on essentials. We will return to this point later, when we discuss our recommended benchmarking methodology.

13 This statement is true if the quarterly CES errors are distributed independently or are positively correlated. With sufficiently strong negative correlation, it is possible for the variance of the revised estimate to actually exceed the variance of the initial estimate, but this case is extremely unlikely. In contrast, there have been times during which the quarterly CES errors have, in fact, been positively correlated. A prime example is the onset of the Great Recession, when CES understated employment losses.

14 The discussion abstracts from the possibility that errors in the CES are negatively correlated. If that were the case, there would be an even greater tendency for the sum of the absolute values of the quarterly March revisions to exceed the absolute value of the annual March revision. This situation is unlikely to occur in practice. Even if it does, it would not mean that the quarterly revisions do not improve the employment estimates throughout the year.

15 See, for example, Franklin D. Berger and Keith R. Phillips, “Reassessing Texas employment growth,” *The Southwest Economy* (Federal Reserve Bank of Dallas, July/August 1993), pp. 1–3; Berger and Phillips, “Solving the mystery of the disappearing January blip in state employment data,” *Economic Review* (Federal Reserve Bank of Dallas, second quarter 1994); and Jeffrey A. Groen, “Seasonal differences in employment between survey and administrative data,” Working Paper 443 (U.S. Bureau of Labor Statistics, 2011).

16 The proposed revision procedure uses only QCEW employment estimates for the final month of any quarter, because these estimates are more reliable than those for the first 2 months of any quarter. Evidence for this is provided by the existence of “seam effects” across quarters: monthly employment changes between the last month of a quarter and the first month of the next quarter are typically larger in absolute value and less likely to be zero than those between the months within a quarter. This suggests that employers filling out the UI reports underlying the QCEW have a tendency to copy backward from the third month of the quarter to the first and second months. See Groen, “Seasonal differences in employment between survey and administrative data.”

Note that $\tilde{r}_{(q)}$, the quarterly growth rate in the benchmarked estimate, is equal to the product of the three monthly growth rates: $\tilde{r}_{(q)} = \tilde{r}_{1(q)}\tilde{r}_{2(q)}\tilde{r}_{3(q)}$. Performing this multiplication with the use of equation (7), one obtains $\tilde{r}_{(q)} = \hat{r}_{(q)} \times \frac{\hat{r}_q^{SA-QCEW}}{\hat{r}_q^{SA-CES}}$, where $\hat{r}_{(q)} = \hat{r}_{1(q)}\hat{r}_{2(q)}\hat{r}_{3(q)}$ is the original CES estimate of the growth rate in quarter q.

17 Deviating from current practice, we have estimated the seasonal factors concurrently. This distinction is not central to our analysis.

18 We have simplified the analysis by assuming that seasonality does not change over time. In future work, we plan to relax this assumption.

19 Berger and Phillips, “Reassessing Texas employment growth,” pp. 1–3.

20 See Stuart Scott, George Stamas, Thomas J. Sullivan, and Paul Chester, “Seasonal adjustment of hybrid economic time series,” *Proceedings of the Section on Survey Research Methods* (American Statistical Association, 1994), https://www.bls.gov/osmr/research-papers/1994/pdf/st940350.pdf; and Keith R. Phillips and Jiango Wang, “Seasonal adjustment of state and metro CES jobs data,” Working Paper 1505 (Federal Reserve Bank of Dallas, 2015).

21 In fact, the benchmarked quarterly growth rates are essentially the QCEW growth rates with the CES seasonality. Of course, a second advantage of our proposed methodology is that QCEW estimates are incorporated sooner. The Federal Reserve Bank of Dallas incorporates the QCEW information as soon as the QCEW data become available, but, like BLS, does so by producing a hybrid series. In a recent article, Thomas Walstrum examines the effect of applying this early benchmarking procedure to update the CES employment estimates for the five states in the Seventh Federal Reserve District. See Thomas Walstrum, "'Early benchmarking' the state payroll employment survey," Employment, Midwest, Seventh District (Federal Reserve Bank of Chicago, June 2015).

RELATED CONTENT

Related Articles

Benchmarking the Current Employment Statistics survey: the past, present, and future, *Monthly Labor Review*, October 2017.

Benchmarking the Current Employment Statistics national estimates, *Monthly Labor Review*, October 2017.

Benchmarking the Current Employment Statistics state and area estimates, *Monthly Labor Review*, November 2017.

Benchmarking the Current Employment Statistics survey: perspectives on current research, *Monthly Labor Review*, November 2017.

Related Subjects

Survey methods | Benchmarking | Seasonal adjustment | Statistical programs and methods | Employment | Experimental methodology | Survey procedures

Monthly Labor Review

November 2017

Benchmarking the Current Employment Statistics state and area estimates

The Current Employment Statistics survey gathers detailed estimates on employment, hours, and earnings for the nation, states, Puerto Rico, the U.S. Virgin Islands, and more than 450 metropolitan statistical areas. This article focuses directly on the current methods used to benchmark state and metropolitan area data and is one of four series articles on benchmarking.

Kirk Mueller
mueller.kirk@bls.gov

Kirk Mueller is a division chief in the office of Employment and Unemployment Statistics, U.S. Bureau of Labor Statistics.

The Current Employment Statistics (CES) survey, conducted by the U.S. Bureau of Labor Statistics (BLS), is a monthly survey of approximately 147,000 business and government agencies, representing 634,000 individual worksites. The CES program provides detailed estimates on employment, hours, and earnings for the nation, states, Puerto Rico, the U.S. Virgin Islands, and more than 450 metropolitan statistical areas (MSAs). The CES survey is widely considered one of the most timely and accurate economic indicators published by the federal government.

Annually, the CES program benchmarks, or reanchors, the sample-based employment estimates for March of each year to the universe counts derived principally from the Quarterly Census of Employment and Wages (QCEW) program. This benchmark process is done to account for survey and other errors that accumulate over the year. The QCEW counts are much less timely than sample-based estimates and provide an annual point-in-time census for employment.

The current research agenda for the CES program includes examining alternative benchmarking approaches for national, state, and metropolitan area data. As background for that examination, this article discusses the current methods used to benchmark state and metropolitan area data and is a companion piece to three other articles on benchmarking in the CES program. In the first of the three articles, Christopher Manning and John Stewart discuss the current benchmark methods and results for national CES data. In the second article, Kenneth Robertson presents the conceptual underpinnings of benchmarking and a high-level overview of recent research. Finally, in the third article, Mark Loewenstein and Matthew Dey author a more technical article in which they discuss the alternative method that has shown the most promise.

Establishing benchmark levels and the postbenchmark period

With each annual benchmark, the standard practice for state and area series is to revise 20 months of not seasonally adjusted data before the normal monthly estimation processes begin on the new levels. For example, with the development of the March 2016 benchmark, levels were reestablished for the April 2015 through November 2016 reference months. December 2016 final and January 2017 preliminary estimates were developed based on those newly established levels.

A snapshot of the QCEW is the starting point for building the CES March benchmark level. Added to that level is an accounting of employment that is covered under the CES definition, but not by state unemployment insurance (UI) tax laws. This second segment of employment is called noncovered employment and is only present in select industries.[1]

The total of the QCEW and noncovered employment (referred to as the population) replaces March sample counts for all national, state, and metropolitan series. The benchmark method differs between the state and metropolitan series and the national series. For the state and metropolitan area series, the employment levels for all months, from the April before the March benchmark through the September after the March benchmark, are also replaced by these population counts. For the national series, only the March benchmark month has its employment replaced by the population counts. The CES program staff adjusts the monthly levels before the benchmark month by applying a linear wedge that is based on the revision to the March level. Months subsequent to March are sample-based estimates from the new March benchmark level.[2]

CES staff generates October and November estimates as sample-based estimates linking from the newly established September level. Note that the original postbenchmark population values of April through September are updated again with newer versions of the QCEW and noncovered employment counts are updated with the next benchmark cycle. As given in the example at the beginning of this section, April 2016 through September 2016 employment data are replaced once as part of the March 2016 benchmark process. The April 2016 through September 2016 employment data are replaced again a year later as part of the March 2017 benchmark process.

In addition to replacing the estimates, BLS also addresses some administrative issues in the QCEW. The presence of noneconomic code changes (NECCs) is one such issue. As an example, each year approximately one-third of all establishments in the QCEW are contacted as part of the Annual Refiling Survey. State Labor Market Information agencies contact companies and ask them to verify their NAICS (North American Industry Classification System) industry classification, ownership, and location. These NECC corrections are implemented in the QCEW with first-quarter data. The effect of changes that represent less than 6 percent of the employment in a series is distributed across 12 months. For changes of 6 percent or more, CES staff will lengthen the number of months that the employment change is distributed across.

Reasons behind the current method

In general, small geographic areas and industries have fewer sample units, which leads to more variability in the estimates. Although error is associated with the administrative data that are based on the QCEW, BLS decided more than 30 years ago that the error in the administrative data was a better alternative to preserving the variability of the estimates as the final benchmarked time series.

Although the sample sizes from when the current benchmarking method was established are not readily available, the current sample sizes in cells illustrate the frequency of series with small sample sizes. CES estimation occurs at the most detailed levels, referred to as basic cells, and higher level industry totals are the sum of these basic cells. The CES program models more than 55 percent of the slightly more than 6,000 basic series, rather than using only the sample to estimate the data. It also models series when the sample is deemed too small,[3] which is generally when the sample has fewer than 30 companies (as defined by a distinct UI account). The CES models combine a sample-based estimate with a trend component to smooth the data.[4]

Complications affecting the current method

Research by the Dallas Federal Reserve has shown that CES benchmarked population data display a seasonal pattern different from the sample-based estimates.[5] The CES program accounts for the differences in the patterns by using a two-step seasonal adjustment process to develop the final seasonally adjusted series. However, BLS seasonally adjusts only about 2,000 state and metropolitan series compared with the 17,000 not seasonally adjusted series. Analyzing employment changes can be complicated, particularly over-the-year comparisons, without seasonally adjusted data. These changes cross over the splice point between population- and sample-based data.[6]

The seasonal differences also complicate interpreting revisions associated with the benchmark process. Tracking the population and sample differences nationally shows that typically the QCEW data grow by approximately a quarter million more than the CES sample-based data from March through September each year. The difference grows greater through the fourth quarter, and then the population data have a much larger seasonal decrease in the first quarter.[7] As a result, QCEW data grow by approximately a quarter of a million less than CES sample-based data from September through March each year. These differences may represent administrative error or reporting error rather than true economic differences.[8]

As mentioned, state and metropolitan data are replaced with population data through September with each benchmark. Historically, BLS has highlighted the March revision as the primary indicator of the accuracy of the CES survey. However, given the seasonal differences described previously, in aggregate under the current method, the expected differences in March should average a downward revision of approximately 250,000.

Results

Table 1 provides a summary of March benchmark revisions at the state nonfarm level from 2003 to 2016. Table 2 provides comparable revisions for MSAs.

Table 1. Summary statistics of March benchmark revisions, statewide total nonfarm and supersectors, not seasonally adjusted 2005−16

Variable and supersector	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Variable														
Average absolute percent revision statewide total nonfarm	0.6	0.4	0.5	0.5	0.4	0.4	0.9	0.4	0.5	0.7	0.4	0.5	0.4	0.4
Mean percent revision statewide total nonfarm	−0.2	0.2	0. 1	0.3	0.0	−0.1	−0.8	−0.1	0.2	0.6	0.3	0.1	(1)	−0.1

See footnotes at end of table.

Table 1. Summary statistics of March benchmark revisions, statewide total nonfarm and supersectors, not seasonally adjusted 2005–16

Variable and supersector	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Standard deviation	0.7	0.5	0.6	0.7	0.5	0.5	0.8	0.5	0.6	0.7	0.6	0.6	0.5	0.6
Minimum	−1.9	−0.9	−1.2	−0.8	−1.5	−1.4	−3.8	−1.3	−1.8	−1.5	−0.7	−1.5	−1.8	−1.6
Maximum	1.4	1.8	1.2	4.2	1.2	1.0	1.1	1.4	1.4	2.2	2.9	2.0	1.3	0.9
Average absolute percent revision statewide by supersector														
Mining and logging	3.8	5.8	6.5	3.4	3.8	4.3	6.0	7.5	3.2	4.7	3.7	2.8	4.2	4.5
Construction	2.6	2.4	2.8	2.7	2.2	2.6	4.0	3.6	3.2	4.4	3.1	3.0	2.6	2.3
Manufacturing	1.4	1.2	1.3	1.7	1.2	1.3	2.2	1.8	1.4	1.5	1.4	1.2	1.3	1.3
Trade, transportation, and utilities	1.0	0.8	0.7	0.5	0.7	0.6	1.6	1.2	0.9	1.1	1.0	0.7	0.6	0.8
Information	2.5	2.5	2.2	1.9	2.2	2.0	3.3	2.3	2.4	3.2	2.2	.02	2.6	3.0
Financial activities	1.7	1.0	1.2	0.9	1.1	1.0	1.6	1.8	1.9	2.2	1.6	2.0	1.9	2.3
Professional and business services	2.1	1.9	1.7	2.1	1.5	1.3	2.2	2.2	1.8	1.9	1.8	1.6	1.6	1.4
Education and health services	1.0	1.1	0.6	0.9	0.7	0.8	0.8	1.0	0.9	1.4	1.6	0.9	0.9	0.8
Leisure and hospitality	1.3	1.4	1.4	1.2	1.1	0.9	1.7	1.8	1.9	2.3	1.4	1.4	1.4	1.5
Other services	2.1	2.0	1.9	1.7	1.5	1.3	1.9	1.9	2.4	2.7	2.1	2.4	2.1	2.4
Government (2)	0.8	0.7	0.6	0.7	0.5	0.6	0.6	0.8	0.7	1.0	0.7	0.9	0.7	0.5

Notes:

(1) Less than 0.05.

(2) Includes federal, state, and local governments.

Source: U.S. Bureau of Labor Statistics.

Table 2. Summary statistics of March benchmark revisions, MSAs, not seasonally adjusted 2005–16

Variable	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
All MSAs												
Number of MSAs	367	367	318	319	378	381	381	381	372	258	387	387
Average absolute percent revision	1.1	1.1	0.9	1.0	1.8	1.1	1.1	1.6	1.2	1.1	1.0	1.1
Mean	(1)	0.2	−0.2	−0.3	−1.4	0.0	0.1	0.4	0.4	0.3	−0.3	−0.2
Standard deviation	1.6	1.5	1.2	1.3	1.9	1.5	1.5	2.1	1.6	1.5	1.3	1.5
Less than 100,000 employment												
Number of MSAs	178	177	117	118	183	188	187	182	181	132	188	189
Average absolute percent revision	1.3	1.3	1.1	1.3	2.1	1.4	1.4	2.0	1.4	1.4	1.2	1.3
Mean	−0.2	−0.2	0.2	−0.4	−1.5	−0.2	−0.1	0.4	0.3	0.3	0.5	−0.5
Standard deviation	1.7	1.8	1.4	1.7	2.3	1.9	1.8	2.6	1.8	1.8	1.6	1.8
100,000 to 499,999 employment												
Number of MSAs	140	141	144	144	138	139	138	141	140	100	147	146
Average absolute percent revision	1.1	0.4	0.8	0.9	1.6	0.9	0.9	1.5	1.1	0.9	0.8	0.9
Mean	0.2	0.2	−0.2	−0.3	−1.4	0.1	0.2	0.4	0.3	0.4	−0.2	0.1
Standard deviation	1.5	1.4	1.0	1.1	1.5	1.1	1.2	2.0	1.4	1.0	1.0	1.2
500,000 to 999,999 employment												
Number of MSAs	25	23	28	28	31	30	31	30	25	11	22	20
Average absolute percent revision	0.6	0.6	0.7	0.6	1.4	0.7	0.8	0.9	0.8	0.9	0.8	0.6
Mean	0.2	0.5	0.2	−0.3	−1.2	0.3	0.5	0.6	0.6	0.3	0.2	0.3
Standard deviation	0.8	0.7	1.0	0.7	1.0	0.8	0.8	1.0	1.1	1.2	0.9	0.8
1 million or more employment												

See footnotes at end of table.

Table 2. Summary statistics of March benchmark revisions, MSAs, not seasonally adjusted 2005−16

Variable	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Number of MSAs	24	26	29	29	26	24	25	26	26	15	30	32
Average absolute percent revision	0.7	0.5	0.6	0.5	0.9	0.5	0.7	0.7	1.1	0.7	0.4	0.4
Mean	0.3	0.1	−0.1	−0.3	−0.9	0.3	0.6	0.7	0.8	0.5	−0.1	0.1
Standard deviation	1.0	0.7	0.8	0.8	0.7	0.6	0.7	0.6	1.4	0.7	0.6	0.5

Notes:

(1) Less than 0.05.

Note: MSAs = metropolitan statistical areas.

Source: U.S. Bureau of Labor Statistics.

At the state nonfarm level, the average absolute March revisions were 0.5 percent between 2003 and 2016. Over the same period, the largest average absolute percent revisions at the supersector level were in mining and logging (4.3 percent) and construction (3.0 percent). The smallest average absolute percent revisions were in government (0.7 percent); trade, transportation, and utilities (0.9 percent); and education and health services (1.0 percent).

The average mean percent revision over the last 14 years is 0.05 percent. However, one should be cautious when interpreting that number. Because the number does not reflect the size of the individual states, it should not be interpreted as an average revision of the sum of states over the period.

Revisions since the 2005 benchmark for MSAs are included in table 2.[9] For comparisons, breakdowns are by size of the MSAs. Across the most recent 13 years, the average absolute revisions have averaged 1.2 percent across all MSAs. As is expected, the larger the MSA, the smaller the absolute percent revisions are. MSAs more than 1 million in employment have averaged an average absolute percent revision of more than 0.6 percent, while MSAs with less than 100,000 employment have averaged slightly more than 1.4 percent.

Conclusions

The current benchmarking method used for state and metropolitan area estimates addresses the concern of excessive variability in the final benchmarked data that might result from estimates produced with small samples. In addition, replacing estimates from April through September following the March benchmark updates the series with the most current population information available. However, replacing the estimates with population values does not directly address many of the issues associated with the differing seasonal patterns between the population and sample data. Current research into alternative benchmarking approaches might yield a method that addresses some of the limits of the present-day benchmark method, while continuing to address concerns related to the volatility of small domain estimates.

SUGGESTED CITATION

Kirk Mueller, "Benchmarking the Current Employment Statistics state and area estimates," *Monthly Labor Review,* U.S. Bureau of Labor Statistics, November 2017, https://doi.org/10.21916/mlr.2017.26.

NOTES

1 The CES Technical Notes contain detailed information on the industries, such as education, religious organizations, and rail transportation, that qualify as potentially having employment that is not covered under unemployment insurance tax law, https://www.bls.gov/web/empsit/cestn.htm.

2 Before the 2010 benchmark, BLS did not replace all state estimates with population data through the third quarter of the postbenchmark period. The number of states in which estimates were replaced through the third quarter increased through the 2000s before reaching the stage in which all third-quarter estimates were replaced with third-quarter population data. With the exception of the 2011 benchmark, when BLS replaced all state estimates with population data only through second quarter, all state estimates were replaced with population data through the third quarter since the 2010 benchmark.

3 To assist in dealing with the variability in the estimates at detailed level, the CES program uses a "top-down" approach to the estimates. This approach forces the detailed estimates to sum to estimates independently produced at the super-sector level for statewide and large metropolitan statistical areas.

4 The CES uses a small domain model as well as versions of the Fay-Herriot model in estimating smaller domains. For more information, see https://www.bls.gov/opub/hom/pdf/ho mch2.pdf.

5 For more information, see Franklin D. Berger and Keith R. Phillips, "Solving the mystery of the disappearing January blip in state employment data," *Economic Review* (Dallas, TX: Federal Reserve Bank of Dallas), April 1994, pp. 53−62, http://w ww.dallasfed.org/assets/documents/research/er/1994/er9402d.pdf.

6 The two-step seasonal adjustment process is explained in detail by Stuart Scott, George Stamas, Thomas Sullivan, and Paul Chester, "Seasonal adjustment of hybrid economic time series," *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 1994, https://www.bls.gov/osmr/research-papers/1994/st940350.htm.

7 Kenneth W. Robertson, "Benchmarking the Current Employment Statistics survey: the past, present, and future," *Monthly Labor Review*, October 2017, https://doi.org/10.21916/mlr.2017.24.

8 Jeffrey A. Groen, "Sources of error in survey and administrative data: the importance of reporting procedures," *Journal of Official Statistics*, vol. 28, no. 2, June 2012, pp. 173−198.

9 The 2004 benchmark introduced a large change in the metropolitan statistical area (MSA) definitions that complicated the examination of benchmark revisions. The MSA redefinitions in 2014 affected the comparisons less than the MSA redefinitions in 2004 and are included in the table.

RELATED CONTENT

Related Articles

Benchmarking the Current Employment Statistics survey: the past, present, and future, *Monthly Labor Review*, October 2017.

Benchmarking the Current Employment Statistics national estimates, *Monthly Labor Review*, October 2017.

Benchmarking the Current Employment Statistics survey: perspectives on current research, *Monthly Labor Review*, November 2017.

A quarterly benchmarking procedure for the Current Employment Statistics program, *Monthly Labor Review*, November 2017.

Related Subjects

Survey methods | Benchmarking | Sampling | Seasonal adjustment | Statistical programs and methods | Survey procedures

Monthly Labor Review

November 2017

Benchmarking the Current Employment Statistics survey: perspectives on current research

In the sample-survey world, benchmarking is an alignment of a survey-based estimate with a population value. Aligning estimates to a high-quality control total improves the quality of the set of estimates for both the controlled and correlated characteristics. This article discusses benchmarking in the Current Employment Statistics (CES) survey, from its roots in 1935 to current benchmark procedures. The general practice of CES benchmarking is briefly described and is accompanied by recent quantitative results at national and state levels. Impediments to improvements in the methodology are discussed and illustrated. The article concludes with a brief overview of methods included in recent research and with an outline of future plans.

Kenneth W. Robertson
robertson.ken@bls.gov

Kenneth W. Robertson is the Assistant Commissioner for the Office of Industry Employment Statistics, U.S. Bureau of Labor Statistics.

What is benchmarking?

A benchmark is generally described as a standard of excellence against which other things are judged. In a statistical sense, a benchmark is a high-quality population value against which a survey estimate may be judged in order to assess the quality of the estimate.

For surveys, a benchmark is something obtained outside the survey. For example, a high-quality sampling frame—a list of all those within a population who can be sampled—usually includes a population measure or census count correlated to a major statistic to be collected. Many surveys may select a sample from a population list (at time *t*), take the requisite time to collect the survey data, and then align the survey estimates with population values from an updated population list (at time *t + 1*) to make the estimates as current, relevant, and accurate as possible when published. This is one type of survey benchmarking.

In the Current Employment Statistics (CES) survey, the U.S. Bureau of Labor Statistics (BLS) Quarterly Census of Employment and Wages (QCEW) serves as both the population list and the primary source of benchmark employment. The QCEW covers about 97 percent of the employment that is in scope for CES estimates.

CES background

The CES survey is a quick-response business survey that provides some of the earliest information about the state of the U.S. economy each month. BLS produces and publishes the initial estimates from the survey about 3 weeks after the reference period. The reference period is the payroll period that includes the 12th of the month. The survey produces employment, hours, and earnings data by detailed industry levels. About 2 weeks later, data are produced for states and metropolitan statistical areas. The national data are revised two times on the basis of additional data collected for the reference period, and the subnational data are revised once. BLS produces about 50,000 data series each month, using the business reports collected for the CES program.

The primary estimator for national employment data is referred to as a weighted link-relative estimator. It takes the ratio of current-month weighted employment to prior-month weighted employment (based on a matched sample of collected data) and multiplies that quotient by the prior-month employment level. An estimated adjustment value is then added to account for the net employment change in businesses not captured by the survey; this estimated adjustment value is called the net birth/death factor. The primary employment estimator,[1] shown just below, provides an efficient estimate of over-the-month (OTM) change and is also referred to as the "change ratio" for month *c*:

$$\hat{R}_c = \frac{\sum_i w_i e_{c,i}}{\sum_i w_i e_{p,i}} \text{ ,(1)}$$

where

i is an establishment that reported both this month and last month,

w_i is the sampling weight associated with establishment *i*,

c is the current month,

p is the prior month,

$e_{c,i}$ is the employment for current month *c* for establishment *i*, and

$e_{p,i}$ is the employment for prior month *p* for establishment *i*.

The weighted link-relative estimator, shown below, is composed of

$\hat{R}_c$ and two additional components.

$$\widehat{E_c} = \widehat{E_p}\hat{R}_c + b_c \text{ ,(2)}$$

where

$\widehat{E_c}$ is the employment estimate for current month *c*,

$\widehat{E_p}$ is the employment estimate for prior month *p*,

$\hat{R}_c$ is the change ratio for current month *c* (equation 1), and

b_c is the net birth/death factor for month *c*.

The primary estimator for state and metropolitan area employment estimates, a derivative of the weighted link-relative estimator, includes a robust procedure for identifying and reducing the influence of unusual reports (or outliers).

Although the link-relative estimator is an efficient estimator for employment change, errors may accumulate over time because each estimate is linked to the one before it. The hope, of course, is that the errors are random and offsetting—that is, some errors on the positive side are balanced by some errors on the negative side. To ensure that the accumulated error does not get too large in any direction, the CES program benchmarks the survey estimate level by using a lagged population value.

Benchmarking is used to align a survey estimate with a census value. This generalized statement is, however, overly simplified in some cases. It assumes that the census value used in the alignment is essentially without error. In reality, the CES sample estimates include sampling and nonsampling errors, and the QCEW census data include nonsampling errors.[2] Therefore, the CES benchmark can best be described as an alignment of CES estimates with an independently derived employment count. Nonetheless, both sources of data are subject to error.

This article describes some of the history associated with CES benchmarking, the current process used, and research efforts to improve that process.

A brief history of CES benchmarking

The CES survey was first envisioned as an employment index for a limited number of industries: boots and shoes, cotton goods, cotton finishing, hosiery and underwear, and iron and steel. BLS started the survey in 1915, and at that time there was no plan in place to benchmark the data against another source of employment information. In 1935, however, BLS implemented the first benchmark procedure after identifying a very large downward bias in the manufacturing index of about 12 percent over the period from 1923 to 1929.[3] This first benchmark incorporated approximated census levels for each of the years 1923, 1925, 1927, 1929, and 1931. The censuses of manufactures and the censuses of businesses provided the data used in this first benchmark.

Although this benchmark procedure improved the quality of the CES data, the procedure did not account for the full scope of the survey because the benchmark source data did not cover the full nonfarm business population. In 1935, Congress passed the Social Security Act, which established the unemployment insurance (UI) program; data became available to BLS from the new program in about 1940. These new UI data included industries not covered in the census of manufactures and the census of businesses, making the UI data the preferred source of benchmark data for the CES program. The UI data from each state are edited by states and collected by BLS through the QCEW program.

Originally, both the national and state data series were benchmarked only to March QCEW data.[4] The benchmark procedure essentially replaced the prior year CES March estimate with the QCEW March level and then used a linear wedge to distribute the correction amount to the 11 months before that March. The implicit assumption with this procedure is that errors accumulated at a steady rate over the year before the benchmark. The benchmark with linear wedge procedure is applied as follows:

$$\hat{E}_m^B = \hat{E}_m + \frac{m}{12}\left(P_M - \hat{E}_M\right), (3)$$

where

$\hat{E}_m^B$ is the benchmarked all-employee level for month *m,*

$\hat{E}_m$ is the not-yet-benchmarked all-employee estimate for month *m,*

m is the month being benchmarked (1 is April, 2 is May, …, 11 is February, 12 is March),

P_M is the March "population" value, which is QCEW employment plus the number of jobs not covered by the UI system (about 3 percent of nonfarm jobs), and

$\hat{E}_M$ is the not-yet-benchmarked March all-employee estimate.

Around 1980, all 12 months of UI data became available to the QCEW program. Also at that time, individual states began transitioning from a March replacement with wedge procedure to a 12-month replacement procedure for benchmarking. Meanwhile, the national procedure remained unchanged. By the late 1980s, all states had switched to the replacement method. State and metropolitan area estimates changed to the replacement method because BLS and state analysts believed the error in the administrative data was preferable to the sampling error that remains in the benchmarked data for months other than March.[5] When states began replacing their CES data with QCEW data for all months, BLS and state analysts assumed that the administrative error in the QCEW was smaller than the sampling error in CES, especially for smaller domains. For the larger domains, specifically for national-level data series, a different assumption was made: the best estimate of over-the-month change was the CES estimate, and a level correction with a linear wedge applied once per year was sufficient to ensure high-quality historical estimates.

Current research is focused on answering two questions. First, is there a procedure that would allow for the use of more frequent alignment to the population data that consistently improves the quality of the CES data? And second, is there a procedure that would allow for more consistency between the benchmark procedures used at the national and state levels?

Current methodology and some qualitative results

The current methodology for national estimates differs substantially from that for state and area estimates.

National methodology

The national procedure takes the prior March QCEW level for each industry and adds employment not covered by the UI system. This noncovered employment level is mostly obtained by taking the difference between the QCEW data and older data from the U.S. Census Bureau *County Business Patterns* and then forecasting that difference forward to the benchmark period.[6] The Railroad Retirement Board is another source of noncovered employment data.[7] This census employment value for the prior March then replaces the CES survey estimate, and the difference between the original estimate and the benchmark value is wedged into the 11 months preceding March (see equation 3).

Estimates for the postbenchmark period must also be adjusted to the new March level.[8] This is done by taking the already estimated over-the-month change ratio and newly forecasted net birth/death factors for each month and applying them to the new level and repeating this process moving forward to the most recent estimate. That is, in equation 2, $\widehat{E_p}$ takes on a new value as you move the new levels forward until you get to the most recent month.

More on the current national methodology is available in an article by Chris Manning and John Stewart and in the national program's annual benchmark article.[9]

State and area methodology

The state and area procedure is simpler in concept than the national procedure, although it is also more resource intensive. This procedure develops a census employment value by following the same process used by the national procedure for March but does that for all 12 months. These census data then replace the CES survey estimates.

The postbenchmark period estimates also receive different treatment for the state and area data. Preliminary QCEW data are used to develop population values for the April-to-September period following the procedure used to develop the March values, and these values replace the CES estimates. For the October-to-December period, the values are updated by using procedures similar to those used for national estimates. That is, for this period, estimates are developed that use (1) the new benchmark level, (2) all data collected for the reference periods following the benchmark, and (3) newly forecasted net birth/death factors to update the estimates. More on the current state and area methodology is available in an article by Kirk Mueller and in the annual CES state-and-area benchmark article.[10]

Some quantitative results

One interpretation of the benchmark revision is that it serves as a proxy for total survey error. However, this interpretation does not consider error in the administrative data that is the source of the benchmark data. Taking errors in both data sources into account, BLS interprets the CES benchmark revision as the difference between two independently derived employment counts, each subject to its own sources of error. QCEW errors accrue from the usual errors associated with large administrative datasets: comprehension and collection errors, data entry and coding errors, imputation errors, and other processing errors. Both BLS and its state partners work hard to minimize those errors, but they still exist to some extent.

Some results for CES benchmark revisions are shown below. These are percentage benchmark revisions for national total nonfarm employment in March, 2005–15:

Year	Percent
2005	-0.1
2006	0.6
2007	-0.2
2008	-0.1
2009	-0.7
2010	-0.3
2011	0.1
2012	0.3
2013	-0.1
2014	0.0
2015	-0.1

National revisions generally are small. Over the period depicted, the 5-year average of the absolute values of the percentage revisions, hereafter called absolute average percentage revisions (AAPRs), has ranged from 0.4 percent in 2010 to 0.1 percent in 2015.

Here is the AAPR for states, for total nonfarm employment in March, 2005–15:

Year	AAPR
2005	0.5
2006	0.5
2007	0.4
2008	0.4
2009	0.9
2010	0.4
2011	0.5
2012	0.7
2013	0.4
2014	0.5
2015	0.4

As expected, average state revisions tend to be a bit larger in absolute value than national revisions.

For metropolitan statistical areas (MSAs), revisions are presented by MSA size as well as total. MSAs with smaller employment levels tend to have smaller sample sizes and hence a higher relative sampling error (see table 1).

Table 1. Absolute average percentage revisions for metropolitan statistical areas (MSAs), for total nonfarm employment, March, 2005–15

MSA employment size	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
All sizes	1.1	1.1	0.9	1.0	1.8	1.1	1.1	1.6	1.2	1.1	1.0
Less than 100,000	1.3	1.3	1.1	1.3	2.1	1.4	1.4	2.0	1.4	1.4	1.2
100,000–499,999	1.1	0.4	0.8	0.9	1.6	0.9	0.9	1.5	1.1	0.9	0.8
500,000–999,999	0.6	0.6	0.7	0.6	1.4	0.7	0.8	0.9	0.8	0.9	0.8
More than 1,000,000	0.7	0.5	0.6	0.5	0.9	0.5	0.7	0.7	1.1	0.7	0.4

Source: U.S. Bureau of Labor Statistics.

The largest MSAs generally have revisions that are near the size of the absolute average state revision. The smallest MSAs have revisions that are larger on average, and a few of the smallest MSAs each year can have large revisions. The range of percentage revisions, as shown in table 2, illustrates this.

Table 2. Benchmark revisions for total nonfarm employment in metropolitan statistical areas (MSAs), March 2015

Item	Number of employees				
	All sizes	Less than 100,000	100,000–499,999	500,000–999,999	More than 1,000,000
Number of MSAs	387	188	147	22	30
Absolute average percentage revision	1.0	1.2	0.8	0.8	0.4
Range of revisions	-6.4 to 6.0	-6.4 to 6.0	-3.7 to 2.8	-1.1 to 1.6	-1.9 to 0.7

Source: U.S. Bureau of Labor Statistics.

Benchmarking by data users

Benchmarking of CES data used to be the sole province of BLS. The process formerly taxed the capabilities of mainframe systems and data processing budgets by processing thousands of data series and updating years of historical data. However, with today's powerful desktop and laptop computers, simple macro-level benchmarking is a process available to almost anyone with the technical savvy to download the appropriate datasets and apply a well-considered update process to the data.

Several of the Federal Reserve regional banks conduct a quarterly benchmarking process for state and area estimates before BLS issues the official annual benchmarks.[11] The process, in general, can be described as being made up of three parts. The first part is the oldest data, comprising data already benchmarked by BLS. The second part is data for which QCEW data are available but have not yet been incorporated into the time series. The third and most recent part is data for which QCEW data are not yet available. The process uses QCEW growth rates (similar to a change ratio) to replace the second part and CES growth rates to adjust the third part.

More specifically,

$$\hat{\ddot{E}}_v = \hat{\ddot{E}}_p \hat{\ddot{R}}_v \,,$$

where

$\hat{\ddot{E}}_v$ is the early benchmarked CES value for month v,

$\hat{\ddot{E}}_p$ is the prior month level, perhaps produced as a result of early benchmarking or a CES value as originally published,

$$\hat{\ddot{R}}_v = \left\{ \begin{array}{c} \widehat{CES}_{E,v} \Big/ \widehat{CES}_{E,v-1} \text{; if part 1—data are already benchmarked by BLS—use this, or} \\ \widehat{QCEW}_{E,v} \Big/ \widehat{QCEW}_{E,v-1} \text{; if part 2—QCEW data are available but not incorporated—use this, or} \\ \widehat{CES}_{E,v} \Big/ \widehat{CES}_{E,v-1} \text{; if part 3—QCEW data are not yet available—use this.} \end{array} \right\},$$

$\widehat{\mathrm{CES}}_{E,v}$ is the published CES all-employee estimate for month v, and

$\widehat{\mathrm{QCEW}}_{E,v}$ is the published QCEW all-employee estimate for month v.

This procedure can also be described as follows:[12]

1. Begin with the CES state time series at some level of industry detail, and identify the parts of the series as described above.
2. Where QCEW data are available but not yet incorporated into CES data (i.e., for part 2), use the QCEW data to calculate over-the-month (OTM) change ratios or links. The calculation is $\text{link}_t = E_t / E_{t-1}$. This calculation will serve as a proxy for the eventually benchmarked OTM changes.
3. Multiply the last BLS benchmarked value by the first QCEW OTM link to replace the CES estimate, then multiply the product times the next link, and so on, until you update through the most recently available QCEW data.

4. For data in part 3, create CES OTM links in the same manner as that used to create QCEW OTM links forward from the latest QCEW replacement value, and apply them from that point forward to the most recent estimate.
5. Calculate the seasonal factors for all three parts of the series described above. Using ARIMA X-12 or ARIMA X-13-SEATS,[13] seasonally adjust the three parts described, and then link the parts as described in (1) to (4). Calculating seasonal factors for the first two parts (the officially benchmarked series and the QCEW links) is straightforward because long time series are available for these data. The third part is more difficult. There are two options for this part: (a) create a time series consisting of historical values of nonbenchmarked CES data or (b) simply use the changes in the official seasonally adjusted BLS data to extrapolate the seasonally adjusted CES data past the early benchmark period. Option (b) is the easiest solution, but it limits potential industry detail to those industries that BLS seasonally adjusts, which is typically the broad industry level.

The intent of *early* benchmarking is to take QCEW data published every quarter and incorporate them into the CES data more often than once per year. The expectation is that this process will reduce the error in the portion of the linked CES data thus updated, providing a data series that will be closer to what BLS will eventually publish after it benchmarks the data. As noted above, some entities have been doing early benchmarking for years. Researchers at the Federal Reserve Bank of Dallas, for example, have been exploring methodologies and applying early benchmarking since at least 1993.[14] Since then, early benchmarking has been undertaken by researchers at other Federal Reserve banks and at several universities.

Impediments to improving the methodology—why is this hard?

Data users might well ask: "Why hasn't BLS started benchmarking more often than once per year?" The answer is twofold. First, there are significant resource considerations. Benchmarking, while technically easy to describe and implement at a macrolevel, requires a considerable amount of BLS resources. BLS must carefully review the initial microlevel inputs for accuracy, develop and review forecasts for some inputs (e.g., *County Business Patterns*), and carefully review the results before publication. More than 50,000 CES series are reviewed as part of the benchmark process. BLS data are considered the gold standard of economic labor market data. This standard is maintained because the agency takes great pains to ensure that the data it releases are as accurate and complete as possible. BLS staffing is currently funded to conduct this "peak-load" benchmarking process once per year. To do this process more than once per year means more funding would be needed. Without additional funding, BLS would need to reduce the number of published data series in order to offset the increased workload.

The second consideration is significantly more technical in nature. BLS has known, at least since 1993 when two Dallas Federal Reserve Bank researchers documented and discussed the issue with BLS, that data from the QCEW and the CES survey have different seasonal patterns.[15] A Response Analysis Survey (RAS) was conducted several times to explore differences between CES and QCEW estimates.[16] The most recent RAS was conducted in 2008, and a number of papers and analyses were based on this survey.[17] The RAS results consistently highlight several common themes.

First, when exploring reports with *different* employment responses for the CES survey and the QCEW, the RAS results show that about one-third of the time the CES report was correct, one-third of the time the QCEW report was correct, and about one-third of the time neither was exactly correct.[18] There are many reasons a report might be in error. For example, different people in the business may respond to the two programs on the basis of data from different internal databases. Also, a report to one requester might include employees that are out of scope (e.g., by including contract employees) while the correct scope is reported to another requester.

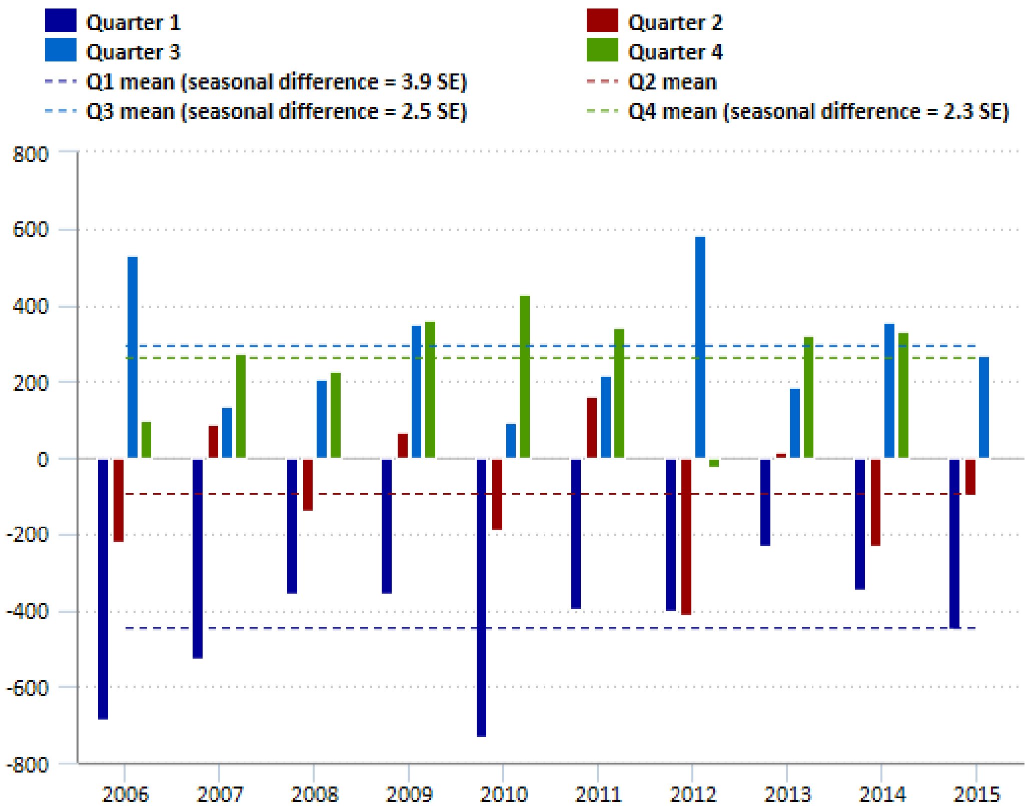
Second, a consistent RAS result over time is that the QCEW has a larger proportion than the CES survey of respondents reporting for pay periods that occur after the reference period. That is, some reports may be for the pay period that follows the one that includes the 12th day of the month, some may be an end-of-month count, etc.[19] A side effect of this pay-period misreporting is that the QCEW would show more employment gain during a seasonal hiring month (such as December) and more employment loss during a seasonal layoff month (such as January). This effect is, in fact, what we see in the QCEW. For example, the QCEW shows more employment gain than the CES survey in December and more employment loss in January.

I note here that this is a minor issue for the QCEW. The goal of the QCEW is to report total employment and wages each quarter. The primary purpose for this data collection is to administer state unemployment insurance (UI) programs. State UI staff collect the data quarterly and provide them to state Labor Market Information (LMI) staff. These data are not intended to be a monthly report of employment change; they are primarily intended to support the collection of UI taxes. After supporting their primary purpose, these administrative data are leveraged to support additional uses. For example, the data can provide additional value for the LMI community by serving as a comprehensive source of employment and wage data by industry and county. The state LMI staff work with BLS to review these administrative data and correct errors and omissions. If we assumed the difference between the CES survey and the QCEW to be pay-period reporting error by the QCEW, then the QCEW is still 99.7 percent correct or better in accuracy in its worst month.[20]

Another interpretation of this is that, for the QCEW, the reports are nearly 100 percent accurate with some reporting-period error. That is, the QCEW reporter is reporting accurate employment monthly counts, just sometimes not for the requested pay period. However, for the CES survey, whose goal is to measure the month-to-month change in employment (for the week that includes the 12th of the month), a 0.3-percent error can lead to huge adjustments to the CES data. For context, the average CES over-the-month change is about 0.6 percent of the employment level. A 0.3-percentage-point error on a 0.6-percent average change suggests that as much as a 50-percent adjustment in the OTM change (not seasonally adjusted) is required for some months to account for the seasonal differences in these data sources. Perception is important, and, historically, the small seasonal errors in the QCEW employment levels have been perceived by some state data users as very large errors in the CES employment estimates even though the QCEW errors actually are small.

Figure 1 summarizes these differences between QCEW and CES data. The differences are depicted on a quarterly basis because the benchmarking procedures being considered use only the third month of each quarter of QCEW data. It is important to note that the data shown in this figure present differences between *benchmarked* CES national data and QCEW data; that is, the CES data series have already been corrected for errors in the level of employment. Therefore, what this figure shows are the differences in CES and QCEW data after the level correction has taken place. Further note that the data labeled as “QCEW (sum-of-state)” are the sum-of-the-states benchmarked CES data. Because this is, by design, QCEW data plus noncovered employment, these data serve as a high-quality proxy for over-the-quarter employment change in QCEW data. The figure graphs the difference in over-the-quarter change between QCEW and benchmarked CES levels. It presents the different quarters highlighted by different colors (bars), as well as the average difference for each quarter over time (shown by the dashed lines). As the figure shows, the average over-the-quarter difference between QCEW and CES data for the first quarter is about −447,000. That is, the QCEW regularly shows a loss of employment that is much larger than the CES loss of employment from December to March. This difference is about 3.9 standard errors (for a CES 3-month change). Given that this large and significant discrepancy is a regular difference between QCEW and CES levels, we interpret this as a normal seasonal difference between these data series, not as sampling error for the CES data. Large seasonal differences occur for other quarters. For example, the third quarter difference is 292,000 (or 2.5 standard errors), and the fourth quarter difference is 261,000 (or 2.3 standard errors).

Figure 1. The difference in the over-the-quarter change between Quarterly Census of Employment and Wages data and Current Employment Statistics data, not seasonally adjusted



Source: U.S. Bureau of Labor Statistics.

One impact of these large seasonal differences is evident in the state CES data because of the replacement of CES data with QCEW data during each annual benchmark. This process results in two different seasonal patterns in every state CES not seasonally adjusted series: a QCEW seasonal pattern from the benchmark point back and a CES seasonal pattern from the benchmark point forward. This issue is corrected in seasonally adjusted data by the use of the two-step seasonal adjustment process.[21] A less obvious but important impact of these large seasonal differences lies with the interpretation of the over-the-year changes. If seasonally adjusted data are available, then an analyst may appropriately interpret the over-the-year change by using the 12-month employment difference in the seasonally adjusted series. However, if a seasonally adjusted series is not available, then the analyst must use caution when interpreting over-the-year change if the analysis crosses the splice point.[22] Here, the value calculated is actually the sum of the over-the-year change plus a value associated with the difference in seasonality between the CES and QCEW components.

These large, regular seasonal differences stymied historical research efforts to define a benchmark procedure that would address the differences in an acceptable way. Basically, historical CES research efforts focused on the exploration of alternative procedures to adjust the not seasonally adjusted CES data to not seasonally adjusted QCEW data. Each of those efforts ended with a recommendation not to implement those methods because CES data quality did not improve. Only after those straightforward efforts failed did we turn to an examination of procedures that more directly address the seasonal differences.

Recent research

BLS has conducted periodic research on CES benchmarking for several decades. Much of that research has explored alternative benchmark procedures that focus on not seasonally adjusted data. For example, research has applied the state replacement method at the national level and similarly has applied the national wedge procedure for states to explore the errors that are incorporated into the CES estimates. After several iterations of this research, we have concluded that neither of these methods is optimal.

What makes a benchmark procedure optimal or what features of a benchmark procedure are most important are useful questions to consider. Originally, the focus was simply on which method was “correct” without defining what that meant. More recently, I have identified a set of features (listed below) that I consider important in determining the usefulness of a new procedure.

1. The procedure must align to the March QCEW level.
2. The procedure must have a well-defined mathematical relationship to the third month of each quarter of the QCEW.
3. The procedure must align the CES data to QCEW data more than once per year.
4. The procedure must address the different seasonal patterns in the data and maintain the CES seasonal pattern.
5. The procedure must address the small-domain problem.
6. The procedure must be explainable from a statistical and economic perspective.

Element 1 requires the alignment of the CES level to the March QCEW level. This requirement is consistent with both the national and state benchmarking procedures, and I consider this a critical requirement for any future procedure. Element 2 refers to a set of procedures that create a relationship between CES and QCEW data for other quarters. One such relationship is the direct alignment to the third month of each quarter that is consistent with the current state procedure. Element 3 is self-explanatory; the procedure must align the data more than once, which means either twice per year or four times per year. Element 4 indicates that the procedure must account for the different seasonal patterns between these data sources and maintain the CES seasonal pattern where the domain is large enough to support its identification. Element 5 indicates that the procedure must account for the small-domain problem. This is the problem that led to the use of the current replacement method for states. This problem is focused on small domains with the smallest samples. Small domains that also have very small samples may have very large relative sampling errors. The problem is that aligning the data once per year or even for the third month of each quarter in these domains can leave very large and uncorrected sampling errors in place for other months. A new procedure must explicitly include a method to mitigate these errors. And finally, element 6 notes that a new procedure must make sense and be relatively easy to explain to data users from both a statistical and economic perspective. While not a criterion for evaluation, a new procedure will be most useful if it can be applied consistently to both national and state data.

We can evaluate alternative procedures against these criteria to see how well they achieve this set of goals. To frame that evaluation, I start with an evaluation of the current procedures.

This is an evaluation of the national procedure:

- ☑ Aligns the CES data to the QCEW data in March
- ☒ Has a well-defined mathematical relationship with the third month of other quarters
- ☒ Aligns CES and QCEW employment more than once per year
- ☑ Maintains the CES seasonal pattern
- ☒ Addresses the small-domain problem
- ☑ Is easily explained

Therefore, the national procedure meets only three of our six criteria.

This is an evaluation of the state procedure:

- ☑ Aligns the CES data to the QCEW in March
- ☑ Has a well-defined mathematical relationship with the third month of other quarters
- ☒ Aligns CES and QCEW employment more than once per year
- ☒ Maintains the CES seasonal pattern
- ☑ Addresses the small-domain problem
- ☑ Is easily explained

The state procedure meets only four of our six criteria.

As stated earlier, neither the national nor state procedures currently used is optimal, and with this set of defined criteria we can evaluate the usefulness of proposed alternatives. I now turn to a brief description and evaluation of procedures reviewed during our most recent round of research. Each of the new procedures is quarterly, and each procedure aligns directly to the March QCEW level, so they meet at least two of our six criteria by definition.

The procedures proposed are (1) Over-the-Year-Change Difference,[23] (2) Weighted Model, and (3) Seasonally Adjusted Difference in Employment Method. These procedures are described below.

The Over-the-Year-Change Difference procedure takes the difference in over-the-year change for the third month of each quarter in CES and in QCEW levels and aligns the CES over-the-year change to the QCEW over-the-year change. It was initially thought that this method created an appropriate relationship between CES and QCEW levels for each quarter. However, further research identified a problem: the seasonal pattern of the first year is "baked in" to the CES survey because this procedure does not account for evolving seasonality. "Baked in" means the seasonality in the start period—that is, whatever year is selected to start the over-the-year difference calculation—becomes the expected seasonality for both series. If CES or QCEW seasonality evolves from what was present in that start period, then there may be residual seasonality errors not accounted for by this method.

The basic equation describing this procedure is shown below. Note that this depiction can be extended to show the linkage to previous adjustments, which ultimately lead to the starting-point issue described above.

$$\hat{E}_t^B = \hat{E}_t + \left((\mathrm{QCEW}_t - \mathrm{QCEW}_{t-12}) - \left(\hat{E}_t - \hat{E}_{t-12}^B\right)\right), \quad (4)$$

where

$\hat{E}_t^B$ is the benchmarked all-employee level for month *t*, where *t* is June, September, or December,

$\hat{E}_t$ is the not-yet-benchmarked all-employee estimate for month *t*, and

QCEW_t are the data from the QCEW (with noncovered employment added) for month *t*.

This is an evaluation of the Over-the-Year-Change Difference procedure:

☑ Aligns the CES data to the QCEW in March

☒ Has a well-defined mathematical relationship with the third month of other quarters

☑ Aligns CES and QCEW employment more than once per year

☒ Maintains the CES seasonal pattern

☑ Addresses the small-domain problem

☑ Is easily explained

Therefore, the Over-the-Year-Change Difference procedure meets only four of our six criteria. Unfortunately, this procedure fails to appropriately account for the seasonal differences in the data.

The Weighted Model procedure develops a weighted function of QCEW and CES levels for the third month of each quarter. The weighting for March is 100 percent of the QCEW level, and the weighting for the other quarters could be determined on the basis of objective criteria using historical data.

The basic equation describing this procedure is as follows:

$$\hat{E}_t^B = \left(w_{1,t}\hat{E}_t\right) + \left(w_{2,t}\mathrm{QCEW}_t\right), \quad (5)$$

where

$\hat{E}_t^B$ is the benchmarked all-employee level for month *t*, where *t* is June, September, or December,

$\hat{E}_t$ is the not-yet-benchmarked all-employee estimate for month *t*,

QCEW_t are the data from the QCEW (with noncovered employment added) for month *t*,

$w_{1,t}$ is the weight applied to the CES data for month *t*, where *w* is between 0 and 1, and

$w_{2,t}$ is the weight applied to the QCEW data, where $w_{2,t} = 1 - w_{1,t}$.

Note that there are many functions that could be used to generate the weights.

This is an evaluation of the Weighted Model procedure:

☑ Aligns the CES data to the QCEW in March

☑ Has a well-defined mathematical relationship with the third month of other quarters

☑ Aligns CES and QCEW employment more than once per year

☑ Maintains the CES seasonal pattern

☑ Addresses the small-domain problem

☒ Is easily explained

Therefore, the Weighted Model procedure meets five of our six criteria. As such, this procedure likely would address the seasonal differences in the data sources and allow for maintaining the CES seasonal pattern. However, we did not pursue this method with extensive research because of another consideration: we were unsure how to explain the results from an economic perspective. That is, what would the result mean? It would not be the CES linked to the QCEW in an easily explained economic perspective. Also, the explanation that the function would be selected to reduce the error in the CES level and to limit change in the CES seasonal pattern was a less clear explanation than we would hope to have for these highly scrutinized data.

The final method included in recent research is the Seasonally Adjusted Difference in Employment Method. This method seasonally adjusts the QCEW and CES levels. It then takes the difference for each of these at the third month of the quarter after aligning the not seasonally adjusted CES level to the March not seasonally adjusted QCEW level.[24]

The basic equation describing the Seasonally Adjusted Difference in Employment Method is as follows:

$$\hat{E}_t^B = \hat{E}_t + \left(\mathrm{QCEW}_t^{SA} - \hat{E}_t^{SA}\right), \quad (6)$$

where

$\hat{E}_t^B$ is the benchmarked all-employee not seasonally adjusted level for month *t*, where *t* is June, September, or December,

$\hat{E}_t$ is the not-yet-benchmarked not seasonally adjusted all-employee estimate for month *t*,

QCEW_t^{SA} is the seasonally adjusted QCEW data (with noncovered employment included) for month *t*, and

$\hat{E}_t^{SA}$ is the CES seasonally adjusted all-employee estimate for month *t*.

This is an evaluation of the Seasonally Adjusted Difference in Employment Method:

☑ Aligns the CES data to the QCEW level in March

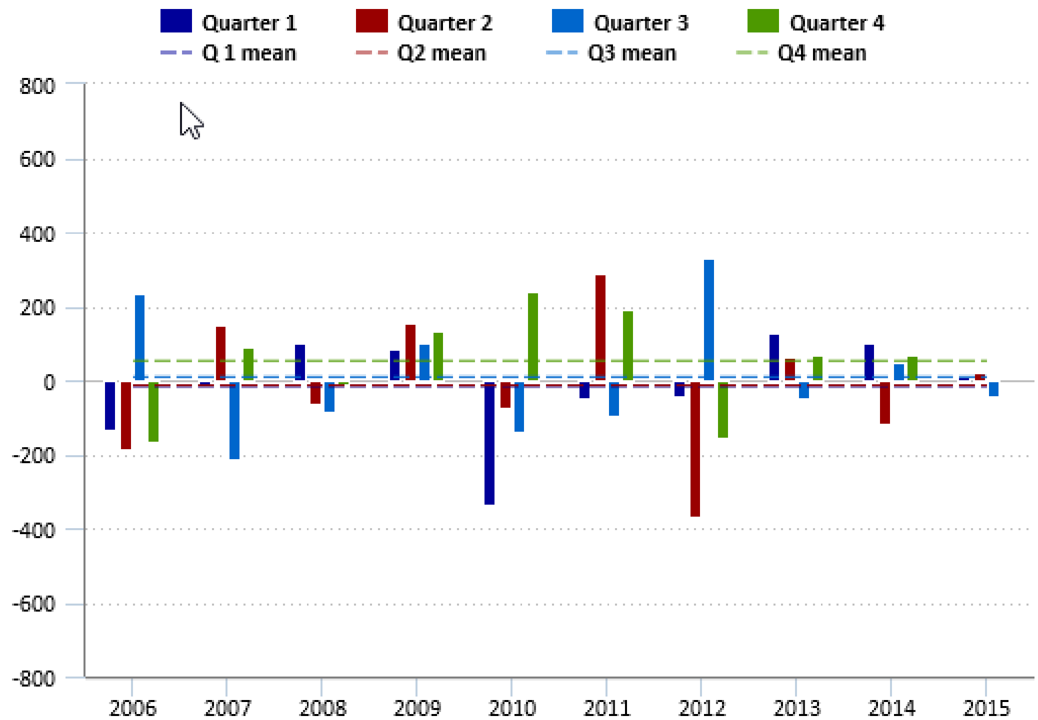
☑ Has a well-defined mathematical relationship with the third month of other quarters

- ☑ Aligns CES and QCEW employment more than once per year
- ☑ Maintains the CES seasonal pattern
- ☑ Addresses the small-domain problem
- ☑ Is easily explained

The Seasonally Adjusted Difference in Employment Method directly addresses and appropriately accounts for the seasonal differences in the source data. This procedure may not directly address the issue of a small domain with small sample, but it does allow that problem to be solved through several potential supplemental procedures. This method also is relatively easy to explain. Therefore, this procedure meets at least five of our six criteria and potentially meets all six. Preliminary results strongly suggest that this method is viable and that it leads to improvements in the quality of CES data. More detail on this method is available in an article by Matthew Dey and Mark Loewenstein, and in a technical paper by the same authors.[25] Furthermore, this method appears able to be developed in a manner that allows a consistent application for both national and state data.

Comparing the difference between figure 1 and figure 2 is informative. As described earlier, figure 1 shows the differences, after benchmarking, between QCEW data and CES data. The differences are large and seasonal. Figure 2 shows differences in benchmarked CES and QCEW data that have been seasonally adjusted. The differences after seasonal adjustment are small and irregular. This simple change illustrates why we expect this method to be successful: it will correct for differences between the two data series after accounting for the different seasonal patterns. That is, this method extracts from the two series the differences that are primarily due to sampling and nonsampling error rather than deviations due to large seasonal differences. Hence, this method provides an appropriate correction to the CES data series.

Figure 2. The difference in the over-the-quarter change between Quarterly Census of Employment and Wages data and Current Employment Statistics data, seasonally adjusted



Note: The Quarterly Census of Employment and Wages data used here are the sum of the states' benchmarked Current Employment Statistics data.
Source: U.S. Bureau of Labor Statistics.

Future research

While the CES program staff (and other sophisticated CES data users) have long understood the substantial seasonal differences between QCEW and CES data, extensive research led us to conclude that there is no good way to address this issue on a not seasonally adjusted basis. Our research shows that a new benchmark procedure must directly address the differences in seasonality between the CES and QCEW levels by aligning data on a seasonally adjusted basis.

Future work on the Seasonally Adjusted Difference in Employment Method falls into two major categories. The first is additional work to verify that the applied results of this method meet the theoretical expectations. The rigorous evaluation of new methods in extensive simulations is the usual practice for BLS. Any changes to this important economic indicator would be made only after thorough research and evaluation has been completed.

The second category of future work is to catalog series in small domains with smaller sample sizes and then evaluate potential methods to produce a benchmarked series for them. The problem of small sample size is most obvious in smaller geographic domains but can also occur in some of the national series with the smallest employment. Two potential methods to explore are (1) a ratio adjustment of not seasonally adjusted data for very small domains to a higher level seasonally adjusted domain for which the seasonally adjusted data meet selected quality characteristics and (2) a method based on a function of the reliability of the seasonal adjustment.

Regarding the first method, it should be relatively easy to design a procedure that replaces CES data with QCEW data for the smallest domains. The procedure would explicitly ratio adjust the QCEW data series and other data series so that they are additive to higher level aggregates because the seasonal adjustment process is more reliable for aggregates. The ratio adjustment would prorate the aggregate CES seasonality down to these smaller domains, with larger shares of that seasonality being allocated to data series with larger employment levels. This process would result in a dataset with the required additivity feature. The process would also highly weight the original QCEW data so that the sampling error component of the small-domain estimate is minimized. The second method is more complex. This method would require the development of a function that weights the seasonally adjusted CES and QCEW levels on the basis of the reliability of the two seasonally adjusted series. If the quality of both adjusted series is felt to be high, then the difference would be fully accepted. If the quality of the seasonal adjustment is low, then the difference might be discounted in a function that places more weight on the not seasonally adjusted QCEW data. An obvious problem with this method is that it will not have additivity to higher level aggregates as a built-in property. In this case, additivity must be provided through a second step, perhaps a ratio adjustment procedure similar to the one described earlier. Other methods for solving this problem likely exist and will be included in future research.

Concluding remarks

The CES survey has a long history, originating in 1915. Benchmarking for CES has nearly as long a history, first occurring in 1935. Over this history, and especially more recently, various user groups have expressed much interest in improving the quality of CES data by improving the timeliness of the benchmark process. As a result, the CES program has returned repeatedly to this research theme.

Earlier research efforts focused on the possibilities of finding an alternative procedure that aligned not seasonally adjusted CES data to not seasonally adjusted QCEW data. These efforts were ultimately unsuccessful because they did not account for the different seasonal patterns in CES and QCEW series, and a procedure could not be found that was convincingly an improvement. Recent research has reconfirmed that there is no obvious procedure to align not seasonally adjusted CES data to not seasonally adjusted QCEW data on a quarterly basis that adequately addresses the known seasonal differences between these independent data series.

The most recent research effort has focused on methods that do account for the different seasonal patterns. This research has found a method that appears to solve this longstanding roadblock by benchmarking to the difference in CES and QCEW levels after seasonally adjusting the two series. This procedure correctly accounts for the different seasonal patterns by removing the seasonality from the data and then benchmarking the not seasonally adjusted data to the difference that remains between the seasonally adjusted series.

This most recent round of research has focused on evaluating both national and state data at the North American Industry Classification System supersector level. Future research will focus on further evaluation of this method to ensure that it meets expectations at more detailed industry and geographic levels. Additionally, future research will evaluate alternative methods to solve the small-cell issues.

CES data are frequently used as an economic indicator. Moreover, these data are used as an input into other economic indicators. Therefore, research must be especially thorough and well documented, and it is typically shared with the economic community for comment before BLS makes significant changes in CES methodology. BLS plans to continue this research, and we will continue to share the results with the public via our advisory groups and through publication in various venues. [26]

ACKNOWLEDGEMENT: I would like to acknowledge that the success of this research is due to the work of a number of BLS employees who have participated in bringing new research ideas to the table. The team includes members from the CES program office, our statistical methods group, our economic research group, and our survey methods research group. The founding members of the team include Kenneth Robertson, Chris Manning, Kirk Mueller, Steven Mance, Greg Erkens, Larry Huff, Mark Loewenstein, Matthew Dey, Polly Phipps, Daniell Toth, and David Talan. Additional members have joined the team since its inception and contributed to these results.

SUGGESTED CITATION

Kenneth W. Robertson, "Benchmarking the Current Employment Statistics survey: perspectives on current research," *Monthly Labor Review,* U.S. Bureau of Labor Statistics, November 2017, https://doi.org/10.21916/mlr.2017.27.

NOTES

1 The BLS *Handbook of methods* (https://www.bls.gov/opub/hom/) is a compendium of detailed information about how BLS collects and prepares its economic data. The second chapter of the handbook (https://www.bls.gov/opub/hom/pdf/homch2.pdf) is devoted to providing detailed information on the methods, procedures, and estimators used by the Current Employment Statistics survey. The handbook discusses in detail the estimator just described and many other aspects of the CES survey.

2 Sampling error occurs because a measurement is made from a sample rather than a population. If a probability-based sample design is used, then the sampling error can be estimated and a confidence interval can be calculated. The confidence interval describes the bounds that the true value of the measured characteristic is expected to fall in, with a corresponding level of confidence. Nonsampling errors include a number of different types of error, such as recall error, classification error, coding error, transcription error, processing error, etc.

3 John P. Mullins, “One hundred years of Current Employment Statistics—an overview of survey advancements,” *Monthly Labor Review*, August 2016, https://doi.org/10.21916/mlr.2016.2.

4 The term QCEW is used for two purposes in this article. When we talk about the program, QCEW refers specifically to the Quarterly Census of Employment and Wages program. When used in the context of CES benchmarking, the term is used to represent the CES benchmark population values, or QCEW plus noncovered employment.

5 In 2003, the CES program redesigned the survey to use modern probability sampling procedures, after which the usual size of revisions was smaller. Prior to the redesign, the state and metropolitan area benchmark revisions tended to be relatively larger than national revisions.

6 The Census Bureau’s *County Business Patterns* program produces data by county with a lag of about 2 years. Information about the program and its data can be found at https://www.census.gov/programs-surveys/cbp.html.

7 Information about the Railroad Retirement Board can be found at https://secure.rrb.gov/.

8 The postbenchmark period is the months following the benchmark, April through October for national estimates and October through December for state estimates.

9 Christopher D. Manning and John R. Stewart, “Benchmarking the Current Employment Statistics national estimates,” *Monthly Labor Review*, October 2017, https://doi.org/10.21916/mlr.2017.25. National annual benchmark information can be found at https://www.bls.gov/ces/tables.htm#benchmark.

10 Kirk J. Mueller, “Benchmarking the Current Employment Statistics state and area estimates,” *Monthly Labor Review*, November 2017, https://doi.org/10.21916/mlr.2017.26. State and area benchmark information for 2016 can be found at https://www.bls.gov/sae/benchmark2016.pdf.

11 Early benchmarking on a quarterly basis is discussed in several articles by members of the Federal Reserve banks, including Keith Phillips (Dallas Federal Reserve Bank at San Antonio) and Thomas Walstrum (Chicago Federal Reserve Bank). See http://www.dallasfed.org/research/basics/benchmark.cfm and https://www.chicagofed.org/.

12 The author would like to thank Dr. Keith Phillips for reviewing and commenting on this section to ensure that the description captures the methodology appropriately.

13 ARIMA X-13 SEATS is publically available from the U.S. Census Bureau at https://www.census.gov/srd/www/x13as/.

14 Franklin D. Berger and Keith R. Phillips, “Reassessing Texas employment growth,” *The Southwest Economy*, July/August 1993, http://www.dallasfed.org/assets/documents/research/swe/1993/swe9304a.pdf.

15 Ibid.

16 Response-analysis surveys comparing CES and QCEW responses were conducted in 1994, 2001, and 2008.

17 Jeffrey A. Groen, “Seasonal differences in employment between survey and administrative data,” Working Paper 443 (U.S. Bureau of Labor Statistics, February 2011), https://www.bls.gov/osmr/research-papers/2011/pdf/ec110010.pdf.

18 By correct, I mean collecting the number of employees that meets the CES definition, not including employees who do not meet that definition, and reporting the data for the specified reference period rather than a different payroll period.

19 The reference period for the CES survey, QCEW, and most federal business surveys is the week or the pay period that includes the 12th day of the month.

20 Note that this assumption is for illustrative purposes only, as it ignores all other errors in the QCEW data and assumes that all error is related to the reporting period.

21 For a description of the two-step seasonal adjustment process, see Patricia Coil, Taylor Le, and TJ Lepoutre, "Revisions in state establishment-based employment estimates effective January 2016" (U.S. Bureau of Labor Statistics),p. 4, https://www.bls.gov/sae/benchmark2016.pdf.

22 The splice point is the point where benchmarked CES data, which is essentially QCEW data for states, meets data that have not been benchmarked, which are purely CES-based data.

23 An earlier, smaller study of several of these methods was conducted by Richard Valliant and Jill Dever in 2008. That research focused on the over-the-year change difference, a seasonally adjusted difference, and a Lowess estimator.

24 To be specific, what is seasonally adjusted on the QCEW side are QCEW data adjusted to meet the CES scope, which includes data for employment not covered by QCEW. These adjusted QCEW data are aggregated to match the industries and geographies published by CES, and those QCEW aggregates are seasonally adjusted in this procedure.

25 Mark A. Loewenstein and Matthew Dey, “A quarterly benchmarking procedure for the Current Employment Statistics program,” *Monthly Labor Review*, November 2017, https://doi.org/10.21916/mlr.2017.28.

26 BLS has two advisory groups involved with this issue, a Technical Advisory Committee (TAC, see https://www.bls.gov/advisory/taccharter.htm) and the Data User Advisory Committee (DUAC, https://www.bls.gov/advisory/duac.htm).

RELATED CONTENT

Related Articles

Benchmarking the Current Employment Statistics survey: the past, present, and future, *Monthly Labor Review*, October 2017.

Benchmarking the Current Employment Statistics national estimates, *Monthly Labor Review*, October 2017.

Benchmarking the Current Employment Statistics state and area estimates, *Monthly Labor Review*, November 2017.

A quarterly benchmarking procedure for the Current Employment Statistics program, *Monthly Labor Review*, November 2017.

Related Subjects

Survey methods | Benchmarking | Employment | Survey procedures | Industry studies

Monthly Labor Review

November 2017

Full employment: an assumption within BLS projections

BLS defines full employment as an economy in which the unemployment rate equals the nonaccelerating inflation rate of unemployment (NAIRU), no cyclical unemployment exists, and GDP is at its potential. The full-employment assumption links BLS projections to an economy running at full capacity and utilizing all of its resources.

Kevin S. Dubina
dubina.kevin@bls.gov

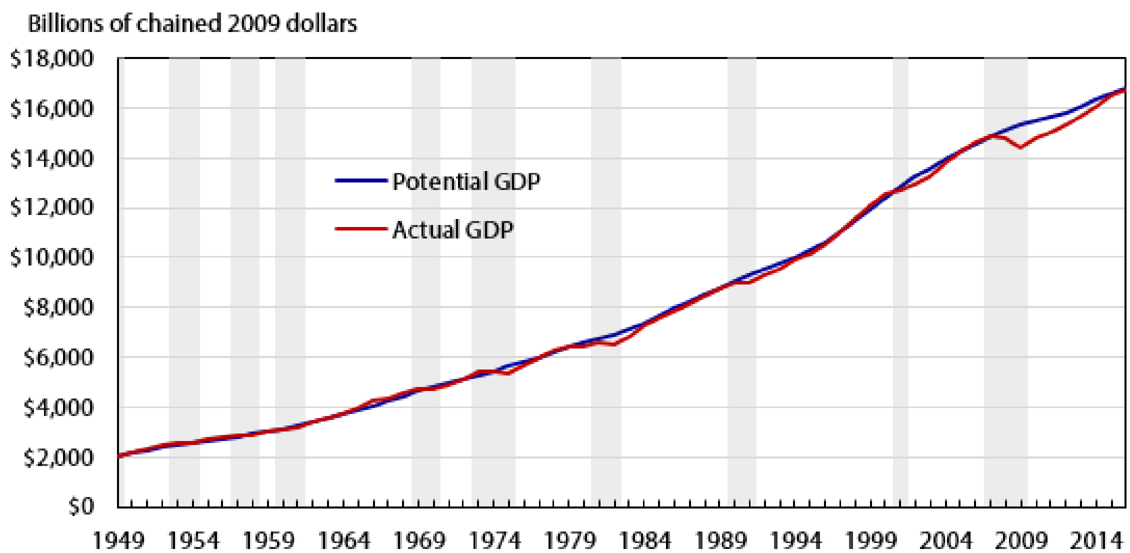
Kevin S. Dubina is an economist in the Office of Occupational Statistics and Employment Projections, U.S. Bureau of Labor Statistics.

The Bureau of Labor Statistics (BLS) develops employment projections of the labor market at the national level covering 10 years into the future. These projections are invaluable for career-planning, training-and-education, and policy-planning purposes. Both jobseekers and those transitioning to a new career are provided valuable information that allows them to prepare for jobs that will be the most in demand. National projections also are valuable inputs for states, assisting them in producing projections that will support the allocation of resources for job training.

The projection process starts with an estimation of the labor force at the 10-year horizon.[1] Given the supply of labor, macroeconomic aggregates—including total employment, output, productivity, prices, interest rates, and many other variables—are projected.[2] This macroeconomic outlook depends on the assumption that the economy will be at full employment in the projection year. At full employment, unemployment will be at the nonaccelerating inflation rate of unemployment (NAIRU) and output will be at potential.[3] In its projections, BLS assumes full employment in order to minimize any effect from cyclical fluctuations, focusing instead on structural changes to the economy. In addition to projecting the macroeconomic outlook, BLS projects employment for hundreds of occupations[4] and industries, as well as output for hundreds of industries.[5] These employment and industry output projections are estimated separately from the projections for the macro model[6] and are benchmarked to the macro model's aggregate projections in order to ensure that BLS publishes a consistent view of the economy at full employment.[7]

A full-employment economy is associated with *potential output*: the sustainable-trend growth of output (usually expressed as gross domestic product, or GDP; see figure 1) that occurs while the economy is operating at a high rate of resource use. Actual output may be above or below potential output at any single point in time. An economy that is operating above potential output is said to be *overheating*, which occurs when aggregate demand exceeds productive capacity—a situation that is unsustainable in the long term. By contrast, being below potential output does not necessarily mean that an economy is in recession: the economy could be expanding, but not yet fully recovered from a recession. In general, the economy bounces back and forth between recessionary and overheating periods, correcting itself when it gets too far from its potential output.

Figure 1. Gross domestic product (GDP) and potential GDP, 1949–2016



Note: Shaded areas indicate recessions, as determined by the National Bureau of Economic Research.
Sources: Gross domestic product: U.S. Bureau of Economic Analysis; potential gross domestic product: U.S. Congressional Budget Office.

There are two types of factors that cause employment levels to change: *cyclical factors* and *structural factors*.[8] Cyclical factors are relatively short-term deviations from potential that have their origin in business cycle fluctuations, the normal up-and-down movements of the economy as it cycles through booms and recessions over time.[9] Anticipating long-term cyclical changes is not supported by economic theory. Structural factors are changes that have their origin in long-lasting, often permanent, changes in the economy, such as automation or demographic changes. These types of changes can be anticipated and are likely to matter to BLS users, because they have long-term effects on prospective careers and certain types of policy. The full-employment assumption puts the focus on longer term employment changes that are due to structural factors.

The full-employment assumption works by benchmarking BLS projections to the same objective level of output capacity—full, or 100-percent, capacity—each publication round. Without the same objective level of output capacity linked to employment levels, structural effects can be masked or incorrectly perceived, based on changes in the business cycle. Setting the objective level to full capacity is the means by which BLS projections exclude cyclicality; levels are therefore determined entirely by structural factors. Moreover, consistently using the same level of output capacity has the added advantage of allowing changes to be evaluated between different projection release dates. Such evaluation can reveal how the economy has changed structurally between two periods. A consistent objective benchmark for BLS projections is an important tool for identifying structural changes to the economy and is implemented via the full-employment assumption. Being open and explicit about this methodology helps users interpret projections advantageously and use them accordingly.

The assumption of a full-employment economy in the target projection year can have important implications for the BLS projected growth rates and requires one to take into account current cyclical economic conditions when utilizing growth rates. Large cyclical deviations in the base year of projections can influence growth rates: under such conditions, some of the projected growth will be due to cyclical factors.[10] In 2010, the year after the most recent recession ended, a year in which the economy was still well below its potential output, BLS published a projected real GDP growth rate of 3.0 percent over the next decade. However, the projected growth rate of potential GDP from 2010 (published by the U.S. Congressional Budget Office) to 2020 was 2.3 percent. This difference implies that 0.7 percent of the published BLS growth rate could be attributed to cyclical factors in the 2010 base value. (See table 1.)

Table 1. Selected levels of gross domestic product (GDP) and their accompanying growth rates (billions of chained 2005 dollars)

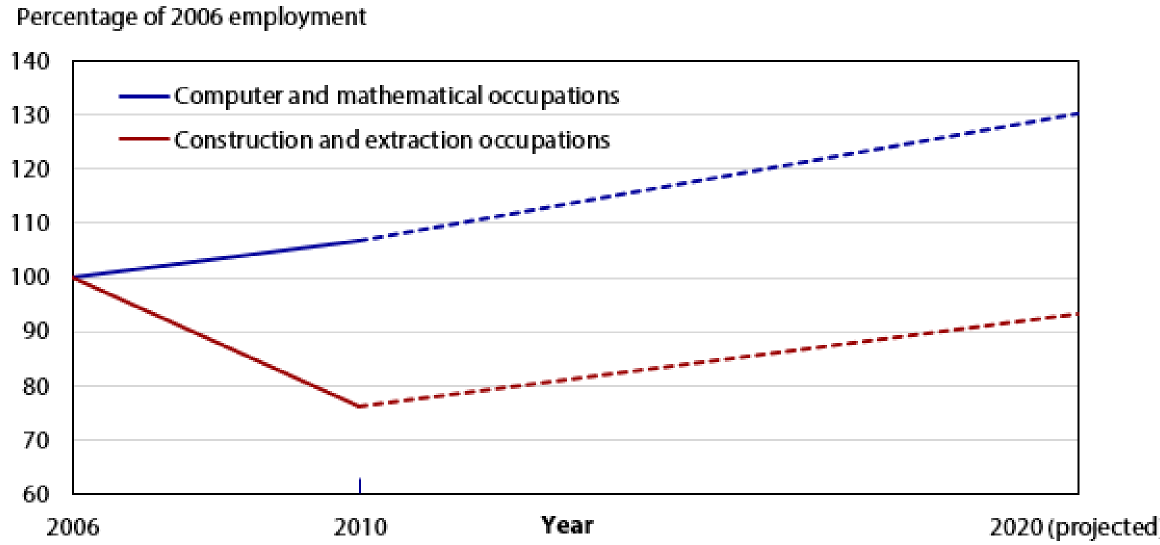
Type of GDP	2010	Projected, 2020	Growth rate
Actual	13,088.0	17,512.9	3.0
Potential	14,017.1	17,512.9	2.3

Note: Potential GDP for 2010 is from U.S. Congressional Budget Office, https://www.cbo.gov/sites/default/files/recurringdata/51137-2011-01-potentialgdp.xls.

Sources: U.S. Bureau of Labor Statistics (2010), U.S. Congressional Budget Office.

The same dynamic occurs in individual occupation and industry projections. Current full-employment estimates at individual occupation or industry levels do not exist. Therefore, it is impossible to estimate precisely how much of a growth rate is attributable to cyclicality. However, it is often clear that some cyclical effects are present. Figure 2, from BLS 2010 projections, demonstrates their presence. Both construction and extraction occupations and computer and mathematical occupations had projected growth rates of 2.0 percent from 2010 to 2020. Upon examination of how much lower the percentage of employment was for construction and extraction occupations in the 2010 base projection year compared with 2006, it becomes clear that most, or even all, of the growth is attributable to cyclical factors. Conversely, employment growth in computer and mathematical occupations over the same timeframe appears to be mostly structural.

Figure 2. Employment trends for selected occupational groups, 2006–20



Note: For each projection cycle, BLS publishes an estimate of GDP for the target year only. The lines shown in the graph connect historical employment trends from 2006 and 2010 with the published BLS projections for 2020 and are not necessarily representative of the expected path of the economy in the intervening years.
Source: U.S. Bureau of Labor Statistics, 2010.

History of full employment

There is no method for directly measuring full employment. Rather, it is a theoretical state within the economy that must be estimated. The idea of full employment has evolved over time. Some unemployment is expected, even desired, in a healthy economy as workers transition to new jobs better matching their skill sets or as they acquire and develop additional skills needed by an employer. These transitions to new jobs and acquisitions of more skills boost the workers' productivity and the capacity of the economy as a whole. Newer concepts of full employment *require* some unemployment to temper inflation.

Earlier views, such as those held by Keynes[11] and Beveridge,[12] defined full employment as an economy in which the only unemployment was due to workers transitioning from one job to another. This kind of unemployment is known as "frictional unemployment." Later, "structural unemployment" was added to the definition of full employment. Structural unemployment is the unemployment of workers who currently lack the skills needed by employers. Both frictional unemployment and structural unemployment involve workers who are searching for a job, but, in order to secure a position, workers who are structurally unemployed must do something in addition to just searching. Some must pursue further education or training, while others must relocate to a different geographical area.

Another, more recent view of full employment links employment to inflation. This theory holds that some unemployment is necessary to control inflation. The relationship between the unemployment rate and the rate of inflation is demonstrated in the Phillips curve.[13] The level of unemployment required to keep inflation in check is the NAIRU, and most economists today mean the NAIRU when referring to full employment.

Just as an equilibrium exists in the goods-and-services market when supply equals demand, there exists an equilibrium in the labor market when labor supply equals labor demand. At this equilibrium, workers are hired at the prevailing wage and inflation is stable. That is the reason the equilibrium is often referred to as the NAIRU. Shocks to the economy knock the labor market out of equilibrium. If the shock is negative, additional unemployment will emerge as labor demand decreases. Wages generally remain sticky and can't adjust, amplifying the decrease in labor demand further. If this change in labor demand is temporary, it is considered cyclical unemployment; if it is permanent, the change is known as structural unemployment. Sources of structural changes in labor demand include technology, outsourcing, and demographic changes to the population, among others.

Regardless of whether a shock is temporary or permanent, the U.S. Federal Reserve has some ability to trade higher inflation for lower unemployment in the short run.[14] Higher inflation effectively lowers wages so that the demand for labor at the new real wage is higher. This relationship between unemployment and inflation is suggested by the Phillips curve. However, the relationship breaks down over time: individuals begin to expect higher inflation because of increases in prices and wages, and these increases must come at an ever-accelerating rate to support an artificially low unemployment rate. Absent this accelerating inflation, the labor force will return to equilibrium at the NAIRU.

The NAIRU is not a constant for all periods and locations. As discussed earlier, if a shock to labor demand is temporary—a cyclical change—unemployment will return to its preshock level. However, if the shock to labor demand is permanent—a structural change—it will result in a change to the NAIRU and the unemployment level will drift to this new equilibrium. Shocks also affect local economies to varying degrees: some areas have benefited more from recent technological advances and have lower NAIRUs, while others have been hurt more by outsourcing and have higher NAIRUs.[15] The degree to which various shocks are structural or cyclical is often a matter of debate, and economists can change their view with time. However, without a doubt, changes to labor demand come from both cyclical and structural factors.

An economy with both the highest level of sustainable employment and stable inflation can still be an economy in which only frictional and structural unemployment exist. Indeed, this is the full-employment assumption that BLS makes: BLS projections assume that the unemployment rate is equal to the NAIRU (implying that inflation is consistent with the Federal Reserve's target value) and any existing unemployment is frictional or structural.

Estimating a full-employment economy

In addition to including the NAIRU, recent full-employment views link a full-employment economy to potential output. Full employment and a corresponding value for potential output cannot be measured directly. Instead, a number of methods are available for estimating historical values and projecting future values. This section summarizes the most common methods.

Growth accounting methods model output as a production function incorporating the various factors affecting output. Input factors often are adjusted to their potential. One common method for adjusting inputs to their potential uses the relationship between inflation and output (Okun's law), as well as estimates of the NAIRU. The most popular growth accounting model is the *Solow model*, which attributes GDP growth to growth in labor, capital, and technological progress. A *labor productivity growth model* is favored by those who believe that capital is too difficult to measure accurately. This kind of model ignores capital and estimates output as a function of labor and labor productivity alone. Growth accounting methods assume that the magnitude of the contribution that each factor makes to growth remains the same over time.[16]

Statistical filtering techniques, such as the *Hodrick–Prescott filter*, extract from the data underlying trends that represent an economy's full employment or potential output. The drawback to these techniques is that the trend extracted is not benchmarked to any external measure of capacity. Consequently, the estimates obtained are not necessarily consistent with stable inflation. In addition, they are subject to end-of-sample problems, in which the trend becomes more responsive to temporary fluctuations in the data toward the end of a sample.[17]

Systems of equations can be specified that estimate variables such as output, employment, and inflation together. Although more complex than either growth accounting methods or statistical filtering techniques, these systems allow the contributions of different factors to GDP to vary over time. Systems of equations often include multivariate time series models, such as vector autoregression (VAR).

Estimating BLS aggregate demand projections

For its macroeconomic projections, BLS uses a system of equations that constitute a structural econometric model of the U.S. economy. The software involved is licensed from Macroeconomic Advisers (MA), LLC. The system includes a growth model, as well as a VAR that is used for capital projections. BLS uses its own labor force outlook and its own assumptions about total factor productivity (TFP) and the NAIRU as an input into the MA model. Labor interacts with capital and a TFP residual to produce output. Therefore, output is determined not just on the basis of labor but also on that of capital services and potential TFP.[18]

Because capital is a factor in potential output, including it is pertinent to the model, despite the difficulty that arises in measuring it. *Capital stock* measures the value of capital in the economy; *capital services*, which stem from capital stocks, measure the contribution of capital to the production process. Using a capital stock measure weights two pieces of capital the same if they have the same market value, even if their contributions to production are unequal. Because it is the contribution to production that influences both output and the amount of labor necessary to produce that output, a measure of capital services is preferred.

TFP is any growth that is not attributable to capital services or labor. TFP is often taken as a measure of an economy's technological change,[19] although other factors, such as economies of scale and increased education in the labor force, can influence TFP. BLS has projected output on the basis of TFP's historic trend, assuming that this trend approximates potential TFP. A growing body of research indicates that current TFP, and therefore likely potential TFP, is lower than it was in previous decades.[20] As a result, BLS places more weight on recent values.

Deciding what input goes into the MA model is just as important as the model itself, to ensure that the projections obtained are as reasonable as possible. Inputs are adjusted on the basis of (1) current economic research, (2) putting them through alternative BLS models to determine a range of projections, and (3) comparisons with models from other government agencies. Projections that take account of the information gleaned from these studies are then compared against the currently specified MA model. Key exogenous data in the model, including the NAIRU and TFP, are overwritten when BLS deems other values more appropriate.

Although BLS is not limited to any one internal model, the most commonly used model is a Cobb–Douglas production function (part of the Solow growth model discussed earlier). This production function is not as complex as the MA model, but it allows BLS to identify how changes to a single input factor affect output. Within a Cobb–Douglas model, output is a function of capital, labor, and TFP. The model can be expressed as

$$Y = AL^{0.7}K^{0.3},$$

where Y = output, A = TFP, L = labor, and K = capital.

The model itself is relatively straightforward. However, identifying what to include for each of the input factors—TFP, labor, and capital—is not. Each input can be estimated any number of ways, all of which likely introduce some degree of measurement error. The labor input has the most objective measurements, but it is still necessary to determine which is most appropriate: employment levels, average hours worked, or some combination of the two. To estimate potential labor input, the NAIRU or other measures of utilization[21] can be used. One BLS approach is

$$L^* = \text{awh}^* \times [(1 - \text{NAIRU}) \times (\text{lfpr}^* \times \text{cnp})],$$

where L^* = potential labor input, awh* = potential average weekly hours, NAIRU = nonaccelerating inflation rate of unemployment, lfpr* = potential labor force participation rate, and cnp = civilian noninstitutional population.

Capital is more difficult to measure, in part because of the subjectivity of depreciation and, more importantly, the differences in the various types of capital. Capital services indexes offer an easy way of approximating capital, and extending their trends is an appropriate method for projecting capital in the future. Still, it should be noted that these indexes introduce their own set of measurement issues. One is that trends can be estimated in various ways; BLS uses a piecewise linear regression to make its estimations.[22]

Like capital, TFP indexes exist and provide a good starting point for projecting potential TFP. Unlike capital, however, which is largely a function of investment during previous periods, potential TFP is determined largely by an economy's technological change and is more subject to sudden shocks as the pace of technological advances increases or decreases. Determining if and how future potential TFP is likely to deviate from its past trend is thus important.

Alternative models, as well as comparisons with other government agencies, are important for a number of reasons. Grounded in economic theory, alternative models allow for a comparative result against which the MA model can be evaluated. Through these models, BLS is able to gain a better understanding of the U.S. economy and the interactions between different variables—an understanding that allows for subsequent sensitivity analysis. Comparisons with outside sources act as a qualitative check confirming that the results are sound.

Conclusion

Like all projections, those produced by BLS are dependent on underlying assumptions about the future. On the basis of how BLS projections are used, such as for career-planning, training-and-education, and policy-planning purposes, structural changes are more important than cyclical changes. Accordingly, BLS projections contain no cyclical deviations; rather, a full-employment assumption is used that projects only structural changes to the economy. Although other definitions of full employment exist, BLS defines full employment as an economy in which the unemployment rate equals the NAIRU, no cyclical unemployment exists, and GDP is at its potential.

The full-employment assumption links BLS projections to an economy running at full capacity and utilizing all of its resources. This—output at 100 percent of capacity—gives users an objective expectation regarding the part of the business cycle on which a particular projected level is based. Because projections are consistently benchmarked to this same objective level, users are provided a means to interpret what is best for them. It is important that interpretations of BLS growth rates be applied in the context of current economic conditions. Although the full-employment assumption ensures that no cyclicality is present in projected *levels*, some cyclical effects may remain in projected *growth rates*, depending on the base year of the BLS publication.

In sum, whereas BLS always assumes an economy at full employment, other assumptions affecting projections are subject to change. Factors such as the labor force outlook, the NAIRU, TFP, and capital services are monitored, evaluated, and changed on the basis of current research—research that takes place both within BLS and through comparisons with other agencies. These assumptions ensure that BLS projections both incorporate long-term structural changes to the economy and reflect current conditions and expectations for the future. The assumptions apply to projections of the aggregate economy, as well to projections of the products benchmarked to it: projections of GDP at a detailed level, and industry and occupational employment projections.

SUGGESTED CITATION

Kevin S. Dubina, "Full employment: an assumption within BLS projections," *Monthly Labor Review,* U.S. Bureau of Labor Statistics, November 2017, https://doi.org/10.21916/mlr.2017.30.

NOTES

1 For data on the labor force, see “Employment projections: labor force projections” (U.S. Bureau of Labor Statistics, October 24, 2017), https://www.bls.gov/emp/data/labor-force.htm.

2 For data on the aggregate economy, see “Employment projections: aggregate economy” (U.S. Bureau of Labor Statistics, October 24, 2017), https://www.bls.gov/emp/data/aggregate-economy.htm.

3 Hereafter, potential output.

4 For data on occupational employment, see “Employment projections: occupational employment projections” (U.S. Bureau of Labor Statistics, October 26, 2017), https://www.bls.gov/emp/data/occupational-data.htm.

5 For data on industry output and employment, see “Employment projections: industry output and employment projections” (U.S. Bureau of Labor Statistics, October 24, 2017), https://www.bls.gov/emp/data/industry-out-and-emp.htm ; for data on interindustry relationships, see “Employment projections: inter-industry relationships (Input–Output Final Demand matrix)” (U.S. Bureau of Labor Statistics, October 24, 2017), https://www.bls.gov/emp/data/input-output-matrix.htm.

6 For additional information about employment projection methods, see *BLS handbook of methods* (U.S. Bureau of Labor Statistics), chapter 13, “Employment projections,” https://www.bls.gov/opub/hom/pdf/homch13.pdf.

7 The most recent BLS projections appear in T. Alan Lacey, Mitra Toossi, Kevin Dubina, and Andrea Gensler, “Projections overview and highlights, 2016–26,” *Monthly Labor Review*, October 2017, https://doi.org/10.21916/mlr.2017.29.

8 Frictional factors are a third type; however, we can effectively ignore them because changes in frictional factors that affect unemployment levels can be attributed to structural or cyclical factors.

9 See Mark Thoma, “Will there be a ‘new normal’ for unemployment?” *MoneyWatch*, cbsnews.com, November 11, 2009, http://www.cbsnews.com/news/will-there-be-a-new-normal-for-unemployment/.

10 BLS could *estimate* current full-employment levels, rather than using actual employment, in determining growth rates. However, this approach would misconstrue the true level of job growth and likely cause confusion among users. In addition, it often takes years to identify structural changes to the economy. Therefore, using an estimate of current full-employment levels would introduce an additional source of error because perceived cyclical changes at the time may later be identified as structural changes.

11 See John Maynard Keynes, *The general theory of employment, interest and money* (London: Macmillan and Co., Limited, 1936).

12 See Sir William Beveridge, *Full employment in a free society* (London: The New Statesman and Nation and Reynolds News, 1944), http://lib-161.lse.ac.uk/archives/beveridge/9A_79_Full_employment_in_a_free_society.pdf.

13 See Kevin D. Hoover, “Phillips curve,” *The concise encyclopedia of economics*, Library of Economics and Liberty, http://www.econlib.org/library/Enc/PhillipsCurve.html.

14 See Jeffrey M. Lacker and John A. Weinberg, *Inflation and unemployment: a layperson’s guide to the Phillips curve* (Richmond, VA: Federal Reserve Bank of Richmond, 2006), https://www.richmondfed.org/~/media/richmondfedorg/publications/research/annual_report/2006/pdf/article.pdf.

15 See Marilyn Geewax, “Why some still can't find jobs as the economy nears ‘full employment,’” *All Things Considered*, National Public Radio, January 31, 2016, http://www.npr.org/2016/01/31/464856256/why-some-still-cant-find-jobs-as-the-economy-nears-full-employment.

16 See “CBO’s method for estimating potential output: an update” (Congressional Budget Office, August 2001), https://www.cbo.gov/sites/default/files/107th-congress-2001-2002/reports/potentialoutput.pdf.

17 See “A summary of alternative methods for estimating potential GDP” (Congressional Budget Office, March 2004), https://www.cbo.gov/sites/default/files/108th-congress-2003-2004/reports/03-16-gdp.pdf.

18 See “Revisions to CBO’s projection of potential output since 2007” (Congressional Budget Office, February 2014), https://www.cbo.gov/sites/default/files/113th-congress-2013-2014/reports/45150-PotentialOutput-OneColumn.pdf.

19 See Richard G. Lipsey and Kenneth Carlaw, “What does total factor productivity measure?” (Vancouver: Simon Fraser University, January 18, 2001), http://www.csls.ca/ipm/1/lipsey-e.pdf.

20 See Robert J. Gordon, “Slower U.S. growth in the long- and medium-run,” *NBER Reporter 2015 Number 1: Research Summary* (Cambridge, MA: National Bureau of Economic Research), http://www.nber.org/reporter/2015number1/gordon.html.

21 For additional information about alternative measures of labor underutilization, see “Alternative measures of labor underutilization for states, third quarter of 2016 through second quarter of 2017 averages,” *Local Area Unemployment Statistics* (U.S. Bureau of Labor Statistics, July 28, 2017), https://www.bls.gov/lau/stalt.htm.

22 For a comprehensive view of an internal BLS model, including formulas, see Maggie Woodward, “A simple model for estimating potential output,” in *Proceedings of the 19th Federal Forecasters Conference: the value of government forecasts* (U.S. Bureau of Labor Statistics, September 27, 2012), pp. 25–33, http://www.va.gov/healthpolicyplanning/ffc/pandp/ffc2012.pdf. For further discussion of piecewise linear regressions, see Robert S. Pindyck and Daniel L. Rubinfeld, *Econometric models and economic forecasts*, 4th ed. (Boston: Irwin/McGraw Hill, 1998).

RELATED CONTENT

Related Subjects

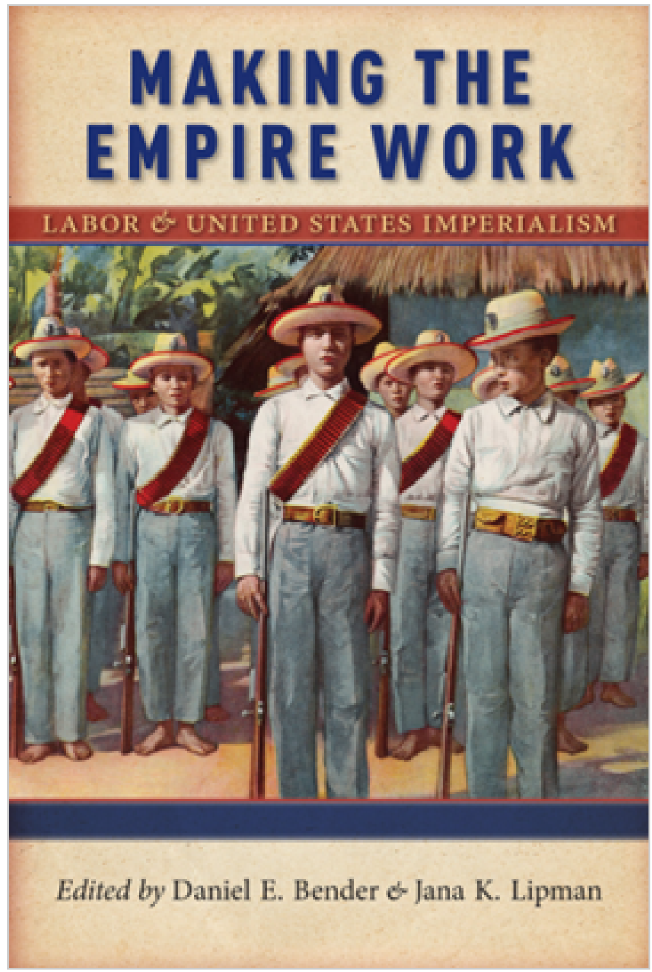
Projections | Labor force | Unemployment | Statistical programs and methods | Employment

Monthly Labor Review

November 2017

Labor in the U.S. empire: a historical perspective

Making the Empire Work: Labor and United States Imperialism. Edited by Daniel E. Bender and Jana K. Lipman. New York: New York University Press, 2015, 384 pp., $35.00 paperback.



Andrew Blank
blank.andrew@bls.gov

Andrew Blank is an economist in the Office of Employment and Unemployment Statistics, U.S. Bureau of Labor Statistics.

U.S. involvement in global affairs increased considerably during the 20th century. Over this period, America fought in two world wars, took part in the postwar reconstruction of Japan and Europe, and promoted democracy in the wake of the Cold War. Yet, the transition into the 21st century saw many around the world reevaluate—and often object to—America's efforts. Since the 9/11 terrorist attacks, the term "empire," with its negative connotation, has been used frequently in reference to the United States. Advocates of stronger U.S. intervention abroad contend that free nations need additional protection from foreign threats. At the same time, critics of U.S. foreign policy suggest that the country's overseas activities are an effort to expand American influence and power throughout the world. Those who generally accept the latter view see the U.S. empire as beginning with the acquisition of the Philippines, Cuba, Guam, and Puerto Rico in 1898, continuing after World War II, and expanding since 2001. However, the authors of *Making the Empire Work: Labor and United States Imperialism*, a 13-chapter anthology, trace the dawn of the empire to earlier times, uniquely claiming that it originated in 1863.

The book, edited by Daniel E. Bender and Jana K. Lipman, examines labor relations in the U.S. empire since the Emancipation Proclamation. In the introduction, editors Bender and Lipman explain that the concept of U.S. empire is defined by America's "geographical boundaries *and* by its labor system." This broad definition includes laborers in U.S. colonies, U.S. corporations abroad, and the U.S.

armed forces, thus expanding traditional notions of U.S. empire. In essence, the authors recount the experiences of forgotten imperial workers who, while neglected in previous research, "made the U.S. empire work."

The volume consists of four thematic parts that outline the paradigm of U.S. empire from the perspective of trained historians. The four parts discuss "Solidarities and resistance," "Intimacies in colonial spaces," "Migration and mobilizing labor for the empire," and "Imperial labor and control in the tropics." Emphasis is given to imperial workers in the Caribbean, Asia–Pacific, and Central and South American regions; West Africa; and the United States. The overall objective of the authors is to challenge how historians have traditionally conceptualized the U.S. empire.

Opening the first part of the book, Julia Greene (chapter 1) examines the origins of the U.S. empire from the 1860s to the 1890s. She connects the U.S. westward military expansion during the Indian wars of the 1860s to the expansion abroad in the Spanish–American War of 1898. Further, she discusses the motivations behind the military's participation in the U.S. empire, noting that, for the majority of U.S. servicemen, serving in the military was a way to enhance one's social status and economic mobility. Shifting to the 1920s, Moon-Ho Jung (chapter 2), Christopher Capozzola (chapter 3), and Kevin Coleman (chapter 4) discuss how local labor engaged in collective resistance against U.S. imperial coercion. Jung evaluates East Asian labor strikes in Hawaii, outlining the struggle of that state's workers against U.S. nationalist policies. Likewise, focusing on the Philippines, Capozzola provides a detailed account of contention and failed wage negotiations between U.S. colonial military officials and Filipino soldiers who served in the ranks of the U.S. military. Next, Coleman examines U.S. corporate expansion in Colombia, detailing the struggle of Colombian labor activists against the United Fruit Company (UFC). In that case, the Colombian government suppressed national labor laws in areas controlled by U.S. "banana imperialists," allowing the UFC to have substantial influence over laborers.

The second part of the book looks at the creation of labor systems in the U.S. empire and at those systems' social effects. Seungsook Moon (chapter 5) argues that U.S. territorial occupations at the end of World War II created high demand for sexual labor among servicemen near U.S. military bases. He conducts a comparative analysis of occupied countries in East Asia and Europe, focusing on sexual laborers as workers in the U.S. empire. Next, Vernadette Vicuña Gonzalez (chapter 6) outlines Hawaii's transition to a formal U.S. territory, which expanded the periphery of U.S. imperial reach. She also examines tourism—an industry that limited economic opportunities for natives in Hawaii. The essays offered by Moon and Gonzales both attempt to display the often ignored shadow economics that comes with imperial rule.

In the third part of the book, Andrew T. Urban, Dorothy B. Fujita-Romy, and Cindy Hahamovitch shift away from the empire-abroad theme, revealing the tumultuous labor experiences of U.S. immigrants. Urban (chapter 7) and Fujita-Romy (chapter 8) illustrate some adversities faced by immigrant workers in California. Specifically, Urban examines how debates about the domestic economic effects of open immigration policies for Chinese migrants in the 1880s led to the exclusion of Chinese workers from the state's labor market. Building on the events recounted by Urban, Fujita-Romy discusses imperial labor mobility in the early 20th century. She describes how the influx of Filipino workers in California in the 1920s was met with violence by locals. Looking at subsequent decades, Hahamovitch (chapter 9) examines the emergence of guest-worker programs in the United States. According to the author, guest workers toiled in low-paying agricultural jobs, facing harsh working conditions in a labor system akin to indentured servitude.

In the final part of the book, Andrew Zimmerman (chapter 10), Jason M. Colby (chapter 11), Augustine Sedgewick (chapter 12), and Lauren Hirshberg (chapter 13) focus on how labor systems were imported and managed in the tropics. Zimmerman, in the only chapter dedicated to West Africa, analyzes the efforts of European colonial powers to implement a region-wide labor system based on that adopted by the United States after the Civil War. Turning to activities by U.S. corporations in the tropics, Colby revisits the UFC case in Colombia, detailing the firm's labor policy to import workers into the South American region from the 1890s to the 1920s. Further, Sedgewick indirectly links U.S. corporations' coffee production in El Salvador to the U.S. empire. In the final chapter, Hirshberg explores worker migration to the Marshall Islands during the Cold War. According to the author, in pursuing its national security objectives, the U.S. military exported high-skilled U.S. laborers to the remote island of Kwajalein, turning natives into domestic servants supporting the local economy established by the newcomers.

Overall, *Making the Empire Work* provides a heterodox perspective in labor history. Defining the U.S. empire by focusing on the structure of labor systems and the workers employed within imperial boundaries (in colonies, corporations, and the U.S. armed forces) goes well beyond traditional conceptions of the term. A major strength of the book is that it sheds light on larger world events through the experiences of specific regions and countries. Generally, the authors provide ample political, economic, and social context in presenting their views. Their rather meticulous, yet complex, analyses would best serve academics who see the United States as an empire.

Although the book's four parts follow a thematic layout, their underlying message is not always easy to discern. Incorporating prologues within each part would have solidified the narrative. It seems to me, however, that the book's major drawback is its depiction of U.S. corporations and the U.S. government as colluding in their external expansion. Several passages also describe U.S. corporations as the *de facto* rulers of the regions in which they reside. Together, these perspectives underplay both the political sovereignty a host nation might have and the autonomy a U.S. corporation might possess.

While the book offers a plausible account of the experiences of workers affiliated with the U.S. labor system, it tends to stretch the boundaries of U.S. expansion. All things considered, I was not overly convinced by the book's theory of labor exploitation and its concept of U.S. empire.

Monthly Labor Review

November 2017

Retirement-claiming decisions: a survey

Graham Boone

For retired workers, it often makes good financial sense to delay filing for Social Security retirement benefits. Full retirement age (the age at which a claimant receives 100 percent of Social Security retired-worker benefits) for most Americans currently eligible for retired-worker benefits is 66. But these individuals have the option of filing as early as age 62, with a 25-percent reduction in benefits, or as late as age 70, with a 32-percent increase in benefits. So retirees filing at 70 could reasonably expect their first check to be about 75 percent greater than it would have been had they filed at 62 (not factoring in cost-of-living adjustments or the effect of income earned after age 62). Despite this financial incentive to delay filing, most people claim Social Security benefits at or before full retirement age rather than waiting until age 70. Why? This is the question addressed in a recent study by John B. Shoven, Sita Nataraj Slavov, and David A. Wise titled "Social Security claiming decisions: survey evidence" (National Bureau of Economic Research, working paper no. 23729, August 2017).

The authors posit a number of potential reasons for why people choose to file for benefits before age 70:

- The immediate need for cash
- The perception that filing at full retirement age is a social norm encouraged by government
- A belief that one will not live long enough to benefit from delayed filing
- Lack of knowledge about the benefits of delaying filing
- Concerns about changes to the system as legislators work to ensure the Social Security program remains solvent

To get a better understanding of what actually drives retirees' filing decisions, the authors conducted a national survey of 1,116 people ages 55 and older.

Respondents were questioned on their financial literacy, knowledge of the Social Security program, general feelings about their health, and their age when they stopped working. Those who already were claiming benefits were asked their age at time of filing, how long after retirement they filed, and how satisfied they were with the timing of their decision to file.

The results of the survey suggest that work status is the most influential factor in the decision to claim retirement benefits. For respondents who claimed before full retirement age, the most popular reason was "I had stopped working" (38.4 percent). For both respondents who waited 6 months or more after retirement age to file and respondents who were 62 and older but had not yet filed, the most popular reason was "I was/am still working" (46.1 percent and 65.3 percent, respectively).

A key finding from the study is that many claimants see filing at full retirement age to be a social norm. Among the respondents who filed within 6 months of age 66, the most popular reason given was “It seemed natural to start my benefit at my full retirement age” (45.7 percent). Nearly 1 in 5 answered, “I wanted to avoid getting a reduced benefit” (notwithstanding the fact that such claimants were getting a reduced benefit compared with what they would have received had they put off filing). Interestingly, those who filed at full retirement were more satisfied with the timing of their claim than both those who filed early and those who delayed, although those who delayed showed much greater satisfaction than those who filed early.

These findings have the potential to help inform policymakers concerned about how and when retirees file for benefits.

Monthly Labor Review

November 2017

The unexplainable, growing black-white wage gap

Eleni X. Karageorge

For the past several decades, black workers have fared worse than white workers in the U.S. labor market. Despite government policies designed to reduce or eliminate racial disparities, black workers continue to experience lower wages and higher unemployment rates than Whites. The latest evidence comes from a recent study by economists at the Federal Reserve Bank of San Francisco, "Disappointing facts about the black-white wage gap" (*Economic Letter*, Federal Reserve Bank of San Francisco, September 5, 2017). Authors Mary C. Daly, Bart Hobijn, and Joseph H. Pedtke show that black workers still earn less than their white counterparts in a worsening trend that holds true even after accounting for differences in age, education, job type, and geography.

In 1979, black men earned, on average, about 80 percent of what white men earned ($15 per hour versus $19 per hour in inflation-adjusted dollars); in 2016, it slipped to about 70 percent ($18 per hour vs. $25 per hour). Although wages for black males increased, the gain lagged significantly behind that of their white counterparts. The gap also widened for black women, though the disparities were smaller. From 1979 to 2016, wages for black women fell from roughly 95 percent of what white women earned ($11 an hour versus $12 an hour) to about 82 percent ($16 versus $20).

The authors found it most notable that a significant portion of the racial wage disparity is attributable to *unexplained* factors, factors that are hard to characterize or measure. While a sizable portion of the gap arises from *explained* factors, including occupation choice, educational attainment, and age—the share due to *unexplained* factors accounts for much of the growth in the wage gap over time. The authors suggested that *unexplained* factors could include discrimination, differences in school quality, and differences in opportunity. In 1979, *unexplained* factors were responsible for about 8 percentage points of the earnings gap for men, about 40 percent of the total earnings gap at that time. By 2016, that share had risen to about 13 percentage points, nearly half the total earnings gap.

Unexplained factors likely also contribute to high black unemployment rates. The study found that even during economic expansions, unemployment rates for Blacks were more than 6 percentage points higher than that of Whites. According to the authors, this gap indicates that "black jobseekers face different job opportunities than their white counterparts."

The research draws focus to the persistence of a racial wage gap even after the passage of the Civil Rights Act, landmark anti-discrimination legislation that first went into law over 50 years ago. This highlights the problem that it's tough to address disparities if it's impossible to measure what's driving them.