Research on Quarterly Benchmarking for the Current Employment Statistics Survey October 2009

Victoria Battista, Chris Manning, Kenneth W. Robertson
U.S. Bureau of Labor Statistics, 2 Massachusetts Avenue NE, Suite 4860, Washington, DC 20212
Battista.Victoria@bls.gov

Abstract
The Current Employment Statistics (CES) survey is a large monthly survey of establishments used to estimate employment, hours, and earnings, by industry, for the nation, states, and metropolitan areas. The program produces estimates each month, about three weeks after the reference period. CES employment estimates are benchmarked on an annual basis to the Quarterly Census of Employment and Wages, which is a near-universe count based on unemployment insurance tax account records. Several alternative benchmarking procedures were evaluated to determine if the quality of the CES estimates could be improved by benchmarking on a quarterly basis. The authors describe the alternative procedures and present the results of the research.

Key Words: Current Employment Statistics, benchmark, wedge, quality, employment, estimation, establishment survey

1. Introduction

The Current Employment Statistics (CES) Survey, conducted by the Bureau of Labor Statistics (BLS), is a monthly panel survey of nearly 390,000 business establishments. The national CES estimates of employment, hours, and earnings are some of the most timely and sensitive economic indicators published by the federal government. They are widely viewed as a key measure of the health of the economy and are closely tracked by both public and private policy makers alike.

2. Background

One of the benefits of CES is the timeliness of the estimates. The survey collects information based on the pay period including the 12th of the month from a scientific sample of establishments. First preliminary CES estimates are published each month about three weeks after the reference period, usually on the first Friday of the following month. In order to incorporate additional sample received after the primary deadline, each estimate undergoes two monthly revisions before being finalized. The second preliminary estimate is typically published about seven weeks after the reference period, and the third preliminary estimate is published about eleven weeks after the reference period. Therefore, for any given reference month, second preliminary estimates are published one month after the initial estimates, and third preliminary estimates are published two months subsequent. On an annual basis, the estimates are "re-anchored", or "benchmarked" to an employment population value based mostly on the March

employment estimate from the Quarterly Census of Employment and Wages (QCEW).[1] The QCEW program provides a nearly complete employment count based on unemployment insurance tax records. The unemployment insurance account data are collected on a quarterly basis, and are available at a seven to nine month lag.

3. Research Question

We conducted research to determine if the quality of the estimates could be improved by benchmarking quarterly rather than annually. To do this, our research re-anchored each calendar quarter to the third month of the most recently available quarter of QCEW data. There are a number of reasons why this is an attractive concept. For example, if quarterly benchmarking corrects one-fourth of our annual benchmark error each quarter, then each monthly estimate would be substantially closer to its final value when initially estimated, and its benchmark revision would be substantially smaller. This would make the data even more valuable to the data users. Also, if this process were to mitigate the risk of a large benchmark revision it would add value to the program. Benchmarking more frequently could potentially reduce the average size of each month's estimation error.

Table 1: Seasonal differences between QCEW and CES, total nonfarm employment (in thousands)

Period	QCEW Quarterly Change	CES Quarterly Change	Difference
2005 Q1	-1836	-991	-845
2005 Q2	2718	2522	196
2005 Q3	126	-394	520
2005 Q4	1020	904	116
2006 Q1	-1070	-925	-145
2006 Q2	2682	2290	392
2006 Q3	-423	-239	-184
2006 Q4	1058	1053	5
2007 Q1	-1596	-1133	-463
2007 Q2	2490	2335	155
2007 Q3	-755	-716	-39
2007 Q4	921	859	62

Note: QCEW adjusted to CES scope.

It has been known for many years that the two programs – the CES and the QCEW – exhibit different seasonal patterns. While the reasons for these differences are not completely understood, the differences are notable (see Table 1). For example, in the first quarter of 2005, the over-the-quarter change in the QCEW is nearly double the magnitude of the CES estimate's over-the-quarter change. The research we conducted initially incorporated only level differences between the CES and QCEW, which did not take the differing seasonality into account. We have long suspected that this level difference approach would detract from the quality of CES estimates; this research provided an

[1] The principal source of CES benchmark data for private industries is the Quarterly Census of Employment and Wages (QCEW). These employment data are provided to State Employment Security Agencies by employers covered by State unemployment insurance (UI) tax laws. BLS uses several other sources to establish benchmarks for the remaining industries partially covered or exempt from mandatory UI coverage, accounting for nearly 3 percent of the nonfarm employment total.

opportunity to explore that assertion. A second approach using differences in the two employment counts' over-the-year change was researched. Our hope was that this second procedure would help mitigate the problem of different seasonality between the two series and yield some tangible improvement over an annual benchmark.

There are also some negative issues to consider related to quarterly benchmarking. First, the CES benchmarking process is resource intensive. The development and review of the benchmarked series on an annual basis takes considerable time, people, and production capability each year. Developing and reviewing estimates four times each year would require additional effort, which would divert resources that could be used to improve other aspects of the program. Second, and perhaps more importantly, quarterly benchmarking means that each estimate would be revised more times. For example, we currently have five iterations for an April estimate - an initial estimate and four revisions. Under a quarterly benchmarking schedule, we would have seven iterations for an April estimate. These additional revisions to the estimate might make this process less attractive to data users unless the quality improvements are substantial.

This paper documents our efforts to evaluate these two quarterly benchmark procedures.

4. Research

4.1 Data

4.1.1 CES

Third release, un-benchmarked CES estimates from April 2003 to March 2007 were used in the research calculations. These estimates are the current annual methodology's starting points for calculating over-the-year change, benchmark differences, and sample links.

4.1.2 Population

Also known as universe employment, population employment mainly consists of data from the QCEW, but also includes an adjustment for industries partially covered or exempt from mandatory unemployment insurance tax coverage. Data were taken from the QCEW employment releases. As is the norm when producing an annual benchmark, the most current versions of QCEW data for the given benchmark period were used in the research.[2] It should be noted that CES and QCEW cover slightly different industries, and therefore, the QCEW was adjusted to CES scope.

4.2 Procedures

4.2.1 How are CES employment estimates created?

The CES survey produces employment estimates each month using a link-relative estimator. The estimator can be depicted as follows:

$$[1]\ \hat{e}_m = \hat{e}_{m-1}\left(\frac{\sum_{i\in m,m-1} e_{i,m}}{\sum_{i\in m,m-1} e_{i,m-1}}\right) + b'_m$$

[2] Note that while the QCEW data used were the most current version for a given benchmark time period, the data might not be the final version because QCEW data are subject to revisions up to a year after the initial release.

where

$\hat{e}_m$ = the current month employment estimate

$\hat{e}_{m-1}$ = the previous month employment estimate

$i \in m, m-1$ = a unit that reported positive employment in both the current and previous month

$e_{i,m}$ = the current month employment for establishment i

$e_{i,m-1}$ = the previous month employment for establishment i

b'_m = the net birth/death factor for month m

As can be seen in equation [1], the link-relative estimator is based on a previous estimate, which is based on yet another previous estimate, going back until you arrive at a base period when a population value has been utilized as the previous month value. Each successive estimate is calculated using a ratio of change in employment reported by the sampled establishments, applying that ratio to the previous month's estimate, and then adding in a net birth/death factor to account for the net of business births and deaths not captured by the sample.[3] The longer you move forward with this estimator without re-anchoring to a population value, the larger the estimation error becomes because the estimation error is the joint error of some number of linked monthly estimates.

4.2.2 What is benchmarking, and why do we do it?

Benchmarking is the process of re-anchoring the CES estimates to a population value, effectively replacing the March CES estimate with a population value and using that new March value to re-estimate the series. A benchmark year is defined as April of the first year to March of the next, although the benchmark process affects estimates beyond that March until the following December. We benchmark in order to keep the estimation error from growing without bound.

Note that whichever benchmark procedure is used, we are assuming that the QCEW is the true population value. However, we know that this is not the case. The benchmark difference is in fact a difference between two independently derived employment values. The CES estimates include error from a number of sources, as do all sample surveys. These errors include both sampling error and non-sampling error. The magnitude of the sampling error is directly related to the size of the sample and the size of the population covered by the sample. The CES sample includes about one-third of the population, which is very large by usual sampling standards. Non-sampling error includes coverage, reporting, non-response, and processing errors. The QCEW employment counts include all of these same errors except for sampling error.

4.2.3 Annual Benchmark

The annual benchmark procedure consists of three major steps. The first step is to determine the difference between the population employment and the CES employment for the benchmark month (March of the previous year for the current annual procedure). We assume in the benchmark process that the error accumulates in a linear manner. In the

[3] For more information on the CES net birth/death methodology, see Kirk Mueller's *Impact of business births and deaths in the payroll survey* (http://www.bls.gov/opub/mlr/2006/05/art4full.pdf).

second step this difference in employment is "wedged" back to April from two years previous (t-2). That is, April $_{t-2}$ gets $1/12^{th}$ of that difference, May $_{t-2}$ gets $2/12^{th}$ of the difference, and so on, until March $_{t-1}$ gets $12/12^{th}$ of that difference. The third step is to "re-project" the seven months of estimates forward from the anchor point based on the new benchmark level. This last step simply substitutes the new benchmark month population value as $e_{i,m-1}$ in equation [1] and moves forward, re-calculating each month from April to October using this new set of linked values. The new October value becomes the value ($e_{i,m-1}$) used for calculating the third preliminary November estimate, the second preliminary December estimate, and the first release of the January estimate.

The benchmark difference can be depicted as follows:

$$[2] \quad d_{t-1,m} = Population_{t-1,m} - CES_{t-1,m}\,.$$

where
t-1 = the benchmark year
m = the month being benchmarked, i.e. March

The wedge process can be depicted as follows:

$$[3] \quad \hat{e}'_{t-x,m} = \hat{e}_{t-x,m} + \left(\frac{q_n}{12}\right) d_{t-1,b}$$

where
$t-x$ = the year, x=(2,1)
$\hat{e}'_{t-x,m}$ = the benchmarked and wedged employment estimate for month m
q_n = 1…12, (April $_{t-2}$ = 1, May $_{t-2}$ = 2, …, February $_{t-1}$ = 11, March $_{t-1}$ = 12)
b = the benchmark month, March of year t-1
$d_{t-1,b}$ = the benchmark difference

In total these three parts of an annual benchmark span more than a year of data. The wedge covers the first eleven months from the April of the year preceding the benchmark to the March benchmark itself. From the March benchmark forward the estimates are re-projected for seven months through October of the benchmark year. The additional months from November to January are then estimated directly using the standard sample-based estimation procedure.

Historically the annual benchmark revision amounts for March have been relatively small, averaging in absolute terms about 0.2% of total nonfarm employment over the past ten years. Benchmark revisions and percent revisions for 1999 through 2008 are shown in Table 2 on the next page.

4.2.4 Quarterly Level Benchmark

A quarterly level benchmark follows steps similar to the annual benchmark, with some modest adjustments. The first step is to determine the difference between the population employment and the CES employment for the benchmark month (the third month of the latest available QCEW data, usually for a month about seven months previous). In the

second step this difference is "wedged" back over the three months of that quarter. Wedging consists of adding 1/3rd of the benchmark month difference to the first month of the quarter, then adding 2/3rd of the benchmark month difference to the second month of the quarter, and finally the full benchmark difference to the last month of the quarter (the benchmark month). The third step is to "re-project" seven months of estimates, the same as in the annual benchmark procedure.

The difference can be depicted using equation [2], where:
m = the month being benchmarked (March, June, September, or December)

Table 2: CES national total nonfarm benchmark revisions, March 1999-2008
(in thousands)

Year	Benchmark Revision	Percent Benchmark Revision
1999	258	0.2
2000	468	0.4
2001	-123	-0.1
2002	-313*	-0.2
2003	-122	-0.1
2004	203	0.2
2005	-158	-0.1
2006	752	0.6
2007	-293	-0.2
2008	-89	-0.1
Average of Absolutes	278	0.2

*Represents the revision to the over-the-year change; however, the revision to the level was different due to the change in federal government definition.

4.2.5 Quarterly Over-the-Year Benchmark

A quarterly OTY benchmark follows steps similar to the quarterly level benchmark, with one change. The difference in the over-the-year (OTY) change of the two series is used to adjust the quarterly benchmark using June, September, and December as their benchmark months. That is, the following difference is developed:

$$[4]\ d_{t-1,m} = \left(Population_{t-1,m} - Population_{t-2,m}\right) - \left(CES_{t-1,m} - CES_{t-2,m}\right)$$

where

$d_{t-1,m}$ = the benchmark difference

t-1 = the previous year
t-2 = two years previous
m = the month being benchmarked, i.e. June, September, or December

The assumption is that the over-the-year change eliminates the effect of within-year seasonal differences. For example, the QCEW experiences a larger seasonal decline in employment in January than does the CES. However, if this is a regular seasonal movement, then you can remove the differing seasonal aspect by looking at these over-the-year differences. Wedging back and re-projecting forward are the same in this procedure as in the quarterly level benchmark. This procedure does not remove the

effects of the independent errors associated with each program, nor does it remove the effects of moving seasonality.

In order to maintain the anchoring affect of benchmarking, the first quarter of each year did not use the difference in the OTY change to define the benchmark amount. Instead the March benchmark was calculated the same way as a level benchmark, fully replacing the March value with the population value and calculating the wedge and re-projection from that point.

4.3 General Expectations

There are some general expectations that can be defined for both alternatives. Suppose for the moment that the benchmark process consisted of a simple level-shift adjustment to the estimates during the re-projection months, and that we didn't replace the net birth/death factors during this period. In this simplistic case the sum of four quarterly benchmark values would always equal the annual benchmark value under either of the two quarterly benchmark procedures. Further, we would then expect that the sum of the absolute values of the four quarterly benchmarks will never be less than the annual benchmark value. In practice, the re-projection of estimates and the replacement of net birth/death factors impacts this relationship, but it can still serve as a general guide for what we can expect from the results.

We expect that the absolute sum of the four quarterly OTY change benchmarks will always be smaller than the absolute sum of the four quarterly level benchmarks, as the latter will include the effects of benchmarking to a population value that includes the effect of differing fixed seasonality.

The results of the research are presented below.

5. Results

The primary objective of benchmarking quarterly is to keep the CES estimates more closely aligned with the universe counts of employment throughout the year, reducing the magnitude of any one particular benchmark. Our model does accomplish this goal. However, the four quarterly benchmarks over the course of a year exhibit some tendencies that may prove less desirable.

The progression of revisions consists of four quarterly benchmarks over the course of the benchmark year. Table 3 on the next page shows the quarterly revisions for both a level difference benchmark (column 4) and an over-the-year difference benchmark (column 6). The level difference benchmarks for each quarter are sometimes within the historical ranges outlined in Table 2, and at other times far outside of this range (such as the -836,000 revision in the first quarter of 2005). The OTY change benchmark is much less volatile, exhibiting quarterly benchmarks that are consistently smaller than historical averages of annual benchmarks with only a couple of exceptions (i.e. the 508,000 revision in the third quarter of 2005).

The 2006 benchmark year is an unusual case because it is a year in which the CES program had an abnormally large annual benchmark. In this simulation, the large yearly benchmark is not spread evenly across quarters. Instead the bulk of its revision is in the third quarter of 2005 for both the level and OTY quarterly benchmark methodologies. Nevertheless, this quarterly revision is smaller than the annual benchmark amount, and

therefore does accomplish our stated goal of quarterly benchmarking resulting in smaller benchmark revisions.

Table 3: Quarterly benchmark revisions compared to annual benchmark revisions (in thousands)

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Benchmark Year	**Period**	**Annual Revision**	**Level Revision**	**Sum of Level Revisions**	**OTY Change Revision**	**Sum of OTY Change Revisions**
2005	2004 / Q2		368		-109	
	2004 / Q3		-20		-4	
	2004 / Q4		407		237	
	2005/ Q1	-158	-836	-81	-268	-144
2006	2005 / Q2		233		-42	
	2005 / Q3		490		508	
	2005 / Q4		162		-12	
	2006 / Q1	752	-202	683	326	780
2007	2006 / Q2		444		169	
	2006 / Q3		-333		-307	
	2006 / Q4		1		-170	
	2007 / Q1	-293	-494	-382	29	-279

By summing the quarterly revisions from each benchmark year, we can compare them to an annual benchmark for the same period. As expected, both procedures produce sums (columns 5 and 7) that approximate the annual benchmark value (column 3).

However, by focusing solely on the sum of the quarterly revisions, which as expected approximates the annual revision, an important aspect of the results is overlooked. In particular, some individual quarterly benchmarks, while not considered large by historical standards, are greater in magnitude than the annual benchmark revision for their reference year. For example, for the 2005 benchmark year, the quarterly level revision for the first quarter of 2005 was -836,000 and for the OTY change revision was -268,000. The quarterly revision under both methods was greater in magnitude than the annual revision of -158,000. In fact, for the 2005 annual benchmark year, three of four quarterly level benchmark revisions are greater in absolute value than the annual revision, and two of four OTY change benchmark values are greater in absolute value than the annual revision.

If these quarterly values are compared to one another, at times they essentially offset each other. For example, under the OTY change benchmark methodology, the fourth quarter of 2004 and the first quarter of 2005 have nearly the same absolute value, but are opposite in direction. In order to get a more complete picture of the magnitude of the revisions involved in quarterly benchmarking, we summed the absolute values of each year's quarterly revisions and compared them to annual revision amounts. Table 4 on the following page shows these comparisons.

In Table 4, rows (1), (2), and (3) are the previously stated values for the annual, summed quarterly level, and summed quarterly OTY change benchmarks. Rows (4) and (5) sum

the absolute values of the quarterly level and quarterly OTY change benchmarks, respectively. Rows (6) and (7) show the absolute value of each year's largest quarterly benchmark difference under both methods.

Table 4: Quarterly benchmark revisions compared to annual benchmark revisions, summary results (in thousands)

Row	Period	2005	2006	2007
(1)	Annual Revision	-158	752	-293
(2)	∑ (Level Revisions)	-81	683	-382
(3)	∑ (OTY Revisions)	-144	780	-279
(4)	∑ | Level Revisions |	1,631	1,087	1,272
(5)	∑ | OTY Revisions |	618	888	675
(6)	Maximum | Level Revisions |	836	490	494
(7)	Maximum | OTY Revisions |	268	508	307

As expected, the absolute sums are always greater in magnitude than the annual revision. Also, as expected, the quarterly OTY change benchmark has smaller absolute sums than the quarterly level benchmarks. While this is a positive summary result, it is negated by the occasional quarterly benchmark that is greater in absolute value than the annual revision.

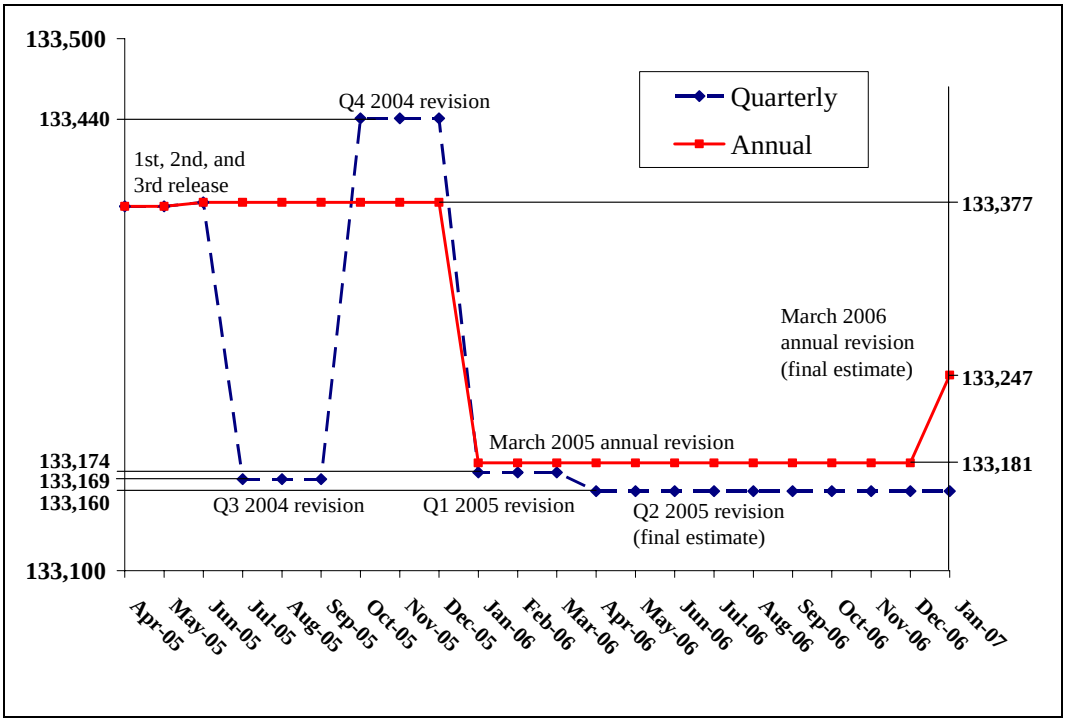
Table 5: Timetable of releases of April 2005 total nonfarm estimates and revisions to those estimates (in thousands)

	(1)	(2)	(3)
	Reference month for release that includes April 2005 estimate or revision	**Annual Benchmark**	**Quarterly Benchmark**
(1)	April 2005 (preliminary April 2005 estimate)	133,374	133,374
(2)	May 2005 (2nd preliminary April 2005 estimate)	133,374	133,374
(3)	June 2005 (3rd preliminary April 2005 estimate)	133,377	133,377
(4)	July 2005 (Q3 2004 benchmark)		133,169
(5)	October 2005 (Q4 2004 benchmark)		133,440
(6)	January 2006 (Q1 2005 benchmark and annual re-projection period)	133,181	133,174
(7)	April 2006 (Q2 2005 benchmark wedge)		133,160
(8)	January 2007 (Q1 2006 benchmark and annual wedge period)	133,247	

As shown in Table 5, under the current annual benchmark methodology, the estimate for April 2005 is revised four times before it is considered final in January 2007 (column 2). After the initial release (row 1), more establishments who failed to report their April data in time for the first release have a chance to report for April and be included in CES estimates in the following two months. As a result, there are two releases of April data containing a more complete sample published with the May and June estimates (rows 2

and 3). The April estimate is also revised during the 2005 annual benchmark as part of the estimate re-projection (row 6) before being finalized during the 2006 annual benchmark (row 8).[4] When benchmarking quarterly, there are several additional revisions to the data. The initial release and the two monthly revisions are the same as in an annual benchmark, but under quarterly benchmarking, each quarterly re-projection or wedge adds another revision to that month's estimate. Consequently, in total there are six revisions for April 2005 using the quarterly benchmarking methodology (column 3).

Chart 1: Revisions to total nonfarm employment for April 2005 (in thousands)



Graphing the effect of these additional estimates compared to the effect of the original five estimates (chart 1) shows the tendency of quarterly revisions to make one month's estimates fluctuate when subjected to repeated revisions. Again the first, second, and third release of estimates are the same for both methodologies. The initial quarterly revision (dashed blue line) of a drop of 208,000 for the third quarter of 2004 is followed by an upward revision for the fourth quarter of 2004 of 271,000, essentially canceling out the third quarter revision. The next quarterly revision, for the first quarter of 2005, again revises the estimate downward by 266,000. The last quarterly revision, for the second quarter of 2005, is more modest with a downward revision of only 14,000.

In contrast, the annual benchmarking methodology (solid red line) resulted in only one large downward revision of 196,000 as a result of the March 2005 benchmark. The second revision resulted in a slight upward revision of 66,000 (as a result of the March 2006 benchmark). The variability of the estimates for April 2005 is far less extreme when using the annual benchmarking methodology than when benchmarking quarterly.

[4] The exceptions to this pattern of estimates and revisions are the months of January, February, and March. These months are revised only three times, as they are only wedged (row 8 above), but never re-projected forward (row 6) during the initial benchmark.

6. Conclusion

Benchmarking quarterly provides some protection against a single large annual benchmark, keeping the estimates somewhat more in line with the population throughout the year. However, several factors counter this periodic benefit. First, during a year with a moderate annual benchmark, the quarterly benchmark process may instill in the estimates one or more quarterly benchmark revisions larger than the annual benchmark revision would have been. Second, additional revisions to the estimates are a disservice to the data user, and should only be contemplated in the event that data quality is enhanced substantially, hopefully with each revision. That does not appear to be the case with either of the quarterly benchmarking procedures explored in this research.

In summary, we find that there is a single potential periodic benefit, and several negative factors frequently associated with quarterly benchmarking procedures for the CES program. At this time, we cannot recommend implementing a quarterly benchmark procedure for the program.

Acknowledgements

The authors would like to thank Hazel Wigdor and Michele Eickman Walker, both of BLS, for their contributions to this research.

Any opinions expressed in this paper are those of the authors and do not constitute policy of the Bureau of Labor Statistics.