
Kenneth W. Robertson and Mike Roosma

Abstract

The CES survey is a large monthly survey of businesses used to estimate employment, hours, and earnings by industry and geographic area. The Business Employment Dynamics (BED) program is a quarterly program that disaggregates population business data into four components: opening establishments; businesses whose employment expanded; businesses whose employment contracted; and closing establishments. The data from the BED program are available about 7 months after the end of the calendar quarter. It is possible to estimate two of these series – the expansions and the contractions – using CES survey reports. Research to produce an experimental CES-BED series has raised a number of estimation and administrative issues. Many of these issues are still being examined. The authors present some preliminary results and discuss some of the issues that still remain to be resolved.

Key Words: establishment survey, business employment dynamics

Background

The Bureau of Labor Statistics’ Current Employment Statistics (CES) survey is a monthly survey which produces estimates of employment, hours, and earnings, for the nation, and for states and metropolitan statistical areas. Estimates are first produced about three weeks after the reference period, making this Principle Federal Economic Indicator one of the timeliest indicators of the state of the economy. The survey collects data each month from a sample of about 150,000 businesses. The probability survey design covers employment in private non-agricultural industries, while a quota sample is collected from government establishments.

The Business Employment Dynamics (BED) data are developed from the Quarterly Census of Employment and Wages (QCEW) establishment level microdata. The QCEW data are based on states’ unemployment insurance records and the BLS Multiple Worksite Report for multi-establishment businesses. These data disaggregate total employment change from businesses which fall into four groups, when comparing employment at the end of one quarter with employment at the end of the previous quarter. The four groups are new businesses (births or openings), businesses whose employment has gone from a positive value to zero (deaths or closings), businesses whose employment has grown from a positive value to a larger positive value (expansions), and businesses whose employment has declined from a positive value to a lower positive value (contractions). The remaining businesses are those which had no change in employment over the quarter.

These BED data are useful in providing insight into which components of the economy are driving changes in employment. For example, employment growth could occur in a number of ways; by an increase in business births, by an increase in business expansions, by a decrease in business deaths, by a decrease in business contractions, or by any combination of these. Knowing the makeup of these four components provides
more information on where we are in the business cycle, and about business cycles in
general. For more detail on the utility of these data, see (6)\(^1\).

**Proposal**

In this document we describe some preliminary research that explores the possibility of using CES data to estimate employment in expanding and contracting businesses. The CES estimate of over the month change in employment can be disaggregated into components. The employment change components estimable from the CES survey data are expansions and contractions. Note that CES data cannot be used to directly estimate births or deaths.

These CES-BED data would provide estimates of the two major components of employment change about seven months in advance of the QCEW-BED. Having such timely estimates of these components would allow government agencies, lawmakers, the financial community, and other data users to have more in depth information on which to base current decisions.

The estimator utilized for CES employment estimates is termed the Link-Relative estimator. The estimator utilizes a ratio – from “matched” businesses who report in consecutive months – of current month employment to previous month employment, and multiplies that ratio by the previous months employment estimate to produce an employment estimate for the current month. In notation form, this estimator is written

\[
\hat{X}_t = \left( \hat{X}_{t-1} - x^*_{t-1} \right) \left( \sum_{i \in m} W_i x_{i,t} \right) \left( \sum_{i \in m} W_i x_{i,t-1} \right) + NB\hat{D}_t + x^*_t
\]

Where:

- \( t \)=current month
- \( t-1 \)=previous month
- \( i \)=a particular business
- \( m \)= the set of matched business that reported for both the current and the prior month
- \( NB \) = Net Birth / Death factor = Births\(_i\) - Deaths\(_i\)
- \( w_i \) = sample weight for business “i”
- \( x_i \) = employment for business “i”
- \( \hat{X}_t \) = employment estimate for current month
- \( \hat{X}_{t-1} \) = employment estimate for previous month
- * indicates businesses with atypical over-the-month employment change

An issue to consider when utilizing CES data for BED estimation is the death employment imputation that is implicitly imbedded in equation [1] above. The estimation process utilizes only those respondents who report positive employment for both the current and the prior month. Because of this matched sample concept, employment for nonrespondents and for business deaths is imputed at the same rate as the over the month change for the matched reporting businesses. The imputation of employment for deaths is necessary because of the surveys inability to account for business births and deaths on a timely basis. This imputation for deaths accounts for – and nearly offsets – the missing

\(^1\) (6) Pivetz, Searson, & Spletzer, “Measuring job and establishment flows with BLS longitudinal microdata”, MLR
employment in births; the remaining net difference between births and deaths is accounted for by the net birth death factor [which is developed using X12-ARIMA].

Because this imputation is implicit in the estimation of employment for CES, if we disaggregate the CES employment change into expansion and contraction categories, the implicit imputation of this birth employment will be included. Therefore, an issue to consider is whether this component can be removed. If the birth component can’t be removed, we must consider how much error it adds to the estimates, and if it makes the total error level unacceptable.

Another issue to consider is shared by the QCEW-BED. This is the dynamic nature of the population. Many businesses each quarter change reporting situations. For example, an establishment may merge with another business, or a large multi-establishment business may divest establishments from the corporation. This dynamic nature makes it challenging to keep track of businesses from quarter to quarter, as well as month to month. The dynamic nature of the population is described in (2)².

Because the CES survey samples and weights establishments at the unemployment insurance (UI) account level, this dynamism brings with it a difference between the QCEW-BED and the proposed CES-BED. The QCEW-BED develops statistics at the establishment (or worksite) level, where the proposed CES-BED statistics would be developed at the CES “report” level. CES “reports” are data that are reported to the CES sometimes at the worksite level, and sometimes at the UI level. Therefore, some level of establishment births and deaths for the QCEW-BED will be counted as UI-level expansions and contractions in the CES-BED statistics.

Data

Matched sample data used to produce national estimates were used to derive CES-based expansion and contraction estimates. National estimates are refreshed each year with the introduction of a new benchmark level and a new sample. The national estimation system (DEEPS) keeps separate microdata files for each year's benchmark and sample. These files typically span a final November estimate in year Y to a December preliminary estimate in year Y+1. The CES BED expansions and contraction estimates in this simulation used the same timeframes and samples as DEEPS.

CES-BED estimation was done at the basic CES estimating cell level for each benchmark/sample file and aggregated to higher levels. Some industries were excluded from BED estimates because of insufficient or little sample to derive a stable or defendable expansion or contraction estimate. These industries include tax preparation services (NAICS 541213), Class I rail lines (NAICS 482110), and churches (NAICS 813110). Each monthly expansion or contraction estimate was anchored to the previous months CES estimate of total employment. Thus, each monthly expansion and contraction estimate was re-anchored by using the official CES estimate of total employees of the prior month. Only matched sample with employment > 0 for both the previous and current month was used. If a CES report was classified as atypical in regular production, it was classified as atypical in CES-BED estimates as well. The results show we were able to separate or disaggregate the over-the-month (OTM) estimate (not seasonally adjusted) into expansion OTM change, contraction OTM change, and net B/D components. The estimates of expansions and contractions including net B/D were essentially equal to published CES OTM change estimates. Slight differences in the estimates are due to late reporters that could not be easily excluded from the DEEPS micro-database. We summed the expansion net change + contraction net change + net

² See (2) Pinkston and Spletzer, “Annual measures of gross job gains and gross job losses”, MLR, page 2
B/D, and compared it to the published CES OTM change. This was done as a quality control check for this CES-BED simulation. In all basic cells the summed values and the published net change were either the same or had very minor differences due to the different microdata utilized.

**Procedures**

The link relative estimator was described above in equation [1]. The estimator uses the employment from businesses that report in consecutive months to form a link ratio, and then multiplies that by the previous month employment estimate.

As we consider how to disaggregate equation [1] to account for BED-type statistics, one component we must consider is the net birth/death factor. It is a component of total employment, and therefore of the over the month change as well. For now we will move ahead, but keep this in mind.

The over the month change in employment can be described as a sum of several components. This is the net change in expanding businesses, minus the net change in contracting businesses, plus employment from births, minus employment from deaths. We can show this as follows.

\[
\hat{X}_t - \hat{X}_{t-1} = \hat{E}_t - \hat{C}_t + \hat{B}_t - \hat{D}_t = \hat{E}_t - \hat{C}_t + NBD_t
\]

Where

- \(\hat{E}_t\) = net change in employment from expanding businesses
- \(\hat{C}_t\) = net change in employment from contracting businesses

This depiction also shows that the net birth/death factor fits in quite well with the disaggregation of over the month change into components.

We can define expansions and contractions in the CES survey as follows:

\[
\hat{E}_t = \hat{X}_{t-1} \left( \frac{\sum_{i=1}^{m} w_i (x_{i,t} - x_{i,t-1})}{\sum_{i=1}^{m} w_i x_{i,t-1}} \right)
\]

\[
\hat{C}_t = \hat{X}_{t-1} \left( \frac{\sum_{i=1}^{m} w_i (x_{i,t-1} - x_{i,t})}{\sum_{i=1}^{m} w_i x_{i,t-1}} \right)
\]

Where

- \(m=q+r+s\)

  - \(q\) is the linked subset of units whose employment grew over the month
  - \(r\) is the linked subset of units whose employment declined over the month
  - \(s\) is the linked subset of units whose employment was unchanged over the month

We can show how this disaggregates the over the month change in employment as follows.
Equations [3] and [4] can be modified as follows to include atypical units.

\[
[6] \quad \hat{E}_t = \left( \hat{X}_{t-1} - x^*_{t-1} \right) \left( \sum_{i \in m} w_i \left( x_{i,t} - x_{i,t-1} \right) \right) + x^*_{t,q}
\]

\[
[7] \quad \hat{C}_t = \left( \hat{X}_{t-1} - x^*_{t-1} \right) \left( \sum_{i \in r} w_i \left( x_{i,t-1} - x_{i,t} \right) \right) + x^*_{t,r}
\]

Where

- \( x^*_{t-1} \) is the employment for month t-1 for the set of atypical units in month t
- \( x^*_{t,q} \) is the net change in employment for the expanding subset of atypical units
- \( x^*_{t,r} \) is the net change in employment for the contracting subset of atypical units

Therefore, the over the month change in employment can be disaggregated into three components, employment change from expanding businesses, employment change from contracting businesses, and the net difference between birth employment and death employment (i.e. the net birth death factor).

Clearly, we can disaggregate over the month employment change into components as shown. However, there is still a problem. As was mentioned earlier, we have not yet accounted for the implicit birth imputation that takes place in the link relative estimator.

Disaggregating the employment into components as shown has not made this issue go away. Equations [6] and [7] still include the implicit birth employment, because business ‘deaths’ employment is included in \( \hat{X}_{t-1} \). Therefore, both of the newly defined components include some error introduced by the implicit birth imputation.

How big is this error component? If we examine the QCEW-BED statistics, we see that death employment typically ranges between 1% and 1.8% of total employment per quarter. Therefore, 0.3% to 0.6% of monthly employment -- \( \hat{X}_{t-1} \) -- is expected to be death employment, and consequently each of the new components are expected to have an error of 0.3% to 0.6% from this source (which is implicitly imputed to account for the largest part of the correspondingly missing birth employment -- the difference between these two is accounted for by the net birth death factor). This 0.3% to 0.6% overestimate [on the magnitude of the components] is not a huge problem, but it would be desirable to eliminate or reduce the size of this error. To put this in perspective we can look at the
QCEW-BED released for September 2007. The expansions value for Total Private was 5.1 percent of total employment. At either the lower or upper bound of the anticipated death imputation error range, the expansions estimate would be unchanged at 5.1 percent of total employment.

One method we might explore to reduce the size of the implicit birth imputation error would be to use a model-based death imputation factor to (attempt to) remove this error source from $\hat{X}_{t-1}$. This could be done simply, as follows\(^3\).

\[ \hat{X}'_{t-1} = \hat{X}_{t-1} \beta_{t-1} \]

Where

\[ \beta_{t-1} = \frac{\text{continuing employment}_{t,t-1}}{\text{continuing employment}_{t,t-1} + \text{death employment}_{t,t-1}} \]

In practice, of course, we would use $\hat{\beta}_{t-1}$, which would be forecast from historical data. In the earlier equations, $\hat{X}'_{t-1}$ would be substituted for $\hat{X}_{t-1}$.

Results

We first developed quarterly estimates of CES-BED, for several reasons. This allowed us to develop a proof of concept and at the same time to review CES-BED estimates against corresponding QCEW estimates. We could not directly compare the monthly CES-BED to the quarterly BED produced by the QCEW program. One cannot simply add monthly expansions to arrive at a quarterly figure of expansions. Businesses with expanding employment in one month could have been contracting or had no employment change in a previous or following month. To simulate CES-BED on a quarterly basis, we replicated a QCEW processes by using only the CES AE ($X_t$) figures of the 3rd month of each quarter. To do this, we carried forward or overwrote CES microdata from the 3rd month of each quarter into the first and second month of each subsequent quarter. For example, a March value was "copied" into April and May. Estimates for April and May would show no change and the estimate from May to June would show the March to June (quarterly) change. We were then able to quickly produce quarterly CES-BED estimates without a major rewrite of the monthly processing logic. For ease of viewing, the contractions data were multiplied by -1 for the chart.

The quarterly results for Total Private are shown below.

\(^3\) Thanks to Kirk Mueller for proposing this adjustment.
Note that these results are a bit surprising. The magnitude of the CES contractions and expansions is in fact somewhat less than the corresponding statistics from the QCEW-BED program. Because of the death imputation the initial expectation was that this would have been reversed. Of course there are other issues that contribute to this result, including the difference between establishment level tabulations and UI (or report) level estimates. For example, employment shifts from one establishment to another within a multi-establishment business would show up as both an expansion and a contraction in the QCEW-BED, but as no change in the CES-BED. An additional issue is the age of the CES sample frame. The CES sample frame is typically nearly two years old by the time we select a sample, initiate new reporters, and begin using those data in estimation. This lag is partially accounted for by doing a birth “supplement” midway through each year. However, for these particular statistics, this omission of new birth businesses and the resultant “churn” in employment they bring may also be contributing to the difference in levels evident here.

These results are strikingly similar when comparing the trends, and tend to tell the same general story on a quarter by quarter basis. Results by industry super-sector were developed, but are not shown here. For most Super Sectors, the trends are similar to those found in the QCEW-BED, and the net difference between the expansions and contractions are similar.

Results for monthly Total Private expansions and contractions are given below.
Note that a simple seasonal adjustment was done to help illustrate the results. This seasonal adjustment does not take into account several adjustments used in practice in the CES employment estimates, including a 4/5 week adjustment to account for the number of weeks in between reference pay periods, and holiday adjustments. Because of these omissions the seasonally adjusted data shown here will not exactly match published CES data.

In the chart above, we have the following series:

- Exp = Net change in employment from expansions,
- SA_Exp = seasonally adjusted expansions,
- Emp = Net change in employment = expansions + contractions + net birth / death,
- SA_Emp = seasonally adjusted net change in employment,
- Con = Net change in employment from contractions, and
- SA_Con = seasonally adjusted contractions.

The vertical line at December 2007 marks the start of the current recession as identified by the National Bureau of Economic Research.

Note that this chart shows that the start of the recent recession – starting in December 2007 – was the result of a noticeable decrease in expansions. About six months into the recession, employment loss due to contractions increased. In more recent months, contractions have returned to a normal level; the continuation of the recession is primarily a result of a reduced level of employment gains from expansions. Because we
have these research data we can characterize this recession in near real time in more detail.

Also note that there is a substantial difference in levels between the quarterly results and the monthly results. As expected, the quarterly results are not the cumulative result of three monthly values. Many offsetting changes will happen across a three month period. The monthly values are, in fact, at about 75 percent of the quarterly values. This additional intra-quarter information on expansions and contractions is another use for these CES-BED data, in addition to the increased timeliness.

**Future Work**

As mentioned earlier, the QCEW-BED data are developed using establishment level data. This provides a way of looking at employment continuity and change that is very understandable to data users. There are, of course, other business entity concepts that can also be utilized to develop employment dynamic statistics. The CES-BED data are developed using a ‘reporter’ level concept. This is a mix between worksite level and the unemployment insurance (UI) account level. A ‘pure’ UI level concept is also an understandable business concept to data users, as it generally relates to a company within a state. However, the mixed level used in the CES-BED process might take some additional explaining for the data user.

As noted earlier, the CES approach means that for some multi-establishment UI accounts, openings and closings within a multi-worksite UI account will appear in the CES-BED as expansions and contractions. For example, if a large multi-establishment retail chain store (that reports to CES at the UI level) opens at a new site this is an expansion of an existing company, and not a new company birth (although it is an establishment birth). At the UI account level, these cases look like any other employment expansion or contraction. Also, as mentioned earlier, offsetting employment shifts at the establishment level within a multi-establishment business will contribute to smaller net changes in growth and decline in the CES-BED statistics. These differences, in addition to the difference in time periods, will add to the difference between the CES-BED and QCEW-BED series. These differences must be documented and understood to a greater degree prior to publishing a CES-BED series. An approach to help understand this better is to develop ‘pure’ UI-based CES estimates of expansions and contractions, and compare them to QCEW-BED data developed at a UI level over the quarter and over the month. While the administrative effects in a monthly QCEW-BED statistic may diminish the utility of the monthly comparison, it might still shed some light on this issue.

The work done to date has shown that we can disaggregate the non-seasonally adjusted CES net change into expansions, contractions, and the net birth/death factor. However, the statistics of interest in the CES program are the seasonally adjusted net change in employment. We need to develop a procedure to make the seasonally adjusted components be equal to the total net change.

Another issue that must be resolved prior to considering development of these statistics for regular release is how to account for rare manual adjustments [overlays] to the estimates, both from a conceptual and from an operational perspective.

As mentioned earlier, the implicit birth imputation adds an error of 0.3% to 0.6% in both of the components; expansions and contractions. Given that this implicit birth imputation error always induces a small overestimate (on the absolute value) of the estimate, we could explore developing forecasts to remove this component. This could be developed and applied as shown in equation [8]. Note, however, that there are some policy issues associated with this that are discussed below after touching on benchmarking issues.
We benchmark the AE (or X) values on an annual basis. Is this a sufficient benchmark for expansions and contractions? Certainly this AE benchmarking will keep the components in line with the CES over the month employment changes. However, it may not necessarily keep the components in line with the QCEW-BED values. We need to explore methods to align the CES-BED with the QCEW-BED. Previous research has suggested that these change statistics will not align for different length measurement periods. Given this previous research, this benchmarking or alignment of CES-BED with QCEW-BED may not be a trivial issue. A major concern is that after benchmarking the CES-BED and QCEW-BED may tell different stories. If possible we need to find a way to align or anchor the estimates to the population.

At the agency level, we have to consider both the positives and the negatives of releasing yet another set of statistics that will compete with an existing Bureau product. On the positive side of course, we would be providing useful information to the public months before they would be able to get it from the QCEW-BED. On the negative side, however, there are a number of issues. Among them is the possibility that some media personnel would use this additional information to write articles berating the program when the expansions or contractions data didn’t meet public expectations. We also need to consider a situation where the CES-BED and QCEW-BED are telling substantially different stories. How would we explore the differences, and as an agency, how would we respond to the resulting criticism.

We need to consider whether this would be produced as a first, second, or third close item. There are substantial resource issues associated with first close items. However, the longer we wait to publish these data items, the less utility they have as a timely economic indicator. We then need to consider how we would revise and seasonally adjust these new series. How large can we expect revisions to be? Would revisions in these series cause a lot of public comment? Does the QCEW-BED program revise their estimates? If so – and if they get public comment on the revisions – then this might provide us with some guidance. We should keep in mind that when we revise AE (X_t) we would also be revising these CES-BED components. A large revision in one is almost certainly going to result in a large revision to the other. This will open up scrutiny and criticism on two fronts. We also need to explore how we would go about analyzing these new data types.

When considering adding a model to remove implicit deaths, we need to keep in mind that this adds another component to estimate and another component for public scrutiny, not only in the context of CES-BED, but in the AE context as well. Given the scrutiny that the net birth-death model currently undergoes, additional modeling information would give the usual “phantom jobs” theorists more to misunderstand and misquote, and more to write disparaging erroneous articles about (and more for BLS to have to respond to).

Modifying DEEPS – the national estimation system –to develop these estimates is not expected to be problematic. It is, however, a resource issue for the DEEPS staff. This is true from both a development and from a production standpoint. Similarly, reviewing, editing, and commenting on these series on a production basis may pose a substantial resource issue for the economists who review estimates on a monthly basis. We need to consider if there are resources available for additional ongoing production before making a decision about whether or not this is something that CES should publish on a production basis.

In summary then, the following issues need to be explored or documented to a greater degree before we can consider making these series part of the CES program.

---

4 See (2) Pinkston and Spletzer, “Annual measures of gross job gains and gross job losses”, MLR
Conceptual Issues
- Differences caused by different time period (i.e. month vs. quarter)
- Differences caused by different business entity concept (i.e. establishment vs. ‘report’)
  - Develop ‘pure’ UI level statistics from both CES and QCEW and compare over the month and over the quarter
- Does the age of the CES sample contribute to the level difference between CES and QCEW BED statistics
  - How should we explore this issue
  - Can we estimate the effect this has on the BED estimates
- Develop and refine procedures to seasonally adjusted BED data
- Develop a procedure to align the seasonally adjusted components with the seasonally adjusted total net change
- Implicit death imputation
- Benchmarking
- Analysis procedures for new data types

Operational issue
- How to distribute AE overlays to CES-BED components
- Resources to develop systems capable of producing these series on a production basis
- Resources to develop and review series on a production basis

Organizational Issues
- How to market new competing statistics that will be different than the existing ones
- How to mitigate the possibility of new criticisms of existing CES AE data that may be raised by additional information available due to this disaggregation
- How to mitigate the possibility of new criticisms of existing CES AE data that may be raised by additional information available if we explicitly model the death imputation employment in order to remove it
- How to mitigate or be prepared for criticisms aimed at both AE and one or both BED components when we have a sizeable AE revision

Conclusions
This paper has explored the development and first results of business establishment dynamics type data from CES reports. We have outlined a number of issues that we have identified that must be resolved before proceeding any further with this project. This paper should be taken as a starting point for additional discussion on this topic. How do we move forward to resolve the issues we have identified? Are there other issues that need to be explored?

In general, the results have been promising. The comparison of the “quarterly” CES-BED with the QCEW-BED shows that some differences exist between the series, but they are telling the same general story. A review of the monthly CES-BED results shows many results that are expected, results that shed additional light on what part of the employment change is driving the total net change in employment in any given month.

These preliminary results lead the authors to conclude that CES reports can be used to develop CES-BED estimates. While there are many complex issues to resolve, none of them appear to be insurmountable. While there are some complex technical issues to resolve, some of the bigger questions relate to policy issues.

The authors recommend that the project continue to move forward to address the issues, with the ultimate goal of publishing these series, at least on an experimental basis.
Any opinions expressed in this paper are those of the authors and do not constitute policy of the Bureau of Labor Statistics.

References
Several papers have been written that illustrate issues and uses of the QCEW-BED data. Among them are:


More information on the QCEW-BED can be found at [http://www.bls.gov/bdm/home.htm](http://www.bls.gov/bdm/home.htm)

More information on the CES program can be found at [http://www.bls.gov/ces/home.htm](http://www.bls.gov/ces/home.htm)