How hard is it to count? Reconciling monthly survey and administrative data: experiences from the CES program November, 2016

Steven Mance[1]

U.S. Bureau of Labor Statistics, 2 Massachusetts Avenue NE, Washington, DC 20212

Abstract

The Current Employment Statistics (CES) program at the U.S. Bureau of Labor Statistics produces monthly nonfarm payroll employment, hours, and earnings estimates at the national, state, and metropolitan level from a fast-turnaround survey of about 146,000 businesses and government agencies representing approximately 623,000 establishments. These data are subject to many common sources of error such as sampling and nonresponse, and the employment data are benchmarked (reconciled) using the Quarterly Census of Employment and Wages (QCEW), which is based on administrative records from the Unemployment Insurance system, available 6-8 months after the sample-based estimates. QCEW employment snapshots are available at a monthly frequency; however, noneconomic changes and differences in coverage, response, and seasonality complicate its use in benchmarking. Each year, National CES data are reconciled to a QCEW-based March level, while state and metro series are replaced using adjusted QCEW values for every month. The respective benefits, complications, and challenges of both methods is recounted. The recent use of QCEW data to reconstruct 23 years of metropolitan employment data due to definitional changes is detailed. Future directions in the use of administrative data for the CES program is explored.

Key Words: Administrative data, benchmarking, time series

1. Introduction

The Current Employment Statistics (CES) program is a large, venerable, fast-turnaround survey at the U.S. Bureau of Labor Statistics that produces some of the most closely-watched economic indicators in the United States. Each month, hundreds of thousands of worksites report information on their payroll employment, hours, and earnings through a variety of methods. These survey responses are then processed by economists and statisticians in the CES program office. BLS staff swiftly prepare the national-level data for publication, packaged with labor force information from the Current Population Survey, and released in *The Employment Situation* report at 8:30 a.m. on the first Friday of each month[2]. At this moment, television reporters for the major U.S. news and financial networks emerge from a press lockup at the Frances Perkins Building in Washington, DC, to broadcast the headline numbers to their viewers; newspaper journalists in the lockup reconnect to the internet and upload their reports; and the data begin to have their effect, moving markets and shaping the national discourse. Roughly two weeks later, data on states and metropolitan areas are released, providing a critical

---

[1] Any opinions expressed in this paper are those of the author and do not constitute policy of the Bureau of Labor Statistics.

[2] There are some exceptions, such as when the first Friday of the month is a Federal holiday, the publication date is usually moved to the preceding Thursday. Publication dates are themselves published well in advance.

look at labor market conditions at the subnational level. As much attention is paid to the monthly survey, producing high quality CES time series is heavily reliant on using administrative data from the Quarterly Census of Employment and Wages (QCEW). The QCEW—also venerable, known as the ES-202 from its inception in the 1930s until 2003, and even larger than CES—takes administrative data from the Unemployment Insurance (UI) system and processes it for statistical uses. It represents the vast majority of jobs in the United States, forms a business register for BLS, serves as a sampling frame for CES and other surveys, and is the basis of the CES universe. A critical aspect of the QCEW is that it collects employment information matching CES concepts, generally available six to seven months after the end of the reference quarter. Since it is a census, data from the QCEW are not subject to sampling error, allowing for CES employment data to be anchored to a firm, sampling-error-free level.

The focus of this paper is on how employment data from CES are "benchmarked"—i.e., reconciled—using administrative counts from the QCEW. The CES data are not simply substituted for the QCEW. A great deal of work is required to take the census levels and form them into a time series fully conforming to CES definitions. As statistical agencies increasingly leverage administrative data, it may be helpful to learn from the experiences CES has had working with the QCEW for over seven decades. Section 2 of this paper gives a necessary overview of the CES and QCEW programs. Section 3 details the benchmarking process. Benchmarking is conducted differently at the national level, compared to state and area data. Section 4 presents an extension of the benchmarking process, where a definitional change required the reconstruction of thousands of time series spanning over two decades. Section 5 briefly discusses future directions—solutions BLS is considering to incorporate administrative data faster while also improving the CES time series.

2. The Current Employment Statistics Survey and Quarterly Census of Employment and Wages

The Current Employment Statistics program produces monthly time series estimates of nonfarm payroll employment, hours, and earnings for the U.S. as a whole, along with the 50 states, the District of Columbia, the territories of Puerto Rico and the Virgin Islands, and 443 metropolitan areas[3]. The survey consists of about 146,000 businesses and government agencies representing 623,000 establishments. Private sector businesses are sampled using a modern probability-based selection procedure, while the government sample uses a quota design (Kelter 2016). The reference period is the pay period containing the 12th of the month, and data are collected in a multi-modal fashion (Johnson 2016). In addition to being available at different geographical levels, CES data are published by industry, using the 2012 North American Industry Classification System (NAICS). At the national level, many estimates are available at the most-detailed 6-digit NAICS level, while industry detail is generally more limited for states and metro areas. The CES program began in 1915, initially publishing hours and earnings data for selected industries on an annual basis, and began publishing a monthly total nonfarm employment value in 1939.

The Quarterly Census of Employment and Wages provides a snapshot of employment and wages for all U.S. jobs covered by the Unemployment Insurance (UI) system often in great industry and geographic detail. Most employers are required by law

---

[3] Metropolitan areas include Metropolitan Statistical Areas, and divisions thereof; New England City and Town Areas, and divisions thereof. Most divisions conform to standards set by the Office of Management and Budget, but CES also produces data on several "non-standard" divisions.

to pay UI taxes and must report total wages along with three months of employment levels on a quarterly basis, allowing QCEW to form a census of all UI-covered employment. Over 9 million establishments report this information in any given quarter, and QCEW staff are responsible for assigning NAICS codes and geographic identifiers for establishments, updating information through the Annual Refiling Survey. Employment covered by the UI system represents about 97 percent of nonfarm payroll employment. Due to the similarity in scope and availability of data, CES uses a version of the QCEW linked in time, based on the Longitudinal Database (LDB) as its sampling frame. Non-agricultural business establishments that have at least some jobs covered by UI are generally eligible for inclusion in the CES, while CES does not sample any worksites without UI-covered jobs.

Both CES and QCEW are "fed-state cooperative programs", conducted jointly by the BLS and individual state workforce agencies. The QCEW traces its origins to the Social Security Act of 1935, a major piece of social welfare legislation, born out of the Great Depression, which contained provisions to assist states in administering their UI systems. Although CES predates QCEW, it could not take its modern form without the administrative data from the UI system.

Total employment ("All Employees" or AE) is estimated from the CES sample using a "weighted link relative"[4]:

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

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

Where:

i = matched sample unit;
j = matched sample unit where the current month is atypical;
w_i = weight associated with the CES report;
$ae_{c,i}$ = current month reported all employees;
$ae_{p,i}$ = previous month reported all employees;
w_j = weight associated with the CES report where the current month is atypical;
$ae_{c,j}$ = current month reported all employees where the current month is atypical;
$ae_{p,j}$ = previous month reported all employees where the current month is atypical;
$\widehat{AE}_c$ = current month estimated all employees;
$\widehat{AE}_p$ = previous month estimated all employees;
and b_c = current month birth/death estimate.

The matched sample represents all businesses reporting positive employment in the current and prior month. At the national level, atypical business reports are designated by industry analysts. At the state and metro area level, a robust estimation procedure is used to flag atypical and other influential reports and reweight them to lessen their impact on the estimates. Since this estimator applies a relative change to the prior month's estimate level, an initialization value is required. This is the "benchmark level", which is derived primarily from the QCEW and is discussed at length in section 3.

---

[4] This formula can be found in the BLS Handbook of Methods, Chapter 2, which provides information CES methodology.

For series with inadequate sample size, primarily in the state and metropolitan domains, small area estimation techniques are used, including variants of the Fay-Herriot model as well as the CES Small Domain Model, all of which use a forecast of the historical benchmarked series or QCEW trend. All of these small area techniques use the weighted link relative as an input, and estimate relative employment change.

Data types other than all employees are estimated using an estimator known as the "weighted difference link and taper." A benchmark source is not available matching CES constructs for these data types, so they will not be discussed further.

The CES sample is not drawn on a real-time basis, which makes business births difficult to capture. In addition, when many establishments go out of business, they simply stop reporting. If a business no longer employs anybody, who would be filling out the CES forms? In order to account for the difficulty in capturing employment from business births and deaths, CES uses the "Net Business Birth/Death Model." In estimation, deaths are implicitly imputed using the matched-sample link. This implicit imputation accounts for missing establishment birth employment which cannot be directly captured in real time. Second, the net birth/death factor is added to the current month's estimate, because the employment attributable to births and deaths is not the same each month but the net difference between the two is fairly stable and seasonal. This factor is created using the Longitudinal Database (LDB), which contains QCEW data linked over time. Essentially, CES estimation procedures are simulated in the LDB, which contains births and deaths and does not contain sampling error, to calculate the historical errors associated with the inability to capture employment from births and deaths in real time. These historical factors are then extrapolated to the current period using an ARIMA forecast (adjusted for outliers) to be applied in estimation[5].

QCEW provides an excellent count of employment, but it is not designed to form a monthly time series. Changes in establishment industry, area, and ownership are usually incorporated in the first quarter of each year, but the prior year's macro data are not reconstructed. This can create level shifts in the series. Administrative error that may be negligible in terms or levels may be quite serious when the value of interest is the over-the-month change. Artifacts of the quarterly reporting frequency can be found in the microdata, such as "stair-stepping", a seam effect where a worksite reports the same number of jobs for every month within a quarter, with jumps between quarters, which may even be evident in the macro level. There are also seasonal differences between CES and QCEW matched microdata and macro seasonal differences attributable to program coverage. Using results of multiple response analysis surveys (RAS), BLS staff attributed seasonal reporting differences largely to reference period error (e.g. QCEW responses are more likely to incorrectly report end-of-month data, or to stair-step) and imputation in the administrative data (although reporting is mandatory, there is some item nonresponse.[6]) These difference add up, and are most pronounced between December and January, where national-level seasonal layoffs in the QCEW are typically hundreds of thousands larger than in CES, and can be seen in most states, areas, and industries. Seasonal differences due to scope are less widespread and generally found in education series. The employment of student workers, not covered by UI, is highly seasonal, with very few on the payrolls between semesters. CES also considers faculty to be employed even if they are not getting paid during summer break. That results in a pronounced difference between CES and QCEW local government education, where the latter shows a summer

---

[5] The net business birth/death model is explained in Mueller (2006).
[6] Groen (2012) examines the results of the RAS to show how differences in reporting between CES and QCEW drive many of the differences between the two series.

jobs loss about 1 million larger than in CES. Many of these concerns can be mitigated, however, which is a major consideration in the CES benchmarking process.

3. The Benchmarking Process

Benchmarking is the process by which the CES employment estimates are reconciled to outside sources, primarily the administrative counts from the QCEW. The benchmarked level creates a starting point to apply the weighted link relative estimator; it keeps the CES trend anchored to a source free of sampling and birth/death error; and the benchmark revision serves as a proxy for total survey error in the CES. Newly benchmarked data are currently released to the public on an annual basis with January estimates. If the QCEW could simply be used as a monthly time series without any concern or complication, then reconciling the CES data to it would be a simple matter— and dozens of economists and statisticians in the CES program office would need not spend the winter months toiling through the benchmarking process—but that is just not so. National and subnational data are benchmarked using different techniques. In this section, the "anchor and wedge" technique used in the national benchmarking processed is described, followed by the "adjusted replacement" method used in benchmarking state and metropolitan area data.

3.1 Benchmarking CES National Data – The Anchor and Wedge

The national benchmarking process relies on forming an independent "anchor"— the benchmark level—for March of every year. The CES estimates are still used to form the month-to-month changes in the time series, but the March values must be aligned with the benchmark. The benchmark error is distributed linearly across the 12 months, a process referred to as the "wedge." After the data have been re-anchored to the new benchmark level, the weighted link relatives from April forward, with updated birth/death factors, are applied to form the "post-benchmark" estimates.

QCEW data form the basis of the March benchmark anchor, representing 97 percent of nonfarm payroll employment. Most of the work for this part is done by QCEW staff at BLS and in the State Workforce agencies, assigning NAICS and ownership[7] codes, collecting data, fixing reporting errors, and imputing missing values. CES takes the QCEW micro data and maps it to its structure, excluding out-of-scope employment such as agriculture and private households. A small percentage of employment in the QCEW has not been assigned a NAICS code. These "unclassified" establishments are mostly new worksites with relatively few workers. CES assumes that employment at these sites is within scope and therefore must be mapped to its series. A simple, proportional method is used whereby each industry receives an unclassified distribution relative to its classified total. For example, if an industry has 5 percent of total private nonfarm employment with NAICS codes assigned, 5 percent of the total private NAICS-unclassified employment will be mapped to it.

Once QCEW-based employment levels are set, CES must account for "noncovered employment," the roughly 3 percent of nonfarm payroll jobs not covered by unemployment insurance. Examples of noncovered jobs include student workers on the payroll of the college or university they attend; hospital interns; elected government officials; railroad workers; employees of religious organizations, such as parochial school

---

[7] Establishments are either considered to be part of federal, state, or local government or of private ownership. Ownership is usually fairly clear-cut, far more so than industry, although sometimes the lines between state and local government can be murky, e.g. public university systems. Tribal authorities are included in local government.

Table 1: Benchmark Wedge Adjustment Process

	Benchmark Levels	Monthly levels, prior to Mar-15 benchmark	Benchmark Revision	Wedge Amt.	Monthly levels, after Mar-15 benchmark
Mar-14	885.5	885.5	-	0.0	885.5
Apr-14	-	892.1	-	-0.6	891.5
May-14	-	897.0	-	-1.2	895.8
Jun-14	-	903.0	-	-1.8	901.2
Jul-14	-	901.3	-	-2.4	898.9
Aug-14	-	902.4	-	-3.0	899.4
Sep-14	-	902.9	-	-3.7	899.2
Oct-14	-	905.9	-	-4.3	901.6
Nov-14	-	909.5	-	-4.9	904.6
Dec-14	-	914.5	-	-5.5	909.0
Jan-15	-	904.8	-	-6.1	898.7
Feb-15	-	906.5	-	-6.7	899.8
Mar-15	900.4	907.7	-7.3	-7.3	900.4

Table 1 provides an example of how the wedge benchmarking process works, using the most recent benchmark of the *Electronic Markets and Agents and Brokers* series. Values are for all employees, not seasonally adjusted, in thousands.

teachers; and, in some states, certain officers of corporations. Unemployment insurance laws vary from state to state, and therefore BLS relies on State Workforce Agencies to identify industries where a substantial number of jobs not covered by unemployment insurance may exist[8]. Railroad workers are covered by an alternate system of unemployment insurance, and CES is able to obtain a count of these jobs from the Railroad Retirement Board. For other industries, CES derives noncovered employment estimates primarily using the County Business Patterns (CBP) and the Annual Survey of Public Employment and Payroll (ASPEP), two programs at the U.S. Census Bureau. CBP derives employment counts using administrative data such as Social Security filings, while the ASPEP is based on a sample of state, local, and federal government entities. Critically, both include employment not covered by unemployment insurance. However, both are available only at a substantial lag (approximately 2 years) and only on an annual basis, with a mid-March reference period. Both CBP and ASPEP are, naturally, subject to non-sampling error[9], and the ASPEP also contains sampling error. For industries where a substantial number of noncovered jobs are believed to be present, CES compares employment levels from CBP and SPEP against the QCEW-based counts to derive an

---

[8] Changes to state unemployment insurance laws are published each year in the *Monthly Labor Review.* For an example, see Lancaster (2016).

[9] CBP and ASPEP are created using the Census Business Register, which presents a source of non-sampling error vis-à-vis the QCEW. This is an unfortunate feature of the decentralized U.S. statistical system, where different statistical agencies, most housed in different departments from one another, are prohibited from sharing a great deal of data with one another. See Fixler and Landefeld (2006) for more on the ramifications. Currently, Census shares NAICS codes with BLS for a subset of reports declared free of federal tax information (a legal hurdle), and BLS provides codes where Census lacks them.

estimate of jobs not covered by unemployment insurance. Due to the fact that this value is substantially out-of-date, the noncovered estimate is trended forward using the growth rate in the QCEW. It is then added to the QCEW-based level to complete the March benchmark level.

After the benchmark anchor is set, the CES time series is wedged back to the previous March value. The wedge process is relatively simple. First, the difference between the previously published March estimate and the benchmark is calculated. This is the "benchmark revision." An assumption is made that the error accrued at a constant rate since the prior benchmark and each month is adjusted accordingly: the previous April is adjusted by one-twelfth of the benchmark revision, the previous May by two-twelfths, and so on. An example of how this process works is presented in Table 1. Noneconomic code changes[10] in the QCEW—which are negligible at high-level aggregates, but not necessarily so for detailed industries—are thus smoothed in the CES series from one year to the next. From the March benchmark level, the previously-estimated relative employment changes are applied, using updated birth/death adjustment forecasts, and additional monthly estimation can proceed.

Although error is assumed to accrue constantly between one March and the following March, which seems like a reasonable approximation, when one considers multiple years of data, this assumption is somewhat more complicated: the time series structure of the assumed error process of the job growth rate is assumed to be constant for 12 months, followed by a sudden level shift. See Figure 1 for an illustration. This can result in a spurious acceleration or deceleration in the published job growth rate. For example, the March 2006 total nonfarm benchmark revision was quite positive at 752,000 (0.6 percent), while the March 2007 revision was moderately negative at -293,000 (0.3 percent). The wedge benchmark results in over-the-month employment changes being adjusted up by 63,000 from April 2005 through March 2006, while the same figures were adjusted down by 24,000 from April 2006 through March 2007. The process inserts a sudden deceleration of 87,000 in the monthly job growth rate in April 2006. Fortunately, affected April data are far enough back that this should not materially affect decisions by government policy makers and businesses. (The full effect of this April 2006 deceleration was not published until February 2008.) Sudden changes in the job growth rate may cause problems in seasonal adjustment, although the shift imputed by the wedge benchmark is small compared to other changes that have been observed, such as steep job losses in late 2008 and early 2009[11]. Other methods of distributing the benchmark error are not without problems. The Denton Method and time series error models could be used to reduce the sudden shifts in implied error. However, these would require additional years of historical data be continually revised[12]. They would also be more complicated than the neat and simple wedge technique.

A feature of the wedge and anchor technique is that, once the not seasonally adjusted data are benchmarked, they are not revised—unless, that is, non-economic changes require a reconstruction. The biggest changes at the national level are generally due to updates to the industry classification system. The largest of these changes came when CES switched from the 1987 Standard Industrial Classification (SIC) to the 2002

---

[10] Non-economic code changes occur when the area, industry, or ownership codes are updated (generally through the Annual Refiling Survey) and do not reflect an actual, concurrent change in economic activity.

[11] For different perspectives on how the 2007-09 Recession affected CES seasonal adjustment, see Kropf and Hudson (2012) and Wright (2013).

[12] For detailed information on these methods, see Dagum and Cholette (2006).

NAICS with the 2002 March benchmark[13]. The SIC was based around the idea of the United States as a goods-producing country, with heavy detail in manufacturing but substantially less detail in service-providing industries. QCEW data were "dual-coded" in the first quarter of 2001, receiving both SIC and NAICS codes. Ratios between the SIC and NAICS industries were obtained and used to create CES time series on a NAICS basis. Take, for example, the SIC 3711 *Motor vehicles and car bodies*. Nearly half of the employment (49.5 percent) was in establishments dual-coded with NAICS 336111 *Automobiles*, with most of the rest in NAICS 336112 *Light trucks and utility vehicles* and NAICS 336120 *Heavy duty trucks*. Proportional amounts were allocated to each series back to 1990. Further historical data was not reconstructed unless there was a very close one-to-one match between SIC and NAICS industries, where histories were reconstructed as far back as SIC histories would allow, and at very high levels of detail (2-digit NAICS sectors, BLS super sectors, and higher aggregates), where series were reconstructed back to 1939. This caution was exercised since 2001 SIC to NAICS ratios would be less and less reliable further back in time, and many high-tech NAICS industries were undreamed of several decades ago.

Figure 1: Error in Monthly Over-the-Month Change Implied by Wedge Process
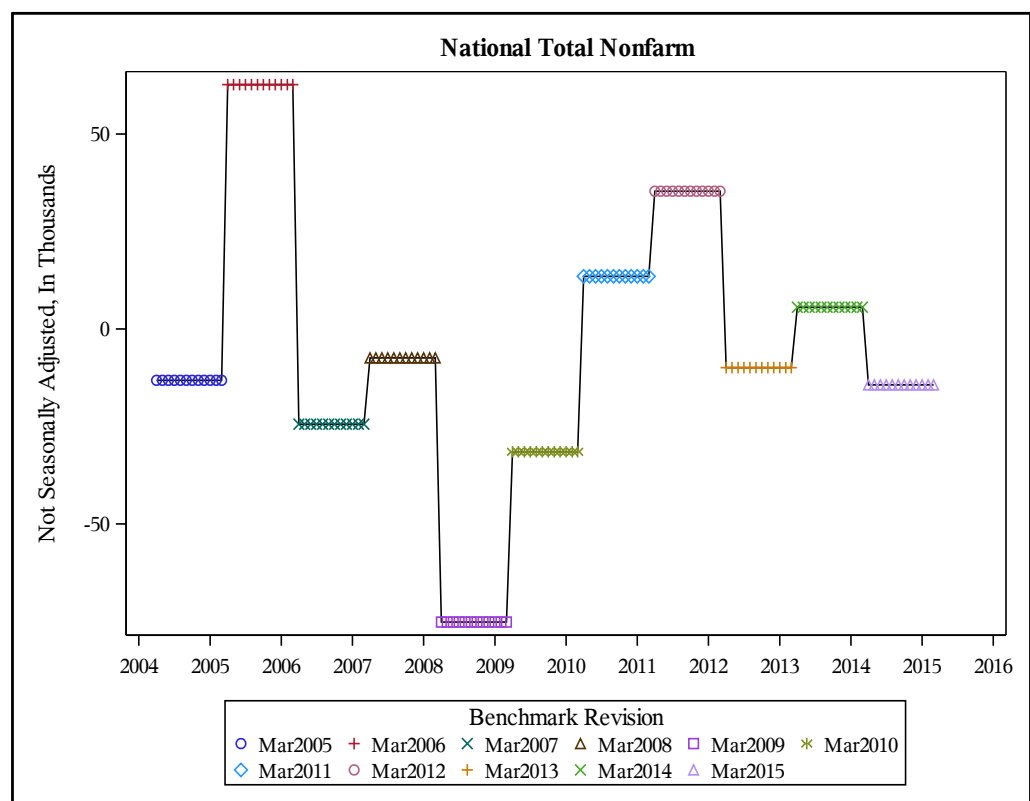
The wedge process adjusts the over-the-month change by a uniform amount between March benchmarks. This assumes that error accrues at a constant rate for 12 month periods, but often implies large shifts in the error rate from March to April.

---

[13] This benchmark also finalized the redesign of CES from a quota sample to a scientific probability-based survey. For a full discussion of changes associated with the 2002 benchmark, see Morisi (2003).

In addition to structural coding changes, such as quinquennial updates to the NAICS, large non-economic code changes in the QCEW microdata sometimes necessitate a reconstruction. A recent example is when about 466,000 jobs related to state-funded home health care for the elderly and persons with disabilities were found to have been incorrectly coded in NAICS 814 *Private households*—which is out-of-scope for CES—into NAICS 62412 *Services for the elderly and persons with disabilities*. CES worked with State Workforce Agencies in the six states where these jobs were found to be miscoded to recreate historical time series. Substantial and apparently spurious jumps in the administrative series were found and smoothed using a trend estimation procedure in X-12-ARIMA[14].

3.2 Benchmarking Subnational Data – Adjusted Replacement

State and metropolitan area data are benchmarked using a different technique from the national data—referred to in this paper as "adjusted replacement." It may be succinctly characterized as deriving benchmark levels for every single month, not just March, with adjustments needed to cope with some of the consequences. Sampling error is not present where this technique has been used, but administrative error is of greater concern than with the anchor and wedge. There are three primary reasons why different benchmark techniques are used for different domains. The first is that state and metropolitan area series tend to have higher relative sampling error than their national counterparts, therefore a technique that purges sampling error carries more benefit. In truth, a super sector estimate for a large state may have a larger sample size and less relative sampling error than some detailed national industries, but the demarcation of benchmark techniques along the national/subnational divide is a natural one. Second, nearly half of state and area series employment estimating cells are modeled, while only 6 national series are modeled. An important component of all CES models (the Small Domain Model and variants of the Fay-Herriot Model) is a forecast of the QCEW. If the initial estimates contain a mere forecast of the QCEW, why not replace with the actualized value? The third reason is historical. Until 2011, when BLS centralized processing, State Workforce Agencies were principally responsible for producing the monthly estimates as well as benchmarking them on an annual basis. In the 1980s, as QCEW (then ES-202) data were made available on a more-frequent-than-annual basis, states began replacing more than just March. The states, it turns out, were laboratories of benchmarking techniques—but methodological decisions for the national data were made separately, and thus different techniques are used.

The core aspects of adjusted replacement were covered in the previous section on the wedge and anchor. QCEW staff provide essentially the same micro files for the state benchmark, except every month's value is used. Noncovered employment values are derived, but for all months instead of just March. Non-economic codes changes in the QCEW are introduced in the first quarter. In national CES series, this is handled implicitly in the wedge. For state and area series, this is handled explicitly with a different kind of wedge. Since the data appear in their new industry or area in the entirety of the first quarter, the sign of the wedge changes from December to January. The mechanics of this process are shown in Table 2. If a non-economic code change results in a substantial proportion of a series' employment being added or subtracted, a longer reconstruction in performed. In many cases, when only one or two firms are responsible for the large change, longitudinally-linked QCEW data are used. A time series of the

---

[14] X-12-ARIMA is a seasonal adjustment and time series modeling program developed and maintained by the U.S. Census Bureau. The current version of the program is X-13ARIMA-SEATS, which CES currently uses for seasonal adjustment.

entering/exiting employment is created as far back as possible and added/subtracted from the published time series. This is not always possible, or the net non-economic code changes come from many firms, and in those cases a longer, historical wedge or a level shift of the entire history is used instead of the micro-level reconstruction.

Noncovered values must be created for all months and are provided by the individual State Workforce Agencies. Some states conduct supplementary surveys, or rely on state or local government payroll counts. States may compare CBP and ASPEP data with the QCEW, using the CES national procedure, although these relationship are more volatile at greater geographic detail. As a fallback, national levels are prorated to states when no better source can be found. The main sources used for national levels, the CBP and ASPEP, are only available for March. While some industries can be assumed to have fairly stable noncovered employment levels, and March can be expected not to differ much from other months, but this is not always reasonable. As mentioned in Section 2, student worker employment falls dramatically every summer, and can be seen by comparing matched CES and QCEW reports for colleges and universities. The difference between the two reports can reasonably be considered mostly to be noncovered student workers. Often times, this microdata-level relationship can be used to provide a reliable, seasonal noncovered value. This is difficult for schools that either do not respond to CES or are not sampled, although in some cases the CES/QCEW ratio at similar colleges in the state may be used.

Table 2. Noneconomic Code Changes in the Replacement Benchmark

	NECC Employment Moving to Series	NECC Employment Leaving Series	Net NECC	NECC Adjustment Amt.
Mar-14	-	-	-	0
Apr-14	-	-	-	71
May-14	-	-	-	142
Jun-14	-	-	-	213
Jul-14	-	-	-	284
Aug-14	-	-	-	355
Sep-14	-	-	-	426
Oct-14	-	-	-	497
Nov-14	-	-	-	568
Dec-14	1044	192	852	639
Jan-15	1260	185	1075	-142
Feb-15	1280	190	1090	-71
Mar-15	1383	198	1185	0

In this example, over 1000 jobs are going into a series, and nearly 200 are leaving due to NECCs. Although jobs do not appear in the replacement benchmark until January, the net effect is taken in December. The adjustment amounts smooth the effect of 852 new jobs appearing in the first quarter of 2015.

The other seasonal difference in scope—faculty employment in the summer months—is adjusted for in local government education series in a similar fashion to the national wedge technique. Instead of March benchmarks, adjusted replacement values for May and September are used as anchors. The CES estimate links are applied from May through September, and a difference is calculated at that point. This discrepancy is then

wedged proportionally over June, July and August. The result is that the series is tied to the administrative count, but that the summer months do not show a severe drop due to a difference in scope.

It should also be noted that the adjusted replacement extends beyond the official March benchmark—the data are constructed for an additional six months, through September. Although it is generally beneficial to replace through the most recent quarter available, the third-quarter QCEW micro data are provisional. Most obvious reporting errors will eventually be cleaned up by QCEW staff, although CES staff at BLS working with their state counterparts have to do this themselves with the most recent data. Reporting error can range from the common "fat-finger" type, where the respondent clearly typed the wrong value, to types perhaps peculiar to the QCEW: respondents who enter zero for every month until the end of the quarter and respondents who report the same employment level within a given quarter, with large shifts between quarters. Reporting changes from one quarter to the next generally present greater challenges than actual reporting errors. Firms will sometimes "break-out" their reporting—going from reporting all of their state's employment together to providing individual worksite detail—or they may "consolidate" it, which is the reverse procedure. If all of the employment in the break-out or consolidation is not found within the same CES time series, this creates a non-economic level shift. If this is found, CES analysts will manually adjust the time series with a wedging technique similar to that used for other non-economic changes, with the values adjusted for the quarter in which the change occurred. In addition to straightforward break-outs and consolidations, more complex transactions may occur in the Unemployment Insurance system. It could be a merger or acquisition, or possibly the revamp of a large company's reporting, but in many cases employment from multiple establishment reports will be transferred to multiple others, with the relationship between the former and the latter quite unclear. In these cases, a great deal of investigation must be done by the CES analysts, often relying very much on their State Workforce Agency counterparts, to clarify what precisely happened.

In order to find large reporting errors and changes, a time series macro-screening procedure is employed. After the initial addition of noncovered employment to the QCEW and the automated smoothing of most non-economic code changes, the resulting series is processed in X-13ARIMA-SEATS[15], a seasonal adjustment and time series modeling program maintained by the U.S. Census Bureau. A seasonal Auto-Regressive Integrated Moving Average (ARIMA) model is fit to the data and "outlier" regression variables are tested against the model residuals. The variables can take the form of an additive outlier (a one-month spike) or a level shift (a permanent change.) Data entry errors often present as the former while reporting changes often show up as level shifts at the start of a quarter. Beyond a certain t-value, the outliers are kept in the model and included in a report for CES analysts to investigate. Often economic events will show up as additive outliers and level shifts, but many non-economic features have been found and resolved due to this screening procedure.

After the benchmark is completed, data are "re-estimated" from the September level. Some additional late-reporting microdata are incorporated in the estimates, new birth/death factors are used, and new forecast values are used in modeled series.

---

[15] X-13ARIMA-SEATS is available for free, along with a great deal of documentation, at: https://www.census.gov/srd/www/x13as/.

Figure 2. Construction of Inputs for the Two-Step Seasonal Adjustment Process
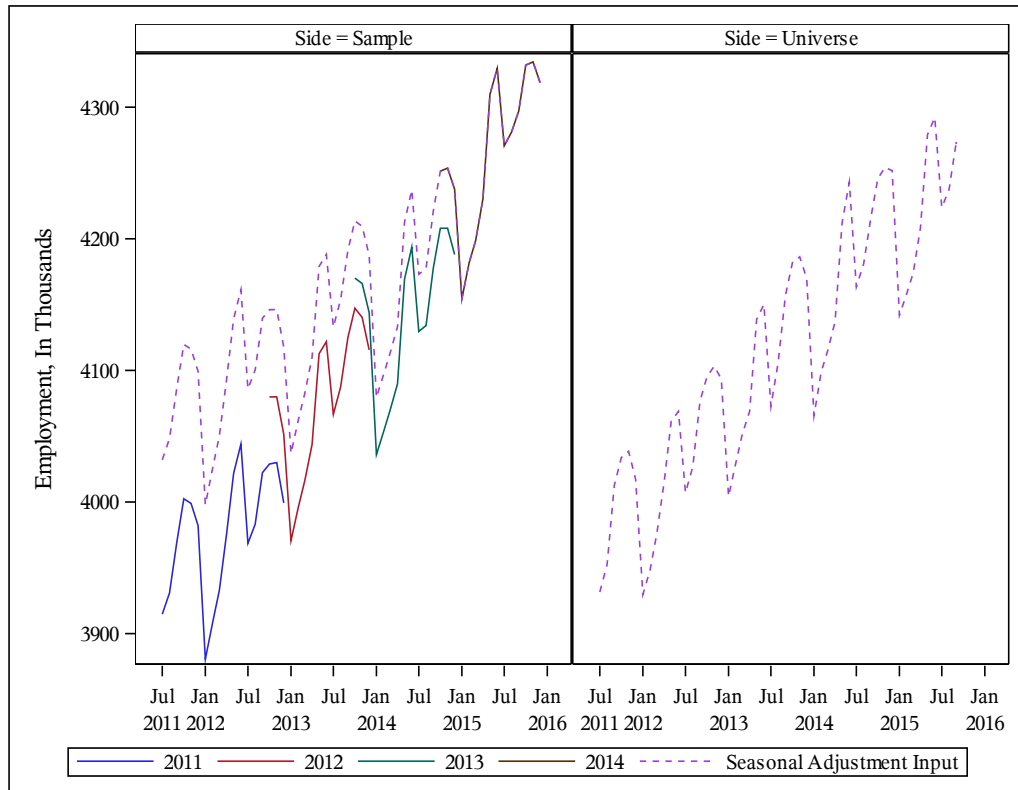
The left panel shows the construction of sample-based inputs for a state-level series into the two-step seasonal adjustment procedure. Four "benchmark years" of sample-based estimates are shown. The series are spliced together to form a set of inputs on the same benchmark year. This series is processed to create seasonal adjustment factors to be applied to sample based data in the coming year (2016 in this case.) The right panel shows the corresponding universe series created through the replacement benchmark method; this historical portion of the published series (September 2016 back in this example) is seasonally adjusted separately. In practice, ten years of data are generally used to create factors for both the sample and universe sides.

3.2.1 The Two-Step Seasonal Adjustment Procedure

In Section 2, seasonal differences between the CES and QCEW were detailed. The adjusted replacement method results in a historical time series with seasonal characteristics drawn from the QCEW, while the most recent values are calculated using the CES sample and reflect the survey's seasonality accordingly. The adjusted replacement was incorporated in all states in the 1980s, as monthly administrative data became available. Before long, practitioners began to notice a "disappearing January blip": over-the-month changes in the state-level seasonally adjusted data would be consistently positive every January only to be revised away (Berger and Phillips 1994). The "two-step" seasonal adjustment procedure was developed by Scott et al. (1994) to deal with the hybrid series. In the first step, the QCEW-based portion of the series, from September of the benchmark year back, is seasonally adjusted independently. The second step requires the creation of a solely-CES time series to be used to create factors for the most recent data. Historical levels are not directly used since estimates were created from different benchmark levels. Instead, the historical relative employment changes are linked

from the current benchmark level to create an input series for seasonal adjustment. This process is illustrated in Figure 2. The "sample-based" time series is seasonally adjusted, and projected factors from this adjustment are applied to the monthly sample-based estimates in the coming year[16].

4. Metropolitan Area Redefinition and Series Reconstruction

In 2013, the Office of Management and Budget (OMB) released a large change to the delineations of metropolitan areas. Metropolitan areas are considered to have a large deal of economic integration and outlying areas are incorporated with core urban areas largely based on commuting patterns. The commuting data have historically come from the Decennial Census, and therefore large changes in metro area delineation have typically come every ten years. These changes included the addition of metropolitan areas that may have fallen below size thresholds in the past; additions or subtraction of counties (or equivalents) from existing areas; and metro areas that were combined together, or were split apart. Some metro areas declined in population and were dropped. In order to remain relevant, CES metro area series had to be adjusted according to the new definitions (or created for new areas) and as long of a time series as possible be maintained.

Fortunately, administrative microdata in the LDB are available back to January 1990 and formed an excellent basis for these changes. The first principle in this reconstruction was that as little should be reconstructed as possible, to minimize the potential for error from very old administrative data. For areas that were changing in some way, only values for the counties being added and dropped were reconstructed, and were added or subtracted from the existing series. When areas merged, the industry-level series were merged when possible. (Often one area had industry detail that the other did not have, and this required an LDB-based reconstruction.) Series for new areas had to be built from the ground-up.

When creating these county-level tallies, establishments in the LDB were assigned their most-recent industry, area, and ownership codes. These codes are sometimes reassigned over the years. In most cases the change is non-economic and the most-recent values are most likely to be correct. Next, noncovered employment values were created. Detailed noncovered values by 6-digit NAICS industry were only available for the most recent few years, but an assumption was made that the ratios of noncovered-to-covered employment are fairly stable over time and that statewide ratios can be applied to individual counties. The difference in scope due to faculty was accounted for in a similar fashion. The summer employment patterns at the statewide level were assumed to be representative of the counties, and the data were adjusted accordingly. In order to smooth over level shifts caused by QCEW reporting changes, CES staff used a file containing predecessor/successor transactions found in the LDB. This file tracked when multiple establishments were consolidated into one, when one establishment was broken out into multiple, and also far more complicated "many-to-many" transactions. CES staff tried to reflect the most recent reporting basis, comparing employment ratios at seam points when reporting changed, prorating some amount of the predecessor establishment's employment to and from series in an automated procedure.

After the automated processes were applied the data underwent a joint review by BLS and state staff. Similar to the process run during the annual benchmark, X-13ARIMA-SEATS was used to screen for level shifts and point outliers in the time

---

[16] Concurrent seasonal adjustment is being considered, but it would employ the same two-step procedure.

series. The main difference was that instead of 18 months of data, 23 years of data had to be considered. Complicated predecessor/successor transactions often required additional work in order to smooth seam months, and research on these relationships could be difficult if they occurred many years ago. Finally, during the March 2014 benchmark processing, data from the historical constructions were lined up against the new benchmark data and any discrepancies were resolved. These tended to be quite small in most cases.

MSA reconstruction also presented a problem for seasonal adjustment. Historical data for the new and changed areas could be reliably adjusted. However, in the two-step process, a history of CES estimates is required to form factors for the most-recent part of the series. These were unavailable. Simulated CES estimates were considered, but in the experience of the program office, there were too many potential sources of serious error if this method was employed. Instead, for changed areas, the old history would be used, but only if the employment revision associated with the redefinition was less than 4 percent. For all MSA-level series that had been published on a seasonally adjusted basis, a comparison of the change in the seasonality for the universe portion of the series was conducted. Above an absolute level change of 4 percent, an unacceptable number of these series showed a noticeably different seasonal pattern, and it could be inferred that a similar change would happen in the sample-based series. This threshold was used to avoid publishing "seasonally adjusted" series that potentially contained a good deal of residual seasonality. An extreme example is the Salisbury, DE-MD, metropolitan statistical area. This area more than doubled in employment, as counties associated with beach resorts in DE and MD were added. Employment in these added counties is extremely seasonal, and the old CES time series bears little relationship to the new one. Currently, series where publication was suppressed are being tracked to determine when they can again be published on a seasonally adjusted basis. For new areas, at least three years of sample-based data will be required before seasonally adjusted data are published.

5. Future Directions

At present, BLS is considering the question, what makes the best benchmarked time series? Both the wedge and anchor technique used on the national data and the adjusted replacement used on subnational series have their benefits and drawbacks. It can indeed still be difficult to count the number of jobs even with an excellent source of administrative data. Ideally, a benchmarked series would incorporate the QCEW more often than on an annual basis. It would avoid administrative breaks and exhibit a seasonal pattern in line with the CES sample. Perhaps, it could be updated on a quarterly basis, keeping the current data from drifting too far away from their eventual benchmark level. In many ways, CES is lucky to have this problem. An administrative source is available in the QCEW, closely matching the CES coverage and construct. Most surveys are not so fortunate. That said, this paper presented complications in reconciling the survey data to the administrative records. Perhaps the seasonal difference between the QCEW and CES could be explicitly modeled and accounted for in the benchmark. End-of-quarter QCEW values, corrected for a seasonal error process, could be used as an anchor, with month-to-month changes still reflecting information from the survey. That said, research on new benchmarking techniques is still very much in a preliminary stage, but is a potential source of improvement to keep the Current Employment Statistics program accurate, timely, and relevant in the early years of its second century.

References

Berger, Franklin D. and Keith R. Phillips (1994), "Solving the Mystery of the Disappearing January Blip in State Employment Data," Federal Reserve Bank of Dallas, *Economic Review*, April, 53-62.

Dagum, Estela Bee and Pierre A. Chollette (2006), "Benchmarking, Temporal Distribution, and Reconciliation Methods for Time Series", *Lecture Notes in Statistics*, Vol. 186. P. Bickel, P. Diggle, S. Fienberg, U. Gather, I. Olkin, S. Zeger (Eds.). New York: Springer.

Fixler, Dennis and J. Steven Landefeld (2006), "The Importance of Data Sharing to Consistent Macroeconomic Statistics," in *Improving Business Statistics Through Interagency Data Sharing: Summary of a Workshop*, Ch. 7. Kuebler, Caryn and Christopher Mackie (eds.).Washington, D.C.: National Academy of Sciences.

Groen, Jeffrey A. (2012), "Sources of Error in Survey and Administrative Data: The Importance of Reporting Procedures," *Journal of Official Statistics* 28, 173-198.

Handbook of Methods, U.S. Bureau of Labor Statistics, Chapter 2. Available at: http://www.bls.gov/opub/hom/pdf/homch2.pdf

Kropf, Jurgen and Nicole Hudson (2012), "Current Employment Statistics seasonal adjustment and the 2007-2009 recession", *Monthly Labor Review*, October, 42-53. Available at: http://www.bls.gov/opub/mlr/2012/10/art4full.pdf

Johnson, Nicolas (2016), "One hundred years of Current Employment Statistics data collection", *Monthly Labor Review*, August. Available at: http://www.bls.gov/opub/mlr/2016/article/one-hundred-years-of-current-employment-statistics-data-collection.htm

Kelter, Laura A. (2016), "One hundred years of Current Employment Statistics—the history of CES sample design," *Monthly Labor Review*, August. Available at: http://www.bls.gov/opub/mlr/2016/article/one-hundred-years-of-current-employment-statistics-sample-design.htm

Lancaster, Loryn (2016), "Changes in federal and state unemployment insurance legislation in 2015", *Monthly Labor Review*, July. Available at: http://www.bls.gov/opub/mlr/2016/article/changes-in-federal-and-state-unemployment-insurance-legislation-in-2015.htm

Morisi, Teresa L. (2003), "Recent changes in the national Current Employment Statistics survey", *Monthly Labor Review*, June, 3-13. Available at: http://www.bls.gov/opub/mlr/2003/06/art1full.pdf

Mueller, Kirk (2006), "Impact of business births and deaths in the payroll survey", *Monthly Labor Review*, May, 28-34. Available at: http://www.bls.gov/opub/mlr/2006/05/art4full.pdf

Scott, Stuart; Stamas, George; Sullivan, Thomas; and Paul Chester (1994), "Seasonal Adjustment of Hybrid Economic Time Series," *Proceedings of the Section on Survey Research Methods*, American Statistical Association. Available at: http://www.bls.gov/osmr/pdf/st940350.pdf

Wright, Jonathan H. (2013), "Unseasonal Seasonals?" *Brookings Papers on Economic Activity*, 65-126.