Forecasting Birth/Death Residuals on a Quarterly Basis November 2011

Nathan D. Clausen

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

Abstract

Beginning in 2011, BLS began updating the Current Employment Statistics (CES) net birth/death model component of the estimation process more frequently, generating birth/death factors on a quarterly basis instead of annually. This allows CES to incorporate Quarterly Census of Employment and Wages (QCEW) data into the birth/death model as soon as it becomes available.

This more frequent updating should help to reduce what is known as the "post-benchmark revision" in the CES series. Because the quarterly updating allows the most recent quarter of available QCEW data to be incorporated immediately, rather than at the end of the year, revisions between the initial birth/death residual forecasts and the revised birth/death forecasts should be reduced. This paper documents research comparing both annual and quarterly methodologies to forecast net birth/death residuals using data from 2003 to 2010. The results show that the quarterly methodology would have led to smaller post-benchmark employment revisions for most years in the study.

Key Words: Current Employment Statistics (CES), Birth/Death model, Birth/Death residual forecast, CES benchmark revision

1. Background on the Current Employment Statistics (CES) Program

The Current Employment Statistics (CES)[2] Survey, conducted by the U.S. Bureau of Labor Statistics (BLS) in cooperation with State Employment Security Agencies (SESAs) collects payroll data each month on employment, hours, and earnings from a sample of nonagricultural establishments. The current CES sample includes about 140,000 businesses and government agencies, representing approximately 440,000 individual worksites. From this data, a large number of employment, hours, and earnings series are prepared and published each month with industry and geographic detail.

National CES estimates of employment are one of the first indicators of current economic conditions each month. Preliminary national estimates for a given month are published

---

three weeks after the reference week (the establishment pay period including the 12th of the month), typically on the first Friday of the following month[3]. Major data users include the Joint Economic Committee of Congress, the Federal Reserve Board, as well as financial markets and major media. In addition, CES employment, hours, and earnings data are inputs to other major economic series including Personal Income, Industrial Production, and the Leading and Coincident Economic Indicators.

1.1 Elements of Employment Estimates

The sample-based estimates are designed to accurately capture the over-the-month change in employment levels. The over-the-month change is then applied to the prior month's level to derive an employment level for the current month. However, the CES sample does not contain information on all of the changes in the universe employment: specifically, the CES estimation methodology is largely able to account for business births by imputing for business deaths[4], but the sample cannot measure the net residual of births and deaths (the birth/death residual). To account for these changes in the universe, CES uses a birth/death model to adjust monthly sample-based estimates.

The CES birth/death model consists of two components.[5] First, the model excludes employment losses due to business deaths. This is accomplished by attributing employment to missing sample reports based on industry trend. Earlier research showed that both business birth and death components of total employment are significant, but the net contribution is relatively small and stable. By imputing for missing reports, CES is effectively able to incorporate births by imputing for deaths. Second, the model forecasts a residual net birth/death component, based on historical net birth/death residuals, to account for the difference between business births and deaths. This net birth/death residual is added to the current month's employment level.

2. Birth/Death Residual Forecasting Methodology

CES relies on the QCEW Longitudinal Database (LDB) as the basis for developing the historical birth/death residuals to be forecasted. The LDB links establishments over time, allowing for the identification of the continuous establishments, establishments that go out of business (deaths), and new establishments (births). To develop the history for modeling, the same handling of business deaths for the CES sample data is applied to the population data. Because the QCEW lags CES estimation by several quarters, CES must forecast birth/death residuals to use in current estimates.

Beginning in 2011, CES started producing birth/death residuals on a quarterly basis. Prior to that, starting with the introduction of the probability based sample, CES

---

[3] The Employment Situation news release schedule, see
http://www.bls.gov/schedule/news_release/empsit.htm.
[4] CES research has shown that while the number of jobs attributed to business births and business deaths are both very large, the difference (net residual) between births and deaths is relatively small.
[5] For a more detailed explanation of this model, see http://www.bls.gov/web/empsit/cesbd.htm.

produced the residuals on an annual basis, scheduled to coincide with the annual benchmark. The benchmark process allows CES to control sample-based estimate levels to population levels.[6] The difference between the sample-based employment estimates for March and the population employment level derived mainly from the QCEW is what defines the annual benchmark revision. During the annual benchmark process, revisions are made differently for the estimates before the benchmark month – what is called the "benchmark" period – and the months following the benchmark month – called the "post-benchmark" period. Error during the benchmark period is assumed to accumulate at a constant rate, so the estimates are revised using a "wedge-back" procedure. Estimates for the post-benchmark period are revised by applying previously derived over-the-month sample changes to the revised March levels. New net birth/death residual forecasts based on the new LDB data are also used for these months.

2.1 Annual Methodology

The CES annual methodology consisted of creating net birth/death residual forecasts for each estimated industry using 5 years (60 months) of residual history, ending in March of the benchmark year. The residual forecasts extend out up to 21 months, from April following the benchmark month to December of the following year. During the subsequent benchmark cycle, forecasts for the post-benchmark period (the last 9 of the 21 months) are revised, incorporating the new benchmark year's data. For example, forecasts made during the 2009 annual benchmark used residual histories from April 2004 through March 2009 and produced forecasted values from April 2009 through December 2010. At the following benchmark for 2010, forecasts for April to December 2010 were replaced with revised forecasts using residual histories from April 2005 through March 2010.

2.2 Quarterly Methodology

The methodology for forecasting birth/death residuals on a quarterly basis is similar to the annual methodology. The first quarterly iteration coincides with the annual benchmarking process and uses 60 months of residual history. The only difference is that residuals are only forecast out 12 months instead of 21 months. The second quarterly iteration includes an additional quarter of residual history (63 months total) and is completed before April of the year following the benchmark. It produces an additional 3 months of residual forecasts, for April through June of the year following the benchmark year. The third and fourth quarterly updates similarly append 3 months of residual history (66 and 69 months respectively), and produce an additional 3 months of residual forecasts each. CES will not revise published residuals for the post-benchmark period until the following benchmark year.

Due to the timing of the quarterly methodology and the fact that it produces residual forecasts only for the subsequent post-benchmark period, it will not have any impact on

---

[6] For more information on CES Benchmark methodology, see
http://www.bls.gov/web/empsit/cesbmart.htm#8.

the CES benchmark revisions because the forecasts will be replaced after the subsequent benchmark. As Figure 1 below shows, initial forecasts made during the 2009 benchmark for April – December 2010 are replaced with revised forecasts made during the 2010 benchmark. Likewise, initial forecasts made during the 2010 benchmark for April – December 2011 are replaced with revised forecasts made during the 2011 benchmark.

Figure 1: Timing of Initial and Revised Birth/Death Residual Forecasts
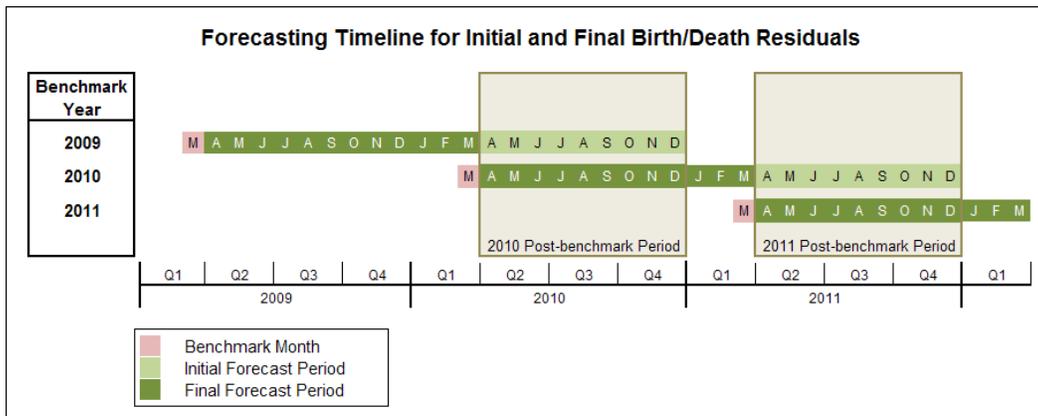
3. Evaluating Residual Forecasts

CES uses the measure of cumulative contribution to evaluate net birth/death residual forecasts during a benchmark period. This measure compares the cumulative contribution of the monthly residual forecasts over the 12-month benchmark period with the cumulative contribution of the actual birth/death residual values for the same period. While there are several sources of non-sample error that determine the size of annual benchmark revisions, the birth/death forecasting error is the only one that is published.

3.1 Evaluating Quarterly Forecasts
A similar cumulative measure is used to evaluate residual forecasts in the post-benchmark period, which is the period of interest for quarterly residual forecasts. Since the actual residual is not yet known for this period, the initial forecasts are compared to the revised forecasts (see Figure 1). The cumulative contribution of the initial monthly residual forecasts over the 9-month post-benchmark period is compared to the cumulative contribution of the revised residual forecasts for the same period using residual histories that are 12 months more current. Table 1 shows the cumulative contributions of the initial and revised residual forecasts using both the annual and quarterly methodologies for the post-benchmark periods from 2003 to 2010.

Table 1: Cumulative post-benchmark net birth/death contribution in thousands

| Post-benchmark Year	Forecast	Annual Methodology	Quarterly Methodology
2010	Initial	759	477
	Revised	537	537
	Difference	**-222**	**60**
2009	Initial	990	730
	Revised	585	585
	Difference	**-405**	**-145**
2008	Initial	1005	1008
	Revised	825	825
	Difference	**-180**	**-183**
2007	Initial	1059	1024
	Revised	883	883
	Difference	**-176**	**-141**
2006	Initial	906	1115
	Revised	1002	1002
	Difference	**96**	**-113**
2005	Initial	866	846
	Revised	817	817
	Difference	**-49**	**-29**
2004	Initial	889	788
	Revised	827	827
	Difference	**-62**	**39**
2003	Initial	768	740
	Revised	695	695
	Difference	**-73**	**-45**

In 6 of the 8 years available for analysis, the quarterly methodology resulted in smaller revisions in the post-benchmark period than the annual methodology. For the 2 years when the annual methodology was better (2006 and 2008), the difference between the two methodologies was not significant. In fact, the differences were not significant in 6 of the years, measuring no more than 4 percent of the revised cumulative post-benchmark contribution. In contrast, the 2009 and 2010 post-benchmark periods (when initial estimates were made with data through March 2008 and 2009, respectively) the quarterly methodology performed 45 percent and 30 percent better than the annual methodology. The markedly improved residual forecasts resulting from the quarterly methodology during the sharp economic downturn in 2008 and 2009, suggests that the CES quarterly methodology implementation will enable it to better reflect economic turning points. Additionally, the quarterly methodology is expected to perform as well as the annual methodology during more stable periods.

4. Quarterly Forecast Analysis

In order to evaluate the quarterly forecasting methodology against the prior annual methodology, a comparison of the cumulative contribution was used. An additional analysis of the quarterly forecasts is also possible by breaking out the cumulative contribution by quarter comparing the results of the annual and quarterly methodology. This analysis shows how the cumulative post-benchmark difference is distributed across the three quarters.

To analyze quarterly contributions and revisions from an annual to a quarterly methodology, annual and quarterly forecasts were generated using QCEW inputs and a generic forecasting specification[7]. Figure 2 shows quarterly contributions based on annual forecasts together with revisions based on quarterly forecasts (adding the annual forecast and the revision results in the quarterly forecast). For most quarters before the 2008 recession, revisions were relatively small and mixed in direction. While the annual forecasts for 2010 are, as a result of the recession, somewhat lower than prior years, the large downward revisions in 2009 and 2010 show the ability of the quarterly methodology to reflect the economic downturn in a timelier manner.

Figure 2: Contribution and Revisions by Quarter
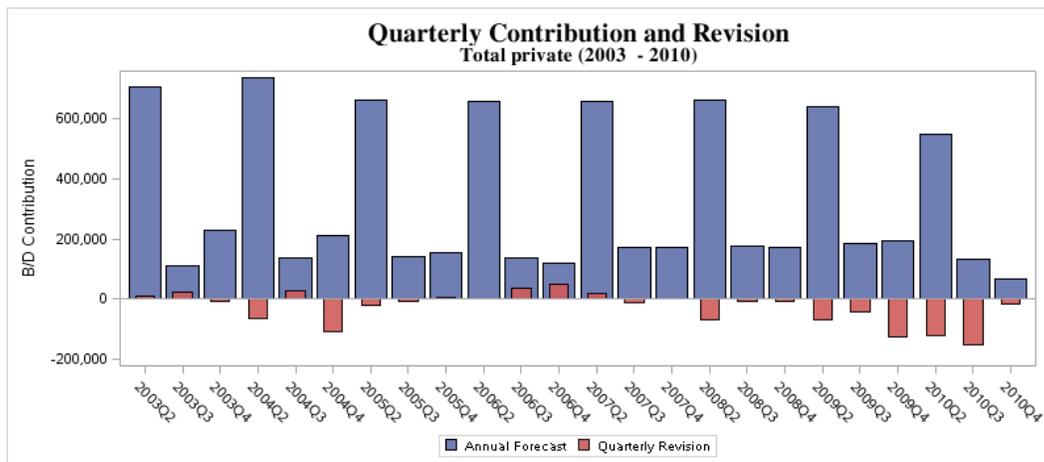
Averages of the annual forecast quarterly contributions and quarterly forecast revisions across the 8 years that CES has used the birth/death model show that the difference between the quarterly and annual models is quite evenly distributed across the three post-benchmark quarters. As shown in Figure 3, the largest revisions were in quarter 2 and the smallest revisions in quarter 3. This corresponds with the largest and smallest quarterly contributions being in the second and third quarters, respectively.

---

[7] The results of this analysis are not directly comparable with published CES birth/death residuals because of the use of a generic specification file and the absence of manual review for outliers. The generic specification file relies on ARIMA auto-model and uses a critical value of 5 for automatic outlier detection.

Figure 3: Average Contribution and Revisions by Quarter
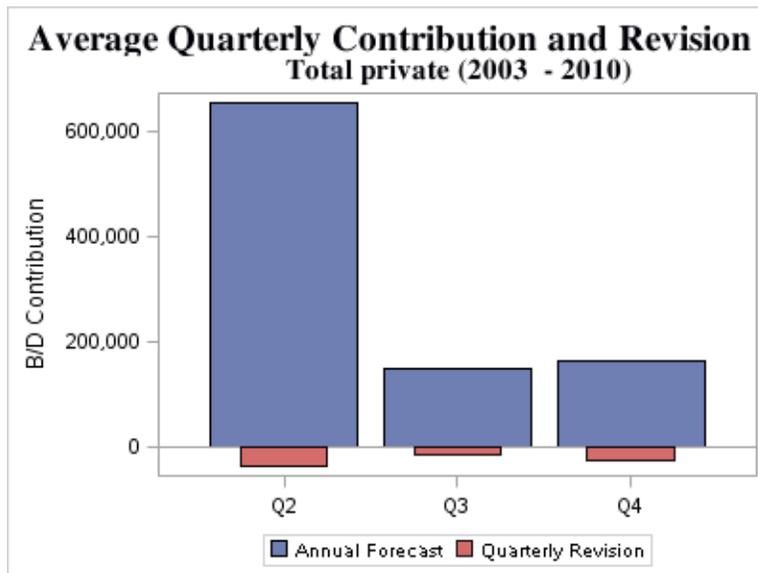
In percentage terms, the largest revisions on average were in the fourth quarter (around -16 percent). However, excluding the years 2009 and 2010, the fourth quarter average revision was only -6 percent. During these 2 years, the revisions for the third and fourth quarters were -61 percent and -56 percent, highlighting the effect of the quarterly methodology at sharp economic turning points.

Figure 4: Percent Revision by Quarter
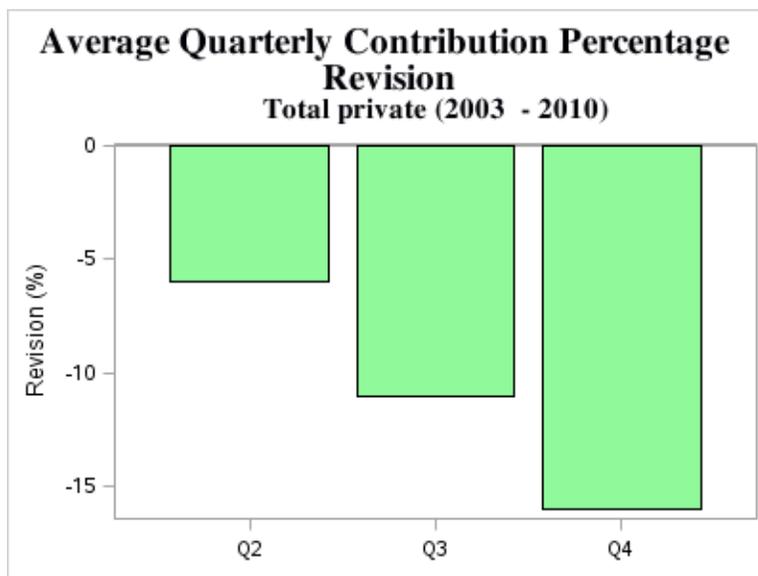
5. Conclusion

While forecasting net birth/death residuals on a quarterly basis would have done little to improve the accuracy of CES net birth/death model prior to 2009, it would have significantly improved the residual forecasts during the most recent recession. Further, this improvement is distributed across the three post-benchmark quarters such that the quarterly methodology does not disproportionally affect any one quarter. CES anticipates that the 2011 implementation of quarterly residual forecasting will reduce post-benchmark employment revisions and improve the stability of its estimates.

Acknowledgements