Back to the Future: Using Current Regression Variables to Forecast Forward from Historical Net Birth/Death Employment

Victoria Battista[1]
U.S. Bureau of Labor Statistics, 2 Massachusetts Ave NE, Washington, DC 20212

Introduction

The Current Employment Statistics (CES) [2] program uses a birth/death model to account for changes in employment that are not captured in the CES sample. One part of this model, the net birth/death residual forecasts, has regularly added between 800,000 and 1,000,000 jobs to the CES employment level on an annual basis from 2003 to 2008. During the most recent recession, the birth/death residual dropped to less than 300,000, a change that was not picked up by the forecasting model and contributed to a very large benchmark revision. CES has been conducting research to determine if the CES birth/death model, based currently on historical time series, could benefit from the incorporation of an additional, more timely regression variable.

1. Background on the Current Employment Statistics Program

The Current Employment Statistics Survey, conducted by the U.S. Bureau of Labor Statistics (BLS) collects payroll data each month on employment, hours, and earnings from a sample of non-agricultural establishments. The current CES sample includes about 145,000 businesses and government agencies, representing approximately 557,000 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.

1.1 Elements of Employment Estimates

The sample-based estimates are designed to accurately capture the over-the-month change in employment measured as the ratio of current month to previous month employment, referred to as the "sample link." The over-the-month link 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. There is an unavoidable lag between an establishment opening for business and its appearance on the sample frame and consequent availability to be sampled. Because new firm births generate a portion of employment growth each month, non-sampling methods must be used to estimate this growth. Inversely, the sample deteriorates over time as businesses fail. To account for these changes in the universe, CES uses a net birth/death model to adjust monthly sample-based estimates.

[1] The authors would like to thank Nathan Clausen, Anne Polivka, Harley Frazis, and Mark Loewenstein for their substantial input to the research.
[2] For more information on the Current Employment Statistics (CES) program, see http://www.bls.gov/ces/home.htm.

2. CES Birth/Death Model

2.1 Birth/Death Methodology

The CES birth/death model consists of two components. First, the model excludes employment losses due to business deaths. This is accomplished by imputing missing sample reports' employment based on industry trend. Earlier research showed that over the period 1995-2007 both business birth and death components of total employment were significant, but the net contribution was relatively small and stable. By imputing for missing reports, CES is effectively able to incorporate births by imputing for deaths. This step accounts for most of the birth/death employment.

The second component is an auto-regressive integrated moving average (ARIMA) time series model designed to estimate the residual birth/death employment not accounted for by the imputation. The historical time series used to create and test the ARIMA model was derived from the Quarterly Census of Employment and Wages[3] (QCEW) universe micro level database, and reflects the actual residual of births and deaths over the past five years. The net birth/death model component figures exhibit a seasonal pattern that can result in negative adjustments in some months, indicating less employment associated with business births than with deaths, and positive adjustments for other months, indicating more birth employment than employment associated with deaths. This birth/death residual is added to the current month's employment level.

2.2 Shortcomings in the Birth/Death Model

Historical time series data used to create current net birth/death forecasts lag behind CES estimates of employment by nearly a year. As with any forecast based strictly on historical data, this lag hinders the model's ability to catch turning points as they happen. Instead, the data reflecting these turning points are only incorporated a year or more after any change in trend actually occurs. Not only does this mean that the model will miss a downward shock like what was seen in the Great Recession (from December 2007 to June 2009 as defined by the National Bureau of Economic Research), but it is susceptible to missing a sharp upturn as well. The 2007 to 2009 recession demonstrated that the basic assumption of the CES birth/death model, that the birth/death residual is relatively stable over time, is less reliable during periods of rapid economic change.

3. Research to Improve the Birth/Death Model

3.1 Potential Improvements

In order to better detect when these changes occur and incorporate turning points in the birth/death model, CES began investigating several methodological changes that incorporate more timely data into the model. Two basic approaches were identified: reducing the lag of the inputs to the model, effectively reducing the length of time between the last month of estimation and the forecasted month, and incorporating auxiliary data in the forecast model that is more current than the QCEW inputs. In addition, the ARIMA model itself is being reviewed to ensure that it is still the most relevant data generating process for forecasts of net business births and deaths.

[3] For more information on the Quarterly Census of Employment and Wages (QCEW) program, see http://www.bls.gov/cew/home.htm.

The first approach focused including each new quarter of input data as soon as it is available instead of waiting to include a full year of inputs annually. This research proved fruitful, and, as of April 2011, the forecasted net birth/death residuals are calculated each quarter (instead of annually) to reflect the most currently available input. The second approach, detailed in section five below, focuses on incorporating exogenous variables in the model for forecasting birth/death residuals to help identify turning points and changes in trend.

3.2 Evaluation Measures

A measure for evaluating alternate birth/death methodologies in comparison to the current methodology was needed. The difference between the initial forecast of a month's net birth/death and its actual value (available a year later) is a logical starting point. At the same time, the context for evaluating the forecast error within the CES program is the effect it has on annual benchmark revisions. If the monthly errors are evenly distributed, with some over and some under the actual values, then the net effect on the benchmark will be small. However, if the monthly errors are biased in one direction, as is the case at turning points, then the cumulative effect on the benchmark can be quite large when compared to the total revision amount. In the most recent case of this effect, during the 2009 benchmark, the cumulative monthly birth/death error was the single largest source of error in the benchmark revision. So, in order to quantify the cumulative effect of monthly errors over a period of time, CES uses the measure of cumulative contribution to evaluate birth/death residual forecasts. 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.

For the purposes of this research, comparisons were made against the currently used model. However, to allow for appropriate comparisons between the regressor and non-regressor forecasted birth/death, all the series including the "current model" were forecast in the same manner. All series were produced as a result of an annual run, and do not reflect quarterly QCEW updates and reviews. None of the forecasts include analyst reviews of the output to identify outliers or level shifts. This separates the effect of regressors from the effect of either quarterly updates or analyst intervention.

3.3 Research Objective

Historically, the current birth/death model has performed well, generally reducing the benchmark revision and having a cumulative contribution that differs very little from the actual cumulative contribution. The current methodology, therefore, serves as a baseline for evaluating forecasts incorporating an exogenous variable. To be considered better than the current methodology, regressor-based forecasts should not only perform better during turning points, but also perform at least as well during more stable periods.

4. Criteria for Regression Variables

For a regression variable to be usable in the production of estimates, it must possess certain characteristics and criteria. The following paragraphs list these conditions with examples of variables that failed to meet the criterion where appropriate. A section of desirable, but less necessary traits describes attributes that an ideal net birth/death regression variable will have.

4.1 Necessary Characteristics

Start date of series. In order to be useful, a series must start on or before April 2002. At least five years of data are needed to calculate birth/death forecasts starting in April 2007. Data starting in April 2002 provides experimental series for the year preceding the recession, the years during the recession, and the years during the recovery period. Over time, as a complete business cycle is available over a more current period, the requirement of data as far back as 2002 can likely be relaxed.

Periodicity of series. The series must be available monthly if not more frequently. In more frequent periodicity series, a rolling average of several weeks or days can be calculated.

Gross Domestic Product is available quarterly one month after the quarter ends from the Bureau of Economic Analysis. These estimates, however, cannot be easily separated into monthly affects. Alternatively, net birth/death can be summed to a quarterly total, and a quarterly model can be estimated. However, the quarterly birth/death forecasts cannot easily be separated into monthly inputs to CES estimates.

Initial UI Claims are available weekly, but at a 2- to 5-week lag, depending on the analysis period for converting weekly data to monthly data. In the initial examination of these data, 3-, 4-, and 5-week rolling averages could be used to approximate a monthly UI Claims number.

Currency of data. The data to be considered must be more recent than the latest available birth/death input data used to fit the model. Data for a regression variable can be lagged in relation to the CES estimates, although the greater the lag in the data the less likely it is that it will be helpful in catching economic turning points in real time.

Industry availability. The series will be most useful if an industry break out by North American Industry Classification System (NAICS) or CES industry is possible.

While using one regression variable can inform the birth/death model at a Total Private level, using the same values across different industries that do not necessarily mirror the changes in Total Private could have adverse affects on the net birth/death forecasts.

Consistent and measurable improvement. Any regressor series and its incorporation into the net birth/death forecasts must demonstrate consistent and measureable improvement to the net birth\death series to be able to justify its use.

A number of the regressors used to augment the net birth/death forecasts showed only minor improvements to the estimates in the recession years. For the pre- and post-recession years, the affect of these regressor series on net birth/death forecasts is dampened, and at times even has an adverse affect on the residuals. The magnitude of this adverse affect must be evaluated in the context of average birth/death errors in a non-recession year. Justifying such slight improvement to data users could be difficult to do, especially if the regressor's inclusion in more stable years could have adverse effects on the forecasts.

4.2 Desirable Characteristics

Parsimony of models. Although it is possible to incorporate a regression variable for just a specific supersector, a regression variable that could operate on and inform all industries is preferable to one that improves only a single sector.

For instance, the number of new housing starts might work well as a regressor for the net of Construction business births and deaths, but not nearly as well for Healthcare Services. Conversely, if no regression variable works well for all series, incorporating a single variable that works extremely well in only one or a few sectors will be considered for those industries, while leaving the other sectors variable-free.

Expectation of continuity. There must be a reasonable assumption that the series chosen will continue to be published at the same frequency, in the same time frame, and using the same or similar methodology to its current series.

If data on the number of Google searches for UI filing information were used, there is no way to guarantee that the hits are being collected uniformly and consistently. Google is constantly evolving the data they collect, the way it is collected, and the formats for distribution of this data. Consequently, it is not a stable enough series to incorporate as a net birth/death regression variable.

Source of data. Ideally, the series chosen will be internal to BLS or internal to another government agency. While external, non-government source material has not been excluded outright from this research, the quality and continuity of internally-generated and government-generated series is easier to ensure and monitor.

Also, producers of externally and publically generated data have the potential to change methodology without informing data users. For instance, if the regression variable chosen is suddenly calculated differently, and the institution that produces that variable has no obligation to inform BLS about the change to methods, it could affect the CES estimates and could even create the appearance of outside manipulation of CES data through the mechanism of the regression variable.

4.3 Additional Considerations

Operational feasibility. Both the timing involved in processing CES estimates and the burden of additional steps created by running concurrent net birth/death forecasts must be considered. The potential benefits of using a regressor must outweigh the costs of additional time and resources devoted to collection of the regression variable, calculation of concurrent net birth/death, and review of the final net birth/death series each month.

Sample link calculations and incorporation as a regression variable illustrate the challenge of operational feasibility. The links must be calculated from the most recent sample to be used as a regression variable in net birth/death and are not available until the week of the CES employment release, a time during the month that is already on a tight schedule. Incorporating additional processing, fact checking, and re-estimation into this timetable could be problematic.

Explainability. The more complex the net birth/death series becomes, the more difficult it is to explain to data users. While this criterion does not alone rule out a regression variable, it does mean that between two variables that result in similar improvements to the net birth/death, the more simply explained variable will be selected.

5. Research Results

The period of analysis for this research used inputs from March 2003 to March 2013. Because five years of inputs are needed for forecasting, four years of results (based on the April to March benchmark period) are available for comparison: 2008-09, 2009-10, 2010-11, and 2011-12. For each period, forecasts with each of the eight selected regressor series (as well as forecasts with no regressor) were compared to the actual birth/death residual values, with the exception of 2013, for which the actual birth/death values are not yet available. Forecasts were generated using the same ARIMA forecasting model that is currently used for net birth/death forecasts at the detailed industry level. The forecasts were then aggregated to the sector level and the total private level. Results are compared and presented at this top industry level.

5.1 List of Variables Used in Analysis

Sector Sample Link. The sector sample link represents monthly change in employment in the CES sample measured as a ratio of current to previous weighted employment, and should inform the forecasts of sharp changes in employment. It is calculated indirectly by aggregating the published over-the-month (OTM) change of industries and removing the birth/death residual. Since not all sectors responded similarly during the recession, separate sample links were calculated for each of the 14 major sectors and used in forecasts of the component industries for each sector. For example, a Construction sample link was used for all Construction industry forecasts, and a Durable goods sample link was used for Durable goods manufacturing industries.

Sector OTM Change. The sector over-the-month change is similar to the sector sample link, except that it is measured in terms of actual employment change (in thousands). Since the sample link varies only slightly around 1, the OTM change may provide a stronger signal.

Total Private Sample Link. The total private sample link is similar to the sector sample link, except that it is calculated across the sample as a whole. It is a more general measure of changes in employment, and may better inform the forecast model of overall economic events separate from changes specific to an industry. Since it is a single top-side number for each month, all detailed industry forecasts were made with the same regressor series.

Total Private OTM Change. The total private over-the-month change is similar to the total private sample link, except that it is measured in terms of actual employment change (in thousands).

CPS Employment Link. The Current Population Survey[4] (CPS) employment link is a proxy for the CES sample links. Since it is based on household data, it differs from the establishment-based CES data both with regard to measurement and employment covered (For example, CPS employment includes unincorporated self employed workers, which establishment surveys estimates of employment such as the CES do not)[5]. To the extent that the differences in the measurement and types of employment covered by the CPS reflect economic movement related to the birth and death of establishments the CPS employment

[4] For more information on the Current Population Survey (CPS) program, see http://www.bls.gov/cps/home.htm.
[5] A comparison of the CES and CPS surveys and published data is available here, http://www.bls.gov/opub/mlr/2006/02/art2full.pdf.

link may be a useful predictor. The link is calculated as a ratio of over-the-month change from the prior month's employment level.

Over-the-year Change in Retail Sales. The over-the-year (OTY) Retail Sales change series is calculated using Census data adjusted for inflation. It represents a general measure of the health of the economy.

Lagged Over-the-year Change in Retail Sales. This series is the same as the previous series, except that it is lagged by one month. Because neither retail sales data nor the current month CPI data used to account for inflation are available in time for use in CES monthly estimation, the concurrent Retail Sales series is not a realistic variable in a production scenario. However, lagging this series one month would allow time for it to be incorporated into the CES estimation process.

UI Claims. The UI claims data are reported weekly and are a measure of the health of the jobs market. The series used in this research is a three week average ending with the week of the reference period.

5.2 Comparative Analysis of Forecasts

A summary of the forecast results obtained using each of the regressor variables is presented in Table 1 below. The actual birth/death values presented at the top of the table include the economic downturn observed in the economy as a whole, beginning from the low point in 2008-09 at 251,000 and slowly climbing in the recovery years. The table also shows that nearly all of the regressors would have improved the forecasts in 2008-09, with most of the cumulative contributions being closer to the 251,000 actual than the 814,000 simulated forecast with the current methodology. The simulated current model may be slightly different than officially published forecasts due to quarterly updating of input data and manual review of the official series.

Table 1. Cumulative Contribution of Birth/Death Residual Forecasts (in thousands)

	Cumulative Contribution			
	2008-09	2009-10	2010-11	2011-12
Actual Birth/Death	**251**	**294**	**703**	**914**
Retail Sales OTY Change	768	457	740	887
Retail Sales OTY Change (lag 1m.)	853	408	758	896
CPS Link	662	557	611	782
Private Sample Link	548	562	391	842
Private OTM Change	559	564	395	820
Sector Sample Link	757	594	466	719
Sector OTM Change	760	597	465	712
Current Model Current Model	**814**	**634**	**471**	**706**
UI Claims	748	769	473	711

To make these forecasts more comparable, Table 2 presents the forecasts relative to the actual, subtracting the forecasted contribution from the actual contribution. Negative numbers represent under-forecasts, and positive numbers represent over-forecasts. The average absolute error (i.e. the average distance from the actual contribution) across all four years is calculated, and the forecasts are ranked from smallest to largest. The top four forecasts using regressors are retail sales over-the-year change, retail sales over-the-year change lagged 1 month, the CPS link, and the private CES sample link . Bearing in mind that the retail sales over-the-year change variable is not actually feasible, the standard deviation of the absolute average error among the three best variables is 20,000. Including all of the feasible variables that were tested increases the standard deviation to 57,000. The top three feasible variables were shown to have a definitively beneficial effect on the forecasts in all years except in 2010-11, when the cumulative error was larger than the current model for the private sample link regressor forecast, and in 2008-09, when the cumulative error was larger than the current model for the retail sales over-the-year change lagged 1 month forecast.

Table 2. Cumulative Forecast Error (in thousands) – Ranked by Average Performance

	Cumulative Birth/Death Error					
	2008-09	2009-10	2010-11	2011-12	Abs. Avg.	Rank
Retail Sales OTY Change	517	163	37	-27	*186*	*1*
Retail Sales OTY Change (lag 1m.)	602	114	55	-18	*197*	*2*
CPS Link	411	263	-92	-132	*225*	*3*
Private Sample Link	297	268	-312	-72	*237*	*4*
Private OTM Change	308	270	-308	-94	*245*	*5*
Sector Sample Link	506	300	-237	-195	*310*	*6*
Sector OTM Change	509	303	-238	-202	*313*	*7*
Current Model	**563**	**340**	**-232**	**-208**	***336***	***8***
UI Claims	497	475	-230	-203	*351*	*9*

The purpose of this analysis is to identify potential variables that can be incorporated into the birth/death forecasting model to reduce forecast error, especially at turning points. Table 3 presents a comparison of the variables considered, showing their relative improvement over the current model for each of the four years considered. Any positive value is an improvement over the current method, and a negative number means that the variable caused forecasts to be farther from the actual. A key observation from this table is that all of the variables except UI Claims performed better than the current model on average.. Including regression variables tended to help the forecasts in most years, but in 2010-11 the private sample link, sector sample link, and sector over-the-month change variables and in 2008-09 the retail sales over-the-year change lagged 1 month were slightly detrimental to the birth/death forecasts. The extent to which the variables hurt the forecasts in some years is, in general, substantially smaller than the extent to which they helped in other years, so the average improvement is positive for all but UI Claims. The retail sales over-the-year change, both current and lagged 1 month clearly show the most improvement to the forecasts, reducing the birth/death error by 150,000 and 139,000 respectively. However, the real time retail sales over-the-year change is not available in time for use in the birth/death forecasts included in the first release of CES estimates. The other top three variables are retail sales over-the-year lagged 1 month, the CPS link, and the private sample link. Of these three, the private sample link is the preferred variable because it is calculated “in house” as part of estimating CES total private employment, although lagged retail sales over-the-year change and the CPS link could feasibly be incorporated into initial CES estimates.

Table 3. Net Improvement Over Current Model Methodology

	Net Improvement Over Current Model				
	2008-09	2009-10	2010-11	2011-12	Ave.
Retail Sales OTY Change	46	177	195	181	*150*
Retail Sales OTY Change (lag 1m.)	-39	226	177	190	*139*
CPS Link	152	77	140	76	*111*
Private Sample Link	266	72	-80	136	*99*
Private OTM Change	255	70	-76	114	*91*
Sector Sample Link	57	40	-5	13	*26*
Sector OTM Change	54	37	-6	6	*23*
UI Claims	66	-135	2	5	*-16*

The improvement in performance gained in 2008-09 is larger than the average performance across all four years for both the CPS link and private sample link. Removing 2008-09 from the average calculation diminishes the average improvement over the current method, and changes the rankings slightly. The average performance for the retail sales over-the-year change regressors is increased by removing 2008-09 from the average calculation. Table 4 shows the average improvement as calculated in Table 3 along with an average improvement excluding 2008-09. The private sample link regressor now only improves the forecasted net birth/death by 43,000. Excluding the key recession year and variables that are not available in real time, the best performing series improve on the current method on average between 198,000 to 43,000 jobs annually. This is a wide range of average improvement that is likely influenced by both the inclusion of the recession years in the historical birth/death input data and by the reflection of the recovery in the more up-to-date regression variable series.

Table 4. Average Forecast Improvement Including and Excluding Key Recession Year (2008-09)

	Average Forecast Improvement			
	With Recession		Excl. Recession	
	Avg.	Rank	Avg.	Rank
Retail Sales OTY Change	*150*	1	*184*	2
Retail Sales OTY Change (lag 1m.)	*139*	2	*198*	1
CPS Link	*111*	3	*98*	3
Private Sample Link	*99*	4	*43*	4
Private OTM Change	*91*	5	*36*	5
Sector Sample Link	*26*	6	*16*	6
Sector OTM Change	*23*	7	*12*	7
UI Claims	*-16*	8	*-43*	8

Another way to evaluate net birth/death forecasts is to compare the cumulative birth/death error to the benchmark revision amounts. The benchmark revision represents the total error in the estimates

accumulated over the course of the benchmark year, a portion of which can be attributed to birth/death error. By subtracting the cumulative birth/death error shown in Table 2 from the benchmark revision amounts[6] for the same year, we can determine if the inclusion of the birth/death forecast helped bring CES employment estimates closer to the universe employment to which we benchmark. Should the birth/death error and other error components be in opposite direction, the birth/death error could mask some of the magnitude of the other errors. Conversely, if the errors are in the same direction, the birth/death could make the overall error larger. In Table 5 the magnitude of the value in each box shows what the benchmark value would have been with the regressor net birth/death adjustment, and the direction shows if the CES estimate using that regressor net birth/death was an under-estimate (positive revision) or an over-estimate (negative revision). The absolute average reduction in benchmark error compares the absolute average of the benchmarks using regressors to the absolute average benchmark error using the current net birth/death methodology.
This way of assessing the impact of the net birth/death and net birth/death with a regression variable gives a slightly different result in the top four regressors. Now the CPS link performs better than the retail sales over-the-year change lagged 1 month, although only by 7,000. Unfortunately, the regression variable that performs the best using this metric is still the unusable retail sales over-the-year change. The 1 month lagged retail sales over-the-year change regression forecast does slightly worse than the real time retail sales, but would have increased the benchmark amount during the recession. Our best option for an additional regressor when compared to the actual birth/death, the private sample link, decreases the size of the benchmark error for both the average and in the recession year.

Table 5. Benchmark Amount with Simulated Regressor Birth/Death

	Cumulative Benchmark Error with Regressor Birth/Death					
	2008-09	2009-10	2010-11	2011-12	Abs. Avg. Reduction in BMK Error	Rank
Retail Sales OTY Change	-856	-201	-107	243	*-115*	*1*
CPS Link	-750	-301	22	348	*-111*	*2*
Retail Sales OTY Change (lag 1m.)	-941	-152	-125	234	*-104*	*3*
Private Sample Link	-636	-306	242	288	*-99*	*4*
Private OTM Change	-647	-308	238	310	*-91*	*5*
Sector Sample Link	-845	-338	167	411	*-26*	*6*
Sector OTM Change	-848	-341	168	418	*-23*	*7*
Current Model	**-902**	**-378**	**162**	**424**	***0***	***8***
UI Claims	-836	-513	160	419	*16*	*9*

Forecasting an accurate net birth/death value does not necessarily lead to more accurate estimates of employment when taking reduction in benchmark error into consideration. For instance, using a lagged retail sales over-the-year change regressor to forecast birth/death results in a larger benchmark in 2008-09, but a smaller average benchmark revision than the current model. The same is true when comparing the 2010-11 forecasts for the private sample link; the regressor birth/death gives a larger value in that year than the current model, but results in a smaller average error across all the years simulated.

[6] Benchmark revision amounts taken from Table 1 in the Benchmark Article (http://www.bls.gov/web/empsit/cesbmart.htm#Table1).

Unemployment insurance claims is the only regression variable that made the benchmark amount larger on average than the current model. Table 6 illustrates the point that excluding the recession changes the results of comparing the birth/death to the benchmark error slightly.

Table 6. Absolute Average Benchmark Improvement with and without the Recession

	Absolute Average Benchmark Improvement			
	With Recession		Excl. Recession	
	Avg.	Rank	Avg.	Rank
Retail Sales OTY Change	*-115*	*1*	*-138*	*2*
CPS Link	*-111*	*2*	*-98*	*3*
Retail Sales OTY Change (lag 1m.)	*-104*	*3*	*-151*	*1*
Private Sample Link	*-99*	*4*	*-43*	*4*
Private OTM Change	*-91*	*5*	*-36*	*5*
Sector Sample Link	*-26*	*6*	*-16*	*6*
Sector OTM Change	*-23*	*7*	*-12*	*7*
UI Claims	*16*	*8*	*43*	*8*

6. Implementation

In addition to the research findings, several other factors are integral to making a decision regarding whether to incorporate exogenous variables into the net birth/death forecasting model. First, any regressor-based model would necessarily mean that CES will need to move to concurrent forecasting of residuals on a monthly basis. The timeliness of the CES estimates, published just three weeks after the week that includes the 12th day of the month, means that exogenous variables will be available only just before estimation begins. The compressed production schedule only allows for several hours, at most, for the new residuals to be generated and reviewed before they are incorporated into estimation. Even a short delay obtaining the exogenous variables would put the CES program at risk of missing the release deadline. This does not make incorporating concurrent forecasting of birth/death impossible; it just puts further strain on an already tight production schedule.

If the private sample link or private over the month change were chosen as a regressor for birth/death, the schedule of producing CES estimates would be easier to control because both of these values are calculated “in house”. The regressors are calculated from the collected sample, and then used as inputs to the birth/death ARIMA model. Once the birth/death estimates have been created and reviewed, the sample links are applied to the previous month’s estimates and the birth/death values added to those estimates. The completed estimates can then be reviewed. Should errors be found in either the birth/death review or the estimate review, the process might have to be restarted to correct the error.

During each month’s production, sample reports relating to the current month and two previous months of past estimates are collected and used in estimation. This means that the sample links and over-the-month change values continue to be revised for two months after the initial release of estimates. Changes to the sample links would necessitate rerunning the birth/death estimates with the revised regressors. This means that during each month, three complete sets of net birth/death forecasts will have to be calculated and reviewed for use in the preliminary, second preliminary, and final CES estimates. These additional burdens on production systems and analysts are not insubstantial.

7. Conclusion

There are clear benefits to including a regression variable as an input to the net birth/death forecasting model in years in which the economy is changing quickly. However, with only four years of regressor-based net birth/death forecasts to evaluate, it is difficult to conclude whether the overall quality of the CES estimates will continue to be enhanced by changing the existing methodology to include an exogenous variable. Additionally, the time and resources needed to include additional steps in the processing of CES estimates with a regressor for birth/death are not inconsequential. Therefore, continuing to monitor how a regressor would affect the birth/death forecasts and the CES estimates until several more years of stable economic activity can be added to the research is advisable in order to make a more informed decision.