

## BLS Benchmark Revisions: Is This Time Different?

In this *Economic Commentary*, we discuss the Bureau of Labor Statistics' (BLS) payroll benchmark revisions, the role of these revisions, BLS methodology, and recent time series. While we do see some large benchmark revisions in recent years, these revisions are not big enough to indicate that something structural about the series has changed. Finally, considering a time series from 2010 through 2025, we find that past benchmark revisions have information that may help predict future revisions.

Roberto B. Pinheiro, Rory G. Quinlan

**Topics** Econometrics, modeling, and forecasting, Labor economics

DOI [10.26509/frbc-ec-202612](https://doi.org/10.26509/frbc-ec-202612) 

*The views authors express in Economic Commentary are theirs and not necessarily those of the Federal Reserve Bank of Cleveland or the Board of Governors of the Federal Reserve System. The series editor is Tasia Hane. This paper and its data are subject to revision; please visit [clevelandfed.org](https://clevelandfed.org) for updates.*

### Introduction

We have recently seen large monthly and benchmark revisions of the Bureau of Labor Statistics' (BLS) payroll employment growth. These revisions have renewed interest in the process of gathering and compiling labor statistics, and they have raised questions about how the accuracy of the recent real-time data compares with historical real-time data. In this *Economic Commentary*, we discuss the annual benchmark revisions of the payroll employment data and describe the reason for the benchmark revision and the process of revising the data. Subsequently, we explore the size of usual revisions and whether we have seen a change in revision patterns in recent years. Finally, we extend the analysis by Haltom, Mitchell, and Tallman (2005) to evaluate if past benchmark revisions are correlated with future benchmark revisions. Overall, our results indicate no clear change in the patterns of revisions in recent years, with no clear indication that something fundamental has changed. Moreover, we still find that past benchmark revisions are correlated with future ones. As a result, examining past benchmark revisions may be helpful in devising a clearer picture of current labor market conditions.

## Benchmark Revisions: Why, How, and History

The monthly payroll employment growth data in the employment situation report are based on a sample of 121,000 businesses and government agencies, representing approximately 631,000 individual worksites (US Bureau of Labor Statistics, 2025c). The active sample includes about one-third of all nonfarm payroll jobs. After an initial data release, the BLS performs three types of revisions. First, there are monthly revisions, usually a result of additional observations obtained from establishments that answered the survey late. There are typically two monthly revisions (first revision and second revision), and the number of responses may increase significantly from the first reading (that is, the original reading) to the third reading.<sup>1</sup> Second, there is an annual benchmark revision. As described in the BLS's *Handbook of Methods* (US Bureau of Labor Statistics, 2025b), annual benchmarks are constructed to realign the sample-based employment estimates for March of each year with the universe employment counts for that month.<sup>2</sup> Population counts are derived primarily from administrative files of employees covered by unemployment insurance (UI), the basis for the Quarterly Census of Employment and Wages (QCEW) database.<sup>3</sup> The benchmark revision factors in the births and deaths of firms and establishments, which in the Current Employment Statistics (CES) sample is usually accounted for by a birth–death model.<sup>4</sup> Moreover, it also corrects sampling errors. Third, there is a seasonal adjustment to the data. As stated by the BLS's seasonal adjustment methodology (US Bureau of Labor Statistics, 2025d), within the calendar year the labor force is strongly influenced by recurring events, such as typical weather, holidays, opening and closing of schools, and so on; to account for this pattern, the BLS applies seasonal adjustment factors. This adjustment removes the influence of these events, making it easier to see changes associated with the broader economy and business cycles. To compute these monthly adjustment factors, the BLS adds the most recent month's observation to its data and uses a modeling procedure to estimate factors.<sup>5</sup> At the end of each year, the BLS adds the past year's observations to its data and re-estimates seasonal adjustment factors for the previous five years. We will turn to this revision in subsequent sections.

Figure 1 presents how nonseasonally adjusted data are revised in the benchmark revision for year  $t$ . Each benchmark revision affects 21 months of nonseasonally adjusted data. The months preceding the March benchmark are readjusted to reflect the benchmark level.<sup>6</sup> Months following the March benchmark are also recalculated by adjusting sample-based links and adding an updated net birth–death firm forecast. Benchmark revisions are announced in early February of the next year, during the release of January's employment situation report. At the same time, the data is adjusted to reflect both benchmark revisions to the nonseasonally adjusted data as well as changes in the seasonal factors.

Figure 1: Benchmark Revision Process, Not Seasonally Adjusted Data

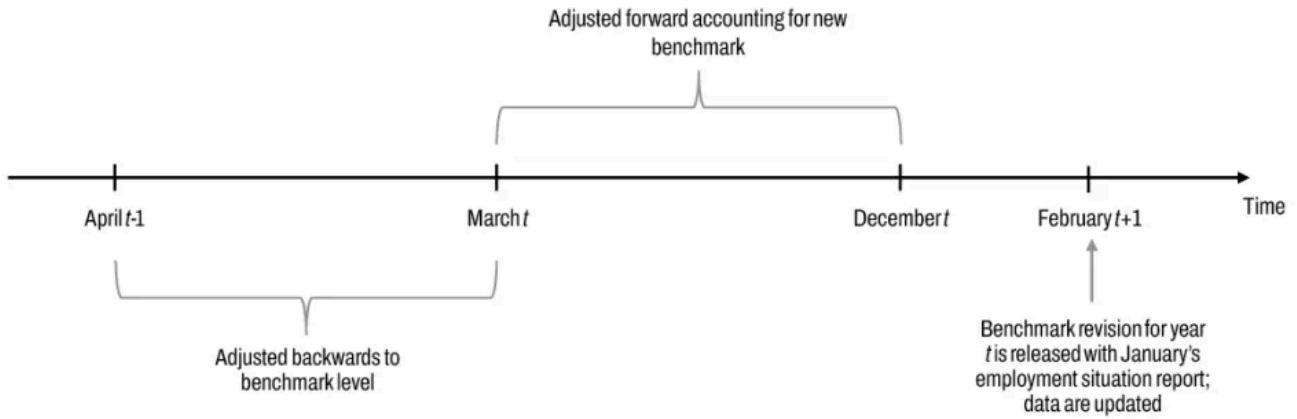
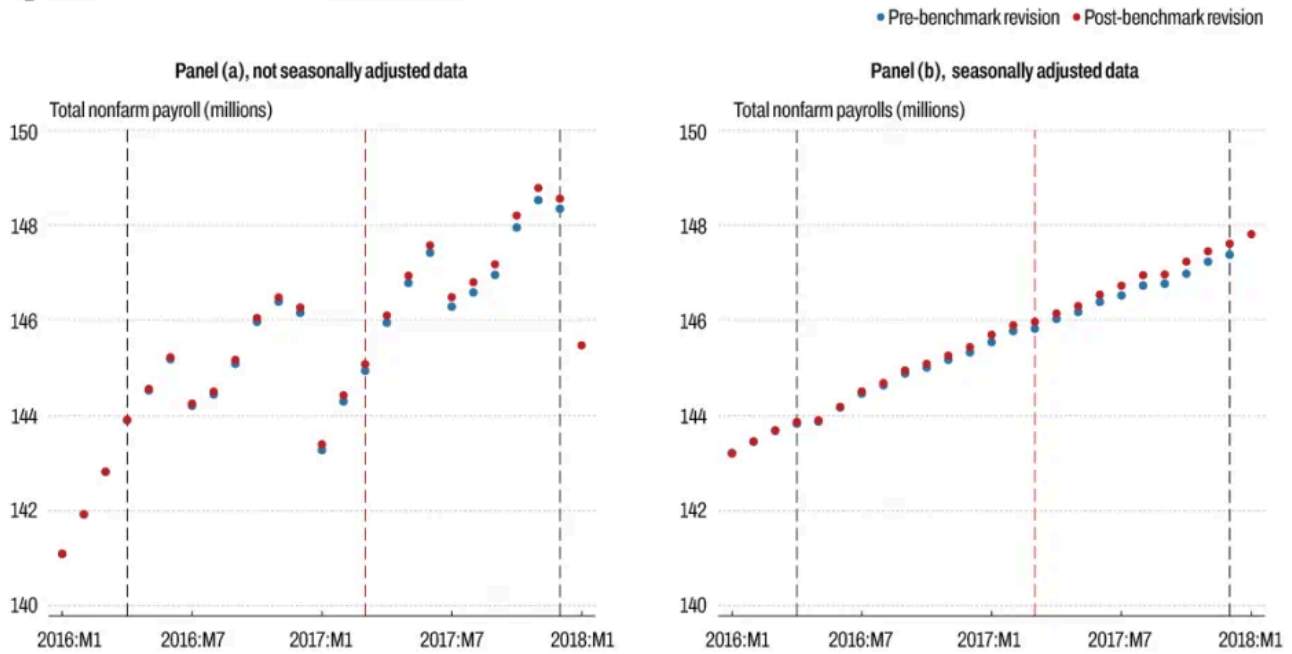


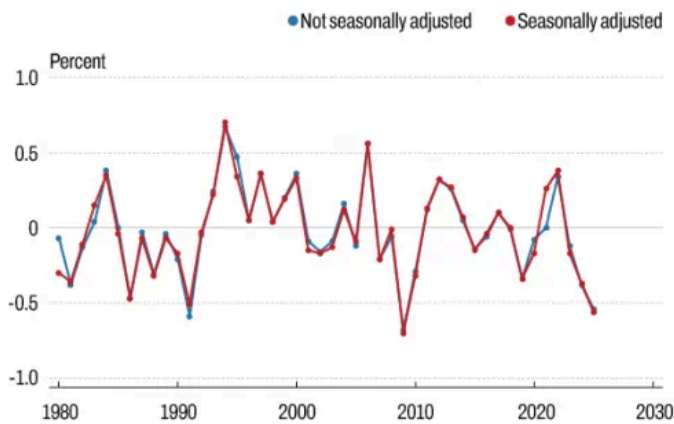
Figure 2, panel (a) illustrates the benchmark revision for March 2017 (highlighted by the vertical dashed red line). The vertical black dashed lines indicate the 21-month range of nonseasonally adjusted data affected by the benchmark revision. However, at the time of the benchmark revision, the BLS also applies updated seasonal factors to the data. As a result, the changes in the seasonally adjusted data go beyond the 21-month window. Figure 2, panel (b) shows the impact of benchmark revisions on seasonally adjusted data.

Figure 2: Benchmark Revision, March 2017



Sources: Philadelphia Reserve Bank, ALFRED, authors' calculations

Figure 3: Change in Payroll Employment Because of Benchmark Revision



Sources: Philadelphia Reserve Bank, ALFRED, authors' calculations  
Last observation: 2025

Finally, Figure 3 shows how much the benchmark revisions changed the level of payroll employment in March of the benchmark year, as a percentage of the originally reported estimate, for the period 1980–2025. Three key messages can be obtained from this figure. First, seasonally adjusted and nonseasonally adjusted series behave similarly in terms of the impact of benchmark revisions on payroll employment. Second, most revisions are within the minus 0.5 percent to plus 0.5 percent range. According to Haltom, Mitchell, and Tallman (2005), the BLS sees this range as the normal (or acceptable) range.<sup>7</sup> Finally, the most recent reading is just marginally outside of the normal range (-0.54 percent) and does not seem to be an obvious outlier compared with the plotted time series. We will corroborate this claim by running a structural break test for monthly changes.

## How big are these revisions?

To evaluate the size and predictive power of revisions, we move from payroll levels (presented in Figures 2 and 3) to monthly payroll growth because growth rates better reflect labor market dynamics and are more relevant for policy decisions. When releasing revisions, the BLS presents the change as both a percentage and in thousands of employees. Since it is more common to talk about revisions in thousands, we primarily present our discussion in terms of thousands to describe the monthly change in payroll employment. We present two measures of revision: first, the benchmark revision that we previously described, representing the difference between the third survey reading and the most recent reading; second, the cumulative revision that shows the difference between the first estimate (original) and the most recent estimate. As a result, the cumulative revision combines the impact of both monthly and benchmark revisions on monthly payroll employment growth. In this analysis, we use seasonally adjusted data provided by the Philadelphia Reserve Bank's Real-Time Data Research Center, supplemented with the St. Louis Reserve Bank's Archival Federal Reserve Economic Data (ALFRED) series to incorporate the most recent benchmark revision.<sup>8</sup> Our main sample spans January 1980 to December 2025 since the most recent benchmark revision, for March 2025, was released in February 2026, not affecting recently released data. Later we consider an extended sample from January 1965 to December 2025.

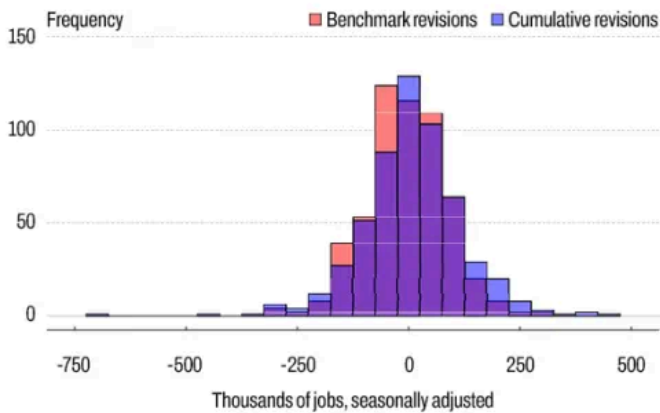
Summary statistics and the revisions' distributions are presented in Table 1 and Figure 4, respectively. We see that, on average, benchmark revisions decreased the month-over-month changes in payroll employment by 4,512, while cumulative revisions increased them by 6,844. Figure 4 shows the histogram of benchmark and cumulative revisions. As we can see, the distribution of benchmark revisions is quite concentrated around the mean, and the tails are not particularly heavy.

**TABLE 1: SUMMARY STATISTICS FOR REVISION VALUES, THOUSANDS**

<b>Revision type</b>	<b>Mean</b>	<b>Median</b>	<b>25th percentile</b>	<b>75th percentile</b>
Benchmark revisions	-4.51	-2.00	-59.00	57.00
Cumulative revisions	6.84	9.00	-58.00	67.00

Note: Cumulative and benchmark revisions series span January 1980 to December 2025 (N= 551).

**Figure 4: Distributions of the Impact of Cumulative and Benchmark Revisions on Monthly Payroll Growth from January 1980 to December 2025**



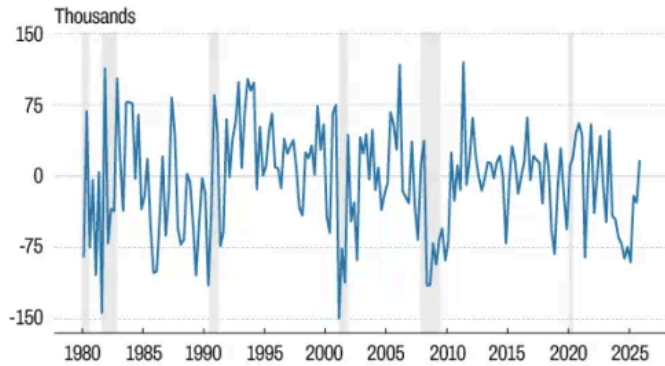
Sources: Philadelphia Reserve Bank, ALFRED, authors' calculations  
Last observation: December 2025

There is still the question around whether something has changed recently. We turn to this question now. Figure 5 displays the time series of benchmark revisions, that is, the most recent estimate of the monthly change in payrolls minus the third estimate (which captures the sum of benchmark and all seasonal adjustment revisions since the third estimate). While Figure 5, based on benchmark revisions to monthly changes, does not seem to present an obvious change in the series, we nonetheless present the results from structural break tests.

As an initial inquiry, we use the supF test from Bai and Perron (1998) that can detect multiple and unknown breakpoints.<sup>9</sup> Because we can conduct this test without having to specify a predetermined break point, it gives us a more open view of the possible structural composition of the data. Additionally, the Bai–Perron test can generate suggested break points that we can then specify as our hypothesized break points in subsequent testing with the standard Chow test (Chow, 1960) that does require a hypothesized break point. Our test does not indicate the presence of a structural break in any point of our main sample (1980:M1–2025:M12). Since the BLS (US Bureau of Labor Statistics, 2025c) suggests that survey revisions may experience a significant structural change just before 1979, we extend our revision series to the period 1965:M1–2025:M12. Our empirical results still show no signs of a structural break in the series of benchmark revisions.

As a robustness test, we perform a standard Chow test with the cumulative revisions from 1980 to 2025 and a hypothesized arbitrary break point of January 2020; this test did not detect a structural change. We also test a break point in mid-1978 because of BLS changes in methodology, but again we find no structural break.

**Figure 5: Benchmark Revisions of Monthly Total Nonfarm Payrolls, Seasonally Adjusted**



Sources: Philadelphia Reserve Bank, ALFRED, authors' calculations

Notes: Most recent estimate minus third estimate. Data are aggregated to quarterly for readability. Shaded vertical bars represent recessions.

Last observation: 2025:Q4

In summary, considering the benchmark revision time series of 1980 to 2025, benchmark revisions do not appear to have undergone a structural change, and even extending this analysis to span 1965 to 2025, there does not appear to be a structural change, according to the standard Chow and Bai–Perron tests. A lack of structural change in the benchmark revision series eases concerns regarding misspecification in our subsequent modeling section.

## Are revision correlations useful information?

Finally, we test whether the benchmark revision series has useful information to help predict future revisions. Both Phillips and Nordlund (2012) and Haltom, Mitchell, and Tallman (2005) find serial correlation for seasonally adjusted payroll employment benchmark revisions. Their result indicates that past benchmark revisions may be informative about future revisions, possibly helping to estimate real-time measures that give a better picture of current labor market conditions. We evaluate if this serial correlation persists in the data as we extend the sample. We follow Haltom, Mitchell, and Tallman (2005) and implement a standard linear regression model to test if all information from past benchmark revisions is uncorrelated with newly released observations at a future benchmark date.

While Haltom, Mitchell, and Tallman (2005) provide full details of the methodology and the empirical specification, we present a brief overview here. The general specification is

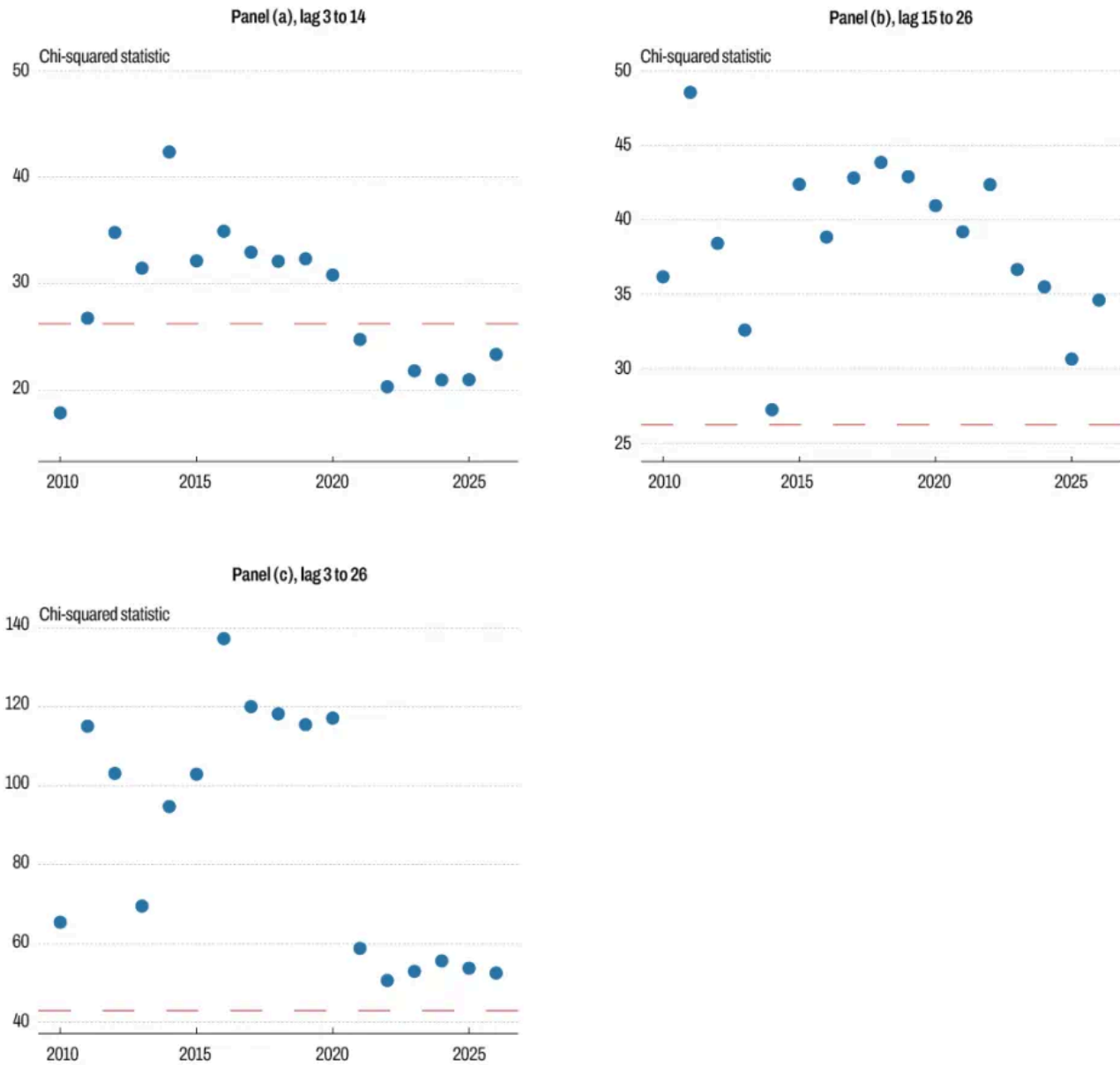
$$YB_t = A + \sum_{i=1}^{13} B_i YP_{t-i} + \sum_{i=1}^{13} C_i Unemp_{t-i} + \sum_{k=m}^n D_k Rvdiff_{t-1,k} + u_t, \quad t = \text{February } 1997, \dots, T,$$

For each year of interest,  $t$ , we regress the newly released observations for payroll employment from the benchmark release ( $YB$ ) on our key variable of interest, the past benchmark revisions ( $Rvdiff$ ).  $T$  represents a benchmark release date and goes from February 2010, February 2011, ..., February 2025 in subsequent regressions. As we show in Figure 2, benchmark revisions affect several months in the time series, and these observations may have predictive power about future revisions. Hence, we include the impact of benchmark revisions on several previous months,  $k$ , in the sample (from three up to 26 months).<sup>10</sup> Finally, we control for the

previous 13 months of unemployment (*Unemp*) and payroll employment vintages available in the months prior to the benchmark release date (*YP*). Since we are testing whether the information from all these different months is uncorrelated with future benchmark revisions, this joint test is carried out through the calculation of chi-square statistics.

Figure 6 shows the chi-square test statistics for each year (blue dots) against the 1 percent critical value (red line). Values above the red line indicate that past benchmark revisions may have valuable information about future revisions. Notice that Figure 6 shows results for models that include a different range of past months affected by previous benchmark revisions. Figure 6, panel (a) includes the most recent months affected by the previous benchmark revision (three to 14 months). Figure 6, panel (b) includes the next 12 months affected by revisions (15 to 26 months). Finally, Figure 6, panel (c) includes all past revisions. Notice that we start with a three-month lag to avoid the impact of monthly revisions.

Figure 6: Chi-Squared Test Statistics for Each Benchmark Year for Each Lag Specification



Sources: Philadelphia Reserve Bank, ALFRED, authors' calculations  
 Note: Panels show horizontal dashed line at 1 percent critical value for significance.

All panels in Figure 6 show that chi-square values in recent years are lower than in earlier years in our series. However, even though the chi-square values are lower, they are still statistically significant, even at the 1 percent critical value level in all years of specifications with more lags panel (c) and further back lags panel (b). This means we can reject the null hypothesis that previous benchmark revisions have no predictive power for future revisions. There is a deviation from this pattern of significance in recent years in specification 1, which includes the first 12 months of benchmark revisions. However, even in this case we retain statistical significance at least at the 10 percent level.

## Conclusion

In this *Economic Commentary*, we show that while recent benchmark revisions may have been elevated compared to the historical mean, there is no clear sign that a structural break has occurred. The recent benchmark revision has been just marginally outside of the range the US Bureau of Labor Statistics considers normal and does not seem to be an obvious outlier. That said, we also show that there is still a correlation between past and future benchmark revisions in recent data vintages. As a result, taking into account the benchmark revisions that occurred in the recent past may help us have a better real-time picture about current labor market conditions. For further work with BLS payroll revisions, Quinlan and Pinheiro (2026) provide complementary analysis focusing on properties of payroll survey revisions and their usefulness in forecasting recessions.

## References

- Bai, Jushan, and Pierre Perron. 1998. “Estimating and Testing Linear Models with Multiple Structural Changes.” *Econometrica* 66(1): 47–78. [doi.org/10.2307/2998540](https://doi.org/10.2307/2998540) [↗](#).
- Bai, Jushan, and Pierre Perron. 2003. “Computation and Analysis of Multiple Structural Change Models.” *Journal of Applied Econometrics* 18(1): 1–22. [doi.org/10.1002/jae.659](https://doi.org/10.1002/jae.659) [↗](#).
- Chow, Gregory C. 1960. “Tests of Equality Between Sets of Coefficients in Two Linear Regressions.” *Econometrica* 28(3): 591–605. [doi.org/10.2307/1910133](https://doi.org/10.2307/1910133) [↗](#).
- Federal Reserve Bank of Philadelphia. “Data Files – Real-Time Data Set (EMPLOY).” Accessed 2026. [philadelphiafed.org/surveys-and-data/real-time-data-research/employ](https://philadelphiafed.org/surveys-and-data/real-time-data-research/employ) [↗](#).
- Haltom, Nicholas, Vanessa D. Mitchell, and Ellis W. Tallman. 2005. “Payroll Employment Data: Measuring the Effects of Annual Benchmark Revisions.” *Economic Review* 90(Q2): 1–23. [fedinprint.org/item/fedaer/30511](https://fedinprint.org/item/fedaer/30511) [↗](#).
- Phillips, Keith R., and James Nordlund. 2012. “The Efficiency of the Benchmark Revisions to the Current Employment Statistics (CES) Data.” *Economics Letters* 115(3): 431–434. [doi.org/10.1016/j.econlet.2011.12.118](https://doi.org/10.1016/j.econlet.2011.12.118) [↗](#).
- Quinlan, Rory, and Roberto Pinheiro. 2026. “BLS Payroll Revisions: Forecasting Recessions.” *Economics Letters* 261(March): 112859. [doi.org/10.1016/j.econlet.2026.112859](https://doi.org/10.1016/j.econlet.2026.112859) [↗](#).
- US Bureau of Labor Statistics. 2025a. “Current Employment Statistics – National: Calculation.” [bls.gov/pub/hom/ces/calculation.htm](https://bls.gov/pub/hom/ces/calculation.htm) [↗](#).
- US Bureau of Labor Statistics. 2025b. “Handbook of Methods.” [bls.gov/pub/hom/](https://bls.gov/pub/hom/) [↗](#).
- US Bureau of Labor Statistics. 2025c. “Nonfarm Payroll Employment: Revisions between Over-the-Month Estimates, 1979–Present.” [bls.gov/web/empsit/cesnaicsrev.htm](https://bls.gov/web/empsit/cesnaicsrev.htm) [↗](#).
- US Bureau of Labor Statistics. 2025d. “Seasonal Adjustment Methodology for National Labor Force Statistics from the CPS.” [bls.gov/cps/seasonal-adjustment-methodology.htm](https://bls.gov/cps/seasonal-adjustment-methodology.htm) [↗](#).

## Endnotes

1. Response rates from the first monthly reading have a current and six-month moving average just above 55 percent; the second reading has response rates for a current and six-month moving average of 90.9 percent and 91.2 percent, respectively; and the third reading has response rates for current and six-month moving average above 93 percent. [Return to 1](#)
2. “Universe” counts are the total, actual, and complete counts of all units in a target population. Since 2004, benchmark revisions are announced and incorporated into the data in February of the next year, with the announcement of the January payroll numbers. Between 1981 and 2003, benchmark revisions were usually announced in June. [Return to 2](#)
3. These files cover about 97 percent of the entire labor force. The remaining 3 percent are constructed from alternative sources. For details about the construction process and the population, please see US Bureau of Labor Statistics (2025b). [Return to 3](#)
4. There is usually a time gap between a new firm’s starting operations and the firm’s being available for sampling. Similarly, it may take some time until the BLS recognizes that a nonresponse is due to a firm’s going out of business. As a result, the BLS factors

in the contributions of these movements on the extensive margin through a birth–death model in order to present payroll statistics in a more timely fashion. However, the model may deviate from the actual net firm birth, which is corrected through the benchmark revision. [Return to 4](#)

5. The BLS uses an X-13ARIMA-SEATS (X-13) program to estimate these adjustments. For further details, reference the BLS’s *Handbook of Methods* at [bls.gov/opub/hom/](https://bls.gov/opub/hom/) [↗](#). [Return to 5](#)
6. Adjustment follows a linear “wedge-back” procedure. The difference between the final benchmark level and the previously published March sample-based estimate is calculated and spread back across the previous 11 months. The wedge is linear; eleven-twelfths of the March difference is added to the February estimate, ten-twelfths to the January estimate, and so on, back to the previous April estimate, which receives one-twelfth of the March difference. [Return to 6](#)
7. From Haltom, Mitchell, and Tallman (2005), “Judging from commentary in BLS Employment and Earnings releases over this period, one may infer that the normal (or acceptable) range of revision is from negative 0.5 percent to positive 0.5 percent for total nonfarm payroll employment.” [Return to 7](#)
8. Data can be found at the Federal Reserve Bank of Philadelphia, Real-Time Data Research Center (accessed 2026). Moreover, since we are using seasonally adjusted data and the most recent reading in this section, the revisions include not only the initial benchmark revision, but also all the following changes because of seasonal adjustments. [Return to 8](#)
9. For additional detail regarding this method, see Bai and Perron (2003). [Return to 9](#)
10. Since we are using seasonally adjusted data, the impact of benchmark revisions goes beyond the 21-month range, as shown in Figure 2, panel (b). [Return to 10](#)

## Suggested Citation

Pinheiro, Roberto B., and Rory G. Quinlan. 2026. “BLS Benchmark Revisions: Is This Time Different?” Federal Reserve Bank of Cleveland, Economic Commentary 2026-12. <https://doi.org/10.26509/frbc-ec-202612> [↗](#)

This work by [Federal Reserve Bank of Cleveland](#) is licensed under Creative Commons [Attribution-NonCommercial 4.0 International](#) 