NBER WORKING PAPER SERIES

MEASURING PAYROLL EMPLOYMENT:
A NOTE ON THE CURRENT EMPLOYMENT STATISTICS

Ryan A. Decker

Working Paper 34924
http://www.nber.org/papers/w34924

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2026

The analysis and conclusions set forth are those of the author and do not indicate concurrence by other members of the research staff or the Board of Governors of the Federal Reserve System. This material was initially prepared for the NBER Advancing Economic Measurement, Fall 2025 conference of November 13, 2025. I thank Tomaz Cajner, Christopher Kurz, Norman Morin, and John Stevens for comments on an earlier draft. Without implication, this paper draws heavily on joint work and discussions—over many years—with Tomaz Cajner, Leland Crane, Adrian Hamins-Puertolas, and Christopher Kurz. In addition, I thank Katharine Abraham, John Coglianese, John Haltiwanger, Chris Nekarda, John Stewart, and David Wilcox for helpful recent or previous conversations. I thank Adrian Hamins-Puertolas and Joshua Hu for generous technical guidance and data provision. I thank staff at the Bureau of Labor Statistics and the Census Bureau for fielding many questions with speed and clarity. Aidan Owens and Mason Thieu provided excellent research assistance. The many acknowledgments above notwithstanding, any errors in empirical exercises, methodological explanations, or judgment are my own. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Measuring Payroll Employment: A Note on the Current Employment Statistics
Ryan A. Decker
NBER Working Paper No. 34924
March 2026
JEL No. C82, E0, J01

ABSTRACT

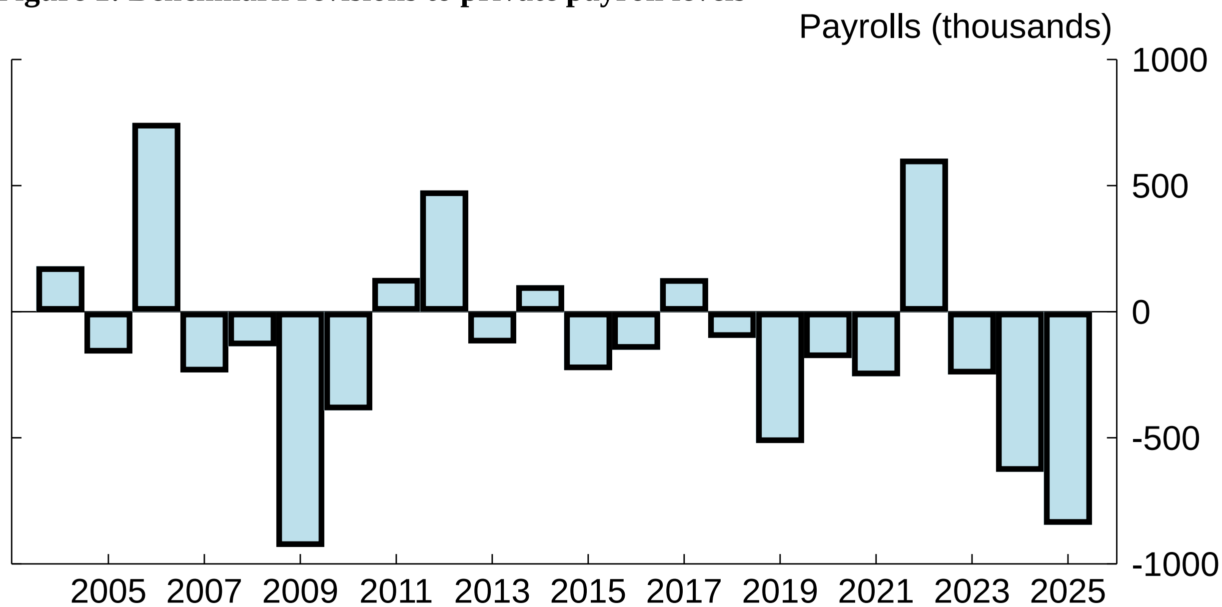
The Bureau of Labor Statistics Current Employment Statistics (CES) is a high-quality, closely watched indicator of labor market health and the business cycle. Recent revisions to CES data have garnered significant attention. I provide a primer on CES methodology and analyze recent survey response and data revision patterns. CES revisions actually tend to be small, but they have increased in recent years. I discuss potential ideas for improving the CES including expanding the data sample with internal or external data, adjusting the estimates with benchmark data more often, and enhancing estimation of establishment birth and death.

Ryan A. Decker
Federal Reserve Board of Governors
ryan.a.decker@frb.gov

1 Introduction

The Bureau of Labor Statistics (BLS) Current Employment Statistics (CES)—the “payroll” or “establishment” survey—is one of the most closely watched U.S. economic statistical releases. CES is used by the National Bureau of Economic Research (NBER) Business Cycle Dating Committee, and it is a critical input to other economic statistics including Gross Domestic Product (GDP), Productivity and Costs, and Industrial Production. In recent years the CES has been subject to larger-than-usual benchmark revisions (figure 1), raising concerns about the accuracy of the CES statistics.

Figure 1: Benchmark revisions to private payroll levels



Note: Private sector, March reference. Difference between benchmark and pre-benchmark estimated level. Some years may include revisions from scope changes.
Source: BLS Current Employment Statistics.

I review key methodological details of the CES, discussing the CES survey itself as well as the famous Net Birth-Death Model (hereafter “NBD Model”), which now accounts for well over half of published payroll gains. I then characterize the CES revisions, which tend to be small relative both to the employment level and to revisions of a comparably broad statistic, GDP. I discuss trends in survey response rates, monthly and annual revision patterns, and the (declining) role of the NBD Model in revisions.

Finally, I discuss potential methodological changes for addressing CES challenges. These include supplementing the survey with other microdata (from both public and private sector sources), increasing the frequency of benchmark revisions, and adjusting the approach to estimating and modeling establishment births and deaths. None of these ideas are entirely new—indeed, most of these ideas have been studied in depth by BLS staff—but I attempt to provide a new perspective and impetus for further research.

Despite recent concerns, the CES is a high-quality, highly accurate economic measurement program with a rigorous scientific methodology and clear, transparent technical documentation for data users. The CES is rightfully relied upon as a critical indicator of the cyclical state of the economy. The product benefits from an extraordinarily timely and comprehensive source of benchmark data, and benchmark revisions tend to be modest. The CES design has improved over time based on rigorous, forward-looking research by BLS staff who have proven adept at adjusting the product in response to macroeconomic developments while maintaining broader methodological consistency. Still, given the critical importance of the CES, recent revision patterns merit examination and discussion.

Throughout the paper, I limit my discussion to the national CES program (i.e., not the state and metro area program), and I focus mostly on private sector employment.

2 The CES

In this section I briefly review key CES design features.[1] Readers already familiar with these details may prefer to skip to section 3, where I review CES accuracy and challenges.

2.1 The CES survey

The CES is based in part on a monthly survey of more than 600,000 establishments at just over 100,000 businesses or government agencies, inclusive of all industries except farms.[2] The sample accounts for roughly one-fourth of total U.S. payroll jobs and one-fifth of private sector jobs—though this share ranges widely across industries—and is drawn from a sampling frame derived from state unemployment insurance (UI) records. The UI records also provide most of the annual benchmark data for CES (see subsection 2.4) and are tabulated publicly in the Quarterly Census of Employment and Wages (QCEW) and the Business Employment Dynamics (BED).

Survey respondents are asked questions about business activity during the pay period including the 12th of the month. The most publicized of these refers to total employee counts, but there are also questions about production or nonsupervisory employees, women employees, payroll, and hours worked including overtime hours. The BLS says that "the total nonfarm employment level is the primary estimate of interest" (Bureau of Labor Statistics 2025a), and some of the other estimates provided in the release are calculated based in part on total employment (e.g., average weekly hours is calculated from total hours—which are directly reported by respondents—divided by total employment).

The BLS uses the survey data to construct, typically at the narrow industry level, a "weighted link relative" (WLR) for measuring employment growth. The WLR is the employment growth rate of establishments that responded to the survey both in the current month

[1] Many details in this section are from the CES Handbook of Methods (Bureau of Labor Statistics 2025a). For simplicity, I use the term "establishment" to be interchangeable with "worksite," though there may be some practical distinctions. A "firm" is a collection of one or more establishments under a common tax identifier.

[2] The survey is not mandatory nationwide but is mandatory in five states plus Puerto Rico. The survey's sample consists of certainty cases—government establishments, large establishments, respondents that use the BLS's Electronic Data Interchange (EDI), and establishments not covered by the UI system—as well as a random sample that is stratified by state, industry, and establishment size. The sample is drawn annually using source data for the first quarter, with a third-quarter update to pick up recent establishment births. The probability sample design of the CES dates back to at least 2003 (with specific industry tests starting earlier); previously the CES employed a quota sample design.

and in the prior month; a common alternative term for the WLR is the "continuer" or "matched sample" growth rate. Construction of the WLR is labor intensive, featuring analyst review of microdata, follow-up discussions with respondents, and rigorous outlier rules.

By design, the WLR does not account for the job creation of new establishments that hired their first employee between the reference periods ("births"); in fact, births cannot appear in the sample for at least a year after the hiring of their first employee. Nor does the WLR account for the job destruction of establishments that close between the reference periods ("deaths"). The challenge with a business survey is that businesses that shut down are not likely to respond to the survey, while nonresponse does not necessarily imply business death since healthy businesses may choose not to respond. This is a problem faced by all business surveys, but the CES addresses it in a unique—and uniquely transparent and well-documented—way.[3]

2.2 Birth and death in CES

The BLS estimates the net employment contribution of establishment birth and death using two steps.[4] First, in the "imputation step," the CES program implicitly assumes that, at the cell level, job creation from births is proportional to the employment size of newly *nonresponding* establishments—assumed to be deaths—in the sample, where the proportion is determined by the growth of continuers in the WLR. The process is *equivalent to assigning the WLR growth rate to nonresponders*. This can be thought of as using deaths to impute births, where the size of births (in current and previous months) relative to deaths is determined by the current growth rates of continuing establishments within the relevant cell. This framing is motivated by research finding strong positive comovement between births and deaths (e.g., Mueller 2006), and by the notion that the relative size of births versus deaths is related to broader economic conditions as proxied by growth of the WLR. In practical terms, the imputation step simply means that nonresponders and first-time responders are excluded from the WLR.

Second, because historically the net job creation of establishment births and deaths tends to exceed what can be accounted for in the imputation step, the CES also includes a forecast from the NBD Model. The model is intended to capture a residual and is estimated at the cell level.

The specific design of the model has changed a few times since its inception in the early 2000s. The consistent feature of the NBD Model is that it features an autoregressive integrated

---

[3] For example, birth and death cannot be measured in real time for the three critical Census Bureau surveys that underly estimates of household and business spending in the National Income and Product Accounts (NIPAs): the Monthly Retail Trade Survey (MRTS), the Quarterly Services Survey (QSS), and the Manufacturers' Shipments, Inventories and Orders (M3). The MRTS and QSS, which are based on representative probability-based samples, have processes for gradually introducing birth and death into their sample with a lag of several months, while the M3 is not a probability-based sample and does not introduce births into the sample in a systematic manner. These surveys each use some combination of imputation (somewhat similar to the CES program's imputation step) and "carry-forward" factors projecting the most recent benchmark revision onto current estimates, but these carry-forward factors only adjust the level of sales, not the growth of sales.

[4] Prior to the redesign of the CES in the early 2000s, the CES program dealt with the problem of birth and death, and other sources of error, using a simpler "bias adjustment" process that projected past benchmark revisions forward into current estimates. After testing in select industries, starting in 2003 a new two-step birth-death adjustment process was implemented for all industries, with the goal of accounting for any jobs created (on net) by establishments that were opened or closed since the initiation of the survey sample. The canonical description of this development is Mueller (2006), though the model is described in detail in Bureau of Labor Statistics (2025a).

moving average (ARIMA) forecast that is estimated on about 5 years' worth of QCEW microdata then projected forward to inform current payroll estimates.

> To develop the history for modeling, the same handling of business deaths as described for the CES sample data is applied to the population data. Establishments that go out of business have employment imputed for them based on the rate of change of the continuous units. The employment associated with continuous units and the employment imputed from deaths are summed. The difference is compared with the actual population level to create the series modeled by the birth/death models. (Mueller 2006, 31)

The simulated net birth-death residuals just described form the estimation data for the ARIMA model. The resulting estimates of residual birth-death job creation are added to the WLR-based jobs estimate at the industry level prior to seasonal adjustment. Sector-level NBD Model contributions are published on the BLS website with each month of data. In addition, aggregate NBD Model "actuals" are published—with monthly detail—with each year's annual benchmark revision, such that data users can compare the model forecasts with the actual NBD Model residuals as calculated with the simulation approach described above.

In the current methodology, the NBD Model combines the ARIMA forecast with signal from the contemporaneous WLR growth rate of a given month; this methodology only became standard with the release of January 2026 data in February 2026 (see Bureau of Labor Statistics 2026b). But for most of its history, the NBD Model has been a pure ARIMA forecast, incorporating no concurrent information, because early BLS research had found that the NBD residual was fairly stable and acyclical. The use of the concurrent WLR in the NBD Model has been found to reduce benchmark revisions on two previous occasions, though it is not without its own challenges; for example, this method might exaggerate monthly noise in the sample-based WLR (a concern also flagged by Reinhart 2026).[5]

The ARIMA component of the NBD Model is a forecast based on data that extend only through the corresponding quarter of the year prior.[6] That is, NBD Model contributions for January through March of a given year are estimated on data through March of the prior year; contributions for April through June are estimated on data through June of the prior year; and so on. NBD Model estimates for the months of April through December (the post-benchmark

---

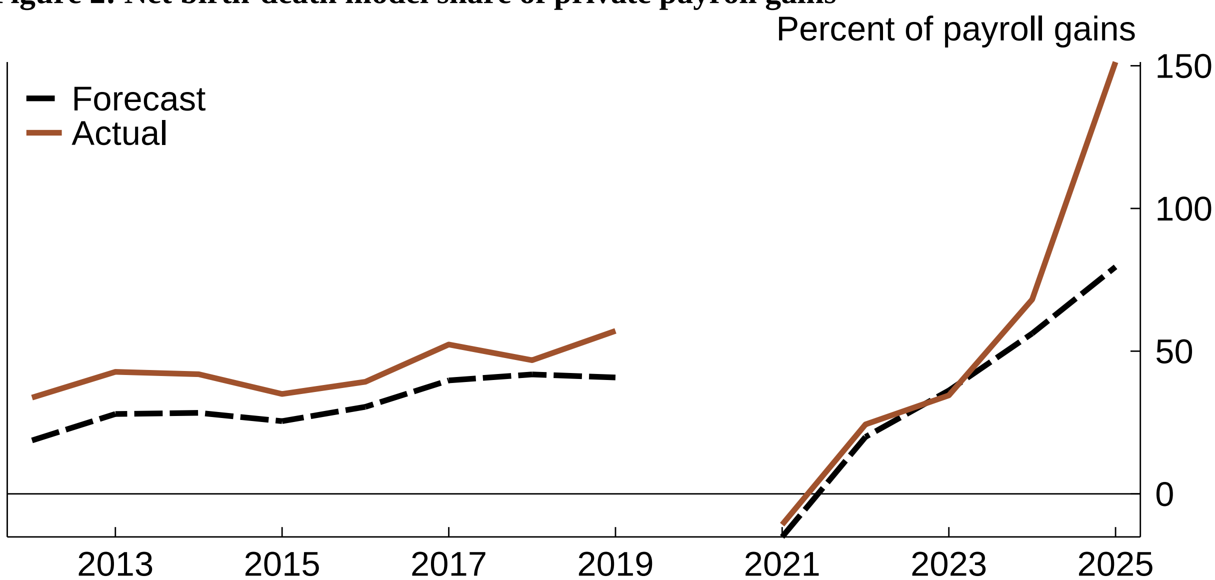
[5] The now-standard methodology that includes the contemporaneous WLR signal has been *temporarily* implemented on two previous occasions. The first was during the early pandemic months of 2020 and discontinued in October 2021 (described in Bureau of Labor Statistics 2021b, 2026a; and Stewart 2021). Later BLS research found that this methodological adjustment, along with certain others at that time, significantly reduced overall CES errors, though with larger errors in some cells. Second, as part of the 2024 benchmark revision published in February of 2025, the BLS applied the method to the revised NBD Model estimates for the post-benchmark period of April through October of 2024 (Bureau of Labor Statistics 2025c), which ultimately reduced the size of the 2025 benchmark revision by about 200,000 jobs (Bureau of Labor Statistics 2026b). At that time, the BLS indicated that the CES program would not use this methodology on an ongoing basis as "there is not sufficient time during the monthly estimate period to incorporate the additional processing needed" (Bureau of Labor Statistics 2026a). Since then, BLS staff have found a way to incorporate this method into monthly estimates without disruption to the usual publication schedule.

[6] Before 2011, the initial model forecast for the 12 months of a given year relied on QCEW data through March of the prior year, such that NBD Model forecasts extended out for up to 21 months beyond available QCEW data. The model was then re-estimated at the time of the annual benchmark revision to provide revised estimates for April through December (the "post-benchmark period"). In response to large NBD Model errors observed during the Great Recession, starting in 2011 the BLS switched to quarterly estimation of the model. See Clausen (2011) and Bureau of Labor Statistics (2020) for detail on the improved accuracy of the quarterly estimation approach.

period) are revised at the time of the benchmark revision (published the following February), generating a second vintage of NBD Model forecasts.

While data users likely think of the CES as being primarily a survey-based statistic with some ancillary birth-death machinery attached, in reality the NBD Model is a significant contributor to the overall payroll estimates. Figure 2 shows annual NBD Model contributions as a percent of CES total private payroll gains, where each year is defined as the 12 months through March. The dashed line refers to the initial NBD Model forecasts divided by private payroll gains as published just before the subsequent benchmark revision, while the solid line refers to NBD Model actuals (as published in benchmark articles) as a percent of final published private payroll gains after all revisions.

Figure 2: Net birth-death model share of private payroll gains



Note: Private sector; annual data with March reference. 2020 omitted. *Forecast* divides first-print model forecasts by eve-of-subsequent-benchmark published payroll gains. *Actual* divides model actuals (from benchmark article) by January-2026-vintage payroll gains.
Source: BLS Current Employment Statistics and author calculations.

Prior to the pandemic, the NBD Model share of private payroll gains was substantial and trending up. During the most recent few years it has risen further such that the model alone accounts for well over half of published payroll gains, no matter how or when these are measured.[7] Focusing on 2025 in particular, the model *forecast* cumulatively accounted for more than 70 percent of published private payroll gains, as published in monthly releases. After the benchmark revision, the NBD Model actuals *more than* accounted for private payroll gains, a puzzling outcome that must be treated with caution.[8]

---

[7] For 2025, the difference between the solid and dashed lines in figure 2 mainly reflects the different denominators used for the two series. The dashed line uses the pre-benchmark data vintage for private payroll gains, while the solid line uses the benchmarked data—which featured a large downward revision as published in February 2026.
[8] This fact does not imply that continuing establishments were a net drag on payroll employment growth for the year; the NBD Model actuals are a residual from a complex simulation which can be somewhat misleading in extreme situations. Per BED data, during the full year ending in March 2025 net establishment entry accounted for roughly half of total private employment gains—the highest share since BED data began in 1994.

2.3 The monthly CES estimate

Within narrow industry cells, the WLR growth rate for continuing establishments is applied to the previous month's employment level, then the contribution of the NBD Model is added. The BLS then applies concurrently estimated seasonal adjustment factors (i.e., the seasonal factors are re-estimated every month). Each month, the BLS publishes CES data for that month along with revisions to the previous two months; these revisions can arise from late-arriving survey responses and revisions to the seasonal factors.

2.4 Benchmarking

Many major statistical products feature high-frequency estimates (e.g., monthly or quarterly) based on imperfect data, which are later "benchmarked" to higher-quality or more comprehensive data. A consideration for CES, in particular, is that the nature of its WLR-based monthly employment estimates implies that errors in monthly data *accumulate* throughout the year. Since the 1940s, the CES has been occasionally benchmarked to data from the UI system, though regular annual benchmarks did not begin until the 1980s (Mullins 2016).[9]

The UI-based QCEW provides the main source data for CES benchmarking.[10] The national CES program focuses on QCEW for the first quarter of a given year; the first quarter is generally considered to provide higher-quality data than other quarters.[11] QCEW is a quarterly data release but features employment data for all three months of a quarter; the third month of the quarter is generally thought to be the most accurate (Mance 2016), and the CES program focuses on a March benchmark. The benchmark is made in non-seasonally adjusted terms. In current practice, the benchmark for March data is published in the following February along with the first release of January data.

The CES program matches CES total employment to the March level from the benchmark data within industries. The revision to the employment level is then smoothed back through the 11 months before March using a linear "wedge-back" procedure. With a new March employment level in hand, monthly data for the post-benchmark period are updated so that the existing WLR-based growth rates are applied to the new March employment level. The NBD Model is re-estimated, and the data are newly seasonally adjusted.

Importantly, the annual benchmark revision to national CES estimates does not incorporate significant new information about *monthly* employment fluctuations, either in the

---

[9] Before 1940, CES was occasionally benchmarked to manufacturing censuses (Mullins 2016, Robertson 2017).

[10] QCEW accounts for more than 95 percent of the CES benchmark. Some businesses that fall under the scope of CES are not present in QCEW because they are not covered by the UI system; these industries include railroads, some insurers, schools, healthcare and social services providers, and religious organizations (for a full list see Manning and Stewart 2017). The CES program fills this gap using other data sources and projections as necessary; for example, timely railroad employment data are available from the Railroad Retirement Board, while data on many nonprofits are only available from the Census Bureau's County Business Patterns (CBP) with a multi-year lag. See Decker et al. (2021) for extensive discussion of the QCEW and the CBP universes.

[11] For example, changes in establishment characteristics (e.g., industry) are typically introduced in the first quarter (Mance 2016), and the first-quarter QCEW receives more revisions than other quarters. The first quarter receives four revisions; the second quarter receives three revisions; the third quarter receives two revisions; and the fourth quarter receives just one revision (Bureau of Labor Statistics 2025b).

months before or the months after March. Revisions to the monthly data before the March benchmark simply reflect the wedge-back procedure and new seasonal factors; post-March months receive new NBD Model forecasts and seasonal factors, along with late-arriving survey responses for the most recent months. To a significant extent, the monthly "wiggles" generated by the survey-based WLR and the NBD Model forecasts live on in the data forever.[12]

What causes benchmark revisions? NBD Model forecast errors are one source of revision, as is "nonresponse error" (i.e., establishments not responding to the survey have different employment growth patterns than those responding). A third source of error is "response error" (sometimes called reporting error), or discrepancies between what survey respondents report in the CES survey and how their establishment appears in the QCEW. BLS staff have found that response error can arise from imputation in the QCEW, data collection timing differences, and inconsistent reporting procedures by establishments (Groen 2012).

3 CES accuracy and challenges

3.1 Accuracy

The accuracy of CES is a perennial topic, perhaps due in part to the ability of data watchers to derive reasonable estimates of benchmark revisions several months in advance using QCEW data and the BLS's own "preliminary benchmark" published in the fall prior to the formal revisions.[13] The size of CES benchmark revisions sometimes garners widespread attention.

The CES benchmark quality and cadence are extraordinary relative to other important business surveys. The QCEW is a near-census of the employer universe. The first publication of first-quarter QCEW data is available in September, a lag of just two quarters, providing both the BLS and external data users with a good guess of the likely benchmark revision. The BLS then thoroughly processes the benchmark data—and observes revisions to first-quarter QCEW—and publishes the full, detailed benchmark revision in February of the following year, less than four quarters after the end of the benchmark reference quarter. To be more concrete: As of February 2026, published CES data have incorporated an employment near-census through March of 2025. At the same time, the MRTS and the M3—both critical inputs to GDP estimates—have incorporated relevant annual data only through 2022.[14] And some important economic indicators, such as unemployment rate featured in the Current Population Survey (CPS), do not even have benchmark revisions because they do not have benchmark data.[15]

[12] Moreover, the benchmark data only provide information on total employment and wages; other CES variables, such as hours worked or the number of production and nonsupervisory workers, are adjusted based on their relationship with total employment or wages as originally estimated in the monthly survey data.

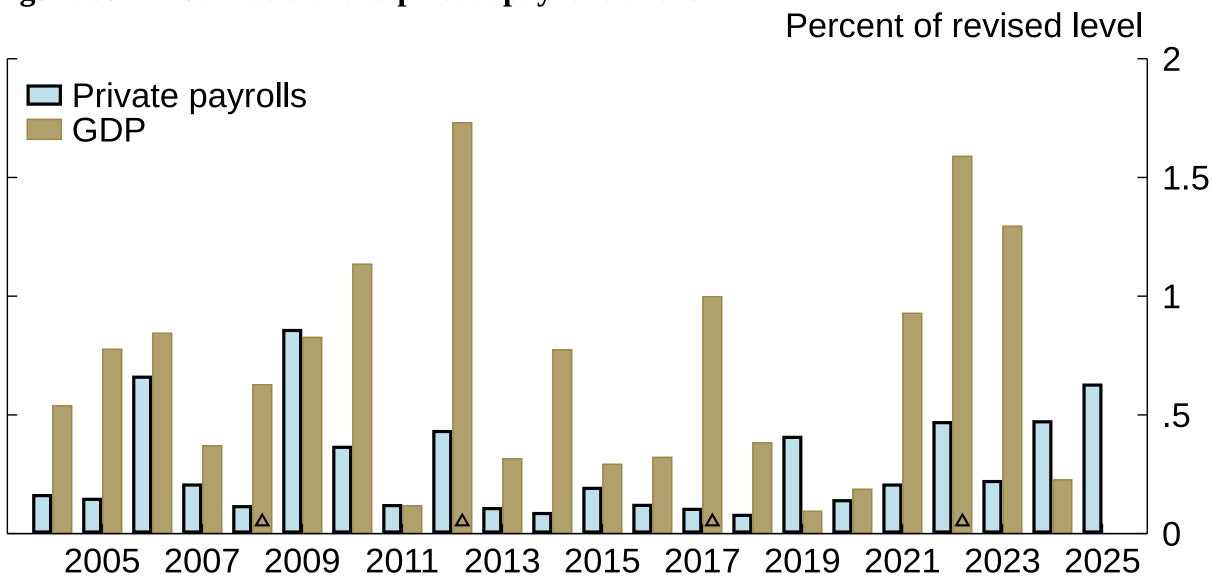
[13] Throughout this paper, I take the view that benchmark revisions are a good measure of statistical product accuracy. This view presupposes that the benchmark source data are more accurate than real-time survey- and model-based estimates. But I acknowledge that benchmark data have their own limitations; for QCEW, recent patterns of revisions have raised some concerns.

[14] Most annual revisions to MRTS and M3 benchmark to *survey* data, with *census* data used only every five years. The third major business survey feeding GDP estimates, the QSS, has had an even longer lag at present.

[15] While the CPS does incorporate new population controls each year which users can project backwards to adjust population weights in published data (Coglianese, Murray, and Nekarda 2025), there are no benchmark data for

"The long-term accuracy of CES estimates is best measured by analyzing annual CES benchmark revisions" (Calvillo and Downing 2016), and "the total nonfarm employment *level* is the primary estimate of interest" (Bureau of Labor Statistics 2025a, emphasis added). The outlined bars in figure 3 report absolute revisions to the *level* of private payroll employment as a percent of employment.[16] The largest revision of the past 20 years occurred in 2009, a benchmark period during which employment had fallen dramatically. Notably, the most recent revision—pertaining to the employment level in March 2025—was the largest since 2009. But looking over the whole period, revisions to the level of payroll employment tend to be small, rarely exceeding 0.5 percent of employment and never reaching 1 percent of employment.

Figure 3: Annual revisions to private payrolls and GDP



Note: Absolute revision to level in March (for private payrolls) or full year (for GDP) as published in subsequent year. Δ denotes GDP comprehensive revision years.
Source: BLS Current Employment Statistics; BEA National Income and Product Accounts. BEA data retrieved from ALFRED, Federal Reserve Bank of St. Louis.

Figure 3 also shows that annual revisions to GDP (as published by the Bureau of Economic Analysis, or BEA) tend to be larger than revisions to CES payrolls. This comparison is not entirely fair, as GDP is a much more complex statistical product than CES and combines numerous separate data sources.[17] But the products are similar in that they are both designed to target most or all of the economy. For the time period shown (2004 onward), GDP absolute

---

actual CPS-based economic indicators like the unemployment rate. Therefore, the fact that the unemployment rate does not revise while CES payrolls estimates do is not a point in CPS's favor.

[16] The figure is limited to private, not total nonfarm, employment, to be consistent with other analysis in this paper. In the chart, each year on the horizontal axis corresponds to the revision to that year's employment level as published early in the following year. For example, the year 2025 on the chart refers to the benchmark revision to the employment level of March 2025 as published in February of 2026.

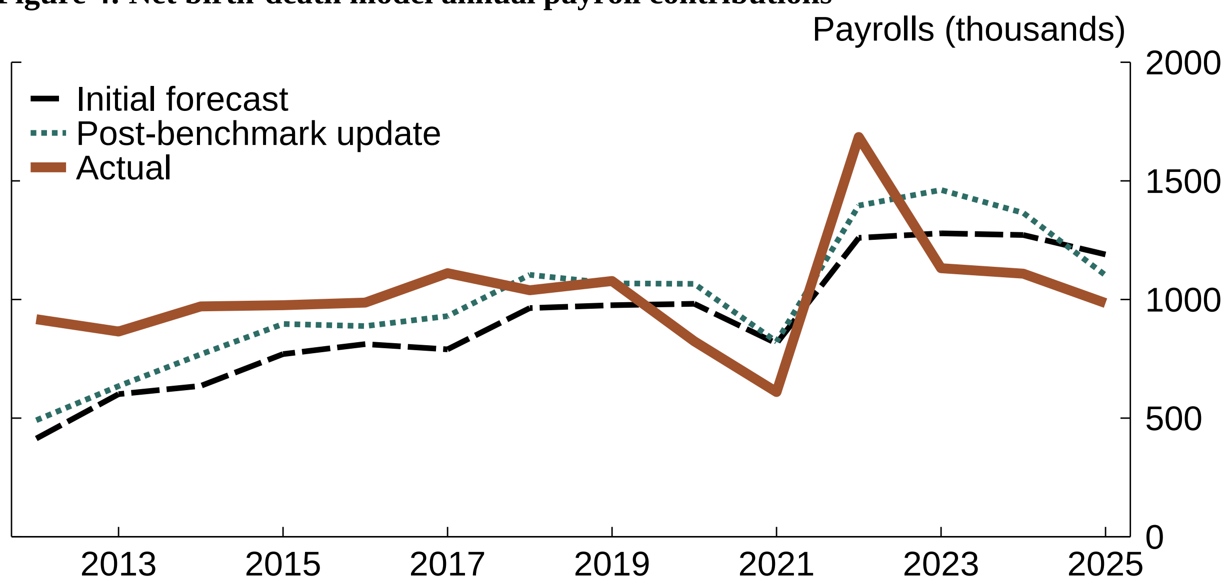
[17] Indeed, the comparison of CES to GDP is not meant to be a criticism of GDP or the BEA. GDP is an immensely complicated product. And high-quality benchmark data for GDP are only available with much longer lags than CES benchmark data. I make the comparison to GDP only because it has a comparable scope to CES (i.e., the entire economy rather than a subet of industries or geographic areas).

revisions averaged 0.7 percent of GDP, more than twice the absolute employment revisions (0.3 percent of employment); this result is similar when limiting to the last ten years and does not depend on whether the 2025 employment revision is included.

The fact that CES revisions tend to be small—both as a percent of employment and relative to GDP revisions—may be underappreciated by data users.

One reason the CES revisions are small is that the NBD Model, for all its various limitations and challenges, is arguably an impressively accurate model. This can be seen in figure 4, which provides measures of net birth-death job contributions, summed to the annual frequency on CES benchmark timing (i.e., April through March). The heavy solid line shows NBD Model "actuals," as reported in the annual benchmark articles. The dashed line reports the initial NBD Model estimates as published in monthly CES data, while the dotted line shows revised NBD Model estimates (where only the post-benchmark period is revised) as published at the time of the benchmark revision. I report NBD Model contributions in thousands of jobs for practical use by readers, but note that private payroll employment started the shown time period around 110 million and ended the period around 135 million, and the gap between the solid line and the dashed and dotted lines tends to be small relative to overall employment.

Figure 4: Net birth-death model annual payroll contributions

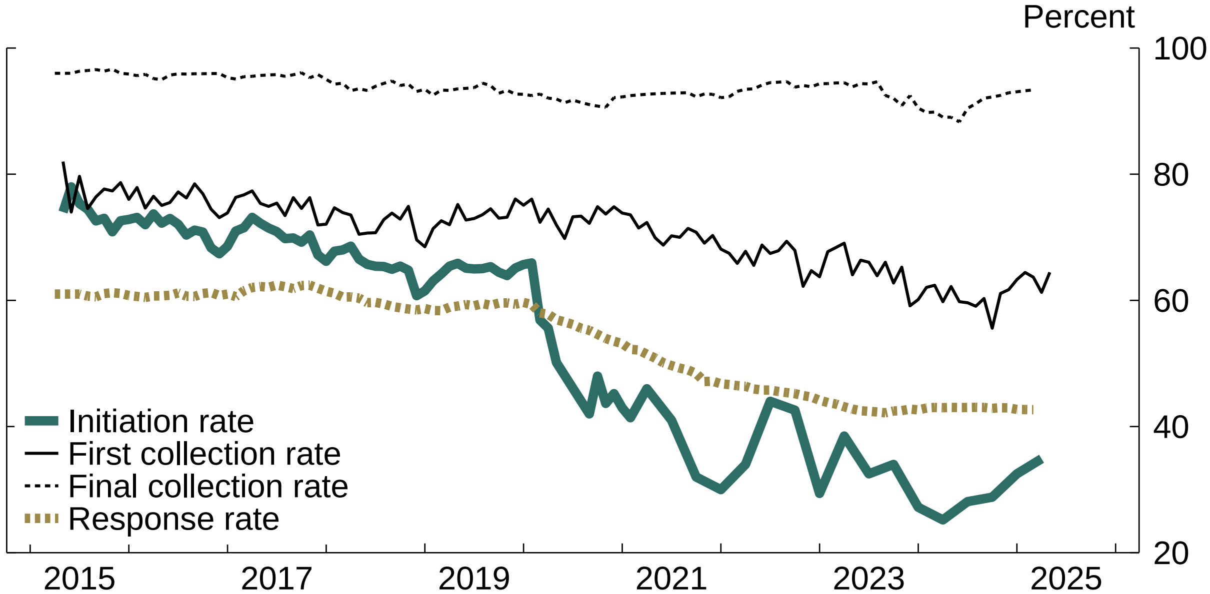


Note: Annual data with March reference. *Post-benchmark update* is revised model forecasts for April through December of prior year (published with benchmark revision to March of prior year) with initial model forecasts for January through March of current year.
Source: BLS Current Employment Statistics and author calculations.

3.2 Survey initiation, collection, and response

Despite the above defense of CES accuracy, the product is not without its challenges. Like other business surveys, CES has seen a marked decline in "response rates." But this is a complicated issue, which I will illustrate using figure 5.

Figure 5: Initiation, collection, and response rates



Note: 3-month moving average where possible.
Source: BLS Office of Survey Methods Research.

The CES program first draws a sample from the universe of UI units and attempts to contact sampled businesses.

- The share of that sample that BLS staff can activate for data collection—agreeing to participate and providing initial data—after several months is called the **initiation rate**, which is the heavy solid line in figure 5. The BLS has struggled to persuade businesses to participate in the CES survey.

- The thin solid and thin dashed lines indicate the first and final **collection rates**. The collection rate is the share of actively participating surveyed locations returning the survey, where the first collection rate pertains to the initial employment estimate for a given month, while the final collection rate pertains to the third release of the month. The first collection rate has trended down since the pandemic, but the final collection rate has been relatively stable. As Wilcox (2025) writes, "BLS is gaining more information than ever before between the first and second estimates of payroll employment." But the collection rate measures survey responses as a share of the actively responding sample—it is a measure of responses *relative to how many responses the BLS expects*. Businesses in the CES sample that refuse to respond to the survey or have not responded for several months do not count in the denominator of the collection rate. This means it is almost inevitable that final collection rates are fairly high.

- The **response rate**—the heavy dotted line in figure 5—measures the share of the total CES sample that is responding to the survey; that is, the denominator of the response rate features all sampled businesses (except known closures), including those refusing to answer the survey or not answering for many consecutive months. While the collection rates indicate that BLS is able to obtain data from nearly all active survey participants, the response rate indicates that fewer businesses targeted by the BLS sample selection processes are actively participating in the survey.

A decline in response rates reduces the effective sample size available for constructing payroll estimates and increases the variance of the estimates. It may also introduce bias if the nonresponders have systematically different employment growth rates than responders (see Davis et al. 2010, discussed more below).

3.3 Monthly revisions

As described above, CES features two separate revision cycles. Monthly data have a 3-month estimation and revision window, such that each month's initial estimate is revised in each of the two following months based on late survey responders and re-estimated seasonal adjustment factors. After the second monthly revision, the revision window is closed until the annual benchmark revision.

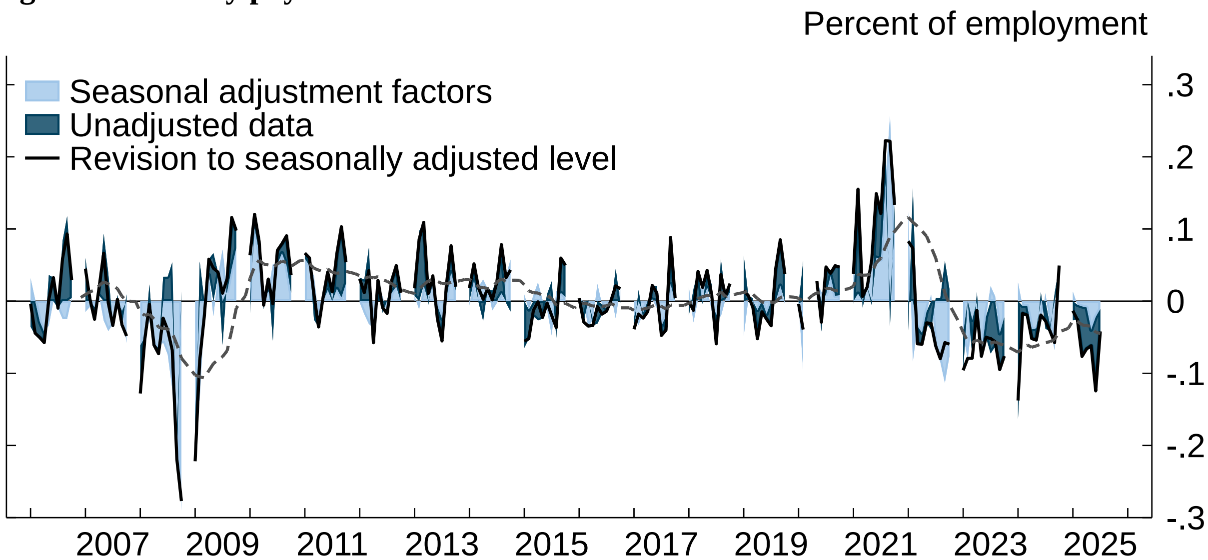
Figure 6 reports revisions to monthly (seasonally adjusted) private payrolls, measured as the difference between the employment levels of the third estimate and the first estimate (abstracting from annual benchmark revisions).[18] The total revision is given by the solid line and is expressed as a percent of third-estimate employment. The shaded regions provide perspective into the sources of these revisions. The lighter shaded area shows the portion of revisions arising from monthly re-estimation of seasonal factors. The darker shaded area shows the portion arising from the underlying non-seasonally adjusted (NSA) data; these revisions result from late survey submission.

Revisions were large during the recessionary period of late 2008 and early 2009 then again in the years since the pandemic. In some periods revisions appear reasonably unbiased and without persistence (2015-2019); but revisions have been persistently negative in some periods, including recent years, as illustrated by the 12-month moving average (dashed line). Large revisions tend to feature both revision sources (seasonal adjustment factors and unadjusted data) moving in the same direction; indeed, among months with larger-than-average revisions, the two sources have the same sign more than two-thirds of the time.

Since the pandemic, revisions have clearly been larger than the subdued levels seen before. Indeed, average absolute revisions during 2020–2025 were more than double those seen during 2015–2019 (0.065 versus 0.028 percent of employment). The 2020–2025 period also saw larger revisions than during 2010–2014 (0.041 percent of employment) but similar revisions to the Great Recession years of 2007–2009 (0.065 percent). Of course, the 2020–2025 period is strongly affected by the large revisions of 2021 and early 2022, but even in 2025 revisions still averaged 0.059 percent of employment.

[18] November and December are omitted because their monthly revision window includes the benchmark revision, which distorts the analysis. March and April of 2020 are omitted due to extreme outliers resulting from the pandemic. I end this exercise in August 2025; data for September through the end of 2025 were affected by a federal government shutdown, distorting the collection, publication, and revision timing of the estimates.

Figure 6: Monthly payroll revisions



Note: Private sector. Level revision from first to third print. November-December of all years and March-April 2020 omitted. Denominator is third-print seasonally adjusted employment level. Dashed line is 12-mo. moving average; omitted months are assigned calendar-year average.
Source: BLS Current Employment Statistics and author calculations.

The increase in average absolute revisions in the wake of the pandemic is driven in part by changing seasonal adjustment factors: On average, revisions to seasonal factors accounted for about two-thirds of overall revisions in 2021–2024, a much larger share than in prior years. But 2025 was different: Seasonal factors accounted for less than 10 percent of overall revisions. In other words, the large first-to-third revisions seen in 2025 were mostly driven by revisions to the underlying data (e.g., due to late survey responders). On the one hand, these revisions imply that BLS is gaining significant new information within the 3-month publication and revision window, as emphasized by Wilcox (2025). On the other hand, these revisions, along with the historically low overall *response* and *initiation* rates, raise concerns about how accurate even the third print might be. To assess that, we must study benchmark revisions.

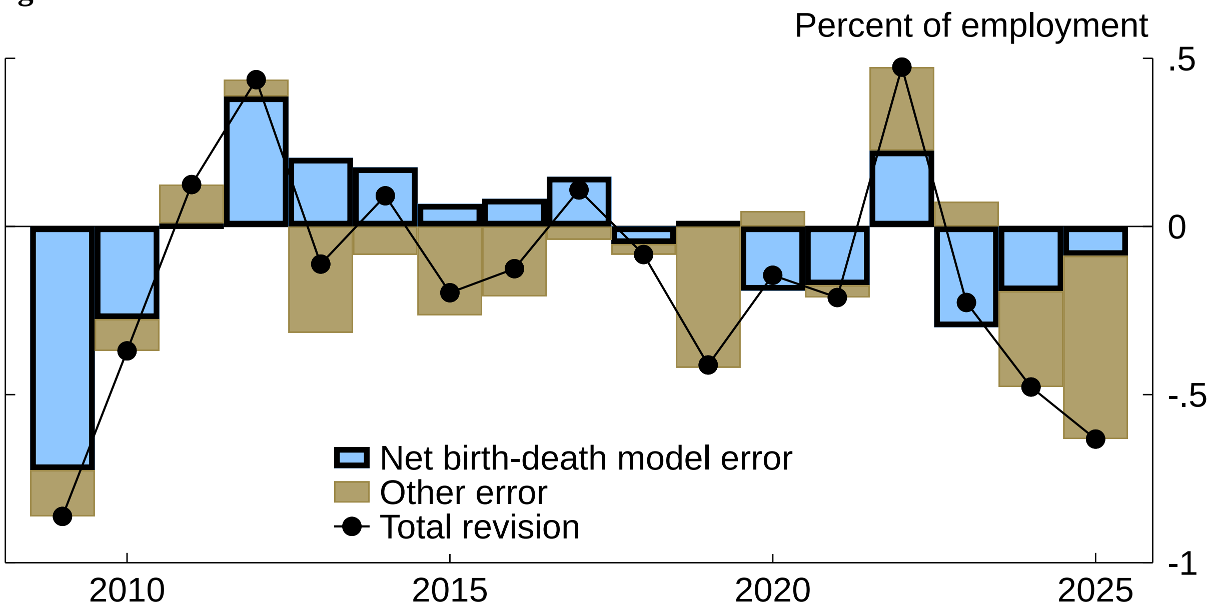
3.4 Benchmark revisions

Figure 7 reports annual benchmark revisions decomposed into two broad sources: NBD Model error and “other error,” that is, response and nonresponse error.[19] Most years feature some of both kinds of error, though in some years these errors work in opposite directions and therefore partly offset each other, at least at the aggregate level. But years with large benchmark revisions

[19] Scope changes and similar adjustments may also appear under “other error.” “Net birth-death model error” refers to the difference between NBD Model actuals (published in benchmark articles) and the second vintage of NBD Model estimates (the dotted line in figure 4), which is incorporated into the data at the time of the prior year’s benchmark revision. Therefore, the outlined bars in figure 7 may understate the total error from early vintage CES data to benchmark revision, if the second vintage NBD Model forecasts are an improvement on the first vintage NBD Model forecasts (the dashed line in figure 4). As a reminder, “nonresponse error” arises from establishments not responding to the survey and having different growth patterns than responders, and “response error” reflects discrepancies between an establishment’s reporting to the CES survey and its record in the QCEW data.

tend to feature both kinds of error contributing in the same direction. For example, consider the seven years during 2008 through 2025 in which the (absolute) benchmark revision was larger than its overall (2008–2025) average of about 0.3 percent of employment; these above-average years are 2009, 2010, 2012, 2019, 2022, 2024, and 2025. In six out of these seven years, NBD Model errors and other sources of error were in the same direction. The one exception is 2019, which featured a large negative "other error" and a negligible, but positive, NBD Model error.

Figure 7: Sources of benchmark revisions



Note: Percent of real-time vintage benchmark employment. *Net birth-death model error* is difference between model actuals and post-benchmark updates (i.e., second vintage).
Source: BLS Current Employment Statistics and author calculations.

Interestingly, the relative role of the NBD Model in large benchmark revisions has declined over time. In the large benchmark revision years of 2009, 2010, and 2012, the NBD Model share of the total revision averaged 82 percent (recall that the NBD Model methodology switched to quarterly estimation only in 2011). In the large benchmark revision years of 2022, 2024, and 2025, the NBD Model share of the total revision averaged 34 percent, with a noteworthy share of just 14 percent in the most recent (2025) revision. It is encouraging—and a credit to BLS staff—that the NBD Model has been so accurate lately.[20]

But large benchmark revisions with low NBD Model error suggest that the CES sample is providing less accurate payroll estimates than in past years, due to some combination of nonresponse and response error. This is particularly concerning in light of the large *monthly* revisions (from first to third print) of recent years shown in figure 6; not only are initial CES

[20] Part of the reason for the good performance of the NBD Model during the 2025 benchmark year was the methodological adjustment made for the second-vintage NBD Model estimates incorporated for the months of April through October of 2024 during the benchmark revision published in early 2025. The BLS later reported that this methodological adjustment improved the NBD Model forecast by more than 200,000 jobs (Bureau of Labor Statistics 2026b).

estimates requiring large immediate revisions, but even the third monthly estimate appears less accurate than in times past.

4 Potential methodological innovations

4.1 Obtaining more data

The large declines in CES initiation and overall response rates seen in figure 5 imply that the effective size of the dataset available for generating CES estimates has contracted. A natural question is therefore: Can the CES program obtain more data? One approach may be to increase survey coverage through a larger survey or higher response rates. An alternative is to obtain supplemental sources of data, either from the public sector or the private sector.

4.1.1 Increasing survey coverage

The most obvious—though perhaps not the most easy or practical—way to improve the quality of the CES survey is to expand the survey. Sampling variance declines with sample size (this has been confirmed empirically for pre-2000s CES data by Krueger and Fortson 2003). Survey expansion has happened on numerous previous occasions. The current sample targets about 630,000 establishments, but this sample has grown considerably since the current sample design methodology was introduced in the early 2000s.[21] Of course, any expansion of the sample would be subject to BLS resource constraints.[22]

An alternative is to increase the existing survey's response rates through mandatory reporting or a lower reporting cost, which the BLS has previously argued would particularly increase coverage of larger units (Bureau of Labor Statistics 2015). CES reporting is already mandatory in California, New Mexico, Ohio, Oregon, South Carolina (for firms with more than 20 employees), and Puerto Rico. Whether these jurisdictions actually achieve higher response rates cannot be determined, to my knowledge, from public-facing information—but CES staff could study the question (and may have already done so).

A natural reaction to low response rates is to acknowledge that responding is costly and to ask whether the BLS could make the reporting process easier for surveyed businesses. There may be avenues to pursue in this dimension, but the BLS has already made significant investments in easing the reporting process. Respondents can report via a web interface, with separate systems for single-establishment and multi-establishment firms. For very large firms (and government agencies) there is the Electronic Data Interchange (EDI), which is optimized for firms with at least 50 establishments and features a streamlined process for collecting data on thousands of establishments. Moreover, BLS staff have studied ways to maximize response rates (e.g., timing of reminders; see Johnson 2016). There may be scope for increasing cooperation

[21] An examination of past CES reports reveals a sample of 300,000 establishments in 2002 (as the transition from the previous sample design methodology was nearly completed), 400,000 in 2003, almost 500,000 in 2012, almost 600,000 in 2015, and nearly its current size in 2016. The 1990s quota-based sample targeted about 425,000 establishments, up from about 160,000 in 1975-1989. See Kelter (2016) for a history of CES sample design.

[22] In 2015, the BLS indicated that increasing the sample size by 85,000 UI accounts would cost $16 million per year (Bureau of Labor Statistics 2015).

with large payroll services providers to facilitate easier reporting, but it is likely that the CES program has picked most of the low-hanging fruit in terms of respondent burden.[23]

4.1.2 Obtaining data from public sector sources

The public sector generates abundant administrative data through day-to-day operations. I provide two potential examples of timely administrative data that might inform payroll estimates.[24]

The first is the tax collection system, in which nearly all employers withhold taxes from employee paychecks for submission to the Internal Revenue Service (IRS). A precedent for using tax collection data in timely payroll estimates comes from the Australian Bureau of Statistics (ABS) in coordination with the Australian Taxation Office (ATO). From January 2020 through March 2025, the ABS published weekly payroll statistics derived from administrative data on tax withholding and related processes; "the ABS receives selected employer and employee level data from the ATO STP system, which are combined with employer and jobholder characteristics from the ABS Business Register and the ATO Client Register" (Australian Bureau of Statistics 2025). The weekly index was discontinued in preparation for incorporation of the tax-based data feed into the headline Monthly Employee Earnings Indicator starting in early 2026. One might imagine studying a similar approach for the U.S., in which the BLS obtains establishment-level tax withholding microdata—the universe or even just a random sample—in a timely manner to be integrated with, or perhaps even to replace, monthly survey microdata.

Another potential example is the National Directory of New Hires (NDNH). The NDNH tracks new hiring of workers for the purposes of facilitating collection of child support and other obligations. The directory relies on state and federal sources, including the UI system. "Employers have up to 20 days from the date of hire, depending on State law, to report" required information (Administration for Children and Families 2025). NDNH microdata may be particularly useful for identifying new businesses shortly after their first formal hire.

4.1.3 Obtaining data from private sector sources

Numerous private companies generate timely and high-frequency payroll data as part of their business operations. The most obvious are payroll services providers.[25]

In a long-term project, a team at the Federal Reserve Board constructed CES-mimicking payroll indexes using timely microdata from a large payroll services provider (Cajner et al. 2018, 2019, 2020a, 2020b, 2022, 2023). The resulting indexes were useful in tracking rapidly moving economic developments in the pandemic and, more generally, have been found to help predict CES benchmark revisions. A key drawback of such data sources is that they may be a highly selected and non-representative sample of businesses: Certain types of businesses choose to engage certain payroll providers. QCEW can provide sampling weights to account for industry,

[23] Payroll companies reportedly already participate in the survey on behalf of clients selected for the CES sample. And as of 2014, about 40 percent of *QCEW* employment was reported by payroll and tax companies (Clayton 2014).
[24] Provision of additional administrative data to the BLS may require policy changes (Casselman 2015). I do not discuss or take a position on policy questions in this paper but simply describe potential ideas.
[25] Guessing from publicly available sources, I estimate that employment services firms ADP, Paychex, and UKG combined cover between one-quarter and one-third of U.S. private payrolls. Of course, there would likely be considerable overlap between these firms' clients and the CES survey.

establishment size, and geography composition, though selection biases would remain. But statistical agency staff are well qualified to combine such microdata with survey data to improve payroll estimates.[26] Reportedly, some payroll providers already contribute CES survey responses on behalf of clients in the sample; it may be that they could do so for other clients as well at relatively low additional cost.

While several private companies publish aggregate employment indexes based on their own data, the most useful way for the BLS to use private (or public) payroll data is through access to the underlying microdata. This is a key lesson of both the review of CES methodology provided above and of actual experiences with private payroll data. The reasons are twofold. First, the development of a payroll index requires many methodological choices: treatment of outliers, benchmarking, estimation of business birth and death or projection of past benchmark revisions forward, seasonal adjustment method, and so on. These methodological choices are best left to statistical agency staff who can weigh all considerations in the broader context of the statistical product's goals. Second, ideally the statistical agencies would obtain data from multiple payroll sources, potentially in combination with survey data; and ensuring consistent methodology across sources is critical for data users to have confidence in resulting estimates.

Importantly, many official data products *already* rely on private sources and have done so for many years (e.g., the NIPAs and Industrial Production). Still, alternative sources of data come with data quality limitations and methodological challenges, as well as various other problems posed uniquely by private sector-generated data (e.g., Abraham et al. 2022). These should not be thought of as prohibitive concerns but instead as tradeoffs that can best be navigated by trained and experienced BLS staff. A critical tool in this regard is the CES program's uniquely timely and high-quality benchmarking process, which facilitates continued evaluation of data accuracy and ensures high data quality over time.

4.2 Quarterly adjustments

The CES is benchmarked to the first-quarter QCEW each year, but the QCEW is published for every quarter; "CES is lucky to have this problem" (Mance 2016). Data users can—and often do—use the quarterly data to maintain a tracking estimate of the likely upcoming benchmark revision each year; anecdotally, such tracking estimates have performed well in recent years. Instead of every data user doing this themselves, the BLS itself could make use of the quarterly QCEW data with the goal of reducing the size of benchmark revisions and providing more accurate within-year employment estimates.

BLS staff have studied the possibility of quarterly benchmark for many years. For example:

- In a 2015 report to Congress (Bureau of Labor Statistics 2015), the BLS explains the potential for more-frequent (semiannual or quarterly) benchmarking to improve estimates—or make them worse—indicating a need for research on the topic. As of 2015, the BLS estimated that semiannual benchmarking would cost $4 million per year, while quarterly benchmarking would cost $7.5 million per year.

[26] Statistical agency staff already have experience working with private payroll data; see Dunn et al. (2024).

- Mance (2016) describes the benchmarking process for BLS state and local statistics which, unlike the national CES program, makes use of all months in QCEW in its annual benchmark; one critical point is that seasonal patterns differ between CES and QCEW. Mance also emphasizes that QCEW itself is subject to revision, creating a tradeoff between timeliness and accuracy of the potential benchmark data.

- Loewenstein and Dey (2017) show that, *ex ante*, whether quarterly benchmarking will improve CES accuracy depends on seasonal adjustment accuracy, the relative variance of QCEW and CES errors, and the persistence of CES errors. The authors propose a method for quarterly benchmarking with a focus on the quarterly *growth* of the *seasonally adjusted* data from both sources, finding that the approach generally does improve CES accuracy.

- Robertson (2017) emphasizes that benchmarking is extremely resource intensive ("More than 50,000 CES series are reviewed as part of the benchmark process"). Robertson outlines criteria for judging quarterly benchmarking methods, explores a few methods, and identifies the method of Loewenstein and Dey (2017) as meeting most or all criteria.

- Robertson (2021) emphasizes the limitations of the annual benchmarking method with its accompanying "wedge back" approach to revising April–February data on the assumption that CES errors accumulate linearly ("We know that this assumption is almost certainly wrong"). Robertson again endorses the Loewenstein and Dey (2017) quarterly benchmarking method: "It is readily apparent that the proposed solution is an improvement over either of the procedures currently used to benchmark CES data."

- In contrast, a 2021 post on the BLS website (Bureau of Labor Statistics 2021a) states, "additional research revealed that all of the proposed [quarterly benchmarking] methods produced numerous critical issues," some of which are listed. "Therefore, the BLS has decided to halt this research and redirect resources to other program initiatives."

Clearly, the question of quarterly benchmarking is complicated. Moreover, concerns about the quality of QCEW itself—particularly in quarters two through four—must not be ignored.[27] On balance, though, the research above suggests that the idea may merit continued consideration.

Even if quarterly *benchmarking* has been ruled out, though, full benchmarking is not the only way to use the QCEW's quarterly census information. Quarterly information could instead be included in a model designed to improve CES estimates without a full benchmark. One example of how this might work is provided by the Federal Reserve Board's monthly Industrial Production and Capacity Utilization (IPCU) statistics. IPCU is benchmarked annually to large Census Bureau surveys or quinquennial censuses. But the IPCU program also ingests data from the Census Bureau's Quarterly Survey of Plant Capacity Utilization (QSPC), which provides industry-level capacity utilization rates. In 2018, the IPCU program implemented a procedure in

[27] It is worth noting that quarterly QCEW data are already relied upon for other purposes. The CES program itself uses the data for NBD Model estimation, and the BED program is based entirely on quarterly QCEW. The BEA uses QCEW wage data in the NIPAs.

which monthly IPCU estimates at the industry level are adjusted by a model that takes signal from QSPC industry utilization rates.[28] In other words, some initial IPCU industry-level estimates are *adjusted* based on the model, not *replaced* or benchmarked by the QSPC data. One might imagine a similar approach in which industry-level CES payroll estimates are adjusted with a model informed by the discrepancy between quarterly QCEW and CES payroll growth.

Another example of using a model to adjust survey-based estimates is provided by Davis et al. (2010). The authors link microdata from the Job Openings and Labor Turnover Survey (JOLTS) to the BLS UI-based microdata and argue—based on indirect evidence—that JOLTS survey nonresponse and weighting schema cause a bias in JOLTS employment growth rates. The authors propose a method for adjusting JOLTS data using differences between JOLTS and the benchmark microdata where available. A similar method might be applied to CES: If QCEW microdata in hand can be matched with CES microdata each quarter, a "nonresponse bias adjustment" could be calculated and, if it is empirically persistent, projected forward.

4.3 Birth and death estimation

The recent methodological change to the NBD Model appears to have improved model performance. Whether this improvement is robust to the full range of macroeconomic environments has yet to be seen. I discuss three potential ideas for further improvements.

First, there might be scope to improve model estimates by exploiting more recent BLS data on establishment birth and death. As currently implemented, the NBD Model is estimated on QCEW microdata through the same quarter of the prior year (now also featuring signal from the contemporaneous sample WLR). For example, the NBD Model contribution incorporated into the July 2025 CES data (published in early August) featured an ARIMA forecast estimated on QCEW microdata through 2024:Q3.[29] But by that time, the public had access to 2024:Q4 data for both QCEW and BED (barely). Presumably the extensive data work and modeling required for NBD Model re-estimation takes considerable time such that full incorporation of 2024:Q4 data into the ARIMA model cannot be done in time for 2025:Q3 CES estimates; but it may be possible that simply including industry- or even aggregate-level net establishment entry data for the most recently available quarter in the NBD Model regressions could improve the forecasts.[30]

Second, further model improvements may be available through the use of contemporaneous external data. Battista (2013) explores several options and finds some value in including recent data on retail sales and CPS employment gains, though these performance improvements are strongest in recessions. A potentially promising additional data source is the timely Census Bureau's Business Formation Statistics (BFS; see Bayard et al. 2018), which counts applications for new Employer Identification Numbers.[31]

---

[28] See Board of Governors of the Federal Reserve System (2018).

[29] Here I describe mid-2025 publication patterns to avoid the distorted publication schedule in late 2025 and early 2026 resulting from a government shutdown.

[30] For example, in exploratory (in-sample) analysis, I regress NBD Model actuals on NBD Model (second-vintage) forecasts along with net job creation from establishment openings minus job destruction from closings in BED. Relative to having only NBD model forecasts in hand, the root mean squared error reduction from having two quarters of BED data is about 10 percent, three quarters is about 25 percent, and four quarters is about 60 percent. But these performance improvements may be (much) smaller under the new NBD Model methodology.

[31] Puzzlingly, in preliminary exploratory analysis I find that including BFS data in a regression model forecasting NBD Model actuals does slightly improve forecast performance, but that BFS data enter the model with a negative

Third, it may be possible to improve the model through direct observation of birth and/or death in real time. Recall that a key challenge in surveys of businesses is that it is difficult for a business to reliably report itself as having just opened or just closed in a survey of that business. Instead of tracking businesses themselves, what if the BLS could track the businesses' customers or vendors? For example, if customer credit card transactions have merchant identifiers, then we might observe the initiation or cessation of transactions at specific merchants. In some industries cell phone location tracking data may indicate business operational status through customer visits, or online business registers (like Google Maps or Facebook) that are organically updated by customers may provide relevant information. These ideas are explored to some extent in Crane et al. (2022) and more fully in Kurmann, Lale, and Ta (2025). What kind of data are most useful would vary widely by industry; for example, customer visit data may be of little use in the construction industry but highly useful for restaurants.

5 Conclusion

The CES is a high-quality statistical program based on sound and transparent methods. On balance the product tends to have high accuracy, but monthly and benchmark revisions have risen in recent years. The large declines in CES survey initiation and response rates demand serious attention and may suggest the survey itself is in need of enhancements.

One of the CES program's greatest assets is its staff and associated body of staff research. Early in the COVID-19 pandemic, the CES program rapidly modified both components of their establishment birth-death estimation method (the imputation step and the NBD Model); these adjustments were enabled in part by years' worth of directly relevant research by staff. When the modifications no longer provided value for the product, BLS staff adeptly reverted to the previous methodology, all while continually explaining their decisions to data users. More recently, after a series of large NBD Model errors, the CES program again adjusted the NBD Model methodology—again informed by extensive staff research—despite the difficulty of doing so in terms of staff time and production schedules.

Another key asset of the CES program is its uniquely timely and comprehensive benchmark data, the QCEW. When considering ideas for adjusting or improving the statistical product—especially ideas that feature external data sources—it is critical for all stakeholders to understand the importance of the benchmarking process which, unlike monthly estimates, simply *cannot be replicated with data sources from outside the statistical agencies*. Benchmarking provides a backstop allowing for innovation in the methods for constructing monthly estimates by providing a measure of data accuracy and ensuring reliability of the product over time.

References

Abraham, Katharine, Ron Jarmin, Brian Moyer, and Matthew Shapiro. 2022. "Introduction: Big Data for Twenty-First Century Economic Statistics: The Future Is Now." In *Big Data for Twenty-First-Century Economic Statistics*, edited by Katharine Abraham, Ron Jarmin,

sign. That is, higher business applications are associated with lower NBD Model actuals, possibly reflecting positive correlation between birth and death. This topic merits further investigation.

Brian Moyer, and Matthew Shapiro, 79:1–22. Studies in Income and Wealth. National Bureau of Economic Research.

Administration for Children and Families. 2025. “National Directory of New Hires.” At https://acf.gov/css/outreach-material/national-directory-new-hires (accessed February 16, 2026).

Australian Bureau of Statistics. 2025. “Payroll Jobs.” At https://www.abs.gov.au/statistics/labour/jobs/payroll-jobs/week-ending-15-march-2025 (accessed February 16, 2026).

Battista, Victoria. 2013. “Back to the Future: Using Current Regression Variables to Forecast Forward from Historical Net Birth/Death Employment.” *Working Paper*. At https://www.bls.gov/osmr/research-papers/2013/st130140.htm.

Bayard, Kimberly, Emin Dinlersoz, Timothy Dunne, John Haltiwanger, Jarmin Miranda, and John Stevens. 2018. “Early-Stage Business Formation: An Analysis of Applications for Employer Identification Numbers.” NBER working Paper no. 24364. At http://www.nber.org/papers/w24364.

Board of Governors of the Federal Reserve System. 2018. “Industrial Production and Capacity Utilization: The 2018 Annual Revision.” At https://www.federalreserve.gov/releases/g17/Revisions/20180323/DefaultRev.htm (accessed February 16, 2026).

Bureau of Labor Statistics. 2015. “A Report to Congress on the Bureau of Labor Statistics’ Current Employment Statistics Methodology for Metropolitan Statistical Areas.” At https://www.bls.gov/sae/additional-resources/bls-report-to-congress-on-ces-methodology-for-metropolitan-statistical-areas-2015.pdf (accessed February 20, 2026).

———. 2020. “Introduction of Quarterly Birth-Death Model Updates in the Establishment Survey.” https://www.bls.gov/ces/methods/ces-quarterly-birthdeath.htm (accessed December 22, 2025).

———. 2021a. “Experimental Quarterly Benchmarking from the Current Employment Statistics Survey.” *News Release*. At https://www.bls.gov/ces/notices/2021/experimental-quarterly-benchmarking-from-the-current-employment-statistics-survey.htm (accessed December 22, 2025).

———. 2021b. “The Employment Situation - October 2021.” *News Release*. At https://www.bls.gov/news.release/archives/empsit_11052021.pdf (accessed December 22, 2025).

———. 2025a. “Current Employment Statistics: Handbook of Methods.” At https://www.bls.gov/opub/hom/ces/ (accessed December 22, 2025).

———. 2025b. “Overview of QCEW Revisions.” At https://www.bls.gov/cew/revisions/ (accessed February 14, 2026).

———. 2025c. “CES National Benchmark Article.” *News Release*. At https://www.bls.gov/ces/publications/benchmark/ces-benchmark-revision-2024.pdf (accessed December 22, 2025).

———. 2026a. “CES Birth-Death Model Frequently Asked Questions.” At https://www.bls.gov/web/empsit/cesbdqa.htm (accessed February 14, 2026).

———. 2026b. “CES National Benchmark Article.” *News Release*. At https://www.bls.gov/ces/publications/benchmark/ces-benchmark-revision-2025.pdf (accessed February 12, 2026).

Cajner, Tomaz, Leland D. Crane, Ryan A. Decker, Adrian Hamins-Puertolas, and Christopher Kurz. 2019. "Tracking the Labor Market with Big Data." *FEDS Notes*. Board of Governors of the Federal Reserve System. doi:https://doi.org/10.17016/2380-7172.2441.

———. 2020a. "Tracking Labor Market Developments During the Covid-19 Pandemic: A Preliminary Assessment." Finance and Economics Discussion Series (FEDS) Working Paper 2020-030. Board of Governors of the Federal Reserve System.

———. 2022. "Improving the Accuracy of Economic Measurement with Multiple Data Sources: The Case of Payroll Employment Data." In *Big Data for Twenty-First-Century Economic Statistics*, edited by Katharine Abraham, Ron Jarmin, Brian Moyer, and Matthew Shapiro, 79:147–70. Studies in Income and Wealth. National Bureau of Economic Research.

———. 2023. "Payroll Employment at the Weekly Frequency." *AEA Papers and Proceedings* 113 (May): 145–50.

Cajner, Tomaz, Leland Crane, Ryan A. Decker, Adrian Hamins-Puertolas, Christopher Kurz, and Tyler Radler. 2018. "Using Payroll Processor Microdata to Measure Aggregate Labor Market Activity." Finance and Economics Discussion Series (FEDS) Working Paper 2018-005. Board of Governors of the Federal Reserve System.

Cajner, Tomaz, Leland D. Crane, Ryan A. Decker, John Grigsby, Adrian Hamins-Puertolas, Erik Hurst, Christopher Kurz, and Ahu Yildirmaz. 2020b. "The U.s. Labor Market During the Beginning of the Pandemic Recession." *Brookings Papers on Economic Activity* Summer: 3–33.

Calvillo, Michael and Tyler Downing. 2016. "One hundred years of Current Employment Statistics—Busting CES Myths." *Monthly Labor Review*, October.

Casselman, Ben. 2015. "Big Government Is Getting in the Way of Big Data." *FiveThirtyEight*, March. At https://fivethirtyeight.com/features/big-government-is-getting-in-the-way-of-big-data/ (accessed February 14, 2026).

Clausen, Nathan D. 2011. "Forecasting Birth/Death Residuals on a Quarterly Basis." *Working Paper*. At https://www.bls.gov/osmr/research-papers/2011/st110020.htm (accessed December 22, 2025).

Clayton, Richard. 2014. "Data Sharing: Progress and Challenges at BLS." *Presentation to FESAC*. At https://apps.bea.gov/fesac/meetings/2014-12-12_clayton.pdf (accessed February 20, 2026).

Coglianese, John, Seth Murray, and Christopher J. Nekarda. 2025. "Harmonized Population and Labor Force Statistics." Finance and Economics Discussion Series (FEDS) Working Paper 2025-057. Board of Governors of the Federal Reserve System.

Crane, Leland, Ryan Decker, Aaron Flaaen, Adrian Hamins-Puertolas, and Christopher Kurz. 2022. "Business Exit During the Covid-19 Pandemic: Non-Traditional Measures in Historical Context." *Journal of Macroeconomics* 72 (June).

Davis, Steven J., R. Jason Faberman, John C. Haltiwanger, and Ian Rucker. 2010. "Adjusted Estimates of Worker Flows and Job Openings in Jolts." In *Labor in the New Economy*, edited by Katharine Abraham, James Spletzer, and Michael Harper, 71:187–216. Studies in Income and Wealth. National Bureau of Economic Research.

Decker, Ryan A., Robert J. Kurtzman, Byron F. Lutz, and Christopher J. Nekarda. 2021. "Across the Universe: Policy Support for Employment and Revenue in the Pandemic Recession." Finance and Economics Discussion Series (FEDS) Working Paper 2020-099. Board of Governors of the Federal Reserve System.

Dunn, Abe, Eric English, Kyle Hood, Lowell Mason, and Brian Quistorff. 2024. "Expanding the Frontier of Economic Statistics Using Big Data: A Case Study of Regional Employment." *BEA Working Paper Series, Wp2024-4*. Bureau of Economic Analysis.

Groen, Jeffrey. 2012. "Sources of Error in Survey and Administrative Data: The Importance of Reporting Procedures." *Journal of Official Statistics* 28 no. 2:173–98.

Johnson, Nicholas. 2016. "Maximizing Web Survey Response: Research from the Current Employment Statistics Survey." *Working paper.* At https://www.bls.gov/osmr/research-papers/2016/st160040.htm (accessed February 20, 2026).

Kelter, Laura. 2016. "One Hundred Years of Current Employment Statistics—The History of CES Sample Design." *Monthly Labor Review*, August.

Krueger, Alan and Kenneth Fortson. 2003. "Do Markets Respond More to More Reliable Labor Market Data? A Test of Market Rationality." *Journal of the European Economic Association* 1 no. 4:931-57.

Kurmann, Andre, Etienne Lale, and Lien Ta. 2025. "Measuring Small Business Dynamics and Employment Withy Private-Sector Real-Time Data." *Journal of Public Economics* 250 (105477).

Loewenstein, Mark, and Matthew Dey. 2017. "A Quarterly Benchmarking Procedure for the Current Employment Statistics Program." *Monthly Labor Review*, November.

Mance, Steven. 2016. "How Hard Is It to Count? Reconciling Monthly Survey and Administrative Data: Experiences from the CES Program." *Working Paper*. At https://www.bls.gov/osmr/research-papers/2016/st160060.htm (accessed December 22, 2025).

Manning, Christopher D., and John R. Stewart. 2017. "Benchmarking the Current Employment Statistics National Estimates." *Monthly Labor Review*, October.

Mueller, Kirk. 2006. "Impact of Business Births and Deaths in the Payroll Survey." *Monthly Labor Review*, no. 5: 28–34.

Mullins, John P. 2016. "One Hundred Years of Current Employment Statistics—an Overview of Survey Advancements." *Monthly Labor Review*, August.

Reinhart, Abiel. 2026. "US Focus: New Birth/Model and Jan Employment." February 13. J.P. Morgan North America Economic Research.

Robertson, Kenneth. 2017. "Benchmarking the Current Employment Statistics Survey: Perspectives on Current Research." *Monthly Labor Review*, November.

Robertson, Kenneth. 2021. "A Better Benchmark Process for the U.S. Current Employment Statistics Survey." *Working Paper*. At https://www.bls.gov/osmr/research-papers/2021/st210070.htm (accessed December 22, 2025).

Stewart, John. 2021. "Birth Death Methodology Adjustments During the Pandemic (Presentation to FESAC)." *Presentation to FESAC*. At https://apps.bea.gov/fesac/meetings/2021-12-10/STEWART-FESAC-Birth-Death.pdf (accessed December 22, 2025).

Wilcox, David. 2025. "BLS Investigation: Challenges? Yes. Rigged Data? No." Peterson Institute for International Economics, Originally Published by *Bloomberg* via Bloomberg Terminal. At https://www.piie.com/blogs/realtime-economics/2025/bls-investigation-challenges-yes-rigged-data-no (accessed December 22, 2025).