Building Labor Market Statistics with Administrative Labor Market Data

Tomaz Cajner, Leland D. Crane, Ryan A. Decker,
Adrian Hamins-Puertolas, Christopher Kurz, Daniel Villar*

September 2026

Abstract

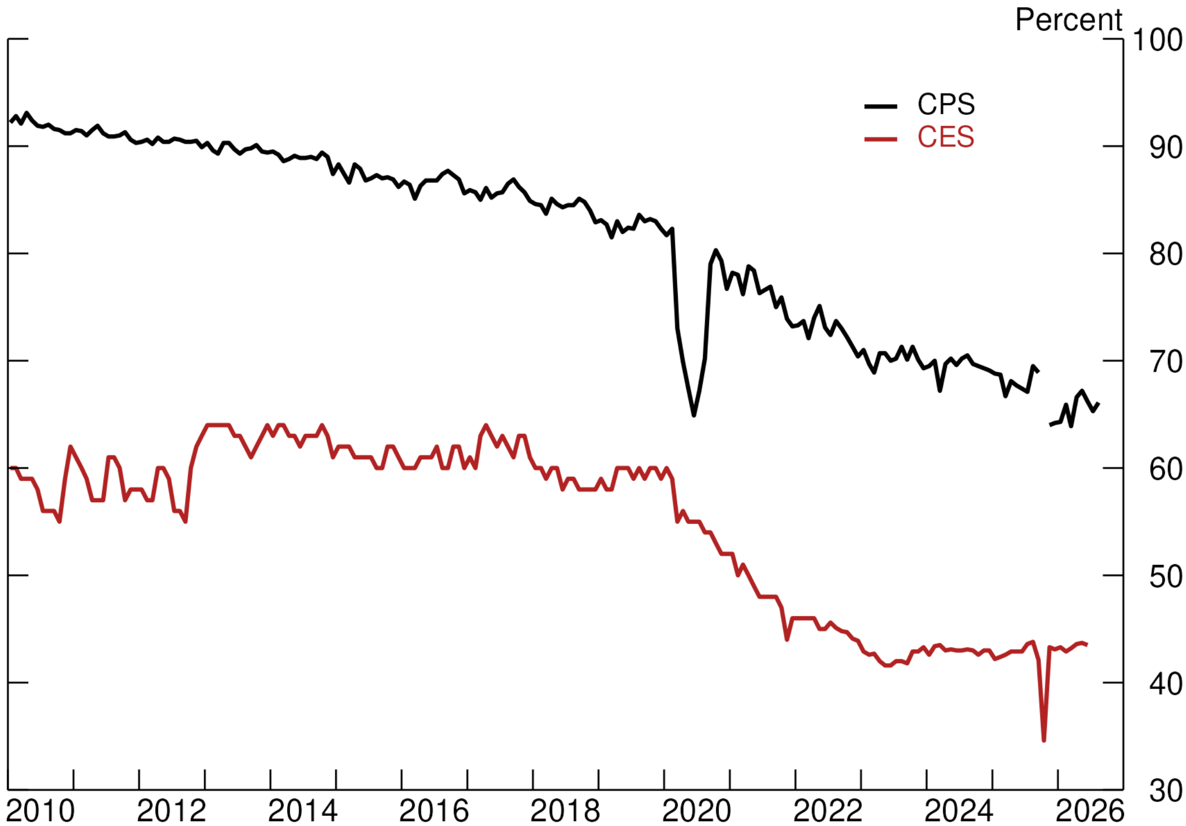
This paper provides a blueprint for building labor market statistics with administrative labor market data based largely on our experience with one large payroll services provider data. We start by reviewing the official labor market statistics, including a discussion of their challenges, most notably falling survey response rates. We then consider tradeoffs associated with both official and private sector labor market data. The main part of the paper provides a discussion of challenges in constructing labor market statistics from payroll provider data. We propose practical solutions for dealing with issues such as accounting for different payroll frequencies, seasonal adjustment, addressing potential sample selection bias, and making adjustments for unobserved business births and death. We focus primarily on constructing employment data but also review wage data. We conclude with a summary of lessons learned.

*The analysis and conclusions set forth here are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors.

1 Introduction

Official labor market statistics in the United States have traditionally been generated through sample-based surveys, such as the Current Employment Statistics (CES) survey or the Current Population Survey (CPS). However, in recent years these statistical products have suffered from declining response rates as shown in Figure 1, which in turn can potentially lead to the selection bias from non-response. Moreover, the CES survey has been subject to relatively large benchmark revisions in recent years.[1]

Figure 1: CES and CPS Response Rate



At the same time, new sources of labor market data have become available in recent years. Most notably, private companies providing payroll processing services have access to data that are well suited to constructing labor market statistics. In this paper, we summarize lessons learned from using one such dataset from a U.S. payroll provider to construct real-time labor market statistics, in particular measures of employment—the primary focus of

[1] See Decker (2026) for discussion; while CES benchmark revision were larger than usual in recent years, in general they tend to be small, typically smaller than 0.5 percent of employment. GDP comprehensive revisions, for example, tend to be much larger yet capture minimal attention; see Figure 11 of Appendix A.

the paper—and wages—which we summarize more briefly.

To build real-time statistics with private administrative labor market data, one needs to deal with several challenges, including accounting for complex payroll payment frequencies, seasonal adjustment, addressing potential sample selection bias, and making adjustments for unobserved business births and deaths. After discussing solutions to deal with each of those problems, we show that indeed one can construct reliable labor market statistics that are almost as accurate as CES. Our results provide a blueprint to build modern labor market statistics in the 21st century. Toward that end, we also discuss possible hypothetical arrangements to produce high-quality statistics with data from multiple providers. Such a project would need to confront many issues around data privacy, information security, and statistical accuracy. Newer technologies—such as secure multi-party computation and autonomous coding agents—may assist here, and we review the related evidence.

Alternative data are not designed for policy-relevant statistical purposes, and unexpected challenges arise in terms of both unexpected methodological shortfalls and unexpected events that change how the data will be used. Throughout the paper, we highlight areas where our thinking evolved as we used the data in real time for informing macroeconomic analysis.

The paper proceeds as follows. Section 2 sets the stage by summarizing the official U.S. labor market data. Section 3 briefly reviews some of the available alternative data from payroll providers. Section 4 goes into detail on our use of the ADP data, showing step by step how to construct payroll employment statistics. Section 5 evaluates the performance of the ADP-FRB employment indicators. Section 6 briefly summarizes parallel work focused on measuring wages with ADP microdata. Section 7 discusses the tradeoffs involved in producing indicators while providing necessary privacy and data security assurances, and ways that technology may help lessen those tradeoffs. Section 8 concludes.

2 Official Labor Market Statistics

In the U.S., the most important source of timely labor market statistics is the employment situation release, published monthly by the BLS.[2] The employment situation release contains information from two separate surveys: CPS, a survey of households, and CES, a survey of business establishments.

[2]Another important timely source is the weekly report on initial and continued claims for unemployment insurance, published by the Department of Labor. However, given that fewer than half of unemployed file for unemployment insurance (see Lachowska, Sorkin and Woodbury, 2025), and given that the share who file varies cyclically, these data provide an incomplete view of the developments in the labor market.

2.1 The Household and Establishment Surveys

CPS—the household survey—is drawn using a stratified probability sample based on the most recent decennial census and subsequent updates to the U.S. housing unit address list. Effectively, about 40,000 employed individuals appear in each monthly survey.[3] CPS employment is relatively noisy at monthly frequency: the estimate of the monthly change in CPS employment has a 90-percent confidence interval of +/− 650,000. Importantly, CPS data do not undergo material *revisions*, as no high-quality "benchmark" source exists for measuring CPS error; the lack of large, noteworthy revisions to CPS is not an advantage of this statistical product.[4]

CES—the establishment survey—is drawn using a stratified probability sample based on the employer records from the unemployment insurance system (see U.S. Bureau of Labor Statistics, 2025; Decker, 2026, for explanations of CES methodology). The survey contains about 120,000 businesses and government agencies each month, representing approximately 630,000 establishments.[5] In terms of employment, the CES sample includes about one-quarter of all nonfarm payroll employees, or one-fifth of private payroll employees, with large variation across industries, ranging from about one-tenth of construction to two-thirds of government. The response rates for CES survey have been declining in recent years and currently hover around 40 percent. In CES, individuals working in an establishment are counted as employed if they receive pay for any part of the pay period including the 12th day of the month. Unlike the CPS, the CES counts *jobs* rather than *employees*. Since newly created or recently closed establishments are prohibitively difficult to survey, the CES sample is supplemented with estimates of the net job creation of establishment births and deaths from the CES Net Birth-Death Model (NBD).[6] While it is tempting to think of the CES as being primarily a survey supplemented with a minor birth/death adjustment, in many recent years the NBD has accounted for half or more of reported payroll gains (Decker, 2026). The CES provides detailed information payrolls, hours worked, and earnings for

[3]The survey contains roughly 60,000 eligible housing units each month, with the response rates declining and currently hovering around 65 percent. The CPS uses a 4-8-4 rotation pattern, where a sampled household is interviewed for four consecutive months, then leaves the sample for eight months, and then returns for four additional months before permanently exiting. The CPS employment concept counts individuals as employed if they worked for pay during the reference week (typically the week including the 12th day of a month), and also if they were temporarily absent from a job. Importantly, the CPS measures the number of *employed persons*, not *jobs*. In addition to employment, the CPS also provides detailed information on individual labor market status, hours worked, and earnings, including a monthly measure of usual weekly earnings for wage and salary workers.

[4]CPS does receive updated "population controls" that are not applied retroactively but can be applied following Coglianese, Murray and Nekarda (2025); this allows for updating CPS statistics with more accurate sampling weights, though it does not address error in cell-level statistics.

[5]The universe of unemployment insurance tax accounts consists of roughly 12.1 million establishments.

[6]See `https://www.bls.gov/web/empsit/cesbd.htm`.

payroll employees in the nonfarm sector. Due to a much bigger sample, CES employment estimates are substantially more precise than CPS estimates; the estimate of the monthly change in CES employment has a 90-percent confidence interval of +/- 122,000.

In addition to near-real-time monthly estimates from CPS and CES, the BLS also publishes employment and compensation data in their Quarterly Census of Employment and Wages (QCEW), which is based on a near-census of employment based on unemployment insurance tax records, covering about 95 percent of U.S. jobs. These data typically have a lag of about 5-6 months and are used to benchmark CES data annually (each March). The "benchmark" revisions to CES improve the accuracy of the statistic over time, and the availability of a high-quality, relatively timely benchmark data source is a strong advantage of the CES.[7]

2.2 Tradeoffs in Official Data

Economists rely on official monthly labor market statistics for several good reasons. First, CPS and CES are based on scientifically designed random probability samples, aiming to achieve appropriate representativeness of the population they intend to measure. Second, CPS and especially CES have relatively large samples; in a handful of states, responding to CES survey is even mandatory. Third, CPS and CES statistics are constructed following established statistical methodologies and procedures that are broadly consistent over time and are clearly described in associated documentation. Fourth, CPS and CES data have long histories which renders them suitable for time-series analysis. Fifth, each year both CES and CPS are updated to ensure they continue to well represent the underlying population: CES is benchmarked to data covering the near-universe of payroll employment, and CPS receives updated "population controls."

However, official labor market statistics also face some limitations and tradeoffs. Most importantly, official labor market statistics are subject to measurement error due to factors such as sampling error, nonresponse, reporting errors, and imperfect seasonal adjustment. With declining response rates, the benefits of random probability samples are receding, and measurement issues can arise due to the selection bias if non-respondents differs from respondents in a non-random way.[8] While some of the measurement issues could potentially be improved with larger initiated samples, increasing the sample size is very costly. Additionally, in the case of CES measuring the employment contribution from business births

[7] Since the QCEW does not cover 100 percent of U.S. jobs, the BLS supplements QCEW with other data and estimates when benchmarking CES.

[8] Relatedly, Davis et al. (2010) provide evidence that nonresponse rates to Job Openings and Labor Turnover Survey (JOLTS) lead to a systematic bias in JOLTS-based estimates of worker flows and job openings.

and deaths has been a longstanding challenge, and imperfect measurement from this component often contributes notably to the benchmark revision. In the case of CPS, the BLS has to apply assumptions for the population growth, and those assumptions can be imprecise, especially when large changes in net immigration occur, as has been the case in recent years.

3 Payroll Employment Data

A viable complement or, potentially, alternative to measuring employment with surveys is to instead rely on data from payroll service providers.

3.1 Potential Value from Alternative Data

Alternative data sources, such as private payroll services provider data, may offer advantages relative to official data. Large private sector samples may complement official data especially when the latter are subject to low response rates and other sources of bias, potentially improving the accuracy of inference about underlying employment trends. Private sector data may also provide information at higher frequency (e.g., weekly instead of monthly), in a timelier manner (e.g., shortly after a reference week), which can be valuable in times of rapid economic developments; we provide an example of this below. Finally, private sector data may be available at narrower levels of industry or geographic detail than official data, which can provide insights on industry dynamics or local economic developments.

It is helpful if an alternative source can predict high-frequency (e.g., monthly) fluctuations in official data. Similarly, there may be value in predicting high-frequency revisions to official data reports (e.g., revisions to CES from the first to the third report for a given month, the focus of Feroli et al., 2026). But predicting the monthly patterns of noisy official data can be difficult, and the value of such predictions may be limited by the quality of the official data themselves. Both the CES and any alternative data source are subject to error that can play a large role in monthly fluctuations. Moreover, monthly fluctuations in CES data forever reflect sampling error, birth-death model estimates, and seasonal adjustment error, as the non-March monthly data are never benchmarked to a source of monthly "true" employment; rather, the March employment level is eventually benchmarked while monthly data are adjusted in a manner that leaves original sources of error largely intact.

Importantly, given the availability of the official employment benchmarking source—the QCEW—alternative payroll data can adopt one of the key strengths of the CES by undergoing periodic benchmarking. Indeed, the availability of a particularly high-quality benchmark source for U.S. payroll data grants data producers and users some flexibility in how best to track monthly employment patterns; equivalently, the benchmark data are a "gold standard" asset of the statistical agencies and *enable* the use of alternative data. Given the permanence of the noise inherent in monthly CES data discussed above, one of the most important criteria by which an alternative source can be judged—particularly related to the objective of improving inference about true employment trends—is whether the alternative source can help predict benchmark revisions to the CES.

3.2 Major Payroll Data Sources

Below we provide an incomplete list of major providers of payroll or similar services in the U.S., which are already using their data to construct employment statistics and/or have had their data used for the purposed of academic research. We compile the information about those payroll providers from publicly available 10-Ks and other public reports, company websites, and web searches.

1. **Automatic Data Processing, Inc. (ADP)** is a payroll and human resources services provider. As of June 2025, ADP served 26 million workers and was responsible for the payrolls for 1 out of every 6 workers in the U.S. (1 out of every 5 workers in the private sector), making them the largest payroll provider in the country. The company publishes a regular payroll employment statistic, the National Employment Report: `https://adpemploymentreport.com/`. Note that while ADP-FRB uses essentially the same raw data as the NER, the methodologies are distinct and the final results generally differ.

2. **Paychex** was founded as a payroll service provider but now includes human resource and benefits outsourcing. As of May 2025, Paychex had approximately 800,000 clients across the U.S. and parts of Europe; some of these clients may be non-payroll clients. The company claims to pay 1 out of every 11 workers in the US. Paychex publicly releases its Small Business Employment Watch, which draws on data from approximately 350,000 clients to estimate employment and wages among small businesses each month: `https://www.paychex.com/employment-watch/`. Paychex data were used in Chetty et al. (2024).

3. **UKG** is a multinational technology company and the result of a merger between Ultimate Software and Kronos Incorporated. As of August 2025, UKG serves about 6.2 million employees in the United States, or roughly 1 out of 20 private employees. UKG was releasing a monthly measure of employment in its Workplace Activity Report, but

the updates to that report stopped in August 2025: https://www.ukg.com/learn/insights/workforce-activity-report/.

4. **Gusto** provides payroll and human resource services for about 400,000 small businesses and publishes a monthly small business employment indicator that is focused on firms with fewer than 50 employees. It is unclear how many employees are covered by Gusto's 400,000 clients. We can construct a rough guess by assuming that all Gusto firms have fewer than 100 employees and, subject to that ceiling, a size distribution similar to the U.S. private sector, which would suggest an average firm size of 8 employees for a total headcount of around 3 million. Gusto data have been used for research projects (e.g., Akan et al., 2025), and the firm provides monthly estimates of net hires constructed from their data: https://gusto.com/resources/gusto-insights/labor-market-trends/.

5. **Homebase** provides time clock and payroll services for small businesses (largely in customer-facing industries) and publishes occasional reports based on their data. Homebase data cover about 100,000 businesses with about 2 million hourly employees. Homebase data have been used in many academic papers (e.g., Bartik et al., 2020; Kurmann, Lalé and Ta, 2025) and Homebase provides monthly labor market statistics constructed from their data: https://www.joinhomebase.com/data/.

6. **Intuit** offers a number of payroll products both online and through its QuickBooks application, targeted toward small businesses. Intuit's payroll services serve at least 1 million employees from more than 250,000 small businesses. The company publishes the monthly Small Business Index, which is regularly benchmarked to the BLS Business Employment Dynamics: https://quickbooks.intuit.com/r/small-business-data/index/. Intuit data were used in Chetty et al. (2024).

In addition to payroll providers, employment measures are also constructed by Revelio (https://www.reveliolabs.com/public-labor-statistics/employment) and the American Staffing Association (https://americanstaffing.net/research/asa-data-dashboard/asa-staffing-index/).

4 ADP-FRB Employment Statistics

In this section we discuss the challenges in constructing labor market statistics from payroll provider data, based on our decade-long experience of working with these data. After reviewing the data and the pipeline at a high level, we discuss each major challenge in detail. We propose practical solutions for dealing with issues such as accounting for different payroll frequencies, seasonal adjustment, addressing potential sample selection bias, and making adjustments for unobserved business births and death. The discussion here builds on our existing research, including Cajner et al. (2018, 2020*a*,*b*, 2022, 2023). Notably, some of our methodological choices changed over time as we learned more about the data and as macroeconomic events unfolded.

4.1 Description of ADP Microdata

ADP (Automatic Data Processing, Inc.) is one of the largest payroll providers in the U.S., processing payrolls for approximately 26 million U.S. workers, which is roughly 20 percent of total U.S. private sector employment. The ADP microdata are organized at the level of Payroll Account Controls (PACs), each of which can correspond to an establishment, a firm, or a part of a firm; we typically treated PACs as establishments in the absence of better options.[9] Each PAC has a record for the end of every pay period containing the date payroll was processed, employment information for the pay period, and time-invariant characteristics such as anonymized identifiers, NAICS industry codes, and geographic location. A distinctive feature of the ADP data is that they include two employment concepts: "active employees", which encompasses all individuals in the payroll system regardless of whether they received pay during a given period, and "paid employees," which includes only those issued paychecks during the pay period. The data are available at weekly frequency since July 2009 (and at a monthly frequency since July 1999), providing a substantially higher-frequency window into labor market dynamics than traditional monthly surveys. It is important to note that identifying information like business name and address are not part of the files.

The ADP data offer several advantages for measuring aggregate labor market activity. First, these data are very timely, with information for a pay period typically available within a week of the pay period's end, enabling near-real-time analysis of employment trends. Second, because ADP captures payroll information throughout the entire month rather than just the reference week containing the 12th (as in the BLS CES and CPS data), the data can better capture month-long employment dynamics and are less susceptible to temporary distortions during specific weeks. Third, the sample is broadly representative of the US economy across industries, firm sizes, and geography, though with modest overrepresentation in manufacturing and large businesses; these compositional differences can be addressed through reweighting. However, the data also present challenges: selection into ADP's client base may introduce biases, true establishment births and deaths cannot be directly observed (as

[9] In U.S. statistical agency parlance, an "establishment" is a single operating location of a business, while a "firm" is a collection of one or more establishments under common ownership or operational control.

changes in ADP's client roster may reflect contract decisions rather than genuine economic entry and exit), and the panel structure requires careful handling of pay frequency heterogeneity across businesses. Aside from our research developing ADP-FRB (cited above), ADP data have been used in many other research projects (see, for examples, Grigsby, Hurst and Yildirmaz, 2021; Autor et al., 2022*a*,*b*; DeFusco, Enriquez and Yellen, 2024; Nezaj, Richardson and Wang, 2025; Quach, 2025, 2026; Brynjolfsson, Chandar and Chen, 2026; Hurst et al., 2026).

4.2 Brief Overview of ADP-FRB Methodology

Processing the raw data into useful aggregate indicators is a multi-stage process: calculating weekly employment growth among continuing payroll units in the data, screening for outliers, applying sampling weights from the QCEW, aggregating to the broad index level, seasonally adjusting, and, finally, benchmarking. We explain each of these steps below.

4.3 Payroll Employment Concept: Active vs. Paid

As noted above, the ADP payroll microdata provide two distinct measures of employment that capture different aspects of the employment relationship. *Active* employment includes all individuals who maintain an active record in a business's payroll system during a given pay period, regardless of whether they received compensation. This encompasses not only workers who were paid but also those on unpaid leave, workers with zero hours, temporarily furloughed employees who remain attached to their employer; the active count may also sometimes include separated workers who have not be removed from the payroll system. *Paid* employment, by contrast, counts only those individuals who actually received a paycheck during the pay period, including regular wages and salaries as well as bonus payments and payroll corrections.

Conceptually, active employment aligns more closely with the household survey (CPS) notion of employment, measuring labor market attachment rather than simply paycheck receipt, while paid employment corresponds more directly to the establishment survey (CES) concept. In typical economic conditions, we prefer active employment for our headline ADP-FRB series for several reasons: it is substantially less volatile than paid employment (which fluctuates due to the timing of bonuses, payroll corrections, and other irregular payments), its seasonal patterns more closely mirror those in the QCEW benchmark data, and it has historically demonstrated superior performance in forecasting official BLS employment statistics (Cajner et al., 2018).

The COVID-19 pandemic, however, revealed a critical limitation of active employment and necessitated a temporary change in our methodology. During the initial months of the pandemic, CPS temporary layoffs—in which workers were separated from payroll but employers intended to recall them—soared to nearly 80 percent of all job losses. Because many employers kept these temporarily furloughed workers in their payroll systems (listed as active but unpaid), the active employment measure failed to capture the full magnitude of the labor market collapse, which became apparent very early in the pandemic. We therefore pivoted to emphasizing paid employment from March 2020 through September 2021, as it better reflected the reality that millions of workers had lost their paychecks even if they technically remained attached to their employers. By mid-2021, as the share of temporary layoffs returned to its pre-pandemic level of approximately 14 percent and the gaps between paid and active employment losses converged, we reverted to active employment as our primary measure. Even while we focused on the paid measure, the active measure continued to provide useful insight about the ongoing attachments between firms and (and the difference between active and paid employment served as a good indicator of workers on temporary layoff).

This experience demonstrates that the optimal employment concept depends on the specific labor market dynamics at play: active employment better captures permanent employment relationships and changes in labor demand during normal times, whereas paid employment provides a more accurate real-time picture during periods dominated by temporary separations and widespread furloughs. While we do not expect such events to be common in the future, the episode demonstrates the broader need to be flexible in the face of a dynamic economy—and the value of having multiple kinds of payroll measures.

4.4 Weekly Employment Growth

Roughly speaking, the ADP research microdata are weekly snapshots of the production system, each recording the most recent payroll event for each business. Appending the weekly snapshots allows one to build the history of payroll events at a firm. Payroll units (PACs) process payrolls on different days of the week, making the raw data fundamentally *daily*. To simplify we first estimate payroll employment as of each Saturday, putting everything on a weekly basis with a common day of the week. This is without loss of generality since there are essentially no PACs paying workers at sub-weekly frequencies.

The raw data only show the date that payroll was processed for a PAC and a pay frequency code (e.g. weekly, biweekly, semimonthly, etc.); though the pay frequency code can be unreliable. After studying various methods for interpolating between pay events, we chose a stepwise interpolation approach with the following assumption: if a PAC processes payrolls on date t and t', then $e(t')$, the reported employment for t', applies to the interval $[t+1, t']$. That is, we treat the PAC's employment as jumping discretely to the new value just after the old payroll event. In this way we can build a daily employment series from a sequence of payrolls and then read off the Saturday values to get our weekly series. This approach exactly matches true employment if—as human resource practices suggest—most hires and separations are timed to coincide with pay period starts and ends, and it aligns conceptually with the CES methodology of measuring total employment for the pay period including a reference date (U.S. Bureau of Labor Statistics, 2026*b*).

A natural alternative would be linear interpolation, assuming employment evolves smoothly along a linear path between consecutive pay dates, which has the advantage of producing less volatile series but rests on the less realistic assumption of continuous within-period adjustment at the PAC level. Importantly, neither stepwise nor the linear interpolation can produce fully real-time estimates: stepwise interpolation requires waiting until the end of a pay period to observe a PAC's employment level, while linear interpolation requires data from both the preceding and the following pay dates. This inherent lag means that measured employment changes may be shifted by one to two weeks relative to when actual hirings and separations occurred, though in practice this timing ambiguity matters less for understanding month-to-month trends than for pinpointing the exact week of labor market turning points.

Using the stepwise interpolation rule, we can construct a weekly series of employment levels $e_{j,t}, e_{j,t+1}, \ldots$ for each PAC j, where $t, t+1, \ldots$ are successive Saturdays. Note that if a PAC did not process payroll between t and $t+1$ then $e_{j,t} = e_{j,t+1}$; this will be important for understanding the real-time data flow later.

4.5 Outlier Detection

Constructing aggregate employment indexes from establishment-level microdata requires careful treatment of outliers, as extreme growth rates at individual establishments can arise from both genuine economic events and data quality issues. Filtering data quality issues is difficult since PAC employment is volatile, with a great deal of inaction as well as fat tails. Particularly large swings in PAC-level employment may occur due to processing errors, payroll corrections for past mistakes (such as paying workers for incorrect hours), or the timing of bonus payments to subsets of workers; none of these reflect true changes in employment. These spurious events are more common in the paid employment series, and less common in the active series.

We use a simple symmetric rule-of-thumb filter based on the "DHS" (Davis, Haltiwanger and Schuh, 1996) growth rate measure, which treats positive and negative changes symmetrically. Specifically, we drop observations where weekly DHS growth rates exceed 1.8 or fall below -1.8, thresholds that correspond roughly to a 100-worker establishment losing 95 employees or a 5-worker establishment gaining 95 employees in a single pay period (such large changes more likely reflect data anomalies than actual employment dynamics).

Even with this filter, the paid employment series is more volatile and generally seems to contain less signal than the smoother active employment series. In addition, we sometimes found it necessary to engage in clerical edits, manually removing outliers that do not exceed the cutoff threshold but have a large weight and a material effect on aggregated indexes.[10] We expect future indicators will need to rely on a mix of automated decision rules and analyst judgement, though automation is obviously preferable wherever possible.

4.6 Weighting and Representativeness

We encountered several fundamental weighting design decisions: which dimensions to include (industry, size, geography, etc.), what level of granularity to use within each dimension, and whether to treat ADP payroll units (PACs) as firms or establishments when mapping to official statistics. Our baseline approach employs a two-dimensional weighting scheme using two-digit NAICS industry codes crossed with six establishment size classes (1-19, 20-49, 50-99, 100-299, 300-499, and 500+ employees). For each cell, we calculate the weight as the ratio of target population employment in that cell. In our baseline approach we treat PACs as establishments and use QCEW for sampling weights, but we also experimented with treating PACs as firms and using Statistics of U.S. Businesses (SUSB) data as our alternative benchmark (the choice between these approaches affects the estimated magnitude of employment changes, especially for small businesses, but produces similar overall trends; see Cajner et al., 2020*b*).

An important decision in our weighting methodology was to exclude geography as a weighting dimension, despite some regional variation in ADP coverage. Including state or regional cells dramatically increase the number of weighting cells, in turn creating severe small cell size problems that would undermine the stability and reliability of our estimates. Weighting reduces the bias of estimators but—by upweighting sparse cells—increases the variance. This is the well-known bias-variance tradeoff for weighting estimators (see Kalton and Flores-Cervantes, 2003). A single large employment change at one of the few ADP

[10] An alternative approach to outliers—particularly among very large businesses—would be to remove the large offending business from the weighting scheme, downweight the sampling weights applied to the main data by the full size of the offending business, then add the offending business to the final aggregated payroll number. We considered implementing this approach, but it presented a range of practical difficulties for automated data production processes.

businesses in a sparsely populated cell could generate wildly implausible swings in the aggregate index. We therefore prioritize the industry-by-size structure, which captures the most economically meaningful heterogeneity in employment dynamics while maintaining sufficient cell sizes to ensure stable estimates. This pragmatic choice sacrifices some geographic precision and leads to some amount of bias, but produces a more robust national employment measure. For the same reason, we use only six size bins for weighting when QCEW data permit up to nine, and we use 2-digit NAICS codes instead of more detailed arrangements. Importantly, our industry-by-size reweighting successfully eliminates most of the bias that would otherwise arise from ADP's compositional differences relative to the U.S. economy.

In short, when deciding the weighting scheme, the research must make decisions that balance representativeness and precision; in our judgment and with the ADP data, our specific industry-by-size reweighting (with no geographic consideration) best addresses this tradeoff. Other datasets may require even coarser weighting schemes.

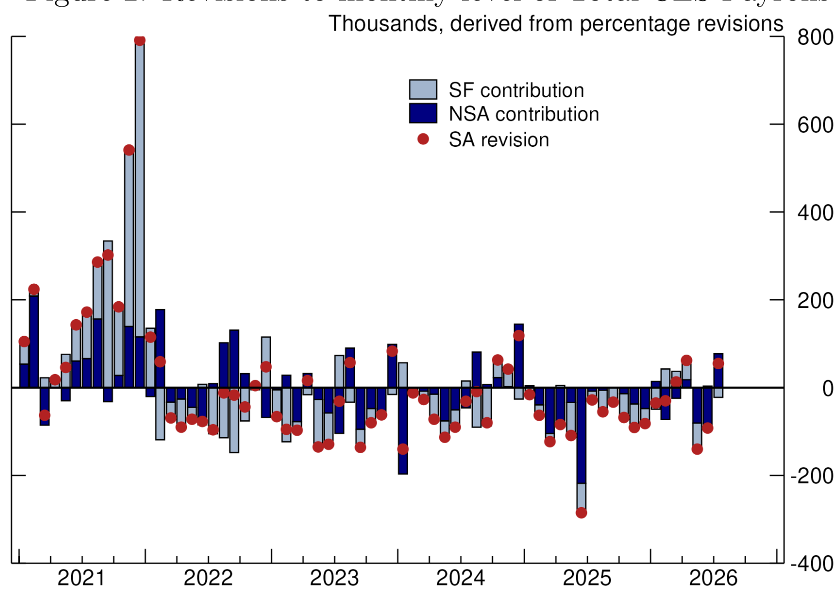
Even ignoring geography, weighting is not trivial. Approximately 10-20 percent of ADP PACs have missing or unreliable NAICS industry codes. This occurs because industry codes are not a critical part of ADP's business operations—a common problem in private sector databases—so collection and assignment of NAICS codes is naturally not the highest priority for the data provider. For these cases we assume that industry is missing-at-random conditional on establishment size. However, this is not generally ideal, as the businesses that are missing NAICS codes tend to be smaller, younger, and faster-growing than the others. The optimal approach would likely involve more data collection to rapidly assign NAICS codes to new firms and check on the characteristics of those left unassigned. Importantly, accurate classification of establishments by industry is not easy; evidence points to sometimes substantial disagreements between successive waves of the same surveys (Kaminski, 2015), between statistical agencies (Fairman et al., 2008), and between agencies and private databases (Barnatchez, Crane and Decker, 2017; Crane and Decker, 2019). Industry codes in private databases will never exactly match those of the statistical agencies. The goal, though, should be to get reasonably close and minimize downstream impact on labor market indicators. In any case, concern about industry code accuracy is another reason to consider weighting at higher levels of industry aggregation, rather than relying heavily on, say, 6-digit NAICS detail.

Mechanically, for each industry-size cell we calculate the employment-weighted average growth rate of the businesses in that cell. Those cell-level growth rates are then aggregated using the QCEW-based industry-size weights, along with the contribution of the missing NAICS group.

4.7 Seasonal Adjustment

Monthly seasonal adjustment is a widely studied problem, with highly developed software packages available (e.g. X-13, TRAMO/SEATS). In the CES, as with most other adjusted data series, the seasonal adjustment procedure eliminates the part of the change in the data series attributable to normal seasonal variation.[11] Importantly, the BLS employs *concurrent* seasonal adjustment, which uses all available data, including those for the current month, in developing sample-based seasonal factors. The concurrent seasonal factor estimates are in part employed to reduce first-to-final revisions in the employment series.[12] To be sure, and as seen in Figure 2, revisions to seasonal factors lead to revisions to CES payrolls, contributing on average a substantial 54 percent of the total revision to the monthly level; the role of seasonal adjustment in first-to-third print revisions has increased in recent years (Decker, 2026).

Figure 2: Revisions to monthly level of Total CES Payrolls



Note: The BLS annual revision interferes with the interpretation of the level revisions for November and December. They cumulate the 1st revision to the 2-month change and the 2nd revision to the 1-month change.
Source: BLS.

Weekly seasonal adjustment of economic time series has received far less attention than monthly adjustment, and the methodological choices of implementation may lead to large variation in any high-frequency employment indicators. The weekly seasonal adjustment problem is in some ways intrinsically more difficult than monthly seasonal adjustment. To

[11] See https://www.bls.gov/web/empsit/cesseasadj.htm.
[12] See https://www.bls.gov/ces/methods/concurrent-seasonal-adjustment-guide.htm.

see this, note that monthly data can be crudely seasonally adjusted with a regression on 12 month-specific dummies. This approach is less productive for weekly data since the shorter interval leads to more noise, making the estimated dummy coefficients less precise. More fundamentally, weekly observations are not exactly periodic: the date of the first Saturday of the year advances by one day each (non-leap) year, and can fall anywhere from January 1st through January 6th. Thus, the day-of-year of n'th Saturday of the year can slip by about a week; this non-periodicity complicates standard seasonal adjustment approaches.

Additionally, weekly data exhibit complex seasonal patterns driven by both slow-moving annual cycles (such as summer hiring in leisure and hospitality) and sharp discrete events tied to specific holidays. Holiday effects are particularly intricate: employment patterns shift not only during holiday weeks themselves (Thanksgiving, Memorial Day, Labor Day, New Year's, Christmas, and July 4th) but also in the weeks immediately preceding major holidays as businesses adjust staffing levels. These issues are also present in monthly data but are usually much attenuated by time aggregation.

A final consideration that is particularly salient for many private sector data sources is that the higher the frequency of the seasonal adjustment, the more time series history is desirable, for standard degrees-of-freedom reasons; many private sector employment data sources have relatively short history.

To address (some of) these challenges, the ADP-FRB weekly series employs a two-component approach based on Cleveland and Scott (2007), the same methodology the BLS has used in the past to seasonally adjust weekly unemployment insurance claims. Up until 2022, we estimated weekly and monthly seasonal factors once a year at the time of the March CES benchmark revision—that is, we used fixed rather than concurrent seasonal adjustment.[13] The pandemic introduced further complications, as the extreme employment swings in 2020-2021 tended to distort estimated seasonal factors, and some underlying seasonal patterns likely changed after the pandemic.[14] As a result, we switched to concurrent seasonal adjustment of the ADP-FRB weekly series, moving the approach closer to the BLS methodology. In the post-pandemic period, the differences between concurrent and yearly seasonal factors were substantial. For example, the April 2024 ADP-FRB estimate of active employment ranges from from -41,000 to 59,000 depending on the use of non-concurrent to concurrent seasonal factors.

The Cleveland and Scott (2007) procedure combines locally weighted regressions applied

[13]This approach mirrored that taken in the estimation of industrial production. See https://www.federalreserve.gov/releases/g17/Meth/MethIP.htm.

[14]Not surprisingly, large swings in employment distort CES seasonal factors in the subsequent years. Wright (2013) describes the issue in the aftermath of the Great Recession and Hudson, Mercurio and Kropf (2022) presents evidence on the challenges of seasonally adjusting CES during and after the pandemic.

to trigonometric functions—which flexibly capture slow-moving, smoothly evolving seasonal patterns—with fixed coefficient regressions that explicitly control for holiday effects. This framework allows seasonal factors to evolve gradually over time while ensuring that discrete holiday-related patterns are consistently identified. Of particular note, we found that the weekly dynamics around the Thanksgiving holiday were particularly important. The holiday itself, combined with the ramp up in holiday season employment in certain sectors and the change in timing from year to year all contribute to produce large, sudden, unpredictable swings in employment in late November and early December. This swing and associated problems are mirrored in early January when the holiday season ends. These dynamics are of course present in monthly data but are smoothed and attenuated by time aggregation.

For the narrow COVID-19 period, we adopted a special treatment: data between late February 2020 and late September 2020 were excluded from seasonal factor estimation; that is, we assumed that data patterns during these months did not provide useful information about changes in seasonal patterns. Further, the input series were rescaled for the post-intervention period to line up with the pre-intervention period, to avoid having the seasonal adjustment algorithm interpret the level shift as a seasonal event. This roughly aligns with the way BLS treats the period in the CES data, excluding spring and often summer. Treating this period as a series of outliers prevents the pandemic-driven swings from distorting the underlying seasonal patterns, at the cost of not allowing the possibility of any genuine persistent shifts in seasonality that may have occurred during that stage of the pandemic. When we compute monthly-average ADP-FRB series from the NSA data we seasonally adjust using the Census Bureau's X-13 ARIMA-SEATS program.

A separate seasonal adjustment decision is: at what level of aggregation should we estimate and apply seasonal factors? The CES seasonally adjusts mostly at the 3-digit NAICS industry level, and its aggregate seasonally adjusted series simply add up industry-level series. This is a complicated approach with the potential to be very labor intensive, given the peculiarities of seasonality across industries. We opted instead to seasonally adjust (a) at the aggregate level, and (b) separately, at the broad sector level. This allows us to have a simple aggregate approach for our main series but to also pay attention to broad sectors, as we found sector-level ADP-FRB indexes to be useful in many environments. The downside of this approach is that our sector-level indexes do not sum neatly to match our aggregate indexes, though the differences were typically minor.

4.8 Benchmarking and Bias Correction

At this stage our data series are still (weekly) indexes. To calculate the level and change in the number of jobs month over month, one could put the indexes on this basis by a simple normalization: fixing the index value to a true job count in a single month and then growing the series forward and backward using the index growth rates. This adjustment is useful, but it does not address fundamental challenges which can cause our estimates of employment growth to differ from truth. First, the continuers-only growth rates do not include the contribution of business birth and death. Second, since the the continuers are actually just continuers *remaining in the ADP client base* (implying that non-continuers are not necessarily true establishment births or deaths), there may be selection bias, as the firms choosing ADP and retaining their contracts over time may differ from the business population in terms of growth, age, size, or other characteristics. Third, sampling error and bias from the weighting may also contribute to error.

Retrospectively, we correct for these sources of error with a benchmark process similar to the BLS CES. We set the series equal to CES employment levels each March (which, in turn, have been benchmarked to, for the most part, QCEW) then linearly wedge back the adjustment to the previous March benchmark, ensuring that year-over-year changes align with the comprehensive QCEW counts while preserving the higher-frequency variation captured in the ADP data.[15] The benchmarking process can only happen in August at the earliest (if we use the preliminary CES benchmark as a target) and even later in the following January if we wait for the final benchmark. Sometimes—especially when the preliminary CES benchmark implied a substantial ADP-FRB revision—we have implemented a preliminary ADP-FRB benchmark in August.

The benchmarking process ensures close adherence of ADP-FRB to "true" payroll gains, at least on a year-over-year basis, up to the most recently benchmarked March level. To adjust the data in the "post-benchmark period"—that is, the observations postdating the March benchmark and extending through the current data—we construct a "forward benchmark" correction.

Our forward benchmark is a very crude analogy to the CES birth-death model but is more similar to the kind of "bias adjustment" CES used prior to the introduction of that model—an approach still used by other business surveys such as the Census Bureau's Monthly Retail Trade Survey and Quarterly Services Survey. The forward benchmark is intended to adjust the real-time data not only for birth and death but also for any other persistent source of measurement error in the ADP-FRB index. Specifically, we use lagged annual benchmark

[15]We benchmark to the CES rather than directly to the QCEW so that we do not have to handle QCEW scope corrections and other issues ourselves.

revisions to identify systematic discrepancies between unbenchmarked ADP-FRB estimates and "true" employment (the latter is effectively determined by QCEW, that is, benchmarked CES). We then estimate a historical regression of these ADP-FRB benchmark revisions on their lagged values, using the estimated relationship to project a constant correction factor for periods beyond the most recent benchmark. For example, following each benchmark, we add a constant adjustment to each subsequent week based on the prediction from our regression model of how much ADP-FRB is likely to diverge from "true" employment. Between 2005 and 2020, this approach reduced our root-mean-squared benchmark miss from 625,000 jobs to 427,000 jobs, demonstrating a substantial improvement in accuracy. The success of this relatively simple correction suggests that the combined effects of unmeasured birth-death dynamics and sample selection biases are reasonably stable and at least to a certain extent predictable over time. This method worked well in most years, though on occasion we made judgmental adjustments to the straight read of the regression model.

Here it is worth stepping back for a broader perspective on the problem we attempt to solve with the forward benchmark, which is similar to the problem the CES confronts with the birth-death model and other business surveys address with some form of bias adjustment. It is difficult or impossible to capture business birth and death in data sources that focus on measuring business activity by observing businesses, either through surveys or through client databases like a payroll database. Data producers must address this issue through some kind of modeling. The CES birth-death model is an especially well-researched, extensive, and well-documented approach to the problem; bias adjustment approaches like our forward benchmark are more simplistic. The only way to avoid the broader problem is to measure business activity in some other way, such as by tracking customers or households that staff such businesses. Notably, both bias adjustment and the more complicated and targeted birth-death model implicitly play the additional role of potentially compensating for other sources of measurement error, though the way the birth-death model is estimated and evaluated does focus on birth and death specifically.

4.9 Real Time Adjustments

The nature of payroll data and the heterogeneity of pay frequencies creates both lags in data arrival and predictable biases in the real time data. To make things concrete, let $\{..., t-1, t, t+1, ...\}$ be a series of Saturdays. The raw data arrive on Saturday, t, and contain the payroll events processed in the interval $[t-1, t)$ As noted in Section 4.4, for each PAC we assign the employment level to the preceding Saturday, $t-1$. For weekly payers this results in complete data, but for biweekly, semimonthly or monthly payers there will be additional missing employment values. We assume employment is constant throughout the pay period and fill in all the missings back to the latest non-missing Saturday.

Thus, as of t we can compute an aggregate growth rate between $t-2$ and $t-1$, but it will be biased. To see this, note that for the biweekly payers that pay in $[t-1, t)$, growth will be zero between $t-2$ and $t-1$. That is, $t-2$ and $t-1$ fall *within the same pay period* and so have the same employment level by assumption. The other half of biweekly payers do not enter the calculation since they last paid in $[t-2, t-1)$, and we won't observe their t value until payrolls for $[t, t+1)$ are available. At that point the estimate of period $t-1$ growth will revise. While weekly payers will provide useful information every week, the zero growth of biweekly payers and businesses using other pay frequencies will bias estimated real time growth towards zero in the most recent weeks of data. The diversity of pay periods makes this bias significant: only about one-fifth of businesses issue paychecks weekly, while almost one-half issue biweekly, one-fifth semimonthly, and one-tenth monthly (Cajner et al., 2020*b*).

The ADP-FRB methodology implements a real-time adjustment procedure that uses historical relationships between preliminary "first print" estimates and final values (once all businesses have reported) to forecast likely revisions to preliminary estimates. Specifically, we track the historical relationship between the first print of the weekly growth rate and the weekly growth rate as estimated with seven more weeks of data (which is nearly identical to the fully-revised value). In a simple model, a bivariate linear regression would give the correct adjustment assuming that weekly payers are noisy, unbiased proxies for the other pay groups. In practice, we find a substantial role for seasonality and holidays (such as Thanksgiving) in the adjustment. We found this adjustment to be fairly reliable, but it performed less reliably during the acute phases of the COVID-19 pandemic when employment changes were unprecedented in speed and magnitude.[16]

4.10 Monthly Aggregation

The weekly data are of interest by themselves, but they can be noisy and are not directly comparable to the (monthly) CES. Recall that the CES reports employment for the pay period containing the 12th of each month, and the CES job gain numbers are the difference between two of these point-in-time estimates. Over the years we have used two different approaches to produce monthly numbers. Initially, we focused on monthly values that averaged all the weeks in the month (Cajner et al., 2018). This approach uses more data, provides a fuller picture of the economic conditions during the month, and—because of averaging—

[16] We exclude growth rate revisions during the COVID-19 episode from real-time adjustment model estimation.

should produce less noisy estimates than point-in-time comparisons. In addition, we found better forecasting results when using the monthly averages. However, this approach has significant drawbacks. First, it is not completely comparable to CES, since it does not use the 12th day reference period concept; it is difficult, in practice, to track alternative indicators that follow fundamentally different timing conventions than the official data that serve as the primary information followed by forecasters and policymakers. Second, the right approach to seasonal adjustment of this full-month object is not clear; the most efficient approach is to time average the NSA weekly data and use X13 monthly seasonal adjustment on the resulting series, but this approach means the weekly data are not consistent with the monthly since they have been seasonally adjusted differently.

Most concerning, in real time we do not have data for all the weeks in the current month, so monthly aggregation requires requires extrapolating the weekly series "in hand" out to the end of the month. As we progress through a month in real time, the monthly average for that month revises as new weeks of data are processed—"first print" estimates revise, and the average of weekly growth in the month is extrapolated to the end of the month. For months with significant within-month seasonality, such as November and December, it was necessary to adjust extrapolated weekly values based on the number of weeks "in hand." In practice real time monthly aggregation was a cumbersome process that reduced interpretability, offsetting the advantages of the full-month employment view.

During the COVID-19 pandemic the weekly data became much more important than they previously had been. We found it easier and more useful to simply seasonally adjust the *weekly* data and then build a *CES-mimicking monthly* index by selecting the weeks that contain the 12th of each month. This approaches loses the efficiency gains from averaging the month but results in weekly and monthly data that are exactly consistent and monthly data that can be compared to CES easily, requiring no special explanations for end users of the data.

5 Performance of ADP-FRB Employment Series

The goal of ADP-FRB is to provide a timely alternative indicator of payroll that is (1) as accurate as possible and (2) independent of the CES and other indicators. Here "indepedent" does not mean the errors have zero correlation; indeed since they are both large samples overlap in the reporting units is inevitable. Rather, we use independent to mean that CES and its lagged values do not enter the calculation of ADP-FRB.

5.1 Annual Benchmark Performance

We do not necessarily expect ADP-FRB to match the high-frequency variation in CES, nor should we, for at least two reasons. First, outside of extreme business cycles, the high frequency variation of both series is dominated by sampling error, seasonal adjustment noise, and other factors. Second, monthly CES data—even after full benchmark revisions—do not reflect "truth" but instead continue to reflect sampling error and birth-death model estimates, since annual benchmarking focuses only on the March employment level and "wedges back" these revisions across months. Performance on the annual March benchmarks is a much more sensible and policy-relevant metric.

As noted above, each year the March value of CES is benchmarked to the QCEW and other data, which serve as the ground truth for the payroll numbers. A preliminary announcement of the benchmark revision value is made in August, and the final benchmarking takes place early the following year. The benchmarking process shows whether CES correctly captured the trend in employment for the previous April-March interval, or if it ran too hot or too cold relative to the administrative data.

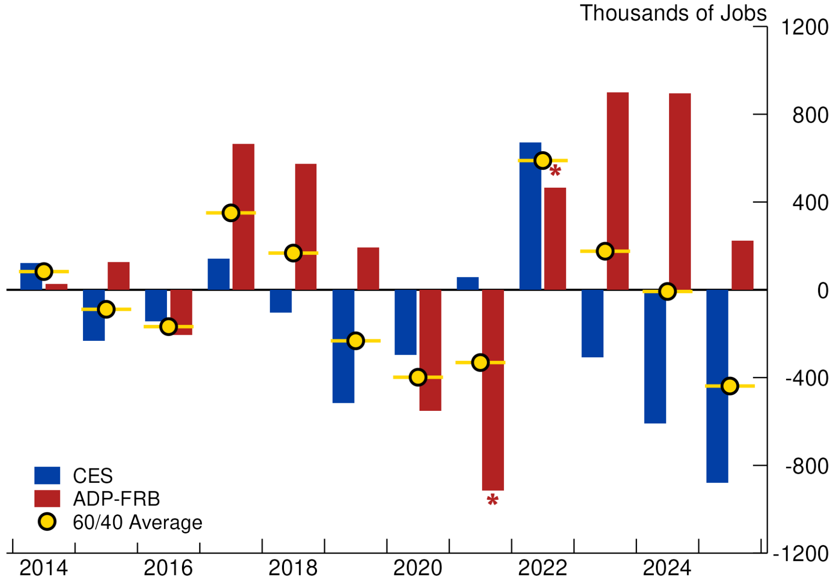
If ADP-FRB has smaller benchmark revisions than CES that is obviously desirable. However, ADP-FRB can be very useful even if the errors are similar in size or larger than CES. Two (semi) independent signals are more useful than one, since the signals can be pooled. In Cajner et al. (2022) we suggest a state-space model as an appealing framework to pool two signals of true employment growth. The Kalman filter approach assumes that true employment growth follows an AR(1) process and that the CES and ADP-FRB values are each equal to the true value plus some (possibly correlated) noise. In practice, we found that our estimated Kalman filter can be well-approximated by a simple 60/40 weighted average of the CES (60 percent) and ADP-FRB (40 percent) estimates.[17] In this section we focus on the 60/40 average since it is more transparent and intuitive, but the results are generally the same or stronger for the actual Kalman filter.

Figure 3 shows the benchmark revisions for both series.[18] To be clear, the red 2025 bar means that the March 2025 ADP-FRB value revised revised up by about 200,000 jobs. There is no consistent pattern, though since the pandemic CES has tended to revise down and ADP-FRB has tended to revise up. ADP-FRB benchmark revisions are often larger than CES. This is not too surprising: while the raw fraction of employment covered is similar to CES, the CES program has more expertise and resources available as well as infrastructure

[17]See Cajner et al. (2022) for details. To obtain the 60/40 weights we regress the smoothed Kalman state on the contemporaneous values of CES and ADP-FRB, under the restriction that the coefficients sum to one. The implied weights vary from sample to sample, but the most recent values were approximately 60/40.

[18]We extend the figure back to 2005 in Figure 12 of Appendix A.

Figure 3: Constant-Scope Benchmark Revisions



Note: The ADP-FRB concept is "active" prior to the COVID-19 pandemic, switches to the "paid" concept for the 2021 revision, and switches back to "active" in October 2021. Benchmark revisions which include the "paid" concept (2021 and 2022) are denoted with an asterisk. ADP-FRB was benchmarked to February 2020 rather than March 2020; 2020 and 2021 ADP-FRB values are annualized. "60/40 Average" is the weighted averages of annual CES and ADP-FRB benchmark revisions.
Source: BLS, ADP, authors' calculations.

for avoiding selection bias. But the ADP-FRB benchmark revisions are generally not *much* larger than CES and, importantly, are often in the opposite direction. This suggests the value of multiple independent, noisy signals. Even if we had a strong prior that CES was a better signal, it would be better to shade the CES estimate towards ADP-FRB than to simply ignore the ADP-FRB. The 60/40 estimate takes this logic further, effectively putting a 40 percent weight on ADP-FRB. The 60/40 average benchmark revision is frequently close to zero and much smaller than that of CES. Since this estimate is a available in real time we could have a very good estimate of the benchmark revision in April, four months ahead of the August preliminary CES benchmark announcement and far ahead of the final benchmark. This conveys a significant benefit to macroeconomic analysts and forecasters.

5.2 Higher Frequency Behavior

Twelve month growth rates are a function of the long-run trend in employment growth. This is important, but policymakers and market participants also care about higher-frequency changes: business cycle turning points, the reaction to policy and natural disaster shocks, etc. Evaluating indicators at this frequency is difficult. High-frequency QCEW data are not considered as reliable and are not seasonally adjusted by BLS (see U.S. Bureau of Labor Statistics, 2026*a*; Dey and Loewenstein, 2017; Groen, 2012, for details), and we lack versions of the data that have gone through BLS's additional cleaning. In the absence of useful administrative benchmark data, a useful exercise is to nowcast monthly CES. Even though forecasting/nowcasting CES is explicitly *not* the objective of these indicators, CES has high frequency signal and forecasting/nowcasting CES would be indicative of signal in the candidate series.

Our starting point is the monthly job gains for CES and ADP-FRB. To remove the lower-frequency signal from both series we run the benchmark adjustment process in reverse: For each series separately, we subtract off the average April-to-March job gains from each observation, such that each series cumulates to zero for each March-to-March period. This means that any statistical relationship between the processed series is going to be at a sub-annual frequency.

Table 1 shows the results on a post-2021 sample. ADP-FRB robustly predicts the high-frequency movemements in CES, with a coefficient implying nearly one-for-one comovement and a reasonably high R^2. We interpret this as evidence that ADP-FRB has significant predictive power for CES, even at higher frequencies where sampling noise and other sources of error are likely to be important. This complements our earlier work in Cajner et al. (2018), showing predictive power for the uncentered ADP-FRB series in real time. The current exercise makes clear that some of this power is in the higher frequency region, rather than all being a function annual-frequency variation.

	(1)
ADP-FRB, centered	1.329***
	(0.278)
Observations	56
R^2	0.618

Standard errors in parentheses

* p< .10, ** p< .05, *** p< .01

Note: Dependent variable is current vintage CES jobs gains, centered by removing the April-March average from each interval. ADP-FRB is is processed symmetrically. Robust standard errors in parentheses. Sample is 2021m1-2025m8.

Source: BLS, ADP NER, authors' calculations.

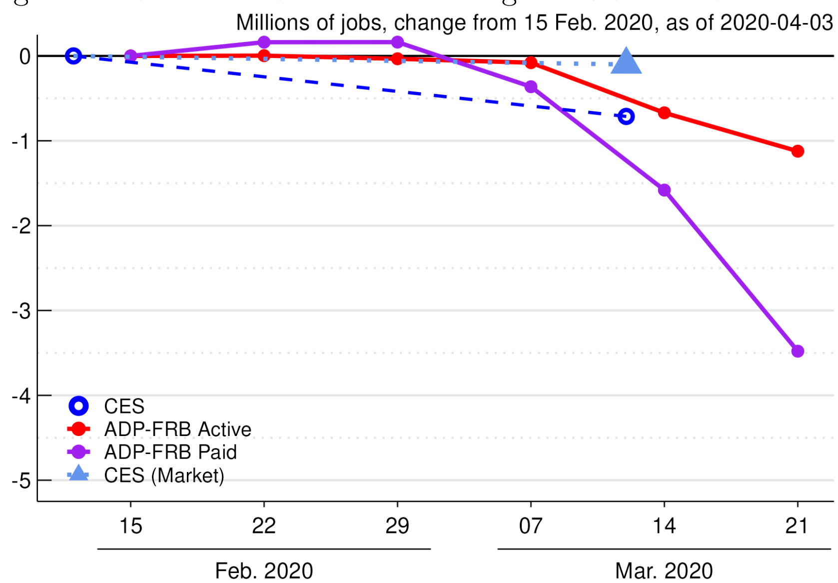
Table 1: Predicting CES with ADP-FRB

Note that we conduct this exercise using seasonally adjusted versions of both ADP-FRB and CES. Seasonal fluctuations are often a dominant source of variation in many economic time series. Using correlations or R^2 to evaluate an alternative data sources's tracking of official data at high frequency may provide a false sense of reliability of the alternative data source, if alternative data source features seasonality that is similar to the official data.

5.3 COVID-19 Pandemic

At times, there are policy-relevant changes in employment growth that unfold over only a couple months. Business cycle turning points can be sharp, and better identification of these events in real time is of great interest. The best recent example is the onset of COVID-19 pandemic. Figures 4, 5 and 6 show how the weekly ADP-FRB data tracked the COVID-19 employment dynamics in real time during the spring of 2020.

Figure 4: Real Time Job Losses during the COVID-19 Pandemic

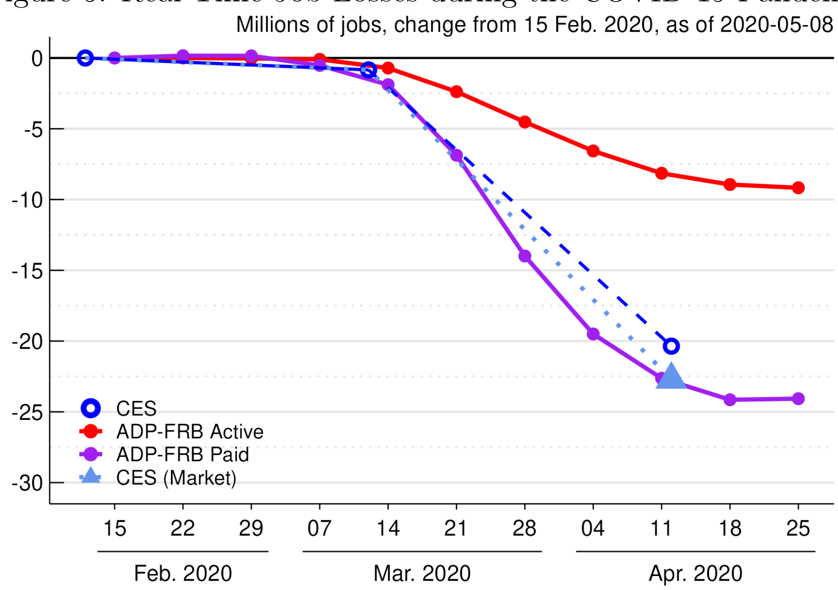


Note: CES dashed line indicates new month of data from 2020-04-03 Employment Situation. ADP-FRB series are real time on the eve of that Employment Situation.
Source: BLS, ADP, authors' calculations.

Figure 4 shows the real-time ADP-FRB data available to us just before the March 2020 employment report published in early April. In particular, these estimates were in hand around Monday, March 30, well before the April 3rd release of the March 2020 CES numbers. At that time the weekly ADP-FRB data extended through March 21st (the ADP series are plotted on their reference dates, Saturdays, and the CES numbers are plotted on the 12th of each month). The red line shows ADP-FRB active employment, which already recorded a decline of more than 1 million jobs. While active employment is generally our preferred measure, at the onset of COVID-19 it was immediately apparent that furloughs would be more visible in the paid series, and that the large nature of the shock would outweigh the usual noise associated with paid employment. The blue triangle shows the eve-of-release market consensus forecast for payroll gains through the March CES reference period (that is, the triangle reports the consensus as of right before the BLS employment situation was published).[19] The consensus forecast did not anticipate a dramatic drop in payrolls.

On April 3, the BLS published the employment situation, which showed a decline of about 700,000 jobs (later revised to about 1.4 million) through the March reference period (i.e., the payroll period including the 12th). The CES print is shown by the hollow blue circle. This was a large surprise relative to the market consensus forecast, but a massive decline in payrolls was easily anticipated based on the ADP-FRB paid index. Moreover, by the time of that release, ADP-FRB paid employment already showed a cumulative decline of about 3.5 million jobs, much of it occurring *after* the CES reference period for March. In this rapidly developing situation, CES data were stale upon release relative to the available ADP-FRB data.

Figure 5: Real Time Job Losses during the COVID-19 Pandemic



Note: CES dashed line indicates new month of data from 2020-05-08 Employment Situation. ADP-FRB series are real time on the eve of that Employment Situation.
Source: BLS, ADP, authors' calculations.

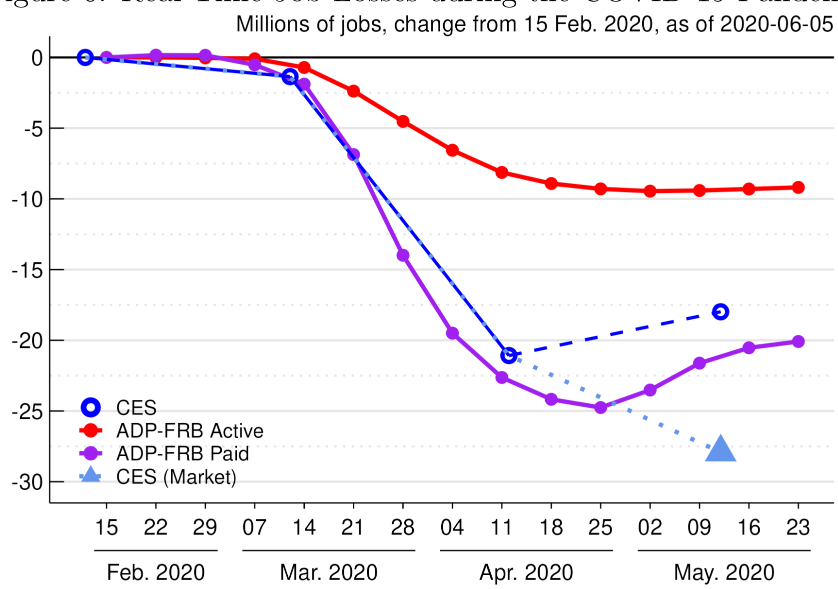
Figure 5 steps forward one month, to the week of the April 2020 employment situation published in early May. Again, ADP-FRB had data in hand covering the April CES reference period and extending beyond it. ADP-FRB paid showed cumulative declines on the order of 25 million jobs, very similar to what CES would report. This time the market consensus forecast was fairly close, guided by unemployment insurance claims data and possibly, by this point in the pandemic, some private sector data sources. But importantly, by this time

[19]There is some mismatch between the market forecast shown in the figure and the CES and ADP-FRB data. The market forecast pertains to total employment, while the other data on the figure pertain to private employment. This is due to data availability at the time of writing, and it does not materially matter for the argument.

the weekly ADP-FRB paid data also showed employment flattening after the reference week, not continuing to plummet.

Stepping forward again, Figure 6 shows that ADP-FRB correctly anticipated the substantial increase in employment between April and May. But the market consensus forecast—many of whose constituents were looking at (other) private sector data by this time, as well as unemployment claims—failed to anticipate the stabilization of payrolls and had expected further massive job losses.

Figure 6: Real Time Job Losses during the COVID-19 Pandemic



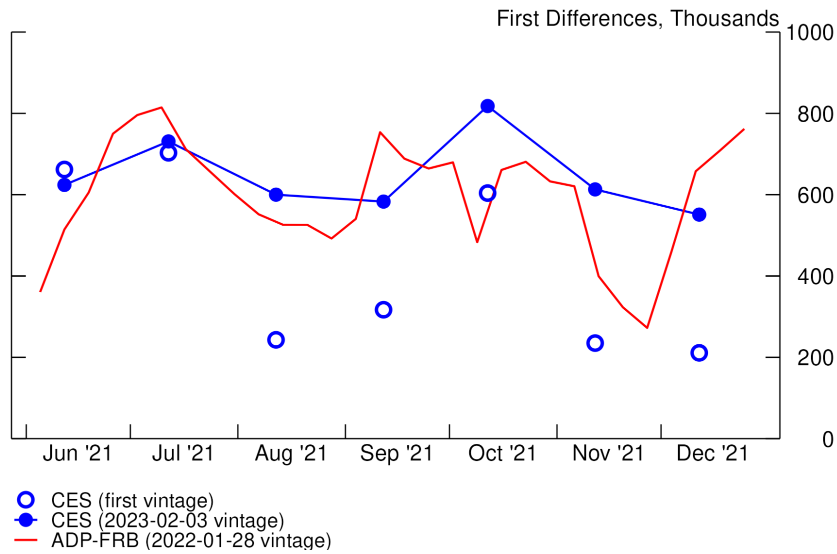
Note: CES dashed line indicates new month of data from 2020-06-05 Employment Situation. ADP-FRB series are real time on the eve of that Employment Situation.
Source: BLS, ADP, authors' calculations.

While we expect shocks as large and as sudden as COVID-19 to be rare, the episode illustrates the value of high-quality alternative data when the unexpected does happen. In earlier work (Cajner et al., 2022), we illustrated a somewhat similar pattern for the fall of 2008, the most acute period of the Global Financial Crisis; while we did not have access to ADP data at that time, we can study the ADP-FRB index for that period.

5.4 CES Revision Episodes

In recent years CES has at times experienced persistent month to month revisions in the same direction. These are distinct from the annual benchmark revisions; instead, the month to month revisions are a function of (1) some CES reports arriving late, and (2) revisions to seasonal factors. These revisions can be frustrating to policy makers and market participants

Figure 7: The Labor Market in the Second Half of 2021



Note: ADP-FRB is the 4-week moving average of the monthly rate. ADP-FRB data is vintaged prior to the 2021 benchmark revision. CES first vintage is the first release for each reference month. CES 2023-02-03 vintage reflects the 2022 benchmark revision.
Source: BLS, ADP, authors' calculations.

since high-frequency revisions may alter the narrative around the labor market. Figures 7 and 8 show how in these episodes, CES tended to revise towards ADP-FRB.[20]

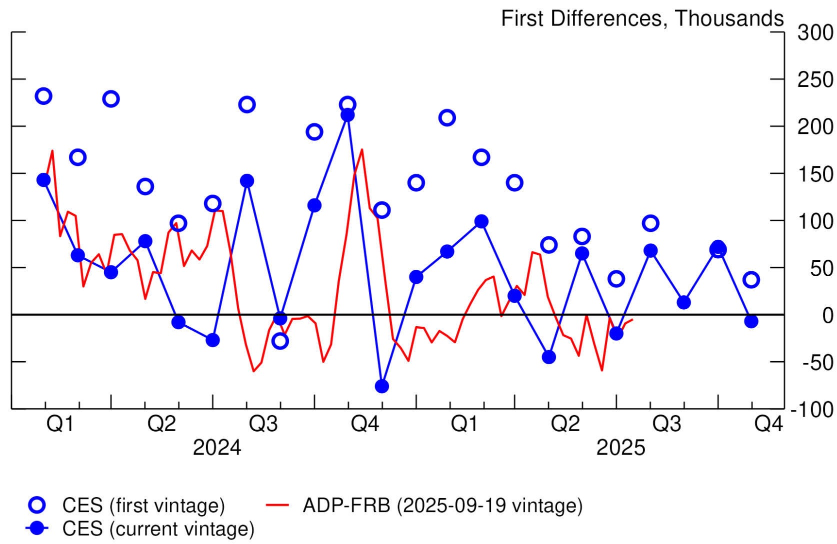
First consider Figure 7, which illustrates a time of great uncertainty about whether the labor market recovery from the pandemic had run out of steam or not. The red line shows the 4-week moving average of ADP-FRB, which suggested continued strong payroll gains.[21] The hollow blue dots correspond to the CES first read for each month, which suggested that the labor market recovery cooled markedly starting late in the summer. The blue line and filled dots show revised CES, which looks more like ADP-FRB. Clearly, CES tended to make large upward revisions towards ADP-FRB over this period, often by more than 100,000 jobs.

Figure 8 describes another time of uncertainty about the labor market, in 2024-2025. The first prints of CES pointed to a continued hot labor market, while ADP-FRB suggested payroll gains had slowed. The revised CES data corroborated ADP-FRB's earlier signal.

In addition to its tracking of aggregate developments, during the pandemic and its aftermath we found the broad sector-level detail of the ADP-FRB data useful. For example, the leisure and hospitality sector was particularly affected by the pandemic, and our ADP-FRB indexes for that sector provided early guidance on sector-level developments.

[20] Appendix Figure 13 repeats Figure 8 but with an additional earlier vintage of ADP-FRB.
[21] For simplicity this is just an early 2021 vintage of ADP-FRB; ADP-FRB does revise some month to month but the revisions are generally small and the pattern is not persistent.

Figure 8: The Labor Market after the 2024 Benchmark



Note: ADP-FRB is the 4-week moving average of the monthly rate.
Source: BLS, ADP, authors' calculations.

Taking stock, over the years we have found that the ADP-FRB payroll indexes are helpful for understanding the underlying strength of the labor market, both at lower frequency—with insight into likely benchmark revisions to official data—and at higher frequency—for tracking monthly patterns in CES and even weekly patterns during periods of rapid change.

6 ADP-FRB Wage Statistics

6.1 Description of Data and Method

While not the main focus of this paper, teams at the Federal Reserve Board have also constructed wage measures from the ADP data. We now briefly summarize these projects.

The ADP research microdata also includes data sets that contain information at the employer-employee pair level. These data sets include information about earnings and hours, among many other employee variables and characteristics. This allows us to create measures of wage growth, both in the aggregate and for different subsets of the workforce.[22]

The high level of granularity in the ADP microdata makes it possible to create different types of wage growth measures. There are also different types of publicly available wage

[22]The employer-employee data are available at a monthly frequency going back to mid-2008, and weekly starting in 2019. We aggregate weekly observations into monthly ones post-2019. As a result, all our series are monthly and run from mid-2008 to July 2025.

growth measures. For example, the BLS' Average Hourly Earnings (AHE)—a product of the CES survey data—consists of an average wage across the workforce employed in any given month. Another kind of measure is produced by the Federal Reserve Bank of Atlanta, whose Wage Growth Tracker (WGT) presents the median of individual year-over-year wage growth rates computed from the CPS.

Here we focus on measures that follow a methodology similar to that of the WGT: we start with hourly earnings at the individual (employer-employee) level in a given month, compute the percent change compared to 12 months prior—again at the individual level—and compute the median. We weight individual observations to match the distribution of industry and employer size among the workforce, based on Census' Quarterly Workforce Indicators (which are based on much of the administrative data underlying the QCEW).

The data also provide details on employee earnings, which allows for indicators covering differential breadth of earnings. We observe an employee's total earnings, including overtime, bonuses, tips, and commissions, from which we compute total earnings per hour using the observed hours.[23] The data also include a variable for an employee's contractual base pay rate. For hourly workers this consists of their hourly pay, while for salaried workers it is their base pay per pay period.

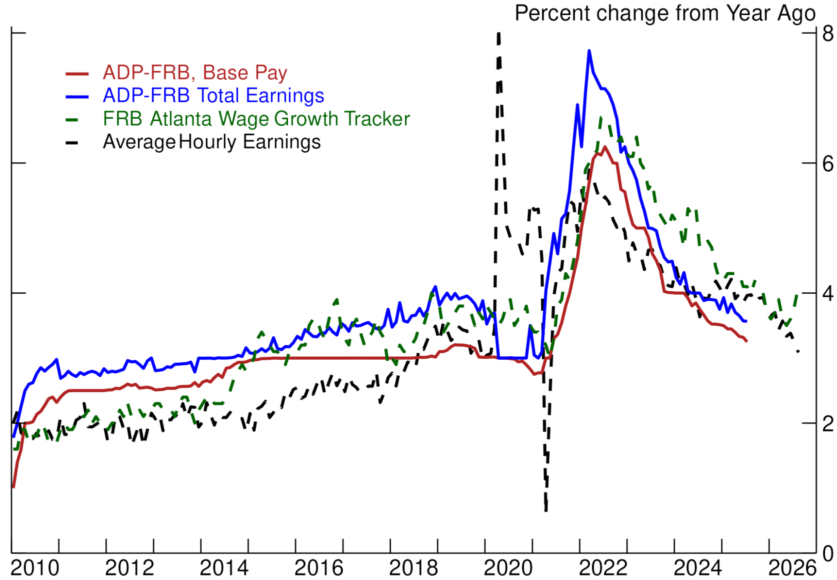
6.2 ADP-FRB Measures and Comparison with Official Measures

Figure 9 shows the two ADP-FRB measures of median wage growth, one based on total earnings and one on base pay, alongside the AHE and WGT measures. All the series show a similar pattern over time: they increased slightly around 2015 and 2016 after a few years of stability, later surged in 2021, and subsequently moved back down. We first highlight that the ADP-FRB measures provide a similar signal on wage growth as the official measures. At the same time, the ADP-FRB measure based on total earnings seems to *lead* many of the others when wage growth is changing—a helpful feature for macroeconomic analysts and forecasters.

We also highlight two important advantages of the ADP-FRB wage measures that make them a useful complement to the official statistics. First, the median wage growth construction implies that changes in the composition of the workforce from month to month do not lead to mechanical changes in the measure. The BLS AHE series is subject to such compositional effects. One stark illustration is at the onset of the COVID-19 pandemic, when the large employment declines were concentrated among low-wage workers (a point made in Cajner et al., 2020*b*, using ADP data). This substantially raised the average wage among the

[23]The data also indicate whether an employee is hourly or salaried. For salaried employees, we assume the equivalent of a 40-hour workweek, unless the hours variable indicates otherwise.

Figure 9: ADP-FRB and Official Wage Measures



Source: BLS, FRB Atlanta, ADP and authors' calculations.

employed workforce and pushed AHE growth to an all-time high. In contrast, the ADP-FRB measures weakened slightly, consistent with the weakness in the labor market.

A second advantage of our measures can be seen in the comparison with the WGT, which similarly controls for the compositional effects. The ADP-FRB series are based on a considerably larger sample: about 13 million monthly individual wage changes, compared to about 2,000 for WGT. In addition, the ADP-FRB measures are based on administratively reported data on earnings, while WGT relies on self-reported earnings and hours, likely leading to greater measurement error. Probably as a result, the ADP-FRB series are smoother and show fewer short-term fluctuations. In order to smooth through noise, the published WGT series consists of a 3-month moving average of median wage growth computed month by month. This alleviates the volatility of the underlying monthly values, but introduces some sluggishness into the published series. In contrast, the ADP-FRB series do not need to be smoothed and reflect each month's median year-over-year wage change.

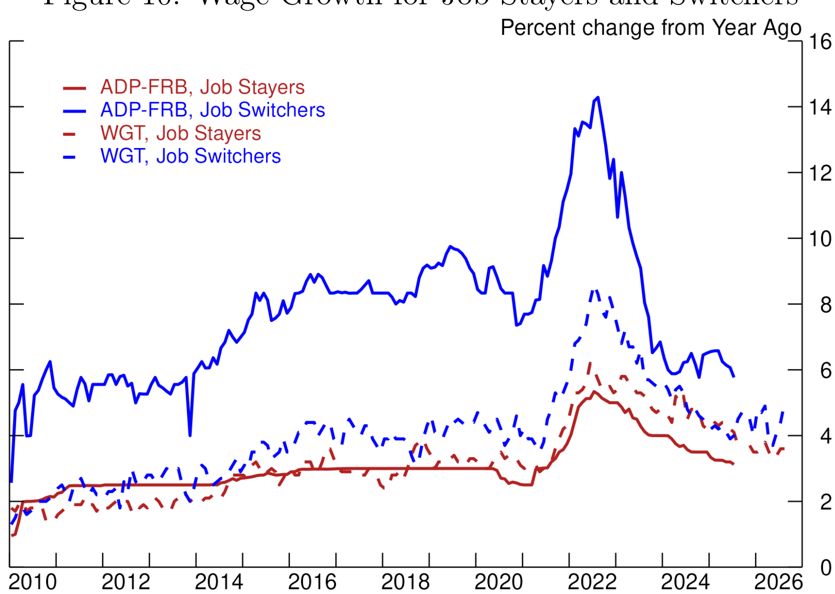
6.3 Job Stayer and Switcher Wage Growth

Given the importance of job switching in the post-pandemic labor market, the wage growth experienced by job switchers has received considerable attention in recent years. The nature of the ADP microdata allows us to identify job switchers and job stayers with great accuracy. The data are at the employer-employee level and include both employer and employee (anonymized) identifiers. This makes it possible to determine when an individual has switched employers (as long as the new employer is also in the ADP sample), and when they have stayed with the same employer (both relative to 12 months prior, the period over which we calculate wage growth).

In contrast, the CPS does not make it possible to confidently identify job switchers across the 12 months over which wage growth is calculated for the WGT. The WGT is still available for job switchers and stayers separately, but the distinction is determined mostly based on whether individuals report a different industry or occupation compared to 12 months prior. This likely introduces error into that determination and in particular could lead to substantially under-counting job switchers.

Figure 10 displays median wage growth for job switchers and job stayers in ADP-FRB and in the WGT. Both types of measures again follow similar qualitative patterns. However, a few important differences stand out. First, the (positive) gap between switcher and stayer growth, which can be thought of as an average job switching premium, is considerably large in ADP-FRB. Second, the swings in switcher growth during and after the pandemic are considerably more pronounced in the ADP-FRB series. According to the ADP-FRB measure, switchers experienced a large surge in wage growth in 2021 and 2022, but by the end of 2023 their wage growth had fallen back to pre-pandemic levels. In contrast the WGT for switchers did not rise as much but remained elevated (relative to pre-pandemic levels) for longer.

Figure 10: Wage Growth for Job Stayers and Switchers



Source: BLS, FRB Atlanta, ADP and authors' calculations.

7 Future Approaches

The success of ADP-FRB makes an encouraging case for the use of private payroll data to produce employment statistics. More accurate and timely indicators seem to be within reach, especially if data from several payroll firms could be used as inputs.

7.1 Information Sharing

There are several possible ways aggregate employment statistics could be compiled. Under the status quo, payroll processors and other firms each independently develop indicators with different scopes, timeliness, and methodologies. Researchers and policy makers then assign weight to each, either judgmentally or using statistical models, or using estimates of sample size and likely overlap. As a result, different estimates can give a different picture of the economy, and even observers taking signal from the same indicators can disagree on the overall picture. Our experience, the many methodological challenges we faced, and our experiences with shifting macroeconomic regimes suggests that decisions about how to construct employment indexes matter greatly, and there is considerable value in having those decisions be transparent to data users and consistent with best practice, to the extent possible. An ideal approach would likely involve direct microdata provision to the statistical agencies, with multiple participating payroll processors.

Alternative approaches could emphasize more standardization and centralization, while respecting the need for maintaining data privacy and protecting competitive information. A coordinated approach to measuring the labor market would benefit from being able to involve the subject matter experts at the agencies and in academia, removing some of the burden from the data providers. Coordination would also help iron out divergence between estimates driven by different methodologies and different implicit assumptions. Importantly, the goal of any cooperation *should not be* to simply make sure there fewer estimates, or that they all tell the same story. The goal should be to improve the total quality of the available labor market data. That could mean improving several independent estimates, or it could mean aggregating the available data into single series.

One alternative approach would be for each data processor to continue processing their own data but share methodological details, anomalies, and analyse with a centralized group. In this way best practices can be propagated and methodologies can be improved. However, this approach does not result in a pooled estimate, so many of the current shortcomings would remain. Explicitly pooling independent estimates using sample weights would require each data processor to share enough information—number of covered employees/firms, for example—for their estimates to be aggregated. This information may be sensitive since it can reveal their market share. On the other hand, large data providers often advertise their approximate market share (see Section 3). It is unclear what data providers would necessarily agree too, but it seems fairly unlikely that many would commit to publicly sharing market share information on an ongoing basis.

One way to address the market share problem would be to have a centralized trusted party receive both the growth rate information and the market share information from each data provider, and only publish the pooled aggregate. This would allow for a truly pooled estimate, and would only require that the data providers trust the compiler with their market share information. A variation of this approach would be to send the trusted parted cell-level aggregates (where cells are some combination of industry bins, size bins, and other characteristics) so weighting and other steps can be done centrally. Developing such a trusted party would clearly have large benefits in terms of allowing for more accurate statistics.

Even if a trusted third party can receive cell-level or top-line aggregates and compile them, it is less likely that data providers would actually send microdata to an external party. For the steps involving microdata a viable approach could mirror that of ADP-FRB: outside researchers and analysts with experience producing economic indicators access the anonymized data and compile the indicators inside the data provider's secure computing environment. In this way data providers retain full control of the microdata but the researchers can examine the data for unexpected issues and verify that the data processing pipeline works as expected. This approach takes some burden off the data providers, who would support the researchers instead of being directly responsible for producing the indicators. And it would ensure that data from all providers was being treated identically.

In summary, it seems there are two major hurdles to producing payroll statistics with private data. First, getting from the raw data to cell-level aggregates requires extensive exploration, adaptation, and processing. This might be done by the data providers themselves, but likely only if there is a great deal of discussion and transparency with the wider community. Allowing outside researchers to do this work instead would ensure it is informed by the state of the art, consistent across providers, and not overly burdensome on the provider. Second, aggregation across data providers requires market share information. Some of this is public, to some degree. But if data providers are to share it consistently they likely need assurances the exact values will not be made public. A trusted third party that can receive these data and compile the public release would be essential for this step.

The optimal degree of coordination is an open question, but our experience with the complexity and difficulty of producing payroll statistics suggests that a relatively high level of coordination is desirable. It seems unlikely that dozens of independent teams will successfully converge on comparable, high-quality methodologies without at a minimum extensive sharing of ideas, approaches, and analyses. Indeed, there is ample evidence that even experienced academic teams frequently arrive at differing conclusions when given the same data and the same question (see Silberzahn et al. (2018), Menkveld et al. (2024), and Huntington-Klein et al. (2021)), though of course a degree of independent innovation is desirable.

7.2 Using Technology to Facilitate Calculation and Aggregation

Clearly, what is feasible for private payroll indicators depends in large part on the requirement to protect the privacy of the raw data and respect the need for data providers to keep some information non-public. In this section, we discuss ways that technology can potentially alter the privacy/data quality tradeoff. These techniques are subsets of privacy enhancing technologies (PETs) or methods of secure multi-party computation. Most borrow ideas from cryptography to enable a particular computation or share a particular piece of information while providing formal guarantees that other information cannot be shared or reconstructed. The theoretical literature on these topics is vast. For examples see Shi et al. (2011) for privacy-preserving aggregation of time series data; and see Fenske et al. (2017) for calculating private set-union cardinality: finding the number of union items in several overlapping private databases, relevant for the case where businesses may appear in multiple data provider's databases. Note that in the literature noise infusion is often used to mask the the data of participants. Given that the large payroll providers have significant market share, the magnitude of noise needed to mask their data may render the data useless.

One interesting real-world example using payroll data is the Boston Women's Workforce Council wage-gap study (Lapets et al., 2015). In this study the firm submits payroll microdata through a web portal. On the client side the payroll data is aggregated into cells, so that no individual worker's data leaves the firm. Next, the firm's aggregates (e.g. total workers, total payroll for female workers, etc.) are masked by adding a random number to each and sent to a data processor. The data processor sums up the masked aggregates they receive from each firm, and passes the totals to a separate "analyzer" organization, who knows the values of the masks. The analyzer removes the (summed) masks and publishes the totals. As a result aggregate statistics can be publicly reported but no one, including the the data processor and the analyzer, observes any firm's totals. The maintained assumption is that the data processor and the analyzer do not collude to share the raw data or the mask values with each other. This example is particularly interesting because it does not require a completely trusted external party to handle aggregation. The data processors just have to be trusted not to collude. In this case the noise is removed before publication, so the published statistics are exact. The noise ensures that no party can learn a contributing firm's exact data, but if a firm is large relative to the sample the published total will reveal something indirectly. See also Bogetoft et al. (2009) for a real world case of secure multiparty computation for clearing the Danish sugar beet market.

Another example is Eurostat (2026), where a system for secure joint computation is outlined. The system relies on trusted secure enclaves: computing environments where the data can operated on before the permitted aggregates outputted. Eurostat (2026) combines this with multiple nodes, so than no single node can access all the sensitive data even if it is compromised. This setup does require trusting the enclave, but the design of the enclave makes it difficult for anyone—even the operator of the enclave—to access the raw data, and the distributed nature of the system provides further protections. See also Mitchell et al. (2024), a demonstration showing how agencies can coordinate to allow a joint computation while protecting their data. In the past secure enclaves have been implemented to facilitate data sharing in Switzerland, see Federal Department of Defence, Civil Protection and Sport (2023).

When it comes to payroll data, secure multi-party computation and enclaves are most relevant for the aggregation cell-level data or topline estimates across payroll providers. As noted above, there is substantial computation and exploration that goes into producing those cell-level totals for a single provider. This work can be done most effectively by a team of experts, who can react to unexpected features of the data and work with the data provider on solutions. An alternative, which does not allow the experts to see the microdata, is to have the outside group write the code which is then executed by the data provider and the results are sent back. Importantly, this type of arrangement is in use at national statistical agencies; see Vilhuber (2025) and Eberle, Müller and Heining (2017). But these are cases where the relevant dataset is exceedingly well-documented and researched, with synthetic test data are made available for code testing. Private payroll data is much more likely to have unanticipated issues, making it much harder for the code to take account of all possible problems.

Other types of arrangements may be feasible in the near future. LLM-based coding agents can, at times, complete an impressive range of complex coding tasks in messy, real-world environments. The success and spread of coding agents suggests that in the future it may be possible to give detailed instructions to an agent then have it operate on data the analyst cannot see and return usable aggregate statistics. In this hypothetical case an agent like Claude Code or Codex would operate inside the data provider's system, using a prompt (or harness) developed by outside analysts. The agent could be sandboxed or restricted to a container to prevent any unexpected impact on the computing environment. The prompt would specify the aggregates to be returned to the analyst, which the data provider could review and approve before anything was sent. It is possible that a capable agent with a well-specified prompt would be able to uncover and understand unanticipated issues in the data and treat them appropriately. Optionally, the agent could also return a qualitative summary of important data issues for further work—such a summary would need to be reviewed by the data provider to ensure no data leakage, but could be a relatively low-cost way to improve the data and code quality.

It is unclear if and when this would become a viable workflow. Current coding agents are very capable but also make amateurish mistakes and are not widely tested on these types of tasks. The workflow also relies on data providers having access to high quality agents inside their computing environment, a challenge for infrastructure, governance, and security.

8 Conclusions

We conclude by summarizing some of the lessons we learned from a decade of producing aggregate payroll statistics from payroll processor microdata. First and foremost, private payroll processors are a promising source of labor market data. The large processing firms cover a substantial share of U.S. employment and—in our experience working with a single provider—the resulting statistics are of high value.

The data are useful for measuring both low-frequency and high-frequency employment changes, and demonstrably add value to what CES already produces. That said high-frequency data are noisy and can be easily affected by outliers; this is likely true even for high-quality official statistics like CES. Predicting monthly statistics may be less useful than predicting slower-moving data, most importantly annual benchmark revisions. Relatedly, the usefulness of private administrative monthly data is conditional on, and enhanced by, the existence of high-quality, timely benchmark data like the QCEW.

Producing such statistics is labor intensive. The measurement problem is complex, requiring care with heterogeneous payroll frequencies, outliers, seasonal adjustment, weighting, and many other factors. It is likely not optimal for each payroll provider to try to produce their own employment report. Payroll providers would benefit from having expert input on how to produce data fit for economic indicators, and outside experts need to understand the idiosyncrasies of providers' data.

Coordination and bringing in outside experts involves two major privacy challenges. First, aggregating a provider's data to the (say, industry x time x size) cell level requires extensive exploration and coding work. This may require allowing vetted experts to work with the microdata. Perhaps in the future this could be assisted by coding agents. Second, aggregating across providers can raise concerns around revealing providers' market shares. This can be addressed either with a trusted computing environment where aggregation can happen, or secure multi-party computation, where totals can be computed without anyone seeing a provider's data. But the technical solutions are not widely adopted elsewhere, and the guarantees may not sound compelling to the data providers. Despite these challenges, the proven track record of private administrative payroll makes them a compelling area for future efforts.

References

Akan, Mert, Jose Maria Barrero, Nicholas Bloom, Thomas Bowen, Shelby R. Buckman, Steven J. Davis, and Hyoseul Kim. 2025. "The New Geography of Labor Markets." NBER Working Paper 33582.

Autor, David, David Cho, Leland D. Crane, Mita Goldar, Byron Lutz, Joshua Montes, William B. Peterman, David Ratner, Daniel Villar, and Ahu Yildirmaz. 2022*a*. "The $800 Billion Paycheck Protection Program: Where Did the Money Go and Why Did It Go There?" *Journal of Economic Perspectives*, 36(2): 55–80.

Autor, David, David Cho, Leland D. Crane, Mita Goldar, Byron Lutz, Joshua Montes, William B. Peterman, David Ratner, Daniel Villar, and Ahu Yildirmaz. 2022*b*. "An Evaluation of the Paycheck Protection Program Using Administrative Payroll Microdata." *Journal of Public Economics*, 211: 104664.

Barnatchez, Keith, Leland D. Crane, and Ryan A. Decker. 2017. "An Assessment of the National Establishment Time Series (NETS) Database." FEDS Working Paper 2017-110.

Bartik, Alexander W., Marianne Bertrand, Feng Lin, Jesse Rothstein, and Matthew Unrath. 2020. "Measuring the Labor Market at the Onset of the COVID-19 Crisis." *Brookings Papers on Economic Activity*, 51(2): 239–268.

Bogetoft, Peter, Dan Lund Christensen, Ivan Damgård, Martin Geisler, Thomas Jakobsen, Mikkel Krøigaard, Janus Dam Nielsen, Jesper Buus Nielsen, Kurt Nielsen, Jakob Pagter, Michael I. Schwartzbach, and Tomas Toft. 2009. "Secure Multiparty Computation Goes Live." Vol. 5628 of *Lecture Notes in Computer Science*, 325–343. Springer.

Brynjolfsson, Erik, Bharat Chandar, and Ruyu Chen. 2026. "Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence." Stanford Digital Economy Lab Working Paper. Revised version; originally released in 2025.

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." FEDS Working Paper 2020-089.

Cajner, Tomaz, Leland D. Crane, Ryan A. Decker, Adrian Hamins-Puertolas, and Christopher Kurz. 2020*a*. "Tracking Labor Market Developments during the COVID-19 Pandemic: A Preliminary Assessment." FEDS Working Paper 2020-030.

Cajner, Tomaz, Leland D. Crane, Ryan A. Decker, Adrian Hamins-Puertolas, and Christopher Kurz. 2022. "Improving the Accuracy of Economic Measurement with Multiple Data Sources: The Case of Payroll Employment Data." In *Big Data for 21st Century Economic Statistics.* , ed. Abraham et. al. University of Chicago Press.

Cajner, Tomaz, Leland D. Crane, Ryan A. Decker, Adrian Hamins-Puertolas, and Christopher Kurz. 2023. "Payroll Employment at the Weekly Frequency." *AEA Papers and Proceedings*, 113: 145–50.

Cajner, Tomaz, Leland D. Crane, Ryan Decker, John Grigsby, Adrian Hamins-Puertolas, Erik Hurst, Christopher Kurz, and Ahu Yildirmaz. 2020*b*. "The US Labor Market during the Beginning of the Pandemic Recession." *Brookings Papers on Economic Activity*, 3–34.

Chetty, Raj, John N Friedman, Michael Stepner, and Opportunity Insights Team. 2024. "The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data." *The Quarterly Journal of Economics*, 139(2): 829–889.

Cleveland, William P., and Stuart Scott. 2007. "Seasonal Adjustment of Weekly Time Series with Application to Unemployment Insurance Claims and Steel Production." *Journal of Official Statistics*, 23(2): 209–221.

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

Crane, Leland D., and Ryan A. Decker. 2019. "Business Dynamics in the National Establishment Time Series (NETS)." FEDS Working Paper 2019-034.

Davis, Steven J, John C Haltiwanger, and Scott Schuh. 1996. *Job Creation and Destruction.* The MIT Press.

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.* Vol. 71 of *Studies in Income and Wealth*, , ed. Katharine Abraham, James Spletzer and Michael Harper, Chapter 5, 187–216. NBER.

Decker, Ryan A. 2026. "Measuring Payroll Employment: A Note on the Current Employment Statistics." NBER Working Paper 34924.

DeFusco, Anthony A., Brandon Enriquez, and Maggie Yellen. 2024. "Wage Garnishment in the United States: New Facts from Administrative Payroll Records." *American Economic Review: Insights*, 6(1): 38–54.

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

Eberle, Johanna, Dana Müller, and Jörg Heining. 2017. "A Modern Job Submission Application to Access IAB's Confidential Administrative and Survey Research Data." Research Data Centre of the German Federal Employment Agency at the Institute for Employment Research FDZ-Methodenreport 01/2017, Nuremberg.

Eurostat. 2026. "JOCONDE: Demonstration Prototype, Test Cases and Test Results." European Commission, Eurostat JOCONDE Deliverable D5.1.

Fairman, Kristin, Lucia Foster, C. J. Krizan, and Ian Rucker. 2008. "An Analysis of Key Differences in Micro Data: Results from the Business List Comparison Project." CES Working Paper 08-28.

Federal Department of Defence, Civil Protection and Sport. 2023. "Successful Cooperation between the DDPS and the Swiss Financial Sector to Protect Privacy." *Press release*, Pilot project with the Swiss National Bank, SIX, Zurich Cantonal Bank, and Decentriq using confidential computing and data clean rooms.

Fenske, Ellis, Akshaya Mani, Aaron Johnson, and Micah Sherr. 2017. "Distributed Measurement with Private Set-Union Cardinality." In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security.* 2295–2312.

Feroli, Michael, Yuriy Gorodnichenko, Fiona Greig, Anil Kashyap, and Nela Richardson. 2026. "Private Canaries: The Value of Private-Sector Data for U.S. Monetary Policy Making." *2026 US Monetary Policy Forum.*

Grigsby, John, Erik Hurst, and Ahu Yildirmaz. 2021. “Aggregate Nominal Wage Adjustments: New Evidence from Administrative Payroll Data.” *American Economic Review*, 111(2): 428–471.

Groen, Jeffrey A. 2012. “Sources of Error in Survey and Administrative Data: The Importance of Reporting Procedures.” *Journal of Official Statistics*, 28(2): 173–198.

Hudson, Nicole, Jeannine Mercurio, and Jurgen Kropf. 2022. “The challenges of seasonal adjustment for the Current Employment Statistics survey during the COVID-19 pandemic.” *Monthly Labor Review*.

Huntington-Klein, Nick, Andreu Arenas, Emily Beam, Marco Bertoni, Jeffrey R. Bloem, Pralhad Burli, Naibin Chen, Paul Grieco, Godwin Ekpe, Todd Pugatch, Martin Saavedra, and Yaniv Stopnitzky. 2021. “The Influence of Hidden Researcher Decisions in Applied Microeconomics.” *Economic Inquiry*, 59(3): 944–960.

Hurst, Erik, Christina Patterson, Nela Thomas Richardson, and Ye Liv Wang. 2026. “Sticky Wage Norms and the Real Wage Cost of Unexpected Inflation.” NBER Working Paper 35624.

Kalton, Graham, and Ismael Flores-Cervantes. 2003. “Weighting Methods.” *Journal of Official Statistics*, 19(2): 81.

Kaminski, Marek W. 2015. “Optimization Targeting of the Annual Refiling Survey.” BLS Statistical Survey Paper.

Kurmann, André, Etienne Lalé, and Lien Ta. 2025. “Measuring small business dynamics and employment with private-sector real-time data.” *Journal of Public Economics*, 250: 105477.

Lachowska, Marta, Isaac Sorkin, and Stephen A. Woodbury. 2025. “Employers and Unemployment Insurance Take-Up.” *American Economic Review*, 115(8): 2529–73.

Lapets, Andrei, Eric Dunton, Kyle Holzinger, Frederick Jansen, and Azer Bestavros. 2015. “Web-based Multi-Party Computation with Application to Anonymous Aggregate Compensation Analytics.” Computer Science Department, Boston University BUCS-TR-2015-009.

Menkveld, Albert J., Anna Dreber, Felix Holzmeister, Juergen Huber, Magnus Johannesson, Michael Kirchler, Sebastian Neusüß, Michael Razen, Utz Weitzel, et al. 2024. “Nonstandard Errors.” *The Journal of Finance*, 79(3): 2339–2390.

Mitchell, Curtis, Sean Lovell, Ben Santos, Massimo De Cubellis, Angela Pappagallo, Ian Munoz, Kate McCall-Kiley, Luke Keller, Stephanie Studds, Beau Houser, Monique Eleby, and Ronald Jansen. 2024. "A New Model for International Privacy Preserving Data Sharing Across National Statistical Organizations." U.S. Census Bureau.

Nezaj, Jeff, Nela Richardson, and Liv Wang. 2025. "US Wage Patterns During and After the Pandemic: Insights from a Novel Data Source." PIIE Working Paper 25-5.

Quach, Simon. 2025. "Wage Hysteresis and Entitlement Effects: The Persistent Impacts of a Temporary Overtime Policy." *The Quarterly Journal of Economics*, 140(2): 1633–1680.

Quach, Simon. 2026. "The Labor Market Effects of Expanding Overtime Coverage." NBER Working Paper 35433.

Shi, Elaine, T.-H. Hubert Chan, Eleanor Rieffel, Richard Chow, and Dawn Song. 2011. "Privacy-Preserving Aggregation of Time-Series Data." In *Proceedings of the 18th Annual Network and Distributed System Security Symposium (NDSS)*. The Internet Society.

Silberzahn, Raphael, Eric L. Uhlmann, Daniel P. Martin, Pasquale Anselmi, Frederik Aust, Eli Awtrey, Štěpán Bahník, et al. 2018. "Many Analysts, One Data Set: Making Transparent How Variations in Analytic Choices Affect Results." *Advances in Methods and Practices in Psychological Science*, 1(3): 337–356.

U.S. Bureau of Labor Statistics. 2025. "Current Employment Statistics: Handbook of Methods." *https://www.bls.gov/opub/hom/ces/*, Current Employment Statistics, accessed September 11, 2026.

U.S. Bureau of Labor Statistics. 2026*a*. "Technical Note for the County Employment and Wages News Release." *https://www.bls.gov/cew/news-release-technical-note.htm*, Quarterly Census of Employment and Wages, accessed September 11, 2026.

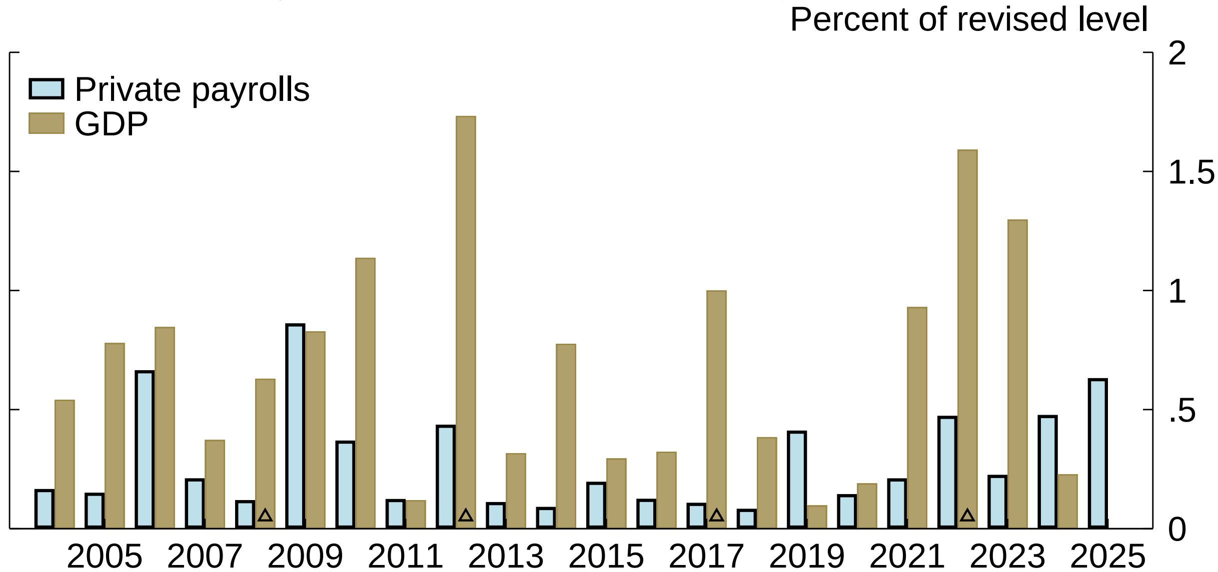
U.S. Bureau of Labor Statistics. 2026*b*. "Technical Notes for the CES-National Benchmark." *https://www.bls.gov/web/empsit/cestn.htm*, Current Employment Statistics, accessed September 11, 2026.

Vilhuber, Lars. 2025. "Using Containers to Validate Research on Confidential Data at Scale." *Harvard Data Science Review*. Special Issue 6: Data Privacy for Social Sciences.

Wright, Jonathan H. 2013. "Unseasonal Seasonals?" *Brookings Papers on Economic Activity*, 44(2 (Fall)): 65–126.

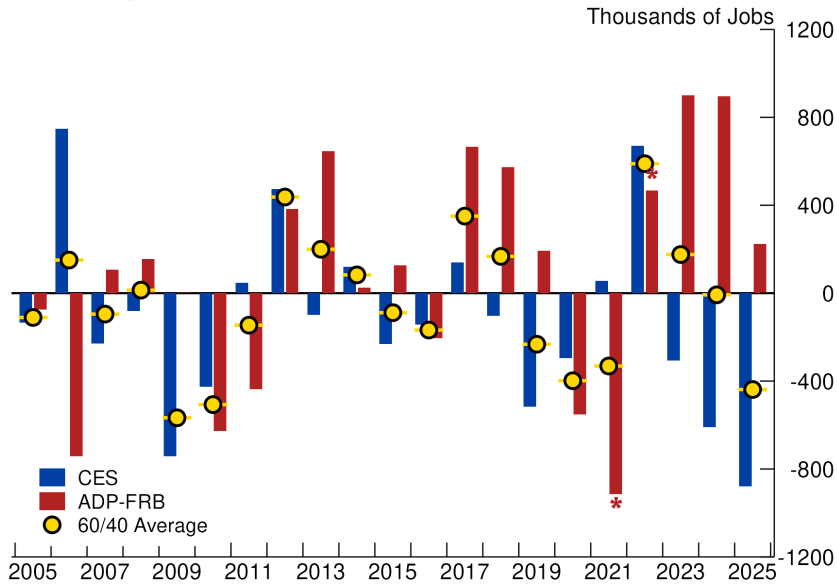
A Appendix: Supplementary figures

Figure 11: Revisions to CES private payrolls and GDP



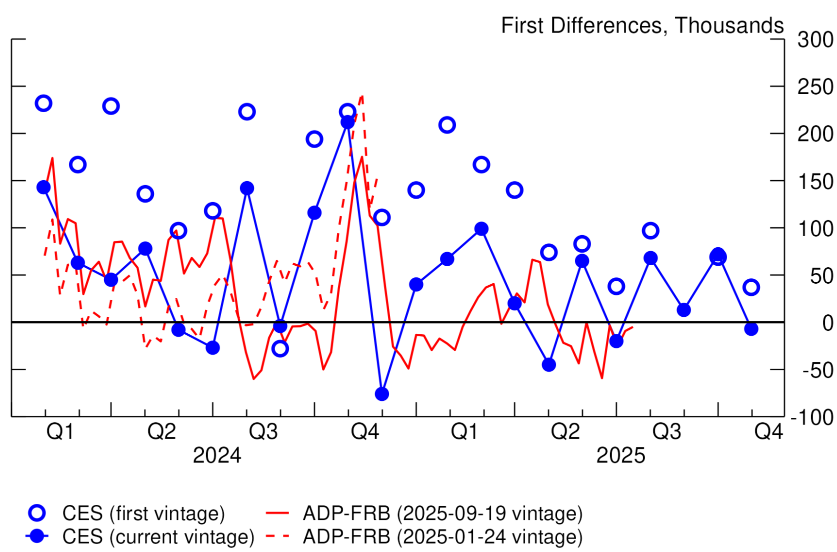
Note: Absolute revision to level in March (for private payrolls) or full year (for GDP) as published in subsequent year. Δ denotes GDP comprehensive revision years.
Source: BLS Current Employment Statistics; BEA National Income and Product Accounts.
BEA data retrieved from ALFRED, Federal Reserve Bank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

Figure 12: Constant-Scope Benchmark Revisions



Note: The ADP-FRB concept is "active" prior to the COVID-19 pandemic, switches to the "paid" concept for the 2021 revision, and switches back to "active" in October 2021. Benchmark revisions which include the "paid" concept (2021 and 2022) are denoted with an asterisk. ADP-FRB was benchmarked to February 2020 rather than March 2020; 2020 and 2021 ADP-FRB values are annualized. "60/40 Average" is the weighted averages of annual CES and ADP-FRB benchmark revisions.
Source: BLS, ADP, authors' calculations.

Figure 13: The Labor Market after the 2024 Benchmark



Note: ADP-FRB is the 4-week moving average of the monthly rate.
Source: BLS, ADP, authors' calculations.