

Abdelwahed, Loujaina; Baldva, Ritesh; Simon, Lisa

Article

Beyond surveys: How alternative data can measure the health of the labor market

EconPol Forum

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Abdelwahed, Loujaina; Baldva, Ritesh; Simon, Lisa (2026) : Beyond surveys: How alternative data can measure the health of the labor market, EconPol Forum, ISSN 2752-1184, CESifo GmbH, Munich, Vol. 27, Iss. 1, pp. 58-65

This Version is available at:

<https://hdl.handle.net/10419/336197>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Big-Data-Based Economic Insights

Beyond Surveys

How Alternative Data Can Measure the Health of the Labor Market

Key Messages

- The private sector offers reliable labor market data that policymakers, investors, and the public can use to guide their decisions
- Revelio Labs provides monthly labor stats from billions of work records, tracking employment, openings, turnover, and pay
- Alternative data is created beyond the goal of measurement, thereby avoiding survey bias and offering timely, cost-effective insights
- Job estimates from profiles need cleaning, title mapping, and models that adjust for sampling bias and reporting lag
- Job openings deduplicated across various sources provide a forward-looking view of demand and make it possible to estimate pay trends



Loujaina Abdelwahed is Head of Research at Revelio Labs. She holds a Ph.D. in economics from the University of Illinois in Chicago.



Ritesh Baldua is Revelio Labs' Chief Technology Officer, leading the engineering efforts and data platforms. He holds a master's degree in computer science from Columbia University



Lisa Simon is Chief Economist at Revelio Labs. She was previously a postdoctoral researcher at the Stanford Graduate School of Business, and she holds a Ph.D. from LMU Munich.

For decades, economists, policymakers, and businesses have relied on national statistical agencies as the main source of labor statistics data. In the US, the Bureau of Labor Statistics (BLS) has been the official source of labor market data since 1884, producing monthly statistics on employment growth, the unemployment rate, labor force participation, wage growth, hiring, separations, layoffs, job openings, and other economic indicators. Like many other statistical agencies around the world, the BLS produces these statistics by means of surveys, conducting multiple household and establishment surveys every month to estimate various indicators.¹

However, survey-based statistics are facing an increasingly difficult environment. Not only are these surveys expensive to run, but response rates to BLS surveys have also fallen to record lows in recent years, eroding the representativeness of the data and amplifying uncertainty around headline numbers (BLS 2025a).² This challenge is not unique to employment-related surveys. Rather, it reflects a broader issue affecting data collection by statistical agencies in the US in general. In addition, statistical agencies in the US have been highly politicized recently. The Trump administration has repeatedly expressed its dissatisfaction with the statistics produced by the BLS, which, however unfounded, erodes public trust in the data. These developments cast doubt on the future of public employment data and highlight the need for reliable alternatives.

Alternative data on the labor market can serve as an indicator of the health of the economy, offering two key advantages. First, updates are immediately visible and can be collected without much delay. Second, the data exists beyond the purpose of macro statistics and hence does not suffer from response biases like surveys (Bogner and Landrock 2016). Alternative data is not produced for the purpose of measurement; it simply exists as a byproduct of the modern economy, which makes it a valuable asset.

Although no dataset is perfect, large-scale, real-time data collected from digital sources and other administrative or transactional records can help fill the gap in the availability of reliable labor market data. Alternative data sources have proven their value as complements to official statistics, offering advantages in terms of timeliness, scale, and gran-

ularity. During the pandemic, when the economy shifted week by week, employment data from firms such as ADP and Indeed, or consumption data from credit card companies, gave policymakers and researchers an early view of downturns and recoveries, well before official releases. Private-sector data provides critical, timely signals when government statistics are delayed or under pressure.

In September 2025, Revelio Labs launched Revelio Public Labor Statistics (RPLS), offering monthly, publicly available data updates on the US labor market. The objective of RPLS is to provide timely data on the US labor market, with a full, accurate view of the workforce. RPLS releases include three core categories of metrics. First, employment levels and monthly changes in employment that provide a real-time view of job gains and losses, comparable to the Current Employment Statistics (CES) employment data. Second, job openings and employee turnover, comparable to the Job Openings and Turnover Survey (JOLTS). And third, wages for new positions, which provide a forward-looking perspective on compensation trends.

Revelio Labs Data

Revelio Labs is a workforce intelligence company that specializes in ingesting publicly available workforce data from online professional profiles, job postings, employee sentiment, and layoff notices. Revelio Labs then standardizes and enhances this data through statistical modeling to provide clean, representative, and reliable workforce data. The data collected is global and can represent almost any company, industry, or labor market in the world. Over time, Revelio Labs has become a trusted source of workforce data for clients in finance, business, and academia.

Revelio Labs provides different data products, such as individual-level data from professional profiles, which includes users' demographics, employment history, and education history. From this data, the Revelio Labs team has developed various models to generate estimates of firm-level headcounts, hiring rates, and attrition rates. The process involves creating sampling weights to correct for any biases that arise from the fact that online professional profiles are not fully representative of the actual workforce. The team has also developed a nowcasting model to correct for lags in reporting, since it takes time for people to update their job transitions on their online profiles.

Revelio Labs also gathers data on job postings from different sources, such as job boards, company career pages, and staffing firms. Collecting data from multiple sources ensures full global coverage of different sectors and occupations in the economy. Revelio Labs has developed mod-

¹ To estimate the unemployment rate and the labor force participation rate, the BLS depends on the Current Population Survey (CPS), a monthly survey of 60,000 households. To estimate employment and wage growth, the BLS runs the Current Employment Statistics (CES) survey, a monthly survey of 630,000 establishments. Data on hiring, separation, layoffs, and job openings come from the Job Openings and Turnover Survey (JOLTS), which surveys 21,000 establishments.

² In March 2025, the response rate of the CPS was 69 percent, while the response rates of the CES and JOLTS establishment surveys were 42.6 percent and 35.2 percent, respectively.

els to standardize and deduplicate job openings across all these sources to create a unified job postings product.

These two data products are the key foundation for our RPLS data releases. Revelio Labs' access to this micro-data, as well as its investment in modeling over the past years, puts the company in a unique position to provide aggregate labor market statistics.

The Case for the Private Sector Providing Public Labor Statistics

Many economists argue that the private sector cannot provide reliable data to measure the economy (Bloomberg 2025). They argue that private companies are not in the business of providing public services, either because they are motivated by profit or because of concerns on continuity; private data providers still need to rely on government data to generate sampling weights to make the data representative of the economy; and private data is not comprehensive enough to cover different sectors and occupations. Yet, we contend that none of these arguments are entirely valid.

First, it is indeed true that private companies are profit seekers. However, there are countless examples of the private sector providing public goods – including open-source software, privately maintained public spaces, and private research grants. Providing public services creates value to private firms that extends well beyond the profit motive. Publishing employment statistics gives a growing company like Revelio Labs welcome visibility, making it part of the cost of doing business. Furthermore, staff economists at private companies have strong incentives to track economic performance, conduct research, and provide thought leadership. Offering public data supports all these objectives.

As far as continuity concerned, it is true that private companies can go out of business or undergo a change in leadership that redefines priorities; however, even government agencies are not immune to such interruptions. For example, in 2025, the US federal government underwent a 43-day shutdown, the longest in history. The BLS did not release key economic data for three consecutive months, leaving researchers and policymakers without timely information.³ Meanwhile, private data providers remained highly incentivized to continue collecting and disseminating data, thus offering valuable signals on the labor market.

³ The data was later released after the shutdown ended, however with great delays and have undergone significant revisions.

Finally, while we indeed use BLS statistics as an input to generate sampling weights, they are not the only factor we consider. Company 10-K filings and people's connectivity on online professional profiles, for example, also serve as excellent predictors of a worker's likelihood of representation in online data. Moreover, the creation of sampling weights does not require constant updates and therefore does not pose a continuous dependency on BLS data.

Measuring Total Employment

Revelio Labs collects and standardizes data from millions of public profiles on professional networking websites such as LinkedIn. These profiles include the entire work history of users, containing information about their employers, their dates of employment, and the job description and responsibilities for each position they have held. The coverage of Revelio Labs' data on professional profiles is global, and within the US, it spans all regions.

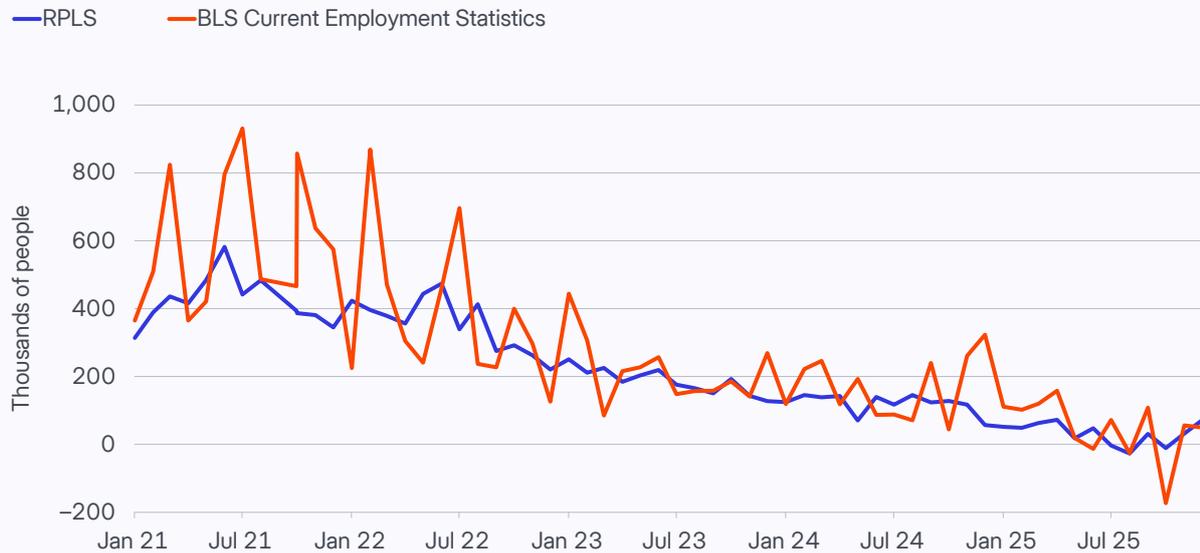
Raw professional profile data is messy: people can have multiple profiles, fake jobs, inactive accounts, missing information, etc. Revelio Labs has implemented a sophisticated treatment pipeline to transform raw profile data into a reliable set of worker records. The process includes detecting and removing fake profiles, consolidating duplicate profiles, and standardizing company names and job titles. More details about these pipelines are outlined in our online data dictionary.

To create a measure of total employment comparable to the BLS, we first construct a sample of workers that resembles those counted by the BLS establishment survey: full-time and part-time workers in the private and public sectors, excluding military personnel, interns, self-employed persons, and workers who are on leave or sabbatical. We further limit the sample to workers affiliated with companies that have at least five observed profiles. The final dataset contains approximately 104 million profiles, covering about 65 percent of the US workforce.

One significant challenge that arises from working with online professional profiles is lags in reporting. These lags arise due to two main reasons. First, employees may not update their profiles immediately after changing jobs, particularly in the case of involuntary separations and layoffs. Secondly, there can be lags in scraping. Scraping providers revisit every profile periodically, so it may take up to a few weeks until we observe individual workers' transitions. These sources of delay contribute to an observable lag distribution that correlates with attributes such as job function, geography, company size, and industry. To address this problem, we have constructed a model of workforce

Figure 1

Month-over-Month Change in Non-Farm Employment

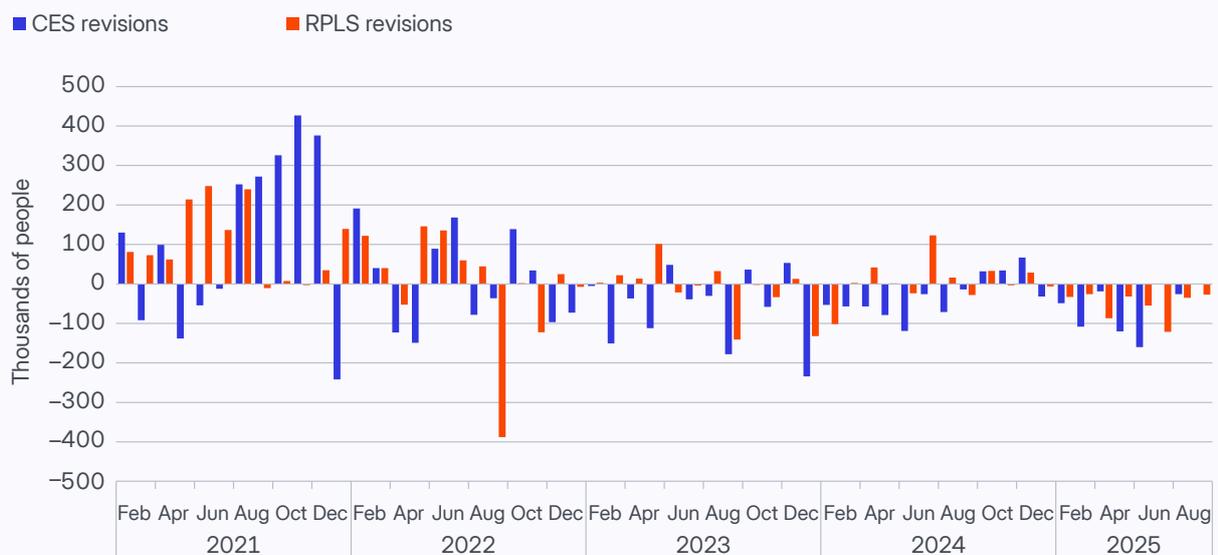


Notes: Estimates of month-over-month change in employment (in thousands of people) from RPLS (blue line) and the BLS's CES (orange line). The BLS series ends in September 2025, since this was the last BLS release data right after the government opened back up in November 2025. Both series are seasonally adjusted.
 Source: RPLS and BLS CES data.

© ifo Institute

Figure 2

Difference in the Month-over-Month Change in Employment Between the First and the Last RPLS and CES Releases



Notes: The difference in month-over-month estimates of total non-farm employment (in thousands of people) between the first data release and the latest data release. RPLS revisions are represented by the orange bars, while CES revisions are represented by the blue bars.

Source: RPLS and BLS CES data.

© ifo Institute

dynamics that predicts the inflows and outflows that will be revealed once all public profiles have been updated. The Bayesian structural time series model accounts for trends, seasonality, cyclical, drift, and new shocks. The model is trained using both the most recent snapshot and a longitudinal sequence of historical snapshots, which enables it to learn how headcount observations evolve over time. The model also explicitly accounts for known events, such as layoffs and hiring surges.

Another challenge that emerges when using online professional profile data is the representativeness of the actual workforce. Profiles on these platforms disproportionately include white-collar occupations, larger firms, and workers in metropolitan areas. To correct for these biases, we construct sampling weights that incorporate information from 10K filings for publicly traded firms and from the BLS’s Occupational Employment and Wage Statistics (OEWS), which includes a nationally representative distribution of employment across occupations, sectors, and geographic areas. More details on sampling weight construction can be found in RPLS’s online methodology documentation.

After we reach our best estimate of total non-farm civilian employment, we adjust the data to correct for any seasonal patterns. Once the headcount is estimated and

seasonally adjusted at the sector-occupation-state level, we aggregate across these categories to reach the total employment in the US. Figure 1 compares month-over-month changes in employment from Revelio Labs to monthly change in payroll employment reported by the BLS. The correlation (0.74) between these two series is high.

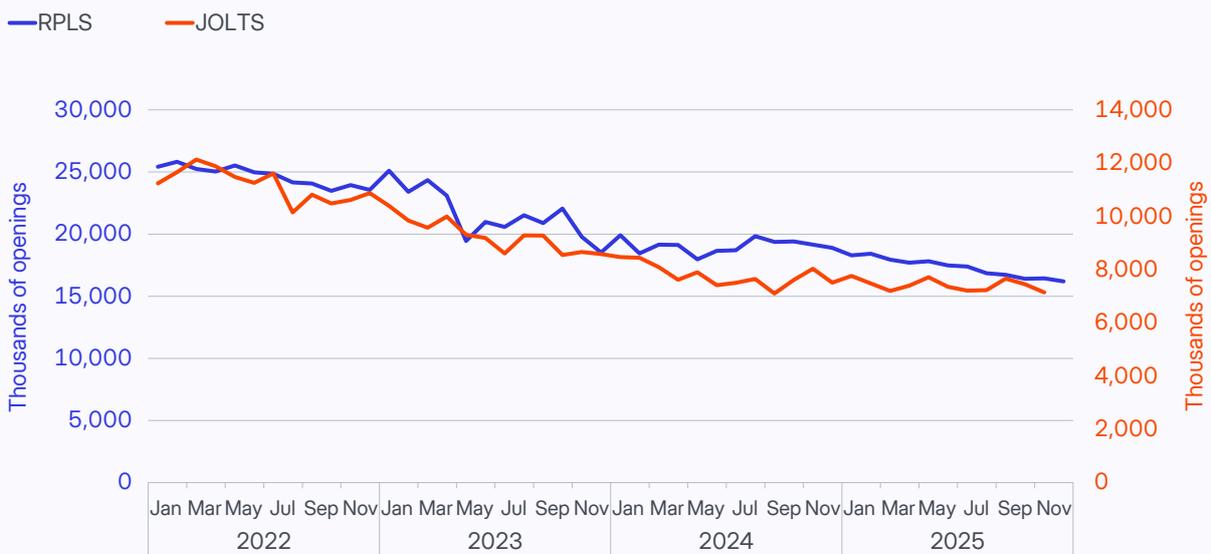
Data Revisions

Data revisions are an unavoidable part of producing macro statistics. They reflect new information, corrections, and methodological adjustments. RPLS updates historical employment estimates as new profile data are ingested and adjustments for representation, reporting lags, and seasonality are made. The BLS revises its estimates of total employment twice in the two months following the initial release. The BLS also revises its estimates of employment as part of its annual benchmark process, incorporating updated unemployment insurance records. These benchmark revisions can substantially change previously published figures.

The BLS has faced criticism for sizable revisions in prior months’ employment figures – revisions that fueled controversy over the reliability of its data. The size of revisions

Figure 3

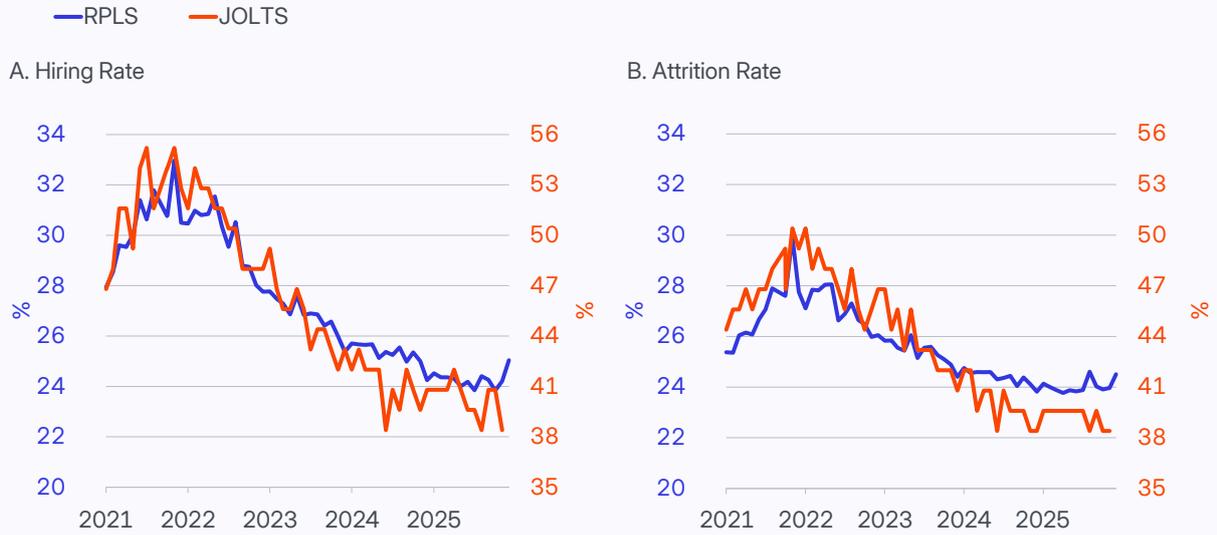
Active Job Postings



Notes: Estimates of active job openings in the US from RPLS (blue line – left axis) and the JOLTS (orange line – right axis). Both series are seasonally adjusted. The latest JOLTS report was released on January 7th of 2026, including data until November 2025. Source: RPLS and BLS JOLTS data. © ifo Institute

Figure 4

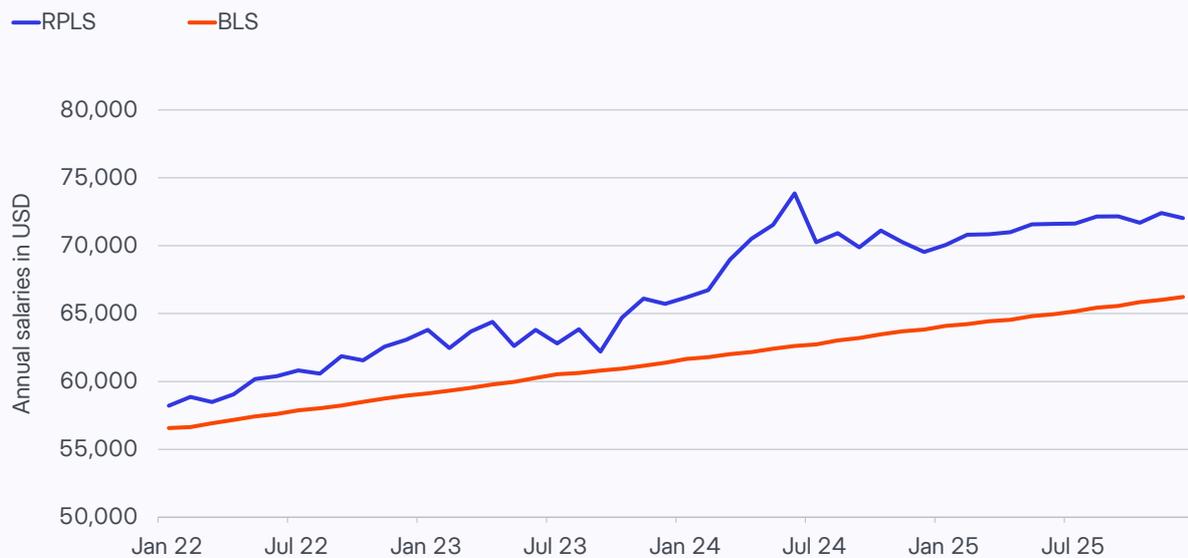
Hiring and Attrition Rates



Notes: Estimates of hiring and attrition rates from RPLS (blue line) and JOLTS (orange line). The hiring (attrition) rate is computed as the inflows (outflow) divided by the total headcount. All series are annualized and seasonally adjusted. The latest JOLTS report was released on January 7th of 2026, including data until November 2025.
Source: RPLS and BLS JOLTS data. © ifo Institute

Figure 5

RPLS Salaries from Job Postings vs. BLS Annualized Wages



Notes: RPLS (blue line) publishes salaries posted in new job postings. BLS (orange line) publishes wages of workers currently in the workforce. Both series are annualized to reflect total annual earnings and are seasonally adjusted.
Source: RPLS and BLS data. © ifo Institute

can be consequential, especially when policy decisions and financial markets rely on preliminary estimates. Below, we compare the magnitudes of RPLS revisions and CES revisions between the first and latest releases of the data over the past four years. RPLS revisions have been roughly half the size of CES revisions in absolute terms.

Job Openings and Employee Turnover

The BLS publishes data on job openings and employee turnover as part of its monthly Job Openings and Turnover Survey (JOLTS) report. JOLTS is a monthly survey of 21,000 establishments that asks about job openings, hirings, quits (voluntary departures), layoffs, and discharges (involuntary separations) (BLS 2025b). JOLTS data is released with significant delays: the release on September 30, 2025, included job openings and turnover data for August 2025.

When comparing RPLS estimates of hiring and attrition to JOLTS estimates, we find that RPLS estimates track JOLTS closely, though with consistently lower levels of both hiring and attrition. For comparability, JOLTS rates are annualized in the figures below by multiplying the monthly rates by twelve. The gap in levels arises from methodological differences. First, JOLTS defines hiring and separations at the establishment level, capturing all flows into and out of payroll jobs, which includes within-firm transitions across establishments, seasonal and temporary hires, and short-term jobs. In contrast, Revelio Labs measures these dynamics at the individual position level, based on observed changes in online professional profiles. As a result, our estimates of hiring and attrition yield systematically lower levels that align with the expectations of HR and corporate leaders.

Salaries From New Job Openings

Job postings are not just a signal of future hiring, but are also a rich source of salary information. The BLS reports earnings gained by workers currently employed, providing a measure of how much workers are earning on average across the economy. In contrast, Revelio Labs tracks salaries from new job postings to provide a forward-looking view of what employers are prepared to pay new hires. Our measure of salary is the nominal annual base salary in USD. About 37 percent of US job postings in our data have a salary value or range.

In order to use salaries from job postings as a proxy for potential earnings of newly hired workers, we apply sampling weights to ensure that the distribution of job postings resembles the national distribution of the workforce. Similar to employment, we construct sampling weights for job

postings using data on employment shares by sector and occupation group from the OEWS. These sampling weights ensure that we measure new salaries across the economy, regardless of demand patterns in a given month. Further, we remove expected seasonal patterns from the weighted annual salary data.

When comparing our estimates of salaries from job postings to annualized earnings from the CES survey, we find that the correlation between the two series over time is 0.91. Salaries from new job postings generally grow faster than the stock of earnings across all workers. The pace of growth of salaries in job postings relative to current workers' salary growth provides an important pulse on the health of the economy. When new salaries grow at a slower rate than current workers' salaries, it is a particularly negative sign for new job seekers.

Policy Implications

In Fall 2025, the US government experienced the longest shutdown in its history – 43 days during which the public did not have access to two consecutive monthly Jobs Reports from the Bureau of Labor Statistics. Yet, the data did not stop flowing. Private-sector data providers, such as ADP and Revelio Labs, continued to publish data showing employment changes, wage changes, and sector-specific shifts. This data provided real-time signals of labor market dynamics when official data was absent, helping policymakers, businesses, and investors maintain continuity in their decision-making.

Having more than one source of economic measurement is highly advantageous. Government institutions may shut down or become politicized or biased. Public labor statistics remain indispensable, and private-sector companies can play a key role in their provision. Rather than viewing private data providers as competitors to national statistical offices, policymakers should view them as vital partners in building a more robust measurement ecosystem. The government should develop a clear framework for the validation of private-sector data, enabling timely and transparent integration during times of disruption.

In the US, the BLS is a century-old institution that has the institutional capacity and expertise to run nationwide surveys and consistently deliver key statistics on the labor market and inflation. However, many economies lack such institutional capacity. Many advanced and emerging market economies face long delays and fragmented data collection processes, which makes producing reliable labor statistics using survey data difficult and costly. However, adoption of and participation on professional networking

sites is growing globally, enabling the labor market to be measured even in countries with fewer sources of survey data. This global coverage of alternative data thus presents a unique opportunity to extend real-time, comparable labor market statistics beyond the US. This is the direction in which we plan to take RPLS: in the near future, RPLS will build on this foundation to expand coverage to other advanced and emerging market countries, adapting our models to local labor markets, helping to fill persistent data gaps, and enhancing cross-country comparisons of labor market data. •

References

Bloomberg (2025), *The BLS Can't Be Replaced by the Private Sector*, <https://www.bloomberg.com/opinion/articles/2025-08-08/the-bls-can-t-be-replaced-by-the-private-sector?smd=undefined> (accessed November 18, 2025).

Bogner, K. and U. Landrock (2016), "Response Biases in Standardised Surveys", *GESIS survey guidelines* 155.

Revelio Labs (2025), *Revelio Public Labor Statistics (RPLS)*, <https://www.reveliolabs.com/public-labor-statistics/> (accessed November 18, 2025).

US Bureau of Labor Statistics (2025a), *Household and Establishment Survey Response Rates*, *BLS Office of Survey Methods Research*, <https://www.bls.gov/osmr/response-rates/> (accessed November 18, 2025).

US Bureau of Labor Statistics (2025b), *Job Openings and Labor Turnover Technical Note*, <https://www.bls.gov/news.release/jolts.tn.htm> (accessed November 18, 2025).