

Measuring What Matters

Context and Recommendations to Improve Reporting on Unemployment Insurance

Matt Weidinger and Amy Simon

April 2024

Key Points

- This report explores the unemployment insurance system's current data and performance management practices, including features designed to inform policymakers and the public about deficiencies in the administration of that system.
- It argues that the program integrity reporting structure is overly focused on established overpayments and reactions without equivalent attention to fraud prevention or resilience measures.
- It suggests areas in which the Department of Labor should modify its current data-reporting requirements to improve transparency and understanding of system performance, particularly around anti-fraud reporting.
- It concludes with a proposal for an aspirational alternative framework, or "scorecard," to better understand and improve system performance in the future.

During the pandemic, improper payments on unemployment benefits led to massive taxpayer losses. And yet recurring questions about the administrative performance of the unemployment insurance (UI) system remain mostly unanswered:

- How do we know that the system meltdown, rampant fraud, and customer service failures that occurred during the pandemic won't happen again when the system is once again under extreme stress?
- How can we better understand the consistent trend of elevated improper payments that has dogged the UI system for years—even when it was not in crisis?
- How do we know which states are operationally prepared for future challenges and which ones remain vulnerable to repeating past failures?

- What do "good states" do differently, particularly in terms of technology and operations?
- Which state-by-state variations are due to specific differences in state law?

Several facts stand out about these concerns. First, these questions are not inherently political; the pain of pandemic failures did not break along partisan lines. Everyone would benefit from making sure the nation's federal and state unemployment benefits system is better prepared for the next time more Americans need timely access to these important, and highly cyclical, benefits.¹

Second, the pandemic experience reflects how a new group of stakeholders—federal taxpayers—bears many of the program's responsibilities. In 1995, a Department of Labor (DOL) Unemployment Insurance Program Letter (UIPL) noted: "The success of the UI system as a whole is to be measured by service to its ultimate customers, principally claimants and employers."² But during the pandemic, American taxpayers bore the burden of several hundred billion dollars of federal general revenue that was directly and brazenly stolen by adversarial nation-states, cybercrime gangs, and sophisticated fraudsters. The opportunity cost of that lost federal revenue was enormous and offsets much of the good that extraordinary federal benefits were able to accomplish.

The interests of federal taxpayers have not traditionally been considered in UI performance operations because past benefits were generally supported by state payroll taxes. However, recent congressional responses to recession and the pandemic emergency indicate that large general revenue–funded federal programs could reappear in response to future recessions. That suggests that the interests of federal taxpayers should be considered more directly and that public transparency on performance is even more warranted. This is particularly true because programs supported by federal general revenue displayed the highest rates of improper payments and fraud during the pandemic.³

There are, of course, ongoing debates about what policy changes are most necessary to improve the system's readiness and performance when it comes to paying benefits. But without reporting reforms, any specific policy change may provide limited evidence of improvement or a decline in such performance. The system's current regime of performance reporting is opaque, and states are not always reporting consistently.⁴ The current reporting structure simply does not encourage the state behaviors or system outcomes that all stakeholders desire.

A Brief History of UI Performance Reporting

At its center, UI reporting measures and structures are designed to shed light on a central question: Does the UI system (and its multiple subprograms) function appropriately for key stakeholders, including benefit claimants and employers?

The nation's unemployment benefits system is a federal-state partnership. The system's administration is federally funded with broad requirements as a baseline, but most tax collection and benefit distribution happen under the terms of state law and via state administration. This creates a somewhat complicated mix of federal and state law and policy, which, in turn, also complicates performance reporting. State laws often have particular quirks or specific definitions that might characterize the same benefit payment with the same facts as an improper payment in one state and not in another state.

Additionally, UI operations function under accumulated layers of statutory, regulatory, and technology policy and other developments stretching back almost 90 years to the program's creation. Temporary federal unemployment benefit programs, which Congress regularly creates during recessions, exacerbate those challenges. These cumulative factors make measuring technical and operational performance a complex endeavor. Although states expend significant time and produce massive amounts of data, these data are not necessarily equally valuable. Over time, reporting burdens are added but rarely reduced.

Amid these competing requirements, there have been cyclical improvements and additions to the UI system's performance reporting regime. The most recent round of revisions was in the early 1990s with the Performance Enhancement Workgroup and later the Performance Enhancement Group. Themes from these and later groups' work, as summarized in a 1995 UIPL, are still relevant:

Even before the [National Performance Review] highlighted the need to "reinvent" government, many within the UI system perceived the need for new and revitalized approaches to ensuring service for the system's customers, animated by a renewed commitment to improving performance across the whole spectrum of UI activities. They shared a definite sense that the commitment to continuous improvement at various points in the system was either missing or had diminished. In certain States, some aspects of performance had lagged for considerable periods of time and corrective action had not proved effective. Federal attention and followup to performance issues varied across the country and by issue. The systems for measuring performance and ensuring program improvement actions needed further integration. In short, there was a need for a more

Table 1. Core Measures, by Area of Administration

Benefits	Program Integrity	Appeals	Tax Administration	Reemployment
 Promptness of first payment Time to determine a claimant's eligibility (nonmonetary determination) Nonmonetary determination quality— non-separations Nonmonetary determination quality— separations 	 Detection of overpayments Measure of improper payments UI overpayment recovery measure 	 Average age of pending lower-authority appeals Average age of pending higher-authority appeals Quality of lower-authority appeals 	 Time to determine new employer status Tax quality (pass/fail only) Effective audit measure 	• Reemployment rate in the second quarter after exit

Note: Program integrity metrics are published annually via the "Integrity Score Card," which gives green, yellow, or red ratings (and trend indicators). See US Department of Labor, Employment and Training Administration, "UI Performance Score Cards," April 18, 2024, https://oui.doleta.gov/ unemploy/score_cards.asp. Source: Authors.

systematic way of looking at all significant areas of performance as a whole, planning selectively and purposefully for change, tracking performance continuously and responding to evidence of both good and bad performance. The context would have to involve a commitment by both Federal and State partners to assume joint responsibility for performance and work together on all phases of performance improvement, from setting priorities and planning to execution and evaluation.⁵

This was also the dawn of the UI Performs system, which still collects much of the data that states report. As described in the Employment and Training Administration (ETA) handbook, this system was designed to provide a better data-collection infrastructure and prioritized sets of reporting requirements, a planning process, and shared commitment from state and federal partners.⁶

These mid-1990s revisions elevated a set of core measures as the key evidence of effective UI operations (Table 1). All states are publicly ranked on these aspects of UI performance.⁷ When states do not meet the acceptable levels of performance, they must address these deficiencies in the State Quality Service Plan (SQSP) report. States also submit most of the underlying non-core measure data to DOL via a public set of monthly, quarterly, and annual reports.⁸

This same set of changes introduced now-standard elements of the UI reporting structure: Benefit Accuracy

Measurement (BAM), Benefits Timeliness and Quality (BTQ), and the biannual SQSP.⁹ These elements were originally designed as continuous improvement exercises; they were to give a wider perspective on system performance than the standard point-in-time data elements used for most of the rest of current reporting requirements.

BAM. To evaluate accuracy in the claims life cycle, states perform an annual statistically representative sampling of claims for the three largest UI programs (Unemployment Compensation, Unemployment Compensation for Federal Employees, and Unemployment Compensation for Ex-Servicemembers). To promote independence and integrity, the state's BAM units are organizationally separate from the UI operations teams. After creating a representative sample of claims, the BAM units evaluate the details of each claim, extrapolate estimates, and use root-cause analysis to determine the error source for making system-wide corrections.¹⁰ Some results of the BAM analysis are used in the core measures metrics as well. These reviews are performed by the state's audit team (separate from the adjudication teams), and the results are then sampled and reviewed again by a combined team of state and federal staff members.

BTQ. This review system, also pulled as a sample from a larger pool of claims, is designed to evaluate the quality of nonmonetary eligibility determinations made

by adjudicators.¹¹ This includes both separation- and non-separation-related issues. As with BAM, the results are also reviewed by both state and federal teams.

SQSP. States must submit biannual comprehensive management plans to the DOL that include both descriptive narrative and corrective action plans, when required. The SQSP is also the vehicle for the annual administrative funding budget and grant documentation. SQSPs and their sub-plans, including Integrity Action Plans, are not made public.

Gaps in Program Integrity Performance Reporting

In the wake of the pandemic experience, there is an additional set of performance or operational quality questions, particularly around program integrity, that cannot be directly answered with current data. The heart of the monthly program integrity reporting is the ETA 227 form (Overpayment Detection and Recovery Activities).¹² Table 2 shows the data elements that this form collects.

While that is an extensive set of data elements, ETA 227 leaves a number of questions unanswered (and unanswerable), including:

- What is the total number of cases at a high level: attempted fraud, impostor fraud, or intentional misrepresentation?
- What percentage of initial claims by program are identified as high risk for identity fraud?
- What is the source of the crossmatch used to establish fraud or an improper payment, if any?
- How many initial claims are backdated more than two weeks? What percentage of backdated claims are valid?
- What is the number and value of claims with suspected fraud but without an established over-payment?
- What share of initial claims originate from international or out-of-state IP addresses?

- Does the state have adequate (according to relevant DOL guidelines) defenses against individual identity theft?
- How long does it take for the state to resolve a case flagged for identity theft?
- What recourse does the state offer legitimate users wrongly flagged for identity-related fraud risk?
- Does the state have adequate defenses against fake employer and business identity theft?
- How many established cases of account takeover schemes, including stale claim takeover, occurred in the past month?
- Does the state have established insider fraud defenses?
- What is the ratio of state staff time spent on fraud investigation, detection, or prosecution relative to its suspected or actual fraud rate?
- How many benefits are paid to "employees" of fictitious employers? What is the ratio of benefits to taxes paid in these cases?
- Does the state have and use a fraud case management or investigation platform?

The current program integrity reporting structure is focused on *established overpayments and reactions* without insight on fraud detection, fraud prevention, or fraud resilience measures. Although measuring prevention is inherently difficult, there are potential proxy measures and certainly ways to better measure preparedness and agility.

Not all of these gaps are due to DOL or state-caused problems. Some of the most recent temporary pandemic programs had very poor statutory design that facilitated fraud and made reporting significantly more difficult and less meaningful. Congress also declined to approve bipartisan program integrity measures proposed as part of almost every DOL budget during the past decade.¹³

Table 2. Data Elements in the ETA 227 Form

Category	Data Element
	Each data element includes the number of schemes, number of cases by program type, and relevant dollar value.
Overpayments Established—Causes	 Fraud (total) Multi-claimant schemes Agency employee benefit fraud High-dollar overpayments from fraud Non-fraud total (by reason) High-dollar overpayments, not from fraud Penalty weeks*
	Each data element includes the number of cases investigated and the number of cases and dollar value broken out by fraud and non-fraud cases.
Overpayments Established—Methods of Detection	 Controllable methods Wage and benefit crossmatch Interstate benefit crossmatch National directory of new hires State directory of new hires Multi-claimant scheme systems Special project** Other Noncontrollable methods Total
	Each data element includes the dollar amount, broken down by cause (fraud or non-fraud), and the program.
Recovery/Reconciliation	 Outstanding at the beginning of the period Recovered (by method) Waived Written off Additions Subtractions Receivables removed at the end of the period Outstanding at the end of the period Recovered for other states
Criminal/Civil Actions	 Fraud cases pending prosecution Fraud cases referred for prosecution Referred fraud cases related to agency employee fraud Cases refused for prosecution Convictions obtained Convictions due to employee fraud Cases referred for civil action Number of civil actions obtained
Aging of Benefit Overpayment Accounts	The monetary value of overpayments, by program and amount of time elapsed

Note: * This is to distinguish between fraud itself and additional claimed weeks that are considered penalty weeks. From the reporting explanation: "Benefits properly paid that were retroactively included in a penalty assessed for a fraudulent overpayment of another week(s) and are, therefore, subject to recoupment." ** "This category is reserved for special detection methods/projects using new methods or technologies." Source: US Department of Labor, "UI Reports Handbook No. 401: ETA 227 Overpayment Detection and Recovery Activities," https://oui.doleta.gov/ dmstree/handbooks/401/401_3c2a.pdf.

What Changes Could Be Made to the Program Integrity Reporting Structure?

Defining success in fraud detection, prevention, and investigation is not straightforward for any organization, public or private, that is fighting ongoing and evolving fraud threats. For many state agencies, success might simply be to adjust a fraudster's cost-benefit ratio enough to deter fraud.

Success is certainly not determined by any one intervention, any one tool, or any one process. It consists of an alert, responsive, and problem-solving culture that takes service delivery and fraud prevention equally seriously.

In an ideal world, state agencies would have a more accurate understanding of the total cost of fraud in their state and be able to make data-driven decisions about the value of various fraud interventions.

Between now and that ideal world, DOL should evaluate the current reporting structure based on three key principles: transparency, incentives, and usability.

- **Transparency.** Does this data element look at root causes or outcomes? Does it provide the right level of granularity to evaluate state performance and compare states? Does it give an accurate sense of state operations?
- **Incentives.** Does reporting this data element encourage the desired outcomes of state preparedness (across multiple domains), improved customer experience, and efficient operations?
- **Usability.** Is publishing this data element useful in proportion to the difficulty of collecting and reporting it? Does reporting this item improve customer experience? Does this item have enough relevance to make a useful comparison across state structures?

DOL should also evaluate when and how there should be consequences for *not* reporting particular information. This is not typically an issue in the regular UI program, but it has been a more frequent problem in temporary programs such as Pandemic Unemployment Assistance, which had fully federally funded benefits and administration but was run by state agencies. Given this funding relationship, DOL could consider imposing more stringent consequences for non-reporting for federally funded benefits compared with statefunded benefits.

On program integrity specifically, the current overpayment-centric reporting system does not get beneath the surface on the state's anti-fraud functionality and preparedness. Given that, we propose requiring data on use case readiness across the fraud life cycle. Although subject matter expert stakeholders should develop the technical specifics of such use cases, as a starting point, Table 3 shows our suggestions for broad categories and sample use cases.

Scorecard Proposal

Any shift to an outcomes-based reporting structure also calls attention to the disparate and dispersed nature of currently available information. Getting an at-a-glance understanding of a state's position or general performance, especially relative to its funding or its peer states, is not straightforward, and indeed, given present technology and system realities, it is decidedly aspirational. However, we believe that developing and maintaining a focus on such capacity will help drive the sort of continuous improvements needed to properly serve future benefit claimants and taxpayers alike. As a starting point for that at-a-glance perspective, we propose that the following four domains of UI performance evaluation be elevated and simplified: benefit administration, tax and trust fund administration, technology administration, and sustainability and resilience (Table 4). The metrics under each domain could come from existing reporting or from analysis based on existing reporting. Some new metrics would need to be developed and tested before public use. The goal is to provide relevant, usable data in context.

By focusing on capabilities and outcomes in public reporting, there is less need to program for or report on the myriad of less-useful, less-used data elements currently being collected. Although apples-toapples comparisons of state UI operations can be problematic, there should be ways to provide more insight. States could be assigned scores, grades, or stoplight color rankings based on the four categories and multiple subcategories. They could be grouped by similar size or economic conditions to

Table 3. Broad Categories and Sample Use Cases

Pre-Claim Use Cases	Initial Claim and Identity Use Cases	Post-Eligibility Determination Use Cases	Other Use Cases
 The agency can distinguish between bots and humans. The agency monitors and responds to user-specific device fingerprint or browser information. The agency can identify and stop most high-risk incoming traffic. The agency identifies and stops credential stuffing attacks without disrupting public access. The agency can scale to meet demand without reducing pre-claim fraud evaluations. The agency can add, update, or modify pre-eligibility behavior or data risk measures within 24 hours. 	 The agency incorporates pre-claim identity and behavior risk into its initial claim, identity, and adjudication processes. If users cannot verify their identities online, the agency offers alternative methods, including human interaction, to complete the process. The agency can identify synthetic identity elements. The agency can crossmatch identity elements with other states and standard administrative data sources. The agency has equivalent security measures for in-person, phone, and online filing methods. The agency uses a multifaceted approach to validate identities, supporting documentation, and initial eligibility. The agency can seamlessly move highrisk identities to a case investigation platform, tool, or workflow for further investigation. 	 The agency can identify suspicious behavior patterns of existing claimants. The agency can validate bank account ownership. The agency can pause payments on the highestrisk cases at the earliest appropriate and legally defensible time. The agency can seamlessly move high-risk behaviors to a case investigation platform, tool, or workflow for further investigation. 	 The agency has a functional insider threat detection program. As part of insider fraud detection, the agency regularly cross-checks internal agency employee and contractor (when relevant) data with claimant and employer identity information and takes action based on results. The agency identifies suspicious claimant or employer behavior patterns and defines ongoing parameters to aid in detection, prevention, and investigation. The agency can and does refer case data and supporting documentation to law enforcement. The agency can communicate accurately and quickly with financial services institutions about suspected fraud, including the ability to reclaim fraudulently paid benefits from financial institutions. The agency can identify, stop, and prevent high-risk employer-side behavior related to fictitious employer accounts. As part of procurement, the agency can require baseline cybersecurity, data security, and insider threat prevention.

Source: Authors.

ensure more fair comparisons. This could be done without disclosing sensitive technology or specific fraud-prevention practices.

On the operational improvement front, these metrics could highlight potential problems for federal support. When states consistently fall in the bottom of their cohort or have consistent downward trends, DOL should be able to mandate technical assistance and additional relevant reporting metrics. Any reporting changes would have significant financial and technical implications for states whose systems would need to be updated or reprogrammed to handle different reporting requirements. To address this, the DOL should consider a onetime infusion of reporting-specific funding to develop, test, and implement the revised structures.

Table 4. Scorecard Proposal

Benefit Administration	Tax and Trust Fund Administration	
 Benefit accuracy Benefit timeliness Appeal quality User experience and accessibility Special programs Reemployment quality 	 Trust fund financial management Experience rating and charging accuracy Appeal quality User experience 	
Technology Administration	Sustainability and Resilience	
 Cybersecurity posture Identity verification Fraud life cycle tools and practices Cloud and scalability Usability 	 Data management, reporting, and performance accountability quality Administration funding trends Procurement quality State law changes 	

Source: Authors.

Conclusion

The current data-reporting structure for the UI system does not align with the expectations of the program or stakeholders' evolving needs. It does not incentivize the behaviors and, more importantly, outcomes that reflect the full definition of mission success. Rather than retaining the current structure and adding new fraud life cycle elements on top, the DOL should consider condensing and reorienting reporting structures around the most important desired outcomes. This will require collaborative work with key legislative, state agency, technology partner, and end-user stakeholders, but a better-functioning UI system would also pay dividends across this stakeholder ecosystem.

About the Authors

Matt Weidinger is a senior fellow and Rowe Scholar in opportunity and mobility studies at the American Enterprise Institute, where his work is focused on safety-net policies, including cash welfare, child welfare, disability benefits, and unemployment insurance.

Amy Simon is the owner and founder of boutique consulting firm Simon Advisory. She previously served as acting deputy assistant secretary for the Employment and Training Administration (ETA) at the US Department of Labor. Her portfolio included unemployment insurance, workforce investment, trade adjustment assistance, and ETA's regional offices.

Notes

1. For a detailed review of unemployment benefit payments, including significant improper payments, during the pandemic, see Matt Weidinger and Amy Simon, "Pandemic Unemployment Fraud in Context: Causes, Costs, and Solutions," American Enterprise Institute, January 29, 2024, https://www.aei.org/research-products/report/pandemic-unemployment-fraud-in-context-causes-costs-and-solutions.

2. Mary Ann Wyrsch, "Unemployment Insurance Program Letter No. 41-95," US Department of Labor, Employment and Training Administration, Office of Unemployment Insurance, August 24, 1995, 5, https://oui.doleta.gov/dmstree/uipl/uipl95/uipl_4195.htm.

3. US Department of Labor, Employment and Training Administration, Office of Unemployment Insurance, "PUA Improper Rate Report," August 21, 2023, https://oui.doleta.gov/unemploy/pdf/Pandemic_Unemployment_Assistance_Improper_Payment_

Rate_Report.pdf; US Government Accountability Office, "Unemployment Insurance: DOL Needs to Address Substantial Pandemic UI Fraud and Reduce Persistent Risks," February 8, 2023, https://www.gao.gov/products/gao-23-106586; and US Government Accountability Office, "Unemployment Insurance: Estimated Amount of Fraud During Pandemic Likely Between \$100 Billion and \$135 Billion," September 12, 2023, https://www.gao.gov/products/gao-23-106696.

4. This was particularly true during the pandemic for fraud-related performance measures. See US Department of Labor, Employment and Training Administration, Office of Unemployment Insurance, "ETA 902P: Pandemic Unemployment Assistance Activities," February 29, 2024, https://oui.doleta.gov/unemploy/DataDownloads.asp.

5. Wyrsch, "Unemployment Insurance Program Letter No. 41-95."

6. US Department of Labor, Employment and Training Administration, Office of Unemployment Insurance, ET Handbook No. 336, 17th Edition: Unemployment Insurance State Quality Service Plan (SQSP) Planning and Reporting Guidelines, July 25, 2005, https://oui.doleta.gov/dmstree/handbooks/336/17th/hb336-chapter1.pdf.

7. US Department of Labor, Employment and Training Administration, Office of Unemployment Insurance, "UI Performs Score Cards," December 2023, https://oui.doleta.gov/unemploy/score_cards.asp.

8. For data maps of each report, see US Department of Labor, Employment and Training Administration, Office of Unemployment Insurance, "Data Downloads—Data Maps," 48, https://oui.doleta.gov/dmstree/handbooks/402/402_4/4024c6/ 4024c6.pdf.

9. The low likelihood of consequences based on State Quality Service Plan reporting has effectively reduced it to an annual paperwork requirement rather than a strategy or continuous improvement tool. See US Department of Labor, Employment and Training Administration, Office of Unemployment Insurance, "Unemployment Insurance (UI) Benefit Operations Self-Assessment Tool: Adjudications and Benefits Timeliness and Quality Reviews (BTQ)," June 2016, https://oui.doleta.gov/unemploy/ui_benefit_operations/self_assessment_tool/june_2016/adjudications_btq_instructions_v2.pdf; and US Department of Labor, Employment and Training Administration, Office of Unemployment Insurance, *ET Handbook No. 336, 17th Edition.*

10. Benefit Accuracy Measurement's fairly small sample size has led to some criticisms that the measure is not reliable. With the advent of data analytics tools and machine learning, it may be practical and possible to drastically expand the scope of claims considered.

11. US Department of Labor, Employment and Training Administration, "Agency Information Collection Activities; Comment Request; Benefits Timeliness and Quality (BTQ) Review System," *Federal Register* 84, no. 67758 (December 11, 2019): 67758–59, https://www.federalregister.gov/documents/2019/12/11/2019-26658/agency-information-collection-activities-comment-request-benefits-timeliness-and-quality-btq-review.

12. US Department of Labor, Employment and Training Administration, Office of Unemployment Insurance, "UI Reports Handbook No. 401: ETA 227 Overpayment Detection and Recovery Activities," https://oui.doleta.gov/dmstree/handbooks/ 401/401_3c2a.pdf.

13. The fiscal year 2025 congressional budget justification for State Unemployment Insurance and Employment Service Operations (SUIESO) has many of these same SUIESO budget proposals as previous Department of Labor budgets. See US Department of Labor, FY 2025 Congressional Budget Justification Employment and Training Administration: State Unemployment Insurance and Employment Service Operations, https://www.dol.gov/sites/dolgov/files/general/budget/2025/CBJ-2025-V1-07.pdf; and US Department of Labor, "Archived Budgets," https://www.dol.gov/general/budget/archive.

© 2024 by the American Enterprise Institute for Public Policy Research. All rights reserved.

The American Enterprise Institute (AEI) is a nonpartisan, nonprofit, 501(c)(3) educational organization and does not take institutional positions on any issues. The views expressed here are those of the author(s).