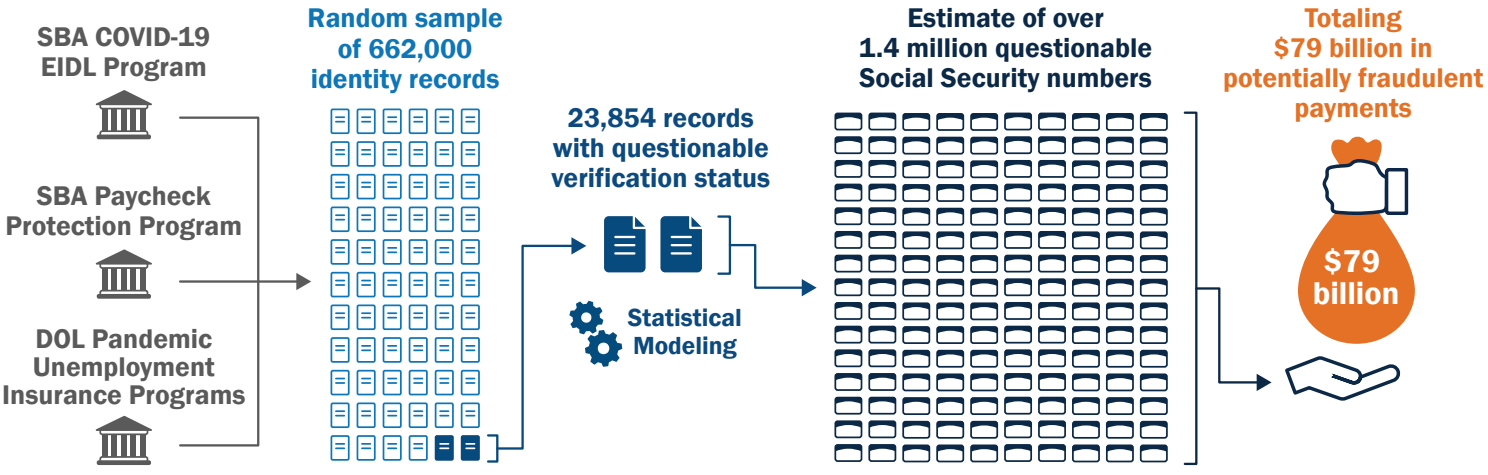
FRAUD PREVENTION ALERT: Pre-Award Vetting Using Data Analytics Could Have Prevented Over $79 Billion in Potentially Fraudulent Pandemic Relief Payments

The PRAC Analyzed Applications Across Major Pandemic Relief Programs to Identify Potential Fraud Using Stolen or Invalid Social Security Numbers

As part of our independent oversight of the $5 trillion in COVID-19 relief programs, the Pandemic Response Accountability Committee (PRAC) conducts in-depth data analysis across multiple federal programs to identify potential fraud, waste, abuse, and mismanagement and to provide agencies with lessons learned to prevent future program fraud.

This PRAC fraud prevention alert focused on some of the largest pandemic relief programs: the Small Business Administration's (SBA) COVID-19 Economic Injury Disaster Loan (COVID-19 EIDL) program and Paycheck Protection Program (PPP), and the Department of Labor's (DOL) pandemic-related Unemployment Insurance (UI) programs. In 2023, the SBA Office of Inspector General (OIG) and the DOL OIG estimated that the total amount of fraud and improper payments for these programs is nearly $400 billion.[1]

In this alert, the PRAC estimates the amount of potential fraud across these programs stemming from the use of stolen or invalid Social Security numbers (SSNs), and illustrates how pre-award vetting using the PRAC's data analytics tools could have mitigated this risk.



The PRAC has previously issued numerous reports detailing the widespread identity fraud that occurred in connection with pandemic relief programs and the need for the federal government to address the issue.[2]

Leveraging our cross-agency partnerships and applying innovative data analytics tools and resources, the PRAC estimates that these pandemic programs disbursed approximately $79 billion in potential fraudulent payments due

1 SBA OIG "COVID-19 Pandemic EIDL and PPP Loan Fraud Landscape" report, and DOL "OIG Oversight of the Unemployment Insurance Program" report.
2 Examples of PRAC reports include Key Insights: Identity Fraud Reduction and Redress in Pandemic Response Programs and Best Practices and Lessons Learned from the Administration of Pandemic-Related Unemployment Benefits Programs.

to the use by applicants of over 1.4 million potentially stolen or invalid SSNs. The findings in this fraud prevention alert overlap with some of the findings of prior DOL OIG and other OIG reports.[3]

PRAC's Efforts to Estimate the Amount of Potentially Fraudulent Payments in the COVID-19 EIDL, PPP, and UI Programs

To conduct our analysis, the PRAC's Pandemic Analytics Center of Excellence (PACE) randomly sampled 662,000 identity records from a population of 67.5 million funded applications across the COVID-19 EIDL, PPP, and Pandemic UI programs.[4] Using our authorities in the Coronavirus Aid, Relief, and Economic Security Act of 2020 (CARES Act), and pursuant to a Memorandum of Understanding (MOU) with the Social Security Administration (SSA), we shared the sampled SSNs and the associated names and dates of birth (DOB) from the 662,000 records with the SSA for verification.

Using those results, our data scientists then employed a statistical sampling method to estimate that 1.4 to 1.5 million potentially stolen or invalid SSNs were used to obtain benefits from the programs.[5] These SSNs were either never issued by SSA or had mismatched identifying information (name and/or DOB) when compared to SSA records, indicating that they were potentially stolen or being used without authorization.[6] The PRAC estimates that $79.41 billion to $79.54 billion was disbursed to applicants using these potentially stolen or invalid SSNs.

As detailed later in this fraud prevention alert, this type of fraud is readily preventable using the authorities and analytics platform that Congress has provided to the PRAC. Moreover, having an identity verification capability using a recognized source such as SSA enables agencies to effectively focus their efforts on a significantly smaller subset of the applications for further examination (approximately 4 percent of the records in our analysis). Additionally, identifying applications that are misusing SSN, name, and DOB information enables the PRAC to use its analytics capabilities to determine if fraudsters have improperly obtained benefits from other federal programs using that identifying information or other linked identifying information (such as IP address, bank account, and address).

Without the PRAC's unique data analytics capabilities and the tools authorized by Congress, we would not have been able to identify this significant amount of potential fraud across these programs. However, in order for these unique capabilities to be effective, the PRAC needs to be able to obtain accurate and complete data from agencies. For example, because PPP applicants were not required to provide DOB information, the PRAC was able to obtain such information for only approximately 13 percent of PPP loan applicants.[7] Requiring applicants to provide data like DOBs is essential for detecting identity fraud, conducting oversight analysis, and strengthening program integrity.

3 For instance, the potentially invalid SSNs we identified for this fraud prevention alert matched DOL OIG's previous analysis. Specifically, 95 percent of questionable SSNs in our random sample for Pandemic UI were previously identified by DOL OIG. Additionally, 23 percent were already included in DOL OIG suspected fraud totals in their prior published reports.

4 An identity record contains key identifying details such as an individual's SSN, first name, last name, and date of birth submitted in their applications for the COVID-19 EIDL, PPP, or Pandemic UI programs. While individuals may have applied to various programs or submitted multiple applications within the same program, we consolidated their information into a single record to be verified by the SSA and avoid duplications. Additionally, a single application can include multiple individuals. For these reasons, we use the term "records" rather than "applications" throughout this fraud prevention alert for clarity.

5 The number of individuals that used these potentially stolen or invalid SSNs to apply for funds across these programs is likely to be less than 1.4 to 1.5 million given that some fraudsters are presumably responsible for the use of more than one of the potentially invalid SSNs.

6 We would expect some false positives in these verification results. Performing an automated match is subject to error, including unrecognized nicknames or misspellings, transposed numbers, and data entry and clerical mistakes. For example, we observed examples in which records were flagged as an incorrect name when an individual's first name and last name were transposed. We cannot estimate the exact proportion of false positives without an in-depth understanding of the verification algorithm used by SSA or/and without addressing the data cleanliness issues in the pandemic program data. To maintain the security of its verification process, SSA does not release the agency's matching logic concerning how it verifies SSNs.

7 Date of birth was not required on the PPP application form; however, PPP lenders may have included the applicant DOB in the loan submission to the SBA.

Further, as evidenced in this fraud prevention alert, applying our innovative tools and the subject matter expertise of our data scientists can enhance information-sharing and identity verification **before** disbursing funds, which can strengthen program integrity, ensuring funds are paid to legitimate applicants, as Congress intends, including during future federal emergency programs, such as natural disasters or financial crises, or regular annual appropriations.

Details of our Analysis

How We Sampled and Verified Records from Pandemic Programs

To conduct this analysis, we randomly sampled 662,000 records from a population of 67.5 million funded applications across COVID-19 EIDL, PPP, and Pandemic UI programs. The sample sizes for COVID-19 EIDL, PPP, and Pandemic UI source tables were set to 256,000, 256,000, and 150,000, respectively. We only considered individual applicants who received funding. Each individual and transaction combination had an equal probability of being sampled from its respective programs.

Using the legal authorities included in the CARES Act of 2020, and pursuant to an MOU between the PRAC and SSA, the PRAC requested that SSA verify the validity of the SSNs supplied in the 662,000 records. The PRAC did not request or obtain SSN data from SSA in connection with this analytics review. Rather, as described below, we requested that SSA simply verify information that applicants had provided in those COVID-19 EIDL and PPP applications and Pandemic UI claims, which SSA performed and then provided us with the results. Specifically, the PRAC provided SSA with the SSN, name, and (where available) DOB information from the randomly selected 662,000 records. It took SSA less than one week to provide these responses to the PRAC for the 662,000 records.

Specifically, for the 662,000 records in our random sample set, we asked SSA to verify the following:

1. Is the SSN valid?
2. If the SSN is valid, does the name associated with the SSN on the application match SSA records?[8]
3. If the SSN is valid, does the date of birth associated with the SSN on the application match SSA records?
4. Is the SSN used on the application associated with a deceased individual?

SSA processed the data and provided the PRAC with an identity verification code for each record (including SSN, name, DOB) they were asked to verify. The codes provided by SSA indicated that the identifying information (SSN, name, DOB) provided by applicants in almost 24,000 of the 662,000 did not fully match the information in SSAs records, indicating potential fraud (See Table 1 below for a breakdown by verification code).

8 As noted above, SSA does not share its matching methodology given the sensitive nature of the data and to maintain the security of its SSN verification process. This includes its matching methodology concerning maiden names, hyphenation issues, or other matching scenarios to determine the verification outcome.

Table 1. Disbursed COVID-19 EIDL and PPP Loans, and Pandemic UI Claims Associated with SSNs with Verification Codes Indicating Potential Fraud

SSA Identity Verification Codes	*PPP*	*COVID-19 EIDL*	*Pandemic UI*	*Total*
SSN Not Issued	1,509	482	139	2,130
SSN with Incorrect Name	11,476	5,603	1,852	18,931
SSN with Incorrect Date of Birth	541	1,803	439	2,783
SSN Not Verified – Other Reason[9]	1	5	4	10
Subtotal – Questionable SSNs	**13,527**	**7,893**	**2,434**	**23,854**
SSN Matched SSA's Records	242,473	248,107	147,566	638,146
Total	**256,000**	**256,000**	**150,000**	**662,000**

Source: PRAC analysis of SSA's verification of the randomly sampled SSNs and associated names and DOBs resulting from SSA's Enumeration Verification System (EVS).

In addition, SSA informed the PRAC that 11,514 of the 638,146 fully verified sample records (the SSN, name, and DOB provided by the applicant fully matched the information in SSA's records) were associated with individuals who were deceased as of the date that SSA responded to the PRAC's request.[10] We have not included those SSNs and the associated loans in this fraud prevention alert because we first need to determine whether those individuals were alive at the time the COVID-19 EIDL or PPP applications or Pandemic UI claims were submitted. Because SSA told us that it is not permitted by law to provide date of death information to the PRAC, determining date of death requires the PRAC to obtain information from the Treasury Department's Do Not Pay service for these deceased individuals. However, the need to confirm date of death does not impact this report's finding that 23,854 records were potentially fraudulent because the identifying information provided by those applicants did not fully match SSA's records.

How We Estimated the Amount of Potential Fraud Within and Across Pandemic Programs

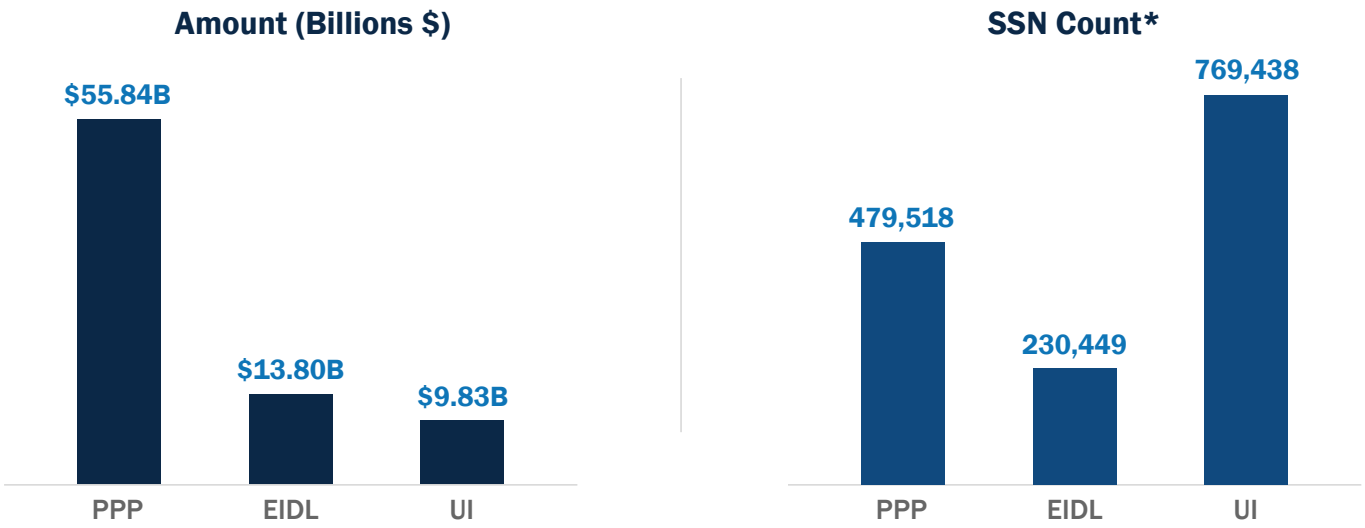
As a result of the SSA's verification process, the PRAC was able to determine that 23,854 of the randomly selected records were potentially fraudulent. We then applied statistical modeling that involved taking repeated random samples, with replacement, followed by a subsequent analysis to infer the entire population for two parameters—total payment amount at risk and potentially stolen or invalid SSN count.

Overall, our statistical modeling estimated that the two agencies disbursed $79 billion to recipients whose applications used potentially stolen or invalid SSNs.[11] Specifically, as indicated in Graph 1 below, the PRAC estimated this resulted in disbursements of $55.8 billion, $13.8 billion, and $9.8 billion in COVID-19 EIDL, PPP, and Pandemic

9 SSA provided no additional information on the verification code "SSN Not Verified – Other Reason."
10 SSA can only provide death indicator information for fully verified SSNs.
11 As noted above, this estimated amount of potential fraud across COVID-19 EIDL, PPP, and Pandemic UI overlaps in part with fraud identified by previous reports issued by SBA OIG, DOL OIG, and the U.S. Government Accountability Office (GAO), which highlighted identity theft as one of the leading risk factors in pandemic fraud. The $79 billion is a result of our statistical estimate and therefore we do not have the actual records to cross reference other OIGs' findings to determine the exact overlap.

UI programs, respectively. While the estimated count of potentially stolen or invalid SSNs used in applications for the Pandemic UI programs is higher than those for the other two programs, the total dollar amount of potential fraud is lower than the other two programs. One possible explanation could be the higher-volume and lower-amount nature of individual Pandemic UI claims when compared with COVID-19 EIDL and PPP applications.

Graph 1. Estimated COVID-19 EIDL, PPP, and Pandemic UI Payments Associated with SSNs with Verification Codes Indicating Potential Fraud



* Note: The amounts and counts shown here represent the population totals estimated from verified samples. The cross-program SSN total is not a direct sum of program SSN counts, as there are SSNs that were used for applications in multiple programs.

Statistical Considerations and Assumptions

Estimating the amount of potential fraud attributed to particular applicants is complex due to scenarios where multiple applicants could appear on the same pandemic loan application, or a single applicant could apply to multiple pandemic programs. Due to this complexity, we estimated the total amounts of potential fraud taking into account the combination of program participation, co-applicants, and funding amounts simultaneously.

Since the underlying true population distributions for each of these are complex, we employed a "Bootstrap Sampling" statistical approach.[12] We transformed transactions to individual applicant records by attributing amounts to applicants proportionally distributed based on the number of co-applicants, then summing funding amounts associated with each applicant.

For example, if a COVID-19 EIDL application with a payment amount of $10,000 had five individual owners, the transformation would create five records listing each individual owner with $2,000 as the funding amount.[13] We then resampled from SSA-verified sample records, with replacement, to reproduce the record-count of the full population and estimate the total payment amount for each verification category. The resulting categorized counts and amounts were used to calculate our population estimates within each program and their corresponding confidence intervals.

By aggregating results across all bootstrap samples, we obtained more robust and reliable estimates, highlighting potential fraud without having to make any rigid assumptions about the data.

12 See An Introduction to the Bootstrap, by B. Efron & R.J. Tibshirani.
13 In COVID-19 EIDL applications, an individual owner is defined as (a) Proprietor, or (b) Limited partner or LLC member who owns 20 percent or more interest and each general partner or managing member, or (c) Stockholder or entity owning 20 percent or more voting stock.

The Amount of Potentially Fraudulent Payments Identified in this Fraud Prevention Alert Likely Would Have Been Higher than $79 Billion Had SBA Required PPP Applicants Provide DOB Information

As noted previously, PPP applicants were not required to provide DOB information and 87 percent of our sampled PPP records did not include DOB information (i.e., 221,951 out of 256,000 PPP records). Of those 221,951 records with missing DOB information, approximately 95 percent were associated with a blank verification code, indicating that the SSN and name on the application matched SSA's SSN and name records. However, for the sampled PPP records that included DOB information, the percentage of records associated with a blank verification code (meaning all three data indicators on the records—SSN, name, DOB—matched SSA's data) dropped to 93 percent.

Had the SBA required PPP applicants provide DOB information in the PPP application, as COVID-19 EIDL and Pandemic UI programs did, we would expect to have identified even more potentially stolen or invalid SSNs, as well as higher potentially fraudulent payment amounts in our PPP estimates, because there would have been one more data point (i.e., DOB) to check against SSA's records.

Potential Synthetic IDs Used to Apply for Pandemic Programs

A portion of the potentially stolen or invalid SSNs we uncovered could be linked with Synthetic Identity Fraud, which involves creating an artificial identity by combining real (often stolen) information, such as an SSN, with real or invented details, such as a made-up name and date of birth. This is different from traditional identity theft, in which fraudsters impersonate a victim using real personal information to apply for pandemic loans or to submit unemployment insurance claims.

It is challenging to quantify how much of the $79 billion we estimated in potential fraud is associated with synthetic IDs, in part because SSA only returns verification codes for each processed record without the underlying name and DOB information in their system of records. Estimations of this type of criminal activity would require knowledge of both the incorrect record and the true data associated with an SSN.

The PRAC's Innovative Graph Analytics Approach to Find Common Connections Between Fraudsters

The PRAC will continue to conduct further analysis using SSA-verified results for lead identification efforts, including by using graph technology to find connections among entities that used shared attributes on applications, such as common SSNs, Tax IDs, phone numbers, and email addresses.

Graph technology offers new, innovative methods of uncovering organized crime, fraud rings, and other complex schemes with a high level of accuracy through advanced contextual link analysis. Our graph analytics capabilities examine relations among entities and determine strength and direction of relationships between entities and their associated attributes. This enables the PRAC to generate actionable leads to share with our federal OIG and law enforcement partners.

Ensuring Program Integrity and Preventing Payments to Fraudsters

In addition to our ongoing law enforcement activities, we urge benefit programs to continue to work with SSA to explore information-sharing agreements that will facilitate pre-award identity verifications across grant, loan, and benefit programs vulnerable to identity-based fraud. Program offices and oversight organizations need accurate and

complete data to ensure that funds are being used as Congress intended. To assist identity verification, we urge program offices to require the collection of DOB information in benefit applications.

When program guardrails were removed during the pandemic, a substantial amount of funds were rapidly disbursed without proper identity verification. Implementing pre-award verification helps streamline the vetting process before disbursal, preventing fraudulent payments from going out and ensuring that funds are disbursed with additional program integrity controls.

Advantages of Implementing SSN Verification Agreement Early to Strengthen Program Integrity – Protecting Taxpayers and Honoring Congressional Intent

SSA is a recognized source to provide identity verification for a variety of federal programs. However, as noted in our January 2023 Fraud Alert, the process to implement new SSN verification agreements among agencies and address legal questions regarding the permissibility of information-sharing can be lengthy.[14]

Prior to the next natural disaster, health crisis, financial failure, or other emergency, it is vital for program administrators to establish SSN verification agreements in a timely manner so that an information exchange can be set up as a part of program integrity controls to prevent fraud, waste, abuse, and mismanagement and protect taxpayer funds from improper payments, as Congress intends in creating such emergency relief programs.

Pandemic Program Background

Pandemic Unemployment Insurance Programs

Reacting to the job losses caused by the pandemic, Congress created five new federally funded, temporary unemployment insurance programs, designed to broaden eligibility, increase benefit amounts, and extend the duration of benefits:

- Pandemic Unemployment Assistance (PUA): provided benefits to individuals who were not eligible for regular unemployment compensation, such as self-employed workers, independent contractors, and gig workers. PUA offered up to a maximum of 79 weeks of benefits.
- Federal Pandemic Unemployment Compensation (FPUC): provided an additional $600 per week to individuals receiving benefits from eligible unemployment compensation programs until July 31, 2020.[15] Later, it was resumed and reduced to $300 per week.
- Pandemic Emergency Unemployment Compensation (PEUC): extended unemployment benefits for those who had exhausted their regular unemployment benefits. With all the legislative extensions, claimants could receive up to 53 weeks of PEUC payments.
- Mixed Earner Unemployment Compensation (MEUC): offered an extra $100 per week on top of the weekly benefit amount and the $300 FPUC benefit to those who earned at least $5,000 in self-employment income in the most recent taxable year before applying for regular unemployment benefits.

14 See PRAC Fraud Alert: PRAC Identifies $5.4 Billion in Potentially Fraudulent Pandemic Loans Obtained Using Over 69,000 Questionable Social Security Numbers, January 2023. According to SSA, it conducts over two billion SSN verifications annually. These exchanges include verifications specifically authorized under statutorily mandated programs (e.g., E-Verify) and consent-based processes (e.g., eCBSV). They also include disclosures made in accordance with the Privacy Act, Section 1106 of the Social Security Act, and SSA's privacy regulations at 20 C.F.R. § 401.150, by which SSA will provide information to another federal agency when the information is necessary to determine eligibility in a health or income maintenance program which is compatible with SSA's Social Security programs.

15 UI programs eligible for FPUC were regular UI, Unemployment Compensation for Federal Employees, Unemployment Compensation for Ex-Servicemembers, Pandemic Emergency Unemployment Compensation, Pandemic Unemployment Assistance, among others.

- Lost Wages Assistance (LWA): was funded by Federal Emergency Management Agency (FEMA) to provide $300 per week to supplement unemployment benefits to eligible claimants.

During the UI pandemic period, March 27, 2020, through September 6, 2021, as defined by DOL, more than $888 billion in total federal and state UI benefits were paid to applicants.[16]

The Pandemic UI programs were susceptible to fraud, waste, abuse, and mismanagement due to the unprecedented volume of claims and the use of antiquated state Information Technology (IT) systems, along with process- and technology-related challenges that state workforce agencies faced to quickly implement the new UI programs.[17]

The expanded benefits required states to implement major changes to their UI IT processing systems. Faced with challenges from IT modernization efforts, states experienced delays in implementing new UI programs and associated antifraud controls, which hampered their ability to quickly and effectively prevent fraud.

DOL OIG issued an audit report in September 2022 highlighting improper payments including fraud in sampled and tested claims.[18] The report found that many states struggled with verifying the identity of claimants, leading to improper and fraudulent payments.

SBA's COVID-19 Economic Injury Disaster Loan and Paycheck Protection Programs

SBA's COVID-19 EIDL and EIDL Advance programs provided funding to help small businesses recover from the economic impacts of the COVID-19 pandemic. SBA has provided over $378 billion in COVID-19 EIDL loans and COVID-19 EIDL advances (grants). The program closed to new applications on January 1, 2022. PPP provided SBA-backed loans to businesses to keep their workforce employed during the COVID-19 crisis. The program ended on May 31, 2021, after more than 11 million PPP loans were approved totaling about $800 billion.

COVID-19 EIDL and PPP were more susceptible to fraud due to the elevated urgency for agencies to provide timely relief to applicants in response to the COVID-19 pandemic. SBA's initial approach to implement these programs quickly made billions of dollars available to millions of borrowers affected by the pandemic, but used few program controls to verify applicants' eligibility prior to disbursing funds.

In June 2023, the SBA OIG released a fraud landscape report offering a detailed estimate of potential fraud within the PPP and COVID-19 EIDL programs.[19] This analysis highlighted how the weakening or removal of critical controls by the SBA enabled fraudsters to exploit these programs more easily.

For information about SBA's COVID-19 EIDL and PPP, see the "Background" section in our January 2023 Fraud Alert.[20]

16 See DOL OIG report, OIG Oversight of the Unemployment Insurance Program, December 15, 2023.
17 See GAO report, Unemployment Insurance - Estimated Amount of Fraud during Pandemic Likely Between $100 Billion and $135 Billion, September 2023.
18 See DOL OIG report, COVID-19: ETA and States did not Protect Pandemic-Related UI Funds from Improper Payments Including Fraud or from Payment Delays, September 30, 2022.
19 See SBA OIG report, COVID-19 Pandemic EIDL and PPP Loan Fraud Landscape, June 27, 2023.
20 See PRAC Fraud Alert: PRAC Identifies $5.4 Billion in Potentially Fraudulent Pandemic Loans Obtained Using Over 69,000 Questionable Social Security Numbers, January 2023.

About the PRAC

The PRAC, created by the CARES Act in March 2020, supports and coordinates independent oversight of the $5 trillion in coronavirus pandemic relief and response programs. Through in-depth data analysis across multiple federal programs, we identify potential fraud, waste, abuse, and mismanagement, collaborating across the government oversight community to identify cross-cutting issues and risks to hold wrongdoers accountable, preserve program integrity, and help prevent misuse of future funds.

The PRAC has the processes, infrastructure, and relationships that enable us to conduct pre-award vetting to identify potential fraud before funds are disbursed.

In addition to SSN verification, the focus of this fraud prevention alert, the PRAC could also leverage existing investigative data on over 100,000 known fraud schemes from across a variety of agencies, including names, bank accounts, dates of birth, internet protocol and email addresses, and other identifiers, to flag new attempts by bad actors to steal American taxpayer dollars. The PRAC could also use its unique graph analytics tools to identify hidden networks and patterns among claims data. This pre-award vetting data could be retained and later be used for future pre-award vetting to serve as an early warning system for emerging fraud threats.

PRAC's Commitment to Quality and Integrity

We conducted this review using agile oversight principles that require we adhere to the professional standards of independence, due professional care, and quality assurance and follow procedures to ensure accuracy of the information presented.

The PRAC relied on SSA's verification results. To maintain the security of SSA's SSN verification process, SSA does not release the agency's matching logic concerning how it verifies SSNs. We believe the evidence obtained provides a reasonable basis for the results of our analysis.

For more on protecting program integrity, read our Chapter 1 of the Blueprint for Enhanced Program Integrity.[21]

For more on preventing payments to fraudsters, read our Chapter 3 of the Blueprint for Enhanced Program Integrity.[22]

21 See PRAC Blueprint for Enhanced Program Integrity Chapter 1: Best Practices for Strengthening Federal Programs.
22 See PRAC Blueprint for Enhanced Program Integrity Chapter 3: Fraud Prevention and Detection.

For more information:

Lisa Reijula
Associate Director of Outreach and Engagement, PRAC
Lisa.Reijula@cigie.gov

Visit us at:

PandemicOversight.gov

Follow us at:

Report Fraud, Waste, Abuse, or Misconduct:

To report allegations of fraud, waste, abuse, or misconduct regarding pandemic relief funds or programs please go to the PRAC website at PandemicOversight.gov.

A Committee of the
Council of the Inspectors General
on Integrity and Efficiency