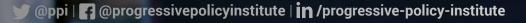
Told radically pragmatic

Closing the Digital Verification Divide

DR. MICHAEL MANDEL PROGRESSIVE POLICY INSTITUTE

SEPTEMBER 2024





Closing the Digital Verification Divide

DR. MICHAEL MANDEL

SEPTEMBER 2024

INTRODUCTION

In the internet era, the digitization of government is essential for the efficient and fair provision of public services. From faster access to unemployment benefits and food stamps to easier taxpayer retrieval of IRS tax records, digitization has the potential to make federal and local government work better, especially for lower-income Americans who need its services the most. It is not an exaggeration to say that "making government work better" requires digitization.

But successful digitization of government was slowed until recently by several factors. First, the "digital access divide" meant that many low-income or rural Americans did not have good enough quality Internet to seamlessly make use of digital government services.\(^1\) As a result, digitization of government ran the risk of widening existing inequities. Moreover, government had to maintain non-digital legacy systems as well as the new digital means of access, driving up the expense of service delivery and undercutting potential cost savings.

True, the original digital access divide has been narrowing. Post-pandemic efforts to bring high-speed broadband internet to everyone, such as the BEAD program, are in the process of successfully reducing the obstacles to access.²

However, government agencies face a more subtle but pervasive issue — what we call the "digital verification divide." Verification is the process by which a user verifies that they are who they say they are. Verification includes identity proofing, in which an individual provides sufficient information (e.g., identity history, credentials, documents) to establish a trusted identity online.³ That's a prerequisite for higher levels of authentication, which verifies the identity of a user, process, or device, in order to

allow access to more protected resources in an information system.

The process of identity proofing and authentication is especially important when users are trying to tap into government systems that contain sensitive personal data, such as individual accounts at the Internal Revenue Service (IRS), the Social Security Administration (SSA), Federal Student Aid (.gov) or Veterans Administration (VA). If government agencies make verification too easy relative to the risk of the transaction, then the wrong people can get access to sensitive personal data. If agencies make verification too hard relative to the risk of the transaction, then it becomes more difficult for constituents to prove their identity, unnecessarily locking them out of services and data that they are entitled to.

A "digital verification divide" is created by two factors that make it harder for low-income and other Americans with sparse document trails to take advantage of digital government. One issue is that low-income and marginalized Americans are less likely to have bank accounts, mortgages, passports, or any of the accumulation of documentation that most people can use to establish their identity and help authenticate themselves for government systems.

The second issue in closing the digital verification divide is that the use of biometrics for identity verification has been mistakenly conflated with the use of biometrics for surveillance and law enforcement, which poses a very different set of technological and implementation challenges. A typical identity verification system might use a face-matching algorithm that does a "1 to 1" comparison between an individual's face and a particular government-issued ID. A law enforcement

application, by contrast, might use a facial recognition algorithm that does a "1 to many" comparison between an individual's face and a database of millions of potential matches.

National Institute of Standards and Technology (NIST) testing has shown a steady reduction in the errors from the sort of face-matching algorithms used for identity verification, with the top-scoring ones performing consistently across demographics. Nevertheless, the continuing debate over the use of biometrics in situations such as surveillance and law enforcement has made policymakers reluctant to mandate biometrics for identity verification.

Closing the digital verification divide should be an important goal of policy, both for equity and efficiency reasons. Enabling government to interact digitally with all citizens in a safe way is essential to move the government into the future. Unfortunately, that progress has been slowed by challenges facing "Login.gov," the widely-used identity proofing and authentication system originally launched by the General Services Administration (GSA) in 2017. The GSA was faced with conflicting demands: On the one hand, guidelines issued by NIST required a physical or biometric component to achieve a high level of assurance needed by federal agencies to ensure legitimate access to restricted information or accounts requiring identity verification. On the other hand, the GSA apparently felt pressure to stay away from biometrics.

The result: the GSA ended up significantly misrepresenting the capabilities of Login.gov to the agencies using (and paying for) the system, according to a report released in March 2023 by the GSA Inspector General. GSA officials claimed



that Login.gov met NIST guidelines which required a physical or biometric component. But "Login.gov has never included a physical or biometric comparison for its customer agencies," according to the report, titled "GSA Misled Customers on Login.gov's Compliance with Digital Identity Standards." Along the same lines, the Treasury Inspector General for Tax Administration came out in September 2023 with a report raising concerns about Login.gov's ability to stop the types of fraud experienced by the IRS and other government agencies — though its deployment across agencies continues to expand.⁵

This policy brief will examine new evidence about how government agencies on the federal, state and local levels can digitize without creating or widening a digital verification divide. First, we note that the digitization of government needs to both boost efficiency and promote inclusion in order to meet its goals. Second, the success of digitization of government requires fair treatment to all individuals who need to log on remotely to public-facing government systems. In practice, this may mean following NIST guidelines that suggest providing an alternative video chat with a "trusted referee" for anyone who chooses and is verifying remotely. Finally, we conclude that with the availability of "trusted referees" or a similar alternative channel, biometric facial verification using leading NIST-tested algorithms can provide a high level of security and strong performance, while closing the digital verification divide.

In particular, an integrated system that includes both biometric face matching and the ability to verify users via alternative channels, such as video chat or in-person, can produce better access to digital government for low-income and other Americans with sparse document trails while limiting fraud. By contrast, an approach that relies only on online records is likely to be both less secure and less inclusive.

Summarizing, this policy brief identifies and names a major roadblock to digitization of government on every level, and explains how following the NIST guidelines helps overcome those obstacles. Indeed, misguided opposition to biometrics as part of a well-constructed digital verification process has been slowing down effective digitization, and widening the digital verification divide.

BACKGROUND

Fundamentally, the process of remote digital verification with a government agency starts with the potential user producing existing documents, such as a passport or driver's license. But the question is how to accurately verify online that these documents are not fraudulent, and being offered up by the person in question. This is also known as "identity proofing." In its Digital Identity Guidelines, NIST wrote that:

Identity proofing is the process by which a Credential Service Provider (CSP) collects and verifies information about a person for the purpose of issuing credentials to that person.

Identity proofing of applicants without requiring them to physically meet in person with CSP personnel is an important but challenging capability. It is important in providing access to CSP services to a larger portion of the population and in reducing the costs to both the applicant and the CSP. It is challenging because many of the identity proofing methods available to the CSP in a face-to-face interaction, such as



detailed inspection of evidence documents, are difficult to perform with comparable security when conducted remotely.

The NIST Guidelines went on to note the need "to strike a pragmatic balance between availability and convenient access to identity proofing services and security of the associated processes."

Remote digital verification typically starts with basic information such as birthdate, Social Security number, or phone number. Beyond those basics, there are four high-level approaches to remote digital verification:

- Passive profiling, which uses algorithmic analysis of massive data sets acquired by data brokers and credit bureaus, or scraped from social media, often without the consent or approval of users.⁷
- Knowledge-based verification (KBV), which asks the applicant for information about their background, with questions generated from data compiled from public records and credit history such as auto loan payment amounts.
- Biometrics, which uses various algorithms to match the face of the applicant with the photo on official documents such as passports or driver's licenses.
- Digital interviews with "trusted referees" who can examine documents remotely against the interaction of the submitter and verify identity.8

It's important to note that these approaches have different plusses and minuses, depending on whether they are used "stand-alone" or in conjunction with each other.

Stand-alone passive profiling — without a

biometric component or digital interview — is relatively cheap and easy to implement, requiring little interaction with the user. But it has several downsides. First, it's harder for low-income and marginalized individuals to be verified, because they have less presence in the credit and other databases being used, and less online visibility. Second, passive profile verification often relies on potentially erroneous information collected without the knowledge or consent of the user. That's a problem if the approach is being used to control access to important government services or if the agency needing verification is concerned about end-user privacy.

Next consider stand-alone KBV, also without a biometric component or digital interview. Like passive profiling, it is highly scalable. But it is generally accepted that KBV by itself is a weak form of identity verification because of the easy availability of personal data. Referring to KBV, the Identity Management Institute writes that:

...this approach for identifying end users is easily compromised and is no longer considered a viable authentication method.

Whether it's based on a static model in which users input answers to questions during account creation or a dynamic approach using random questions pulled from a set of known data about a user, [this approach] fails to provide the level of protection necessary for modern systems and networks.⁹

The second downside of stand-alone KBV is more subtle. One way to reduce fraud with KBV is to ask more difficult and obscure questions. But that discriminates against low-income and marginalized Americans who may not have easy access to the required documents.

Biometrics, broadly speaking, represents a third stand-alone approach to identity verification. For example, an algorithm can be used to match the face picture on an individual's driver's license, passport, or other official ID to an immediate video or photo ("selfie") of that individual. That's combined with a "liveness test," so that it's harder to fool the system with a picture of a face or an Al-generated deepfake.

It's important to note that such a "facial verification" process is much simpler and less controversial than "facial recognition."

Facial recognition technology, which is often associated with law enforcement and border control applications, starts with a single picture of an unknown face and tries to find an individual in a large database that's the closest fit. Facial verification, by contrast, compares a known individual's face with their own ID picture, a simpler task, and one less subject to bias issues.

However, there are still two important downsides to a stand-alone biometric approach. First, even as face-matching algorithms improve for all demographic groups, it is also possible that some subpopulations (race, gender, age) may see higher rates of false matches, meaning they are matching when they should not be.¹⁰ Second,

some people still won't be comfortable with the use of facial verification.

Finally, the fourth stand-alone approach to remote digital identity verification would be digital interviews with "trusted referees"— another way of saying an interaction with a real person. The NIST Guidelines note that "the use of trusted referees is intended to assist in the identity proofing and enrollment for populations that.... would be challenged to perform identity proofing and enrollment process requirements."¹¹ Such populations include, but are not limited to, disabled individuals, elderly individuals, and unbanked individuals with little or no credit history.

However, stand-alone digital interviews have the downside of losing the benefits of automation and economies of scale. Digital interviews are therefore more expensive on a per-applicant basis, and require detailed operational planning to handle large volumes of users.



TABLE 1: ADVANTAGES AND DISADVANTAGES OF FOUR APPROACHES TO STAND-ALONE VERIFICATION*

	COST/ SCALABILITY	FRAUD	INCLUSIVITY BIAS
STAND-ALONE PASSIVE Profiling	Highly scalable.	Medium fraud potential	Harder for low-income and other individuals with sparse document trails to be verified
STAND-ALONE KNOWLEDGE- Based Verification	Highly scalable.	High fraud potential	Harder for low-income and other individuals with sparse document trails to be verified
STAND-ALONE BIOMETRICS (FACE MATCH WITH DOCUMENT PLUS LIVENESS TEST)	Highly scalable, except for people without access to current technology.	Low fraud	Concerns about biased algorithms
STAND-ALONE "TRUSTED REFEREE" (DIGITAL VIDEO INTERVIEW)	More expensive on a per applicant basis	Low fraud	Able to handle individuals who have difficulty navigating online enrollment, or prefer not to.

^{*}Assuming basic information such as birthdate or Social Security number Source: Progressive Policy Institute

DIGITAL VERIFICATION DIVIDE IN STAND-ALONE AND INTEGRATED SYSTEMS

How can we reduce the digital verification divide, and make public-facing government IT systems accessible and secure to all potential users? Real-world examples suggest that passive profiling effectively locks many people out.

Consider, for example, the process of individuals getting verified for online access to their IRS data, such as tax return transcripts. In 2016, the IRS introduced its Secure Access system, which relied mainly on passive profiling, based on financial information from credit bureaus and data brokers. But the system didn't work well for everyone. Overall, former IRS Commissioner Charles Rettig testified to Senate appropriators in 2022 that access rates for the IRS under its former system without biometrics or trusted referees were only 40%. 13

In particular, passive profiling didn't work well for Puerto Rico, a region with a high poverty rate and low credit use. According to Puerto Rico's Resident Commissioner to Congress, from 2016 to 2022, only 24% of Puerto Rican taxpayers were able to verify their identities using the IRS Secure Access system. But after the IRS modernized its system to offer multiple pathways to verification, access rates in Puerto Rico jumped to 79% — a three-fold increase.

These figures point out the fundamental flaw of an identity verification process based only on passive profiling. The better alternative is an integrated system that utilizes several approaches to verification, as recommended by the NIST Guidelines.

Let's start with the "trusted referee" channel. While expensive to operate on a per-applicant



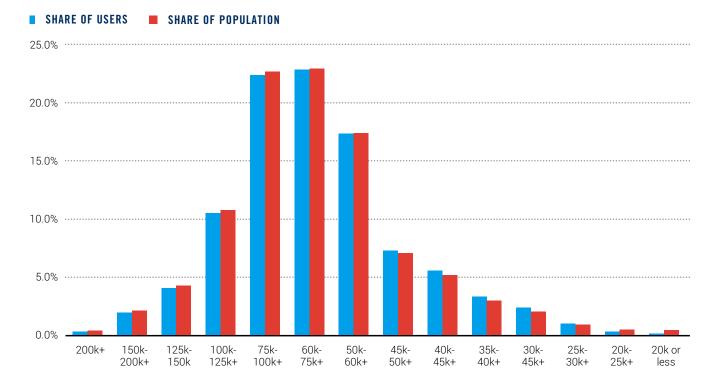
basis, the "trusted referee" channel is essential for closing the digital verification divide. Or, to put it a different way, the manual component to verification makes the automated components fairer.

To show this, we use aggregated data provided by ID.me, a leading identity verification firm, based on its anonymized database of 66 million verification attempts since 2018, primarily for government agencies including but not limited to Treasury, IRS, SSA, VA, HHS, and state unemployment agencies.¹⁷ "Anonymized" means that no demographic, income, or other personal data was collected and retained. The only data retained was the zip code and the result of the verification attempt. These two pieces of data

were paired with zip code level demographic data from the American Community Survey, conducted annually by the Census Bureau. This zip code level demographic data included median income and the share of the population in the zip code that was Black or African American, Hispanic, Asian, or white.¹⁸

ID.me's data set is broad enough to be roughly representative of the U.S. population. Indeed, when zip codes are stratified by income, we find that the share of users closely matches the share of the population (Figure 1). When stratified by race or ethnic makeup, the match is not quite so tight, but still very good.

FIGURE 1. ID.ME USERS VERSUS POPULATION, BY ZIP CODE STRATIFIED BY MEDIAN INCOME



Data: ID.me, Census



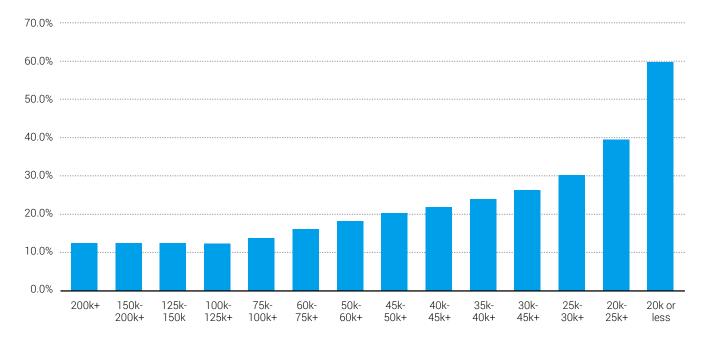
The aggregate data allows us to compare the performance of the trusted referee channel for zip codes with low, medium, and high median income. It turned out that the trusted referee channel was much more important for low-income zip codes.

For example, for applicants living in the highest income zip codes, only 13% used the trusted referee channel to get verified. For zip codes with median income between \$50,000 and \$60,000, 18% of applicants used the trusted referee channel. For zip codes with a median

income between \$30,000 and \$35,000, 26% of applicants used the trusted referee channel (see Figure 2).

For applicants living in the lowest income zip codes, as many as 60% relied on video interviews with trusted referees. That means attempting to do identity verification without a "trusted referee" option or the equivalent is likely to be biased against low-income users. This is extremely important for government agencies in the process of digitization.





Data: ID.me, Census

BIOMETRICS

The next question: Does the use of biometrics introduce bias into an integrated verification process? Or, to put it another way, does the use of biometrics help expand or close the digital verification divide?

ID.me's data suggests that the appropriate use of biometrics can help close the digital verification divide, especially when combined with digital interviews as needed. Following NIST Guidelines, the biometric step consists of checking someone's face against the picture on their driver's license, passport, or other official ID.

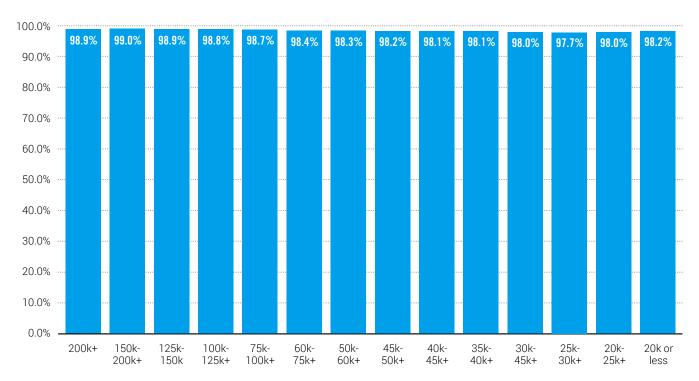


That's combined with a "liveness test," so that it's harder to fool the system with a picture of a face.

The data shows very little variation in biometric pass rates by income when used as part of an integrated process (Figure 3). In other words,

roughly 98% of applicants from low-income zip codes pass the biometric matching step, compared to 99% of the applicants from high-income zip codes. This suggests that in this context, biometrics are not biased by income.

FIGURE 3. SELFIE MATCH PASS RATES, BY ZIP CODE STRATIFED BY MEDIAN INCOME



Data: ID.me, Census

Equally important, biometrics show little bias relative to race and ethnicity when used as part of an integrated process. For example, applicants from zip codes with a low percentage of blacks or African Americans passed the

biometric screening at a 98.6% rate, while applicants from zip codes with a high percentage of blacks or African Americans passed the biometric screening at a 98% rate (Figure 4).



100.0% 98.6% 98.4% 98.3% 98.1% 98.0% 97.9% 98.1% 98.1% 98.0% 90.0% 80.0% 70.0% 60.0% 50.0% 40.0% 30.0% 20.0% 10.0% 0.0% 41-50 0-10 11-20 21-30 31-40 51-60 61-70 71-80 81-100

FIGURE 4. SELFIE MATCH PASS RATES, BY ZIP CODE, STRATIFED BY SHARE OF POPULATION IDENTIFYING AS BLACK OR AFRICAN AMERICAN

Data: ID.me, Census

Two points are important to emphasize here, First, the data is analyzed on the zip code level rather than the individual level. So conceivably, an individual-level analysis might show more bias. Second, the data were generated as part of an integrated process with the option of trusted referees. Taken together, these results suggest that biometrics, as part of an integrated verification process, helps narrow the digital verification divide.

ONLINE RECORDS

The ID.me data also covers the use of online records to assess identity. Once again, it must be stressed that unlike stand-alone passive profiling and KBV, these results represent the outcome of an integrated verification process.

Note that a significant share of the population may not have access to the sort of financial records used by identity verification. For example, according to the FDIC, an estimated 4.5% of U.S. households (approximately 5.9 million) were "unbanked" in 2021, meaning that no one in the household had a checking or savings account at a bank or credit union. About 50 million Americans have thin, incorrect, or unscorable credit, meaning they would have trouble passing records-based checks.

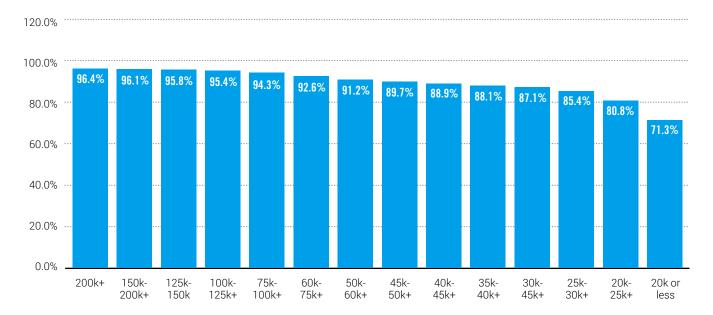
How does that show up in the data? Applicants from low-income zip codes have a much lower pass rate from an online records check (Figure 5). Note that the data in this figure is the result of an integrated verification process which includes biometrics and a digital interview step, if needed.



So as part of the integrated process, applicants that fail the online records check can still be verified as part of the alternative channels.

By contrast, stand-alone passive profiling and KBV have no safety net. Moreover, as noted earlier, both passive profiling and KBV tend to be biased against low-income or marginalized applicants.

FIGURE 5. ONLINE RECORDS PASS RATE, BY ZIP CODE STRATIFIED BY MEDIAN INCOME



Data: ID.me, Census

CONCLUSIONS: CLOSING THE DIGITAL VERIFICATION DIVIDE

Governments are faced with the mandate to do more with less. The obvious solution is to digitize expensive and slow legacy processes. That, in turn, requires both investment in better IT systems and the shift to online access for constituents.

This report focuses on a simple but often overlooked step — identity verification of users, who often come from low-income or other groups who have sparse document trails. Since users cannot get government services without being verified, digitization requires

verification processes that are both secure and demonstrably inclusive.

The analysis presented in this paper supports these conclusions for closing the digital verification divide.

- The goals of inclusion and efficiency are not in conflict.
- The success of digitization of government requires fair treatment of all groups.
 For remote verification, that may mean providing an alternative video chat with a "trusted referee" for anyone who chooses, or the equivalent.



3. Done well, biometric facial verification using leading government-tested algorithms can provide a high level of security and strong performance, while closing the digital verification divide.

ABOUT THE AUTHOR

Dr. Michael Mandel is Vice President and Chief Economist of the Progressive Policy Institute.



References

- 1 "State & Local Government Technology Outlook," American City and County, February 22, 2023, https://www.americancityandcounty.com/2023/02/22/state-local-government-technology-outlook/.
- 2 "Broadband Equity Access and Deployment Program," Broadband USA, accessed August 2024, https://broadbandusa.ntia.doc.gov/funding-programs/broadband-equity-access-and-deployment-bead-program.
- 3 "Identity Proofing," Computer Security Research Center (NIST), accessed August 2024, https://csrc.nist.gov/glossary/term/identity_proofing.
- 4 "GSA Misled Customers on Login.Gov's Compliance with Digital Identity Standards," Office of Inspector General, U.S. General Services Administration, March 7, 2023, https://www.gsaig.gov/content/gsa-misled-customers-logingovs-compliance-digital-identity-standards.
- 5 "Key Events of the IRS's Planning Efforts to Implement Login.Gov for Taxpayer Identity Verification," Treasury Inspector General for Tax Administration, September 27, 2023, https://www.tigta.gov/sites/default/files/reports/2023-10/20232S070fr.pdf.
- James L. Fenton, Michael E. Garcia, and Paul A. Grassi, "NIST Special Publication 800-63: Digital Identity Guidelines," National Institute of Standards and Technology, October 16, 2023, https://pages.nist.gov/800-63-3/sp800-63-3.html.
- 7 "Prepared Remarks of CFPB Director Rohit Chopra at the White House on Data Protection and National Security," Consumer Financial Protection Bureau, April 2, 2024, https://www.consumerfinance.gov/about-us/newsroom/prepared-remarks-of-cfpb-director-rohit-chopra-at-the-white-house-on-data-protection-and-national-security/.
- James L. Fenton and Paul A. Grassi, "NIST Special Publication 800-63A: Section 5.3.4 Trusted Referee Requirements," National Institute of Standards and Technology, October 16, 2023, https://pages.nist.gov/800-63-3/sp800-63a.html#trustref.
- 9 "Knowledge-Based Authentication Weaknesses," Identity Management Institute, accessed August 2024, https://internation-weaknesses/. identitymanagementinstitute.org/knowledge-based-authentication-weaknesses/.
- 10 "Face Recognition Technology Evaluation (FRTE) 1:1 Verification," National Institute of Standards and Technology, accessed August 2024, https://pages.nist.gov/frvt/html/frvt11.html.
- 11 "SP 800-63A: Use of Trusted Referees," National Institute of Standards and Technology, accessed August 2024, https://pages.nist.gov/800-63-3-Implementation-Resources/63A/referees/.
- 12 "Providing Secure Access to IRS Taxpayer Information," United States Digital Service, 2016, https://www.usds.gov/report-to-congress/2016/irs-secure-access/.
- 13 Natalie Alms, "IRS Leader Explains Why the IRS Went to ID.Me," Nextgov/FWC, May 4, 2022, https://www.nextgov.com/digital-government/2022/05/irs-leader-explains-why-irs-went-idme/366516/.
- 14 Sangeetha Malaiyandi, "Financial Struggles in Puerto Rico Bite Deeper than the Rest of the United States," Consumer Financial Protection Bureau, July 27, 2023, https://www.consumerfinance.gov/about-us/blog/financial-struggles-in-puerto-rico-bite-deeper-than-the-rest-of-the-united-states/.
- "IRS Improves Online Identity Verification for Puerto Ricans in Response to Rep. Jenniffer González Inquiry," Congresswoman Jenniffer González-Colón, June 8, 2023, https://gonzalez-colon.house.gov/media/press-releases/irs-improves-online-identity-verification-puerto-ricans-response-rep-ienniffer.



- Douglas W. O'Donnell to the Honorable Jenniffer González-Colón, "Douglas O'Donell IRS Response ID ME," *Department of the Treasury*, December 16, 2022, https://www.dropbox.com/s/2b9ibge68dr6c9g/12.06.2022 Douglas ODonell IRS Response ID ME.pdf?e=1&dl=0.
- 17 A description of ID.me's methodology plus the aggregated data can be found at https://network.id.me/inclusion-and-access-research-methodology/.
- 18 "The Public Health Disparities Geocoding Project," Harvard T.H. Chan School of Public Health, accessed August 2024, https://www.hsph.harvard.edu/thegeocodingproject/.
- 19 1. "2021 FDIC National Survey of Unbanked and Underbanked Households," FDIC, last updated July 24, 2023, https://www.fdic.gov/analysis/household-survey/index.html.
- 20 "More than 45 Million Americans Are Either Credit Unserved or Underserved; Approximately 20% Migrate to Being Credit Active Every Two Years," TransUnion, April 7, 2022, https://newsroom.transunion.com/more-than-45-million-americans-are-either-credit-unserved-or-underserved---approximately-20-migrate-to-being-credit-active-every-two-years/.

ppi

The Progressive Policy Institute is a catalyst for policy innovation and political reform based in Washington, D.C. Its mission is to create radically pragmatic ideas for moving America beyond ideological and partisan deadlock.

Founded in 1989, PPI started as the intellectual home of the New Democrats and earned a reputation as President Bill Clinton's "idea mill." Many of its mold-breaking ideas have been translated into public policy and law and have influenced international efforts to modernize progressive politics.

Today, PPI is developing fresh proposals for stimulating U.S. economic innovation and growth; equipping all Americans with the skills and assets that social mobility in the knowledge economy requires; modernizing an overly bureaucratic and centralized public sector; and defending liberal democracy in a dangerous world.

© 2024
PROGRESSIVE POLICY INSTITUTE
ALL RIGHTS RESERVED.

PROGRESSIVE POLICY INSTITUTE 1919 M Street NW, Suite 300, Washington, DC 20036

Tel 202.525.3926 **Fax** 202.525.3941

info@ppionline.org progressivepolicy.org