

No. 26A308

In the Supreme Court of the United States

U.S. DEPARTMENT OF HOMELAND SECURITY, ET AL.,
Applicants

V.

LEAGUE OF WOMEN VOTERS, ET AL.
Respondents.

Application to Stay the Order of the
United States District Court for the District of Columbia

**BRIEF ON BEHALF OF *AMICI CURIAE* UCLA SOCIAL SCIENTISTS
IN SUPPORT OF RESPONDENTS**

Xavier Becerra
UCLA Voting Rights Project
Senior Advisor and Voting
Rights Counsel
108 W. 2nd Street, #310
Los Angeles, CA 90012
310-924-9310

Chad Dunn
UCLA Voting Rights Project
Legal Director
Bernadette Reyes
Counsel of Record
Sonni Waknin
Los Angeles, CA 90095
310-400-6019
bernadette@uclavrp.org

Counsel for Amici Curiae UCLA Social Scientists

TABLE OF CONTENTS

TABLE OF AUTHORITIES	iii
INTEREST OF AMICI	1
SUMMARY OF ARGUMENT	4
ARGUMENT	4
I. The SAVE System Ignores Well-Known Problems in Large-Scale Data Matching... ..	4
A. Data Matching Large-Scale Government Administrative Records Produces Unreliable Results.	5
B. A Comparison of the Department of Homeland Security and the U.S. Census Bureau Highlights Likely Shortcomings of SAVE.....	11
CONCLUSION.....	18

TABLE OF AUTHORITIES

Cases

<i>Arcia v. Sec’y of Fla.</i> , 772 F.3d 1335 (11th Cir. 2014)	9
<i>Brakebill v. Jaeger</i> , No. 1:16-CV-008, 2016 WL 7118548 (D.N.D. Aug. 1, 2016)	3
<i>Fair Fight Action, Inc. v. Brad Raffensperger</i> , 593 F. Supp. 3d 1320 (N.D. Ga. 2021)	8
<i>Fish v. Kobach</i> , 309 F. Supp. 3d 1048 (D. Kan. 2018)	3, 8
<i>Fla. State Conference of the NAACP v. Browning</i> , 522 F.3d 1153 (11th Cir. 2008)	9
<i>Ibrahim v. Dept. of Homeland Sec.</i> , 669 F.3d 983 (9th Cir. 2012)	11
<i>Veasey v. Perry</i> , 71 F. Supp. 3d 627 (S.D. Tex. 2014)	3, 9, 10

Other Authorities

DEBORAH WAGER, et al., ESTIMATING RECORD LINKAGE FALSE MATCH RATE FOR THE PERSON IDENTIFICATION VALIDATION SYSTEM (U.S. CENSUS BUREAU, CARRA WORKING PAPER No. 2014-02, 2014)	17
DEBORAH WAGNER & MARY LAYNE, THE PERSON IDENTIFICATION VALIDATION SYSTEM (PVS): APPLYING THE CENTER FOR ADMINISTRATIVE RECORDS RESEARCH AND APPLICATIONS’ (CARA) RECORD LINKAGE SOFTWARE (U.S. CENSUS BUREAU, CARRA WORKING PAPER No. 2014-01, 2014)	17
Huimin Zhao, <i>Semantic Matching Across Heterogeneous Data Sources</i> 50 COMM’N OF THE ACM (Jan. 2007)	11
Lindsey Wilkinson, <i>‘Atrocious’ tech hurt ICE investigation branch efforts to counter fentanyl influx</i> , FedScoop (Aug. 25, 2026)	12
Mary Pat Dwyer & Rachel Levinson-Waldman, <i>A Realignment for Homeland Security Investigations</i> , Brennan Center for Justice (2023)	12
Nahid Anwar, et al., <i>Detecting Anomalies in Voter Registration Data</i> , 25(2) ELEC. L. J.: RULES, POL., AND POL’Y 113 (2026)	7
Olivier Binette & Rebecca C. Steorts, <i>(Almost) All of Entity Resolution</i> , STAT.ME (2022) at 6	15
RECORD LINKAGE & MACHINE LEARNING, U.S. CENSUS BUREAU (2026)	15
Stephen Ansolabehere & Eitan D. Hersch, <i>ADGN: An Algorithm for Record Linkage Using Address, Date of Birth, Gender, and Name</i> , 4(1) STAT. AND PUB. POL’Y, 1 (2017)	7, 10

STEPHEN E. FEINBERG, <i>Homeland Insecurity: Data Mining, Privacy, Disclosure Limitation, and the Hunt for Terrorists</i> , in TERRORISM INFORMATICS: KNOWLEDGE MANAGEMENT AND DATA MINING FOR HOMELAND SECURITY 205 (Hsinchun Chen, et al. eds., 2008)	12
Ted Enamorado et al., <i>Using a Probabilistic Model to Assist Merging of Large-Scale Administrative Records</i> , 113(2) Am. Pol. Sci. Rev., 353 (2019).....	7
U.S. ELECTION ASSISTANCE COMM’N, REGISTER TO VOTE IN YOUR STATE BY USING THIS POSTCARD FORM AND GUIDE, https://www.eac.gov/sites/default/files/eac_assets/1/6/Federal_Voter_Registration_ENG.pdf	14
U.S. SOCIAL SECURITY ADMINISTRATION, LETTER AGREEMENT PROVIDING FOR INFORMATION SHARING BETWEEN THE DEP’T OF HOMELAND SEC. (DHS), U.S. CITIZENSHIP AND IMGR. SERVS. (USCIS), AND THE SOC. SEC. ADMIN. (SSA) REGARDING CITIZENSHIP (May 15, 2025).....	13
Wendy R. Weiser, et al., <i>Making the List: Database Matching and Verification Processes for Voter Registration</i> , BRENNAN CENTER FOR JUSTICE (March 24, 2006)	5
WILLIAM E. WINKLER, OVERVIEW OF RECORD LINKAGE AND CURRENT RESEARCH DIRECTIONS (U.S. CENSUS BUREAU, Working Paper No. RRS2006-02, 2006) ..	15
Yves Thibaudeau, <i>A Review of Modern Multinomial-Derived and Partition-Based Record-Linkage Methods</i> 17 WIREs COMPUT. STAT. e70015 (2025),	16
Zhenbang Wang et al., <i>Regularization for Shuffled Data Problems via Exponential Family Priors on the Permutation Group</i> , PMLR 206 (2023)	8, 15

INTEREST OF AMICI

Matt A. Barreto is a Professor of Political Science and Chicana/o and Central American Studies at UCLA. He is the faculty director of the UCLA Voting Rights Project and has published numerous social science articles on the topics of complex data matching of voter information. His academic published work regularly uses state voter file databases and has pioneered research methodologies to more accurately understand information on voter files. He has authored expert reports and testified as an expert witness in dozens of matters related to voting rights issues, specifically about database matching.

Gary M. Segura is Professor of Public Policy and Chicana/o Studies, and the former Dean of the Luskin School of Public Affairs at UCLA. His work focuses on issues of political representation and social cleavages, the domestic politics of wartime public opinion, and the politics of America's growing Latino minority. His work has been published in the *American Political Science Review*, *American Journal of Political Science*, *Journal of Politics*, *Political Research Quarterly*, and the *Annual Review of Political Science*. He has provided expert testimony on discrimination in voting rights cases and civil rights cases.

Efrén Pérez, is a Professor of Political Science and Psychology at UCLA and director of the UCLA, Race, Ethnicity, Politics and Society (REPS) Lab. His scholarship area is political psychology, with a focus on racial and ethnic politics, language and political thinking, implicit political cognition, and the measurement of political concepts. He is the author of numerous articles in leading political science

journals and has been published in *the American Journal of Political Science*, *The Journal of Politics*, *Political Analysis*, *Political Behavior*, *Politics, Groups and Identities*, and *Political Psychology*. He has also published a book, *Unspoken Politics: Implicit Attitudes and Political Thinking*.

Lorrie Frasure is a Professor of Political Science and African American Studies at UCLA. She held the Ralph J. Bunche Chair for African American Studies at UCLA. Dr. Frasure's research focuses on racial/ethnic political behavior, African American politics, and state and local politics and has extensive experience working with voter files and database matching. She has published numerous political science papers and a book, *Racial and Ethnic Politics in American Suburbs*.

Natalie Masuoka is an Associate Professor of Political Science and Asian American Studies and currently the Chair of the Asian American Studies Department at UCLA. Dr. Masuoka's research includes studying racial and ethnic politics, immigration, and political behavior and public opinion. She has written multiple books on racial politics in the United States and has extensive experience working with voter files and database matching. Dr. Masuoka has written numerous social science papers on race and politics and has been cited over one thousand times.

Chris Zepeda-Millán is an Associate Professor of Political Science, Public Policy and Chicana/o and Central American Studies at UCLA. His research has been published in top political science and interdisciplinary academic journals, such as the *American Journal of Political Science*, *Political Research Quarterly*, *Politics*,

Groups and Identities, Critical Sociology, the Chicana/o Latina/o Law Review, Social Science Quarterly, and the Journal of Ethnic and Migration Studies.

The UCLA Social Scientists, as *amici*, are well-positioned to opine on the issues before the Court because of amici's work as social scientists studying voting issues and particularly database analysis. *Amici* have collectively performed hundreds of large-scale analyses of voter data, including database matching analysis. These analyses have been published in the leading journals in their fields (e.g. *The American Political Science Review*) and have been provided as testimony in, and accepted as reliable by, both federal and state courts. *See Veasey v. Perry*, 71 F. Supp. 3d 627, 662 (S.D. Tex. 2014); *Fish v. Kobach*, 309 F. Supp. 3d 1048, 1059 (D. Kan. 2018); *Brakebill v. Jaeger*, No. 1:16-CV-008, 2016 WL 7118548, at *4 (D.N.D. Aug. 1, 2016).

SUMMARY OF ARGUMENT

The Systematic Alien Verification for Entitlements (SAVE) program is rife with error. Decades of well-established social science practice and research confirms this. Should the Court agree to grant Applicants' stay application, swaths of eligible voters across the country run the risk of disenfranchisement.

In this brief *Amici*, Social Scientists seek to provide this Court with important context for the decision it will issue. *Amici* will highlight the fundamental flaw in SAVE and dispel the claims of heightened accuracy lobbed by Applicants in their request to this Court.

ARGUMENT

I. The SAVE System Ignores Well-Known Problems in Large-Scale Data Matching.

In the fields of law, computer science, and social science, the limitations of large-scale databases and data matching have been known for years. The Department of Homeland Security's (DHS) use of SAVE ignores these flaws and even expands them. Applicants point to the modified SAVE system's ability to process bulk requests, more than a million each day, as a significant improvement to the prior system's capabilities. Appl. for Stay at. 2.

However, it is this very "improvement," in addition to the numerous databases that SAVE purports to use to match individuals, that renders the system even more prone to inaccuracies. Regardless of how many searches can be processed, the underlying disparate databases are plagued with inaccuracies and errors. Absent the use of proper methodologies, bulk queries improve nothing.

Such inaccuracies have existed in other matching systems utilized by DHS, who continually fail to prioritize data system improvement and methodology. This starkly contrasts another federal agency, the United States Census Bureau (Census Bureau) who regularly studies, publishes, and improves on methods of accurately maintaining and matching large scale government databases.

Large-scale database matching is a complex, multi-faceted science, one that requires careful methodology and safeguards to ensure accuracy. DHS' hastily put together system lacks these.

A. Data Matching Large-Scale Government Administrative Records Produces Unreliable Results.

It is common knowledge in the social sciences that large scale data matching can be dangerously unreliable, absent proper safeguards and methodology. Numerous studies illustrate the difficulty of performing such comparisons and achieving reliable results when dealing specifically with voter registration lists and like records.

To start, voter registration databases themselves can be prone to inaccuracies due to typos, omissions and the like.¹ The breadth of these errors expand when combining data from different sources, which themselves are collected and recorded

¹ Wendy R. Weiser, et al., *Making the List: Database Matching and Verification Processes for Voter Registration*, BRENNAN CENTER FOR JUSTICE (March 24, 2006) at 4, <https://www.brennancenter.org/our-work/research-reports/making-list-database-matching-and-verification-processes-voter> (last visited September 14, 2026).

using their own unique processes and standards. A 2006 report illustrated the problem thusly:²

The Problem(s) with Matching: Examples of Potential Database Errors		
Name of Registrant		
Source of Error	On Voter Registration Form	In Database
Typos	Pierce	Peirce or Pearce or Perce or Pierrce
Transliteration	Mohammad	Muhammed
Marriage	Mary Pierce (née Owens)	Mary Owens or Mrs. Martin Pierce
Nickname	Sam Pierce	Samuel Pierce
Transposed field	Bao Lu	Lu Bao
Double names	"Mary Ann" (first) "Pierce" (last)	"Mary" (first) "Ann" (middle) "Pierce" (last)
Hyphenated name	"Mary" (first) "Owens-Pierce" (last)	"Mary" (first) "Owens" (middle) "Pierce" (last)
Punctuation	al-Amin	al Amin
Date of Birth		
Source of Error	On Voter Registration Form	In Database (Voter, DMV and/or SSA)
Typos	01/03/05	02/03/05 or 1/00/05 or 1/03/05 or 11/03/05
Transposed field	01/03/05	03/01/05 or 05/01/03
Invented default	01/03/05	01/01/05 (submitted only as January 2005)

The same study warned of possible disenfranchisement on a mass scale if one were to rely solely on this type of data matching.³ And even “the most sophisticated matching technologies—which are not being used in most states—will leave many eligible voters unmatched.”⁴

Countless other studies have agreed. Working with election officials in Idaho, computer and political scientists agreed that voter registration databases are rife with “various anomalies such as missing values in the required fields, abnormal age

² *Id.*

³ *Id.* at 5.

⁴ *Id.*

entries, unspecified gender types, non-unique driver's license numbers, and formatting errors.”⁵

Data matching experts Stephen Ansolabehere & Eitan D. Hersh, who have testified in countless voting cases on this issue, similarly reiterated the ability of such errors to add up over multiple databases, noting, “[e]ven when the incidence of key-punching errors is low, the probability of at least one error in a series of multiple indicators is not trivial.”⁶ Identifiers such as Social Security Numbers are not immune to such errors; as Ansolabehere and Hersh point out, “SSN9 fields sometimes contain 111111111, 123456789 or 999999999, which are invalid SSN9s.”⁷

Several other social science reports expand on the effects such database errors have on accurate matching. Political and data scientists Ted Enamorado, Benjamin Fifield, and Kosuke Imai note a common problem where “a unique identifier that unambiguously links records is often unavailable, and data may contain missing and inaccurate information.”⁸ The results of this problem is severe, especially for large-scale administrative models, and “deterministic merge methods cannot quantify the uncertainty of the merging procedure and instead typically rely on arbitrary thresholds to determine the degree of similarity sufficient for

⁵ Nahid Anwar, et al., *Detecting Anomalies in Voter Registration Data*, 25(2) ELEC. L. J.: RULES, POL., AND POL’Y 113, 113 (2026).

⁶ Stephen Ansolabehere & Eitan D. Hersh, *ADGN: An Algorithm for Record Linkage Using Address, Date of Birth, Gender, and Name*, 4(1) STAT. AND PUB. POL’Y, 1, 4 (2017).

⁷ *Id.*

⁸ Ted Enamorado et al., *Using a Probabilistic Model to Assist Merging of Large-Scale Administrative Records*, 113(2) Am. Pol. Sci. Rev., 353, 353 (2019).

matches.”⁹ The results “fail to account for the uncertainty of the merging procedure, yielding a bias due to measurement error.”¹⁰

In yet another example, statisticians Zhenbang Wang and Martin Slawski collaborated with U.S. Census Bureau data scientist Emanuel Ben-David, and describe another issue regarding integration of different sources of data, stating:

[A] rather common scenario in data integration problems in which different pieces of information about a shared set of entities reside in multiple data sources that need to be combined in order to perform a given data analysis task. The process of identifying matching parts across two or more files is often far from trivial in the absence of unique identifiers, and has thus grown into a vast and active field of research known as *record linkage*.¹¹ (internal citation omitted)

Legal challenges involving voting rights issues have similarly found database matching inaccurate. *See Fair Fight Action, Inc. v. Brad Raffensperger*, 593 F. Supp. 3d 1320, 1330-1331 (N.D. Ga. 2021) (Where Plaintiff’s expert testified that “an individual can be flagged for verification failure even for ‘a minor mistake on a registration form, such as a one-character difference in spelling or spacing in a name, an apostrophe, a misplaced hyphen or a typographical error,’” and a representative of the Georgia Secretary of State testified to a social security match process error rate of “20 or 30 percent”); *Fish v. Kobach*, 309 F. Supp. 3d 1048, 1085 (D. Kan. 2018) (Plaintiffs’ expert Dr. Hersch finding that that “the number of purported incidents of noncitizen registration found by Defendant is consistent with

⁹ *Id.*

¹⁰ *Id.*

¹¹ Zhenbang Wang et al., *Regularization for Shuffled Data Problems via Exponential Family Priors on the Permutation Group*, 206 (2023), at 1
<https://proceedings.mlr.press/v206/wang23a/wang23a.pdf>.

the quantity of other low-incidence idiosyncrasies in [statewide voter registration database] and in voter files more generally, and is suggestive of administrative errors.”); *Arcia v. Sec’y of Fla.*, 772 F.3d 1335, 1341 (11th Cir. 2014) (Finding standing of naturalized citizens in a case challenging Florida’s removal of suspected noncitizens from voter rolls, as “there was a realistic probability that they would be misidentified due to unintentional mistakes in the Secretary's data-matching process.”); *Fla. State Conference of the NAACP v. Browning*, 522 F.3d 1153, 1176 n.5 (11th Cir. 2008) (Barker, J., dissenting) (“The matching scheme in itself is problematic because of the numerous administrative and technological barriers, such as computer glitches or human error that make the possibility of non-matches for qualified voters a strong possibility.”)

In cases where experts have been called on to perform this type of large-scale data matching, they have done so using extreme care and protocol to ensure maximum accuracy. For example, when matching multiple databases to determine voters who might lack a specific identification required to vote, plaintiffs expert Stephen Ansolabehere, referenced *supra*, performed a detailed scientific process. He first “scrubbed” the voter registration list, removing inactive, duplicate, relocated, deceased or otherwise exempt voters using algorithms specifically designed to consider common errors in spelling and naming conventions found in these lists. *Veasey v. Perry*, 71 F. Supp. 3d 627, 659-60 (S.D. Tex. 2014). Another expert confirmed these results utilizing a similar method and both “had to write codes so that the fields of the respective databases were compared correctly, even though the

databases were formatted differently.” *Id.* The reviewing court placed value on the techniques of “coding and algorithms used in the matching methodology that were consistent with the demands of the scientific field.” *Id.*

As demonstrated *supra*, social scientists and statisticians with years of training have published repeatedly on the complex issues that arise when performing the task that Applicants claim to do. To be clear, voter file and voter registrations are not so inaccurate that the lists themselves are useless for their purpose, nor do they lack any probative value for research and legal cases involving voting rights.¹² However, these lists must be used with care, utilizing the latest techniques of data science, statistics and computer science, that have been constantly studied and improved over the past few decades.

What is dangerous is handling voter lists, or any large swathes of personal and governmental databases in a way that boasts accuracy but ignores proper science and methodology; that prioritizes output over accuracy and that runs the risk of labeling someone as a criminal in the event of even the smallest of errors; that is exactly what DHS is doing with SAVE. Applicants provide little factual evidence that the new system is *more accurate*, only that it can process *more* records at once. There is no indication that there have been attempts to validate,

¹² The work of Dr. Stephen Ansolobehere referenced in this brief is just one of many examples of how this data can be properly used. Recognizing limitations of and errors present in databases, including voter lists, he and others created an algorithm to minimize the effect of these common issues and heighten accuracy of matching. See *Veasey v. Perry*, 71 F. Supp. at 659-60 and Stephen Ansolobehere & Eitan D. Hersch, *ADGN: An Algorithm for Record Linkage Using Address, Date of Birth, Gender, and Name*, 4(1) Stat. and Pub. Pol’y, 1, 1-8 (2017). Indeed, almost all other social scientists referenced in this paper, as well as *Amici*, have conducted and published peer reviewed research describing methods in minimizing error in the type of matching discussed here.

clean or update the accuracy of the sixteen databases that SAVE searches across using carefully studied and scientifically accepted methodology, actions the experts agree are necessary.

B. A Comparison of the Department of Homeland Security and the U.S. Census Bureau Highlights Likely Shortcomings of SAVE.

An examination of DHS practices regarding maintenance and use of large-scale databases compared to that of another entity responsible for large scale data matching, the U.S. Census Bureau, provides insight into the lax attitudes that likely went into the creation and use of SAVE.

DHS Database Maintenance

Published research has documented that DHS databases are unreliable for matching specific individuals.¹³ DHS records often pull together a combination of private, commercial, and government databases which are regularly inconsistent, and not up to date.¹⁴ Large-scale data collection by agencies like DHS can be prone to error, with experts noting that a “reasonable fraction of the data in their files will

¹³ Case law has also provided examples of the difficulties of data mismanagement in dealing with issues of homeland security. In *Ibrahim v. Dept. of Homeland Sec.*, a database run by the FBI but supported by numerous agencies including DHS was rife with error due to matching issues. *See* 669 F.3d 983, 990 (9th Cir. 2012). There, a “No Fly List” was cut in half after it was discovered that thousands had been misidentified and mistakenly placed on the list because of “misspellings and transcription errors in the nomination process, and because of computer algorithms that imperfectly match travelers against the names on the list.” *Id.*

¹⁴ Professor Huimin Zhao from the University of Wisconsin, Milwaukee opined on the likelihood of error in situations like this: “consider an example of two security-related databases [] owned by different organizations. They contain several common attributes and overlapping records, though they also include discrepancies and data errors.” Huimin Zhao, *Semantic Matching Across Heterogeneous Data Sources* 50 COMM’N OF THE ACM. RSCH. AND ADVANCES (2007) 45, 46. Zhao expanded on this, noting “[m]ost operational databases include wrong data, spelling errors, and different abbreviations. Human names are often misspelled, mistaken for similar-sounding names (such as Keafer and Keefer), or substituted with nicknames (such as Andy and Andrew).” *Id.* at 47.

be the result of inaccurate and faulty matches.”¹⁵ Indeed, large databases managed and gathered in pursuit of homeland security “depend heavily on matching and record linkage methods that intrinsically are statistical in nature, and whose accuracy deteriorates rapidly in the presence of serious measurement error.”¹⁶

More recently, DHS attempted to compile large databases through the short-lived implementation of data analysis and information sharing Repository for Analytics in a Virtualized Environment platform better known as RAVEn.¹⁷ Through this program DHS would “ingest data from tens of thousands of sources, then make it searchable, shareable between agents, and graphable by values like time and place.”¹⁸ However, despite data coming from a wide array of sources, such data was not required to be verified for accuracy.¹⁹ This system fell into disuse and was described as “atrocious” and as hindering, rather than helping, investigations.²⁰

This failure illustrates DHS’ pattern of problematic implementation of complex technological systems. Those close with the agency note that “developing or

¹⁵ STEPHEN E. FEINBERG, *Homeland Insecurity: Data Mining, Privacy, Disclosure Limitation, and the Hunt for Terrorists*, in TERRORISM INFORMATICS: KNOWLEDGE MANAGEMENT AND DATA MINING FOR HOMELAND SECURITY. Eds. 205 (Hsinchun Chen, et al.) (2008).

¹⁶ *Id.* at 213.

¹⁷ Mary Pat Dwyer & Rachel Levinson-Waldman, *A Realignment for Homeland Security Investigations*, Brennan Center for Justice (June 29, 2023) at 12-13, <https://www.brennancenter.org/our-work/research-reports/realignment-homeland-security-investigations> (quoting Caroline Haskins, *Google, Amazon, and Microsoft Have Taken Advantage of a Commonly Used but Little-Known Tool to Quietly Enter Dozens of Contracts with ICE and CBP*, Business Insider (Oct. 4, 2021)) (last visited September 14, 2026).

¹⁸ *Id.* at 13.

¹⁹ This was consistent with other reports that conclusions that Homeland Security Investigations would produce memos targeting and labeling immigrants as gang members based on “speculative evidence, such as unverified information drawn from state and local gang databases, which can contain false and inaccurate information.” *Id.* at 6.

²⁰ Lindsey Wilkinson, *‘Atrocious’ tech hurt ICE investigation branch efforts to counter fentanyl influx*, FedScoop (Aug. 25, 2026), <https://fedscoop.com/ice-hsi-technology-fentanyl-investigation-inspector-general-report/>

changing IT systems at ICE and DHS has long been an arduous, slow process.”²¹

And that despite an influx of funding the agency lacks “a focus on good governance and genuine systems improvement.”²² An inspector agreed, recommending that DHS when dealing with such vast tools, “develop and implement a process that evaluates capabilities of tools prior to deployment...”²³

Regarding SAVE in particular, DHS’ clear disregard for “good governance and genuine systems improvement” is shockingly obvious.²⁴ First, information regarding SAVE since its update has been released “piecemeal” leaving much of the public in the dark about the scope of the update, and importantly, any specific methodology being used to ensure accuracy of the information it outputs. *See* Reply in Support of Pls.’ Mot. For a Stay Under 5 U.S.C. § 705 and a Prelim. Inj. at 4 (D.D.C. Oct. 24, 2024), ECF No. 41. Second, despite claims of improved accuracy, by allowing multiple and different types of identifier numbers, as well as expanding databases referenced, including Social Security Administration data, DHS has compromised the accuracy of SAVE.²⁵

As mentioned *supra*, the lack of a single, common unique identifier makes accurate data matching more prone to error. Searching different identifier numbers across different databases does not necessarily improve SAVE’s accuracy and can

²¹ *Id.*

²² *Id.*

²³ *Id.*

²⁴ *Id.*

²⁵ *See* U.S. SOCIAL SECURITY ADMINISTRATION, LETTER AGREEMENT PROVIDING FOR INFORMATION SHARING BETWEEN THE DEPT OF HOMELAND SEC. (DHS), U.S. CITIZENSHIP AND IMGR. SERVS. (USCIS), AND THE SOC. SEC. ADMIN. (SSA) REGARDING CITIZENSHIP (May 15, 2025), at 1-2, https://www.ssa.gov/foia/resources/proactivedisclosure/2025/May%2015,%202025%20SSA-DHS-USCIS%20Agreement_Redacted.pdf (last visited September 14, 2026).

increase the chances they land on an incorrect match. Further, the expansion of databases to include Social Security Administration data may have little to no effect in identifying noncitizen voters as some jurisdictions do not require provision of Social Security Numbers to register to vote.²⁶ Instead, as will be discussed *infra*, the expansion of databases used for matching *heightens*, rather than improves, the risk of matching error. And given the limited information DHS has publicly provided about SAVE, these issues may only be the tip of the iceberg.

U.S. Census Bureau

An examination of the U.S. Census Bureau shows a contrasting approach to large databases matching. Comparison of the two agencies provides a shocking glimpse into the truly lackadaisical approach DHS has taken in managing their databases. The Census Bureau has pioneered extensive research on database matching. Even so, in contrast to DHS boasting about the accuracy of the updated

²⁶ *See Generally* ELECTION ASSISTANCE COMM'N, REGISTER TO VOTE IN YOUR STATE BY USING THIS POSTCARD FORM AND GUIDE https://www.eac.gov/sites/default/files/eac_assets/1/6/Federal_Voter_Registration_ENG.pdf (last visited September 14, 2026).

SAVE, the Census Bureau acknowledges that, despite methodological advances, matching and record linkage remains an area prone to error and imprecision.²⁷

The Census Bureau maintains its own matching system called *BigMatch*, one of the most accurate data matching systems to date.²⁸ After the creation of this system, the Census Bureau continued to work to improve methodology.

For example, by utilizing a new method in record matching proposed by expert statisticians involving partitioning accounts meant for matching, the Census Bureau was able to improve accuracy and streamline the process.²⁹ The Census Bureau continues to build on this and other theories to “take advantage of the ever-

²⁷ Record Linkage is defined by the Census Bureau as “the means of combining information from a variety of computerized files.” WILLIAM E. WINKLER, U.S. CENSUS BUREAU, OVERVIEW OF RECORD LINKAGE AND CURRENT RESEARCH DIRECTIONS (2006), <https://www.census.gov/library/working-papers/2006/adrm/rrs2006-02.html> (last visited September 14, 2026). Other sources sometimes use the term interchangeably with data matching. See Olivier Binette & Rebecca C. Steorts, *(Almost) All of Entity Resolution*, STAT.ME (2022) at 6 (“Record linkage is broadly used in the literature and refers to merging together multiple databases and removing duplicate records across the databases.”); see also Zhenbang Wang et. al., *Regularization for Shuffled Data Problems via Exponential Family Priors on the Permutation Group*, 206 PMLR 1 (2023) (“The process of identifying matching parts across two or more files is often far from trivial in the absence of unique identifiers, and has thus grown into a vast and active field of research known as record linkage”). For the purposes of this brief, *Amici* refer to record linkage generally as a necessary tool and step for accurate matching analogous to that necessary for SAVE.

²⁸ See U.S. CENSUS BUREAU, RECORD LINKAGE & MACHINE LEARNING (2026), <https://www.census.gov/topics/research/stat-research/expertise/record-linkage.html> (last visited September 14, 2026) (“Multiple evaluations at the Census Bureau ... have shown the record-linkage software developed in CSRM/SRD, such as BigMatch, perform better than open-source competitors in general.”)

²⁹ U.S. CENSUS BUREAU, RECORD LINKAGE & MACHINE LEARNING (2026), <https://www.census.gov/topics/research/stat-research/expertise/record-linkage.html> (last visited September 14, 2026) (“Multi-file record linkage, as proposed by Sadinle and Fienberg proceeds, from information on the constructs underlying all logically possible assignment of multi-record matches as a whole...This raises the potential of retrieving a full spectrum of logically valid record matches between the records of a Census, post-enumeration survey and internal or third-party administrative files simultaneously, rather than piecing together initially independent record pairs, as is mostly done at this time.”)

increasing computational power made available and to expand the latest scientific advances in an equally functional set-up.”³⁰

Despite advances made, the leading expert within the Census Bureau, Dr. Yves Thibaudeau, notes continued difficulties in matching records across multiple databases still exist.³¹ The Census Bureau, even with its careful and constant review of methodology and techniques (such as the data partitioning), acknowledges that there is still work to be done. For example, Dr. Thibaudeau explained that linkage becomes more difficult as the number of fields and databases matched increases.³² According to Thibaudeau, even trying to match four fields across four databases can exacerbate existing problems.³³

Further, in stark contrast to the approach taken by DHS with SAVE, to respond to the complications of matching persons across various information sources absent unique identifiers, the Census Bureau created a system entitled the

³⁰ *Id.*

³¹ Yves Thibaudeau, *A Review of Modern Multinomial-Derived and Partition-Based Record-Linkage Methods* 17 WIREs COMPUT. STAT. 1, 12 (2025), <https://doi.org/10.1002/wics.70015> (“Despite these very encouraging advances, genuinely large applications (linking databases of millions or billions of records) remain beyond the scope of dynamic Bayesian record linkage at the time of this writing.”)

³² Notably, here SAVE attempts to link at least *five* fields (first name, last name, date of birth, social security, reason for verification) across *sixteen* database systems. Appl. for Stay at 6-7.

³³ Yves Thibaudeau, *A Review of Modern Multinomial-Derived and Partition-Based Record-Linkage Methods* 17 WIREs COMPUT. STAT. 1, 13 (2025), <https://doi.org/10.1002/wics.70015> (“The difficulty is the likelihood [equation] becomes analytically complex, as the dimension of the clusters and the number of comparison fields grow... For example, linking four files over four comparison fields entails a likelihood of supporting 24 latent classes and 2208 parameters. The observed sufficient statistic has 7,962,624 cells.”)

Person Identification Validation System to “ascertain unique person and address identifiers.”³⁴ The use of these systems is subject to ongoing review and testing.³⁵

DHS and the Census Bureau show opposite approaches to utilizing large scale databases for matching purposes. While the DHS appears to have vast amounts of funding to improve its methodology, it appears to employ a “throwing at the wall to see what sticks” approach to how to use such data. This is demonstrated by its: disregard for errors and accuracy in data; frequent and unexplained (or simply not thought out) changes in programs and data practices; and the clear lack of information on its methodology and testing on how the agency manages and utilizes the vast array of information it collects. By contrast, the Census Bureau, though it has created a well-thought-out and accurate matching system, recognizes the inherent shortcomings in the science of data matching and continues to research and improve their systems. A cursory glance at the public website managed by the Census Bureau shows a plethora of research on current practices and ongoing projects and studies meant to improve current practices.

³⁴ U.S. CENSUS BUREAU, DEBORAH WAGNER & MARY LAYNE, THE PERSON IDENTIFICATION VALIDATION SYSTEM (PVS): APPLYING THE CENTER FOR ADMINISTRATIVE RECORDS RESEARCH AND APPLICATIONS’ (CARA) RECORD LINKAGE SOFTWARE (2014) AT 2, <https://www.census.gov/content/dam/Census/library/working-papers/2014/adrm/carra-wp-2014-01.pdf> (last visited September 13, 2026).

³⁵ See U.S. CENSUS BUREAU, DEBORAH WAGER, MARY LAYNE AND CYNTHIA ROTHHAAS, ESTIMATING RECORD LINKAGE FALSE MATCH RATE FOR THE PERSON IDENTIFICATION VALIDATION SYSTEM (2014), <https://www.census.gov/content/dam/Census/library/working-papers/2014/adrm/carra-wp-2014-02.pdf> (last visited September 14, 2026).

The contrast between the two is significant. Both agencies manage unfathomable amounts of data that affect the daily lives of Americans; however, only one of them attempts to be guided by the latest social science best practices.

CONCLUSION

The current iteration and use of SAVE stands in direct defiance of proper database use dictated by years of published and recognized social science. Its continued use threatens to harm voters nationwide. For the foregoing reasons, *Amici* respectfully request this Court deny Applicants request for a stay.

September 15, 2026

Respectfully submitted,
UCLA Voting Rights Project
Chad W. Dunn
Legal Director
Bernadette Reyes
Counsel of Record
Sonni Waknin
3250 Public Affairs Building
Los Angeles, CA 90095
310-400-6019
chad@uclavrp.org
bernadette@uclavrp.org
sonni@uclavrp.org

Xavier Becerra
UCLA Voting Rights Project
Senior Advisor and Voting
Rights Counsel
108 W. 2nd Street, #310
Los Angeles, CA 90012
310-924-9310
Scheduling.becerra@gmail.com