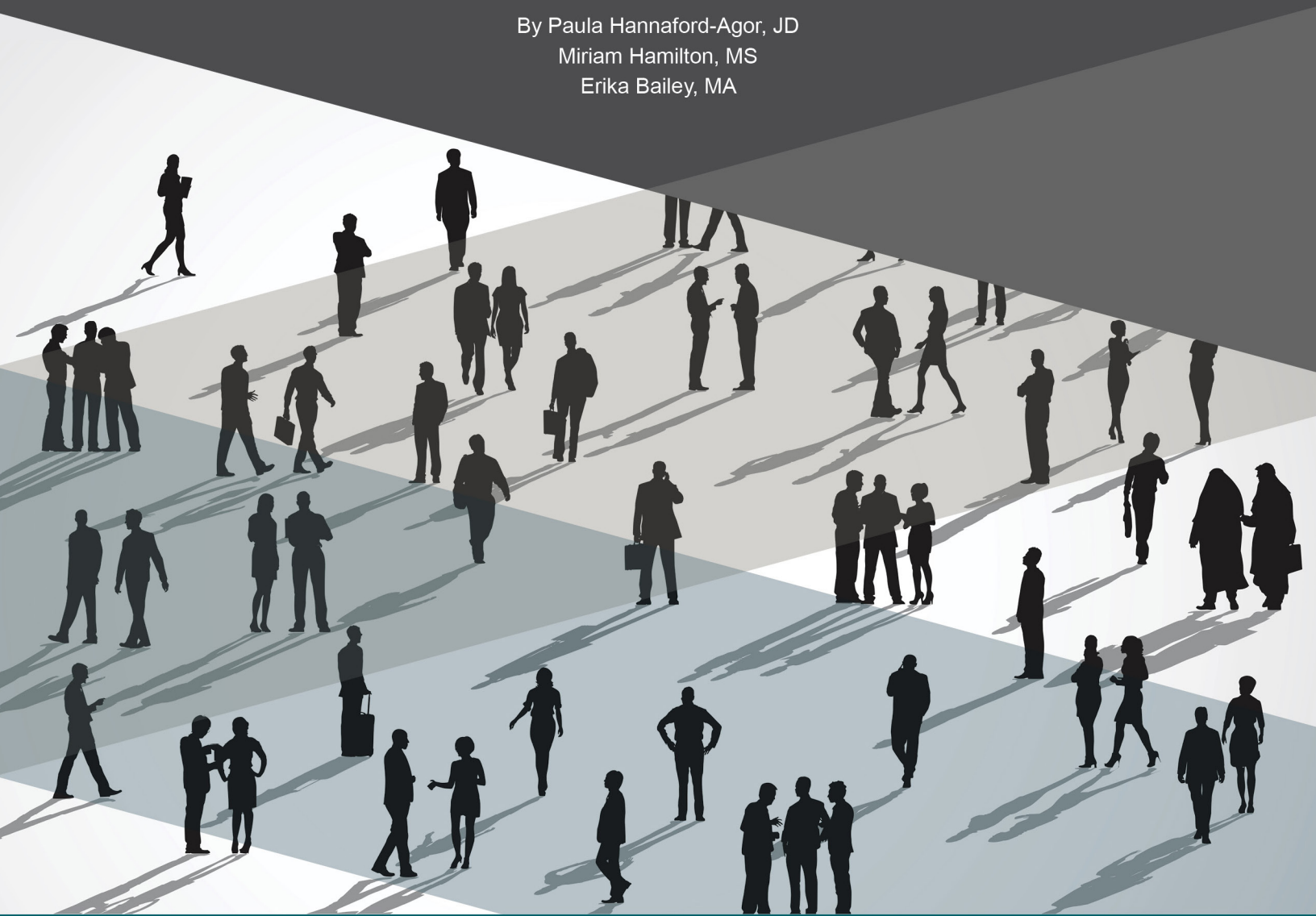
SEPTEMBER 2022

Eliminating Shadows and Ghosts

Findings from a Study of Inclusiveness, Representativeness, and Record Accuracy in Master Jury Lists and Juror Source Lists in Three States

By Paula Hannaford-Agor, JD
Miriam Hamilton, MS
Erika Bailey, MA



Center for Jury Studies
NCSC National Center for State Courts
SJI State Justice Institute

Contents

Introduction

State courts have expressed heightened interest in assessing and improving the demographic representation of jury pools over the past several years. This concern became more urgent in 2020 in light of events causing the American public to question fairness and equality in the criminal justice system. Underrepresentation of people of color may result from multiple factors in the jury selection process, including nonresponse, undeliverable, disqualification, and excusal rates. Thus, it is imperative that the first step, the master jury list from which the names of prospective jurors are initially selected, be as representative as possible. A poorly created master jury list will only compound problems with representation in subsequent steps. Developing and maintaining master jury lists that are broadly inclusive of the jury-eligible population, geographically and demographically representative of their communities, and containing accurate and up-to-date records is critical for safeguarding public confidence in the courts. These three standards are the cornerstone of an ideal master jury list.

An inclusive master jury list is one that includes every citizen eligible for jury service within the geographic jurisdiction served by the court. The concept of inclusiveness is closely related to, but not synonymous with, representativeness. In general, the more inclusive a jury list is, the more representative it is. The standard recommended by the NCSC is that the master jury list consist of unique name and address records for at least 85% of the total adult population within the jurisdiction.[1] To achieve this objective, most states use two or more juror source lists to create the master jury list in a process that involves standardizing the name and address records, merging the files, and identifying and removing duplicate records.[2]

At the same time, the master jury list should not be overinclusive. That is, it should not have more records than adults who reside in the jurisdiction, which can occur due to multiple records for the same person (shadows) or stale records for individuals who no longer live in the community (ghosts). Excessive numbers of shadows and ghosts make it difficult to assess how well the master jury list reflects geographic and demographic representation. It also leads to operational inefficiency. Although it is not feasible to identify every duplicate record and impossible to update the records in real time for prospective jurors who move to a new address, courts should strive to maintain inclusiveness rates as close to 100% as possible. Inclusiveness rates greater than 110% indicate substantial room for improvement.

---

[1] G. Thomas Munsterman, Jury System Management 4-5 (NCSC 1996). The most common combination of juror source lists is the list of registered voters, the list of licensed drivers, and the list of state identification cardholders.

[2] *See Comparative Data* on juror selection and service terms at https://www.ncsc-jurystudies.org/state-of-the-states/jury-data-viz.

Representativeness refers to whether the master jury list reflects the demographic composition of the community, especially with respect to race and ethnicity. In most states, juror source lists do not include race or ethnicity as data elements in the files provided to the court. Consequently, geographic representativeness as measured by whether the proportion of records on the master jury list closely mirrors the proportion of residents living in towns, ZIP Codes, or census tracts is often used as a substitute measure. Underrepresentation of geographically defined communities on the master jury list may signal underrepresentation of distinctive demographic groups that reside in those communities.

The Sixth Amendment to the U.S. Constitution requires that juries be selected from jury pools that reflect "a fair cross section of the community." In *Duren v. Missouri*, the U.S. Supreme Court described three criteria that a criminal defendant must satisfy to show a prima facie violation of the fair cross section requirement: (1) the group alleged to be excluded from the jury pool is a "distinctive" group in the community; (2) the group's representation in the jury pool is not fair and reasonable given the group's representation in the community; and (3) underrepresentation of the distinctive group results from systematic exclusion.[3]

Courts recognize several methods of measuring underrepresentation. The two most widely recognized methods are absolute disparity, which measures the difference between the percentage of a distinctive group in the jury pool and its percentage in the community, and comparative disparity (also called relative disparity), which measures the percentage by which the number of distinctive group members falls short of their number in the community. Comparative disparity is especially important to consider for small minorities, where the absolute difference would otherwise appear inconsequential. The U.S. Supreme Court explicitly declined to establish bright-line numerical thresholds for absolute and comparative disparity,[4] but the converging consensus in case law is that absolute disparities greater than 10% and comparative disparities greater than 50% may be sufficient to show a prima facie violation of the fair cross section requirement. It should be noted, however, that comparative disparities can be misleading when the percentage of the distinctive group in the community is very small (e.g., less than 2%).

In addition to inclusiveness and representativeness, record **accuracy** is an essential characteristic of an effective master jury list. A list that is 100% inclusive and 100% representative would still be ineffective if a large proportion of its name and address records were inaccurate or out-of-date. Inaccurate records decrease the cost-effectiveness of jury operations by incurring printing, postage, and staffing costs for mailings that do not reach the intended recipients. The presence of shadows and ghosts contributes to over-inclusiveness and makes it difficult to assess representativeness, particularly if the inaccurate records disproportionately reflect distinct racial or ethnic groups. State courts cannot control the quality of address records maintained by the state and local government agencies that provide the juror source lists, but they can mitigate their impact by implementing effective duplicate recognition methods, updating the master jury list at least annually, and contracting with commercial vendors to access the U.S. Postal Service (USPS) National Change of Address (NCOA) database to update address records for individuals who have moved to a new address.

[3] *Duren v. Missouri*, 439 U.S. 357 (1979).

[4] *Berghuis v. Smith*, 559 U.S. 314 (2010).

Objectives and Findings of the NCSC Master Jury List Project

In 2021, the National Center for State Courts (NCSC) obtained grant funding from the State Justice Institute to assess the inclusiveness, representativeness, and accuracy of the juror source lists for Missouri, New Jersey, and Tennessee and the resulting master jury lists for Missouri and Tennessee. Each jurisdiction volunteered to be part of the project. Missouri employs two juror source lists to create its master jury list: the list of registered voters, which is maintained by the Secretary of State (SOS list), and the list of licensed drivers and state identification cardholders, which is maintained by the Department of Revenue (DOR list). In New Jersey, three source lists are used to create the master jury list: the list of registered voters (DOS list), the list of licensed drivers and state identification cardholders (MVC list), and the list of state income tax filers (DOT list).[5] In Tennessee, the list of licensed drivers and state identification cardholders maintained by the Tennessee Department of Safety and Homeland Security (DOSHS list) is the sole juror source list used as the master jury list. The NCSC obtained data for the juror source lists from all three states and the combined master jury list from Missouri.[6]

Inclusiveness

For its analyses of inclusiveness, the NCSC compared the total number of records on each juror source list and master jury list with the total number of adult residents in each county reported by the U.S. Census Bureau on the 2020 Census. Table 1 describes the number of counties for each source and master jury list with inclusiveness rates below 85%, from 85% to 100%, from 101% to 110%, and greater than 110%. A majority of counties in Missouri and Tennessee had inclusiveness rates greater than 110% on their respective master jury lists. In New Jersey, the inclusiveness rate also exceeded 110% for the list of licensed drivers and state identification cardholders.

Table 1: Number of Counties, by Inclusiveness Rates

	Missouri			New Jersey			Tennessee
Inclusiveness	Registered Voters	Licensed Drivers/ State ID Cardholders	Master Jury List	Registered Voters	Licensed Drivers/ State ID Cardholders	State Income Tax Filers	Licensed Drivers/ State ID Cardholders
Less than 85%	48	7	0	11	0	15	2
85% to 100%	56	36	4	10	2	6	5
101% to 110%	9	48	4	0	5	0	19
More than 110%	2	24	107	0	14	0	69
Total Counties	115	115	115	21	21	21	95

[5] Consistent with statutory requirements, the New Jersey Judiciary creates a single jury list by combining (sorting and merging) source records supplied annually by the three sources. Through this annual process, New Jersey applies an algorithm to identify potential duplicates, prioritize the most current and reliable record, and eliminate other records for a prospective juror. Jury and technical staff work together to hone and test the algorithm and to visually assess lists to spot check and further reduce unrecognized duplicates. The resulting single jury list, while drawing from the underlying source records, thus is fundamentally different from any of the lists supplied by the statutorily designated sources.

[6] The New Jersey master jury list became available after the study team had started analysis of the source lists. For purposes of this exercise the project team analyzed only the component sources. All descriptions of and conclusions regarding those lists must be considered only as part of an academic analysis and not as a reflection of the list actually used by New Jersey to summons jurors. No inference should be drawn that any over-inclusiveness or under-inclusiveness in those lists correlates to the final New Jersey jury list.

Unrecognized duplicate records are a possible contributor to over-inclusiveness. Table 2, below, shows illustrations of duplicate records. In each instance, missing information or non-standardized formatting indicated in red font caused the computer algorithms to fail to recognize duplicate records. Spelling or other data entry errors, inconsistent use of punctuation, and extraneous spaces are also common causes of unrecognized duplicates.

Using matching criteria that focus on fewer data elements minimizes the risk of unrecognized duplicates but increases the risk that a unique record will be mistakenly identified as a duplicate. For example, if the matching criteria in Table 2 did not include the middle name, the computer algorithm would have recognized that the first two records are identical; adding the middle name to the matching criteria caused the computer algorithm to identify these as two unique records. To the extent that agencies that provide the juror source lists employ different data entry and standardization protocols, unrecognized duplicate errors will occur.

Table 2: Factors Contributing to Unrecognized Duplicate Records

Surname	Suffix	First Name	Middle Name	Street Address	City
DOE	JR	JOHN		34 MAIN STREET	PLAINSBORO
DOE	JR	JOHN	JAKE	34 MAIN STREET	PLAINSBORO
KUMAR		ASHOK	G	12-21 12TH STREET	PLAINSVILLE
KUMAR		ASHOK	G	12-21 12TH STREET APT 1 FL	PLAINSVILLE
MOE		MARTA B		52 DORCHESTER DR	PLAINSCITY
MOE		MARTA B		52 DORCHESTER DRIVE	PLAINSCITY
SMITH		JUAN		6872 3RD AVE	PLAINSTOWN
SMITH		JUAN		6872 THIRD AVE	PLAINSTOWN

The NCSC employed a variety of data matching criteria to investigate the extent to which unrecognized duplicate records contributed to over-inclusiveness on the juror source and master jury lists. As expected, matching on all available data elements yielded very few instances of unrecognized duplicates (ranging from no instances for most lists to 0.9% of records on the Missouri master jury list). Reducing the number of matching criteria to surname, first name, and 5-digit ZIP Code increased the proportion of possible duplicate records, but not enough to account for the extent of over-inclusiveness on the Missouri and Tennessee master jury lists. The reduced matching criteria identified 13.5% of the New Jersey licensed driver/state identification cardholder list as possible duplicate records, three-quarters of which were due to different county designations assigned to records with the same street address. Duplicate records are typically identified and corrected in the summoning process. However, this was likewise insufficient to account for the extent of over-inclusiveness on that list.[7] In all three states, therefore, the presence of stale records for persons no longer living at those addresses is also a likely contributor to over-inclusiveness.[8]

[7] In New Jersey, unrecognized duplicates on the reduced matching criteria accounted for 3.0% of the registered voters list and 3.4% of the state income case filers list. In Missouri, unrecognized duplicates using the same criteria comprised 5.5% of the master jury list, but only 268 records on the list of licensed drivers/state ID cardholders, and 180 records on the list of registered voters. On the Tennessee licensed driver/state ID cardholder list, the rate of unrecognized duplicates was 0.07%, most of which were individuals with multiple types of records (driver's license, driver's permit, state ID card).

[8] For New Jersey, such over-inclusiveness relates only to the underlying sources and not the resulting single jury list, which was not analyzed.

Accuracy

To investigate this possibility, and to assess the accuracy of address records generally for different types of juror source lists, the NCSC contracted with Anchor Computer, Inc. (Anchor), an NCOA^{Link} Full Service Provider,[9] to verify the accuracy of mailing addresses. For budgetary reasons, the NCSC randomly selected approximately 61% of the records from each juror source list and delivered the datasets to Anchor for NCOA processing. Anchor matched the names and addresses of the juror source list records to verify their accuracy and provided updated addresses for individuals who had moved to a new address within the previous 48 months. Records for individuals who moved more than 48 months before NCOA processing would not be identified with an updated address; instead, the original incorrect address would remain on the list returned from Anchor. In addition to new addresses, Anchor provided detailed reports describing the number of records processed; the timeframe of the person's most recent move, if any; and to where the person moved (within ZIP Code, within state, out of state).

Table 3 describes the proportion of records that were identified as persons who filed a forwarding address with the USPS within 12 months, 18 months, and 48 months of the NCOA processing date.[10] Each of these timeframes is important. The 12-month timeframe is important because the USPS will forward first class commercial mail, including jury summonses, to the new address for up to 12 months if a forwarding address has been filed. The 18-month and 48-month timeframes correspond to the differential in access between NCOA^{Link} Limited Service and NCOA^{Link} Full Service Providers.

Table 3: Results of NCOA Processing

Percent Updated Addresses Identified in NCOA Database ...	Missouri		New Jersey			Tennessee
	Registered Voters (n= 2,552,015)	Licensed Drivers/ State ID Cardholders (n= 3,020,730)	Registered Voters (n= 3,676,900)	Licensed Drivers/ State ID Cardholders (n= 5,325,439)	State income tax filers (n= 3,571,233)	Licensed Drivers/ State ID Cardholders (n= 3,853,689)
within 12 months	5.9%	6.6%	5.6%	4.7%	5.4%	4.7%
within 18 months	7.7%	8.8%	7.3%	6.2%	8.6%	6.3%
within 48 months	11.0%	9.9%	8.9%	9.3%	10.8%	10.1%

The percentage of updated addresses varied across the six juror source lists from 8.9% for the New Jersey registered voters list to 11.0% for the Missouri of registered voters. State migration rates — that is, the percentage of the population that moves to a new address each year — account for some of this variation. Table 4 describes the percentage of the adult population that reported moving to a new address in the previous 12 months on the 2020 decennial census, and the percentage who moved to a new address within the same county.[11] The first Census-based percentage provides an important

[9] The USPS also offers three classes of licenses to commercial vendors to access the NCOA database. NCOA^{Link} Full Service Providers are licensed to access up to 48 months of NCOA records; NCOA^{Link} Limited Service Providers are licensed to access up to 18 months of NCOA records; and NCOA^{Link} End User Mailer Providers are licensed to incorporate the interface software necessary to access up to 18 months of NCOA records.

[10] In all three states, mover rates were higher in urban areas compared to suburban and rural areas.

[11] American Community Survey 5-Year-Estimates (2016-2020), Table B07001 (Geographical Mobility in the Past Year by Age for Current Residence).

contextual frame of reference for comparing list accuracy across jurisdictions. The second highlights the practical benefit of NCOA processing — namely, that in all three states, approximately half of the updated addresses were likely for individuals who moved to a new address in the same county and would thus still be eligible for jury service if the jury summons were delivered to the correct address. In addition, migration rates for Whites were lower than for other racial and ethnic groups in all three states.[12] Consequently, NCOA updates would improve representation in the jury pool by ensuring that people of color, who move at higher rates, receive their jury summons at the correct address.

Table 4: Geographical Mobility in the Past Year

	Missouri	New Jersey	Tennessee
Percent of Population that Moved from a New Address in the Past Year	14.5%	10.3%	14.2%
Percent of Population that Moved to a New Address in the same County	7.4%	5.1%	7.6%

Source: Table B07001, American Community Survey 5-Year Estimates (2016-2020)

Examining the ratio between state migration rates and the rate of updated addresses after NCOA processing indicates that the type of list — registered voters, licensed drivers/state ID cardholders, or state income tax filers — is not necessarily a consistent indicator of record accuracy. For example, Missouri residents moved 41% more frequently than New Jersey residents (14.5% / 10.3% = 1.41). The Missouri list of licensed drivers/state ID cardholders had 40% more updated addresses from NCOA processing than the New Jersey MCV list (6.6% / 4.7% = 1.40). This suggests that the difference in list accuracy can mostly be explained by the difference in underlying state migration rates rather than by differences in how the respective state agencies maintain those lists. In contrast, the state migration rate in Tennessee was very similar to that of Missouri (14.2% versus 14.5%), but the rate of updated addresses in Tennessee after NCOA processing was identical to New Jersey's rate (both 4.7%), which suggests that this Tennessee list is more accurate than both the Missouri and New Jersey licensed driver/state ID cardholder lists, especially after accounting for differences in migration rates. This is particularly noteworthy given the renewal periods for drivers' licenses across the three states (four years in New Jersey, six years in Missouri, and eight years in Tennessee).[13]

[12] In Missouri, the percentage of Whites who moved from another location in the previous year was 14% compared to 19% for Black/African Americans, 18% for Native Americans, 20% for Asians, 26% for Hawaiian/Pacific Islanders, 15% for Other Race, 19% for 2 or more races, and 16% for Hispanic/Latinos. In New Jersey, the migration rate for Whites was 9% compared to 13% for Black/African Americans, 12% for Native Americans, 14% for Asians, 8% for Hawaiian/Pacific Islanders, 13% for Other Race, 12% for 2 or more races, and 12% for Hispanic/Latinos. In Tennessee, the migration rate for Whites was 14% compared to 16% for Black/African Americans, 19% for Native Americans, 19% for Asians, 17% for Hawaiian/Pacific Islanders, 16% for Other Race, 19% for 2 or more races, and 17% for Hispanic/Latinos. American Community Survey, Table S0701, Geographic Mobility by Selected Characteristics, 2020 5-Year-Estimates.

[13] All three states require licensed drivers to notify the respective agencies of their new addresses shortly after moving (30 days in Missouri, 60 days in New Jersey, and 10 days in Tennessee). However, it is clear from the percentage of updated addresses identified for individuals who moved at least 18 months before NCOA processing that sizeable numbers fail to comply with these requirements.

Representativeness

For the assessment of representativeness, the NCSC employed the jurisprudential framework outlined in *Duren v. Missouri*. Specifically, it measured the absolute and comparative disparities of distinctive racial and ethnic groups between the jury-eligible population and the master jury list. Few juror source lists include data about the race and ethnicity of prospective jurors, usually because the agencies that maintain those lists do not collect demographic information themselves. To compensate for the lack of accurate data, researchers often use geocoding techniques to infer the race and ethnicity of the populations they study.[14] The geocoding models were based on data from the U.S. Census Bureau for the adult population[15] of each ZIP Code Tabulation Area (ZCTA) in the respective states aggregated to the county-level.[16] The estimated county-level race and ethnicity percentages on each list were then compared to the jury-eligible (adult citizen) population in each county.

The Tennessee licensed driver/state ID cardholder list, used as the sole source list for the master jury list, included self-reported race as one of the data elements, which the NCSC used for its representation analysis. However, the racial classification employed by the Tennessee Department of Safety and Homeland Security differs from that of the U.S. Census Bureau, making it impossible to perform an apples-to-apples comparison. The Tennessee dataset, for example, includes Hispanic/Latino as a racial category rather than a separate ethnic category. In addition, Hawaiian/Pacific Islander and multiracial are race categories in U.S. Census Bureau definitions but not in DOSHS definitions. It is likely that multiracial persons would select the race with which they most closely identify, but it is unclear whether most Hawaiian/Pacific Islanders would select Asian or Other as their race under the DOSHS classification. Likewise, it is probable that the Census-derived Hispanic-inclusive percentages would include a higher count for White and Other than in the Tennessee list, as Hispanic individuals often chose one of these two races when asked to choose a racial identity other than Hispanic.[17]

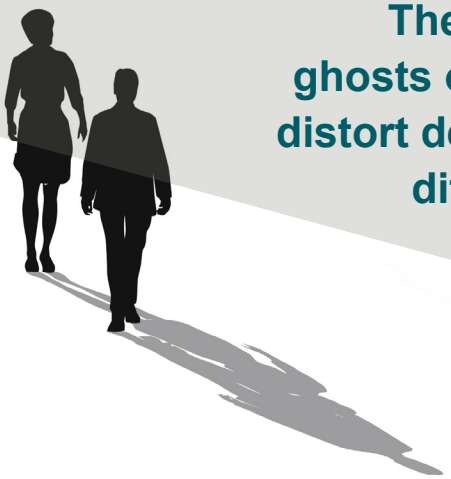
---

[14] Geocoding is a technique used to determine the probability that an individual possesses a defined characteristic (e.g., race/ethnicity, gender, education, income) based on the demographic characteristics of the geographic area (state, county, township, census block/tract) where the individual resides. Because racial and ethnic populations in the United States tend to congregate in the same neighborhoods within local communities, geocoding is a useful technique for estimating the demographic characteristics of a random sample of individuals drawn from those communities when the actual characteristic is unknown.

[15] At the time the analyses were conducted, the U.S. Census Bureau had not yet posted detailed datasets for the 2020 decennial census. Representation analyses employed estimated population totals from the American Community Survey (ACS) 5-Year Average for the Years 2015-2019.

[16] ZIP Code Tabulation Areas (ZCTAs) reflect the geographic boundaries for U.S. Postal Service ZIP Codes. No corresponding ZCTAs exist for ZIP Codes reflecting Post Office boxes or large volume customer addresses, such as commercial, government, or university ZIP Codes., which consequently were excluded from analysis. The geocoding models also excluded records for individuals with missing, incomplete, or out-of-state ZIP Codes.

[17] The New Jersey Supreme Court in *State v. Dangcil*, directed the Administrative Office of the Courts to collect voluntary juror demographic information as to race, ethnicity, and gender. Such data collection has started in three counties and will expand statewide in 2023. Going forward, New Jersey thus will be able to analyze actual demographic data and will not need to rely on the indirect geocoding process.



The presence of shadows and ghosts on the master jury lists can distort demographic representation differently in different states.

The *Duren* framework specifies that the relevant jury pool for the purpose of fair cross section analyses is the jury-eligible population residing within the geographic boundaries of the jurisdiction served by the court. Thus, the relevant jury pool in state general jurisdiction courts is the county. In an ideal master jury list, comparative disparity for each racial and ethnic category should be at or close to zero, but it is perhaps unreasonable to expect that a single source list or even a combination of source lists perfectly reflects the demographic composition of the community evenly across an entire state. Table 5 shows the number of counties in each state at different levels of comparative disparities for the racial and ethnic groups that are most often the subject of fair cross section challenges.

Overall, the representation of these groups is relatively close to U.S. Census Bureau estimates. All counties in each of the states had comparative disparities of less than 20% for Whites. On average, 89% of counties had comparative disparities of less than 20% for Blacks/African Americans, with 5% of counties underrepresenting and 6% overrepresenting Blacks/African Americans. Native American representation does not fare quite as well with only an average of 72% of counties showing less than 20% comparative disparity, an average of 15% showing underrepresentation, and 12% showing overrepresentation.[18] Asian and Hispanic populations were more likely to be overrepresented on the lists, but this was expected because records of non-citizens are included on the lists of licensed drivers/ state ID cardholders.[19] On average, only 3% of counties showed underrepresentation of Asians and 4% showed underrepresentation of Hispanic/Latinos.

[18] Negative and positive comparative disparities indicate underrepresentation or overrepresentation, respectively. Comparative disparities can be misleading when the percentage of the distinctive group in the community is very small. Native Americans comprised less than 2% of the population in 113 out of 115 counties in Missouri, 94 out of 95 counties in Tennessee, and every county in New Jersey. States with large comparative disparities should engage in a close, county-specific examination of the data to assess the practical and legal implications of underrepresentation or overrepresentation for distinctive groups.

[19] In all three states, the NCSC found statistically significant correlations between non-citizenship rates for Asians and Hispanic/ Latinos and overrepresentation compared to the jury-eligible population.

Table 5: Number of Counties, by Comparative Disparity					
	Missouri	New Jersey			Tennessee
	Master Jury List	Registered Voters	Licensed Drivers/State ID Cardholders	State Income Tax Filers	Master Jury List
Comparative Disparities for Whites					
Less than -50%	0	0	0	0	0
-50% to 21%	0	0	0	0	0
-20% to 20%	115	21	21	21	95
21% to 50%	0	0	0	0	0
More than 50%	0	0	0	0	0
Comparative Disparities for Blacks/African-Americans					
Less than -50%	2	0	0	0	9
-50% to 21%	3	0	1	0	7
-20% to 20%	95	21	19	21	66
21% to 50%	8	0	1	0	10
More than 50%	6	0	0	0	3
Comparative Disparities for Native Americans					
Less than -50%	3	0	0	0	25
-50% to 21%	4	1	0	1	29
-20% to 20%	88	19	18	19	13
21% to 50%	10	1	0	1	6
More than 50%	4	0	1	0	14
Comparative Disparities for Asians					
Less than -50%	1	0	0	0	2
-50% to 21%	7	0	0	0	4
-20% to 20%	29	4	8	5	11
21% to 50%	23	15	11	13	3
More than 50%	42	2	2	3	67
Comparative Disparities for Hispanic/Latinos					
Less than -50%	3	0	0	0	2
-50% to 21%	4	0	0	0	9
-20% to 20%	60	13	6	11	25
21% to 50%	26	8	13	10	6
More than 50%	21	0	2	0	51

Although the overall picture of representativeness appears positive, it should be noted that the NCSC assessment of demographic representation was greatly complicated by the over-inclusiveness of the lists. The demographic characteristics of shadows and ghosts might not be randomly distributed, which could affect the overall demographic composition of the jury pools in individual counties. Moreover, the impact of shadows and ghosts could move in opposite directions and possibly cancel each other out.[20]

To investigate this question, the NCSC examined the relationship between over-inclusiveness and overrepresentation of different racial and ethnic groups, which yielded mixed results. In Missouri, for example, there was a positive correlation between inclusiveness and Hispanic/Latino representation[21] — that is, the more over-inclusive the master jury list for each county, the more overrepresented were Hispanics; there was no relationship for other racial and ethnic groups.

In New Jersey, inclusiveness for the record sources correlated with different racial groups on some but not all lists. On the state income tax list and the registered voter list, inclusiveness was positively correlated with representativeness for Black/African Americans.[22] Recall, however, that the majority of counties on both lists were underinclusive and Black/African Americans were underrepresented. As inclusiveness increased, Black/African American representation likewise increased. On the list of licensed drivers/state ID cardholders, however, inclusiveness was positively correlated with White representativeness. Whites were underrepresented in most counties, but the disparities for Whites tended to be smaller in counties in which the list was more over-inclusive.[23] This list was the most over-inclusive, mostly due to unrecognized duplicates related to multiple county designations for individuals living at the same address.

Both Black/African American and White representation were correlated with inclusiveness in Tennessee, but in different directions. Inclusiveness was positively correlated with representativeness for Blacks/African Americans, but negatively correlated with White representation.[24] That is, as the inclusiveness rate increased in a county, Black representation increased and White representation decreased in relation to the jury-eligible population as reported by the U.S. Census Bureau.

---

[20] Some fair cross section challenges have been raised alleging that unrecognized duplicate records disproportionately reflect Whites on the proposition that Whites are more likely than other racial and ethnic groups to appear on the juror source lists used to create the master jury list. *See, e.g., California v. Reggie D. Cole*, No. CF8268 (Superior Court of Imperial County, California); *Arizona v. Edward Vincent Martinez* et al., No. CR2017-150971-001 (Superior Court of Arizona in Maricopa County). Conversely, non-White populations typically have higher migration rates, so stale addresses may disproportionately reflect those populations.

[21] Pearson's $r = .21$, $p = .029$.

[22] DOT List Pearson's $r = .47$, $p = .032$; DOS List Pearson's $r = .66$, $p = .001$.

[23] MVC List Pearson's $r = .46$, $p = .036$.

[24] Black Pearson's $r = .31$, $p = .003$; White Pearson's $r = -.51$, $p < .001$.

Conclusions

The findings from these analyses illustrate the challenges that state courts face in their efforts to create inclusive, representative, and accurate master jury lists, particularly with respect to selecting the optimal number and types of juror source lists. The conventional wisdom for the past 50 years has been that reliance on a single source list, especially the registered voters list, was a substantial cause of underrepresentation for people of color. Since the 1970s, state courts have increasingly moved toward the use of multiple juror source lists — predominantly a combination of registered voters and licensed drivers/state ID cardholders — to ensure that the master jury list reflects the demographic characteristics of the community. Several states, including New Jersey, now use three or more juror source lists, including state income tax, public welfare recipient lists, and unemployment compensation recipient lists. As these analyses demonstrate, however, source list supplementation may have reached the limits of its effectiveness.

- **Substantial over-inclusiveness was more common than under-inclusiveness for many counties in each of the three states. Each type of source list introduced varying numbers of shadows and ghosts.**

The resulting over-inclusiveness can be as problematic as under-inclusiveness. The presence of shadows and ghosts on the master jury list can distort representation in different ways. In some states, it may mask underrepresentation of distinctive groups; in others, it may cause concerns about underrepresentation that does not really exist. This risk increases as each additional juror source list is merged with others to form the master jury list. Instead of adding more source lists, state courts must become more selective about the quality of source lists employed and the care with which the master jury list is created and maintained. They should use only as many source lists as necessary to achieve inclusiveness at or near 100%. The choice of which juror source lists to use should be determined by the highest quality lists available in terms of accuracy, which may differ from state to state, as illustrated by the results of NCOA processing.

- **Different combinations of source lists ultimately may not achieve acceptable demographic representation for all counties across the entire state.**

Distinctive groups may be underrepresented in some counties, overrepresented in others, and adequately represented in the rest. In those instances, state court policymakers must decide whether a uniform statewide policy should prevail or whether individual counties may be authorized to supplement local master jury lists to achieve desired representation for one or more distinctive groups. If the latter, local court policymakers should do so only after determining that additional supplementation is likely to achieve the desired demographic composition and that the court has the capability to mitigate the risks associated with over-inclusiveness.

- **The most accurate type of source list may differ from state to state.**

Until relatively recently, it was widely assumed that state drivers' license lists had more accurate address records than registered voters. Most states required drivers to renew their licenses every four years. Because a valid driver's license serves as an economic lifeline for many individuals, incentives for keeping their license active and in good standing were comparatively higher than for notifying voter registrars about address changes between elections. Moreover, incentives to notify state motor vehicle agencies of address changes were also much higher when drivers' licenses were used to verify identification and current address. In contrast, federal legislation requires rigorous documentation that a person no longer lives at an address before their voter registration record can be deactivated or removed. Consequently, the accuracy of address records on registered voter lists tends to deteriorate in the interval between presential elections, especially for individuals who vote infrequently.

Some of these underlying assumptions may no longer be valid. During the economic recession in 2008, for example, many states increased the renewal period for drivers' licenses as a cost-saving measure.[25] In addition, the frequency with which people pay for goods or services with a personal check, which often required identity verification, has largely evaporated with the proliferation of options for electronic payment (e.g., credit/debit cards, PayPal, Venmo, ApplePay). Yet the rates of updated addresses after NCOA processing for the three licensed driver/state ID cardholder lists were relatively close in spite of differences in renewal period and state migration rates. In the meantime, voter turnout has increased over the past two decades, especially for presidential elections,[26] and state voter registrars have increased efforts to verify the accuracy of voter addresses to confirm eligibility to vote.[27] The New Jersey voter registration list had the lowest rate of updated addresses after NCOA processing while Missouri had the highest.

---

[25] Arizona has the longest renewal period: 12 years. See https://www.iii.org/state-drivers-license-renewal-laws-including-requirements-for-older-drivers for a state-by-state list of license renewal requirements.

[26] See https://www.fairvote.org/voter_turnout#voter_turnout_101 for a graph of voter turnout rates from 1789 to 2020.

[27] National Conference of State Legislators, Voter Registration List Maintenance (available at https://www.ncsl.org/research/elections-and-campaigns/voter-list-accuracy.aspx#resources).

Several states have recently sought legislation to require the use of state income tax lists as a juror source list because taxes must be filed annually, giving state revenue agencies more frequent notice of address changes. It was therefore even more surprising that the New Jersey state income tax filers list had the second highest updated address rate among the juror source lists (10.8%). In fairness to the New Jersey Department of Taxation, however, the list of state income tax filers used in the NCSC master jury list project was delivered to the NJ Administrative Office of the Courts in January 2021, well before the July 15, 2021 deadline for filing 2020 NJ state income tax returns.[28] Thus, most of the address records would have reflected 2019 tax returns that were filed in 2020, approximately 18 months before they were submitted for NCOA processing. A more accurate comparison therefore would be the percentages of updated addresses for the 12 months and 18 months preceding January 2021 (4.7% and 5.0%, respectively), which are an improvement over other juror source lists both in New Jersey and in Missouri and Tennessee.

The time lag for the New Jersey income tax list illustrates an important consideration for state courts about the optimal timing for receipt of the source lists to maximize their "ripeness" for use in the master jury list. State income tax lists are probably at their most accurate on the date of the state income tax filing deadline, which ranges from April 15 to May 15. Voters lists tend to peak in the first week of November before a presidential election. Most states sync renewal of drivers' licenses to the drivers' birthdays, so there are no obvious external dates on the calendar that would cause a peak in the accuracy of those lists. Ideally, the state should time its request to the source list agency to ensure receipt just as the list has reached its accuracy peak, and then create and distribute the master jury list as quickly as possible so that courts will realize the benefits of more accurate address records. States using multiple juror source lists should set the timeline based on the most accurate source list. This may be complicated, however, by statutory requirements that establish a different deadline (January 1, September 1) for creating the master jury list that do not correspond with the optimal timeframe for using the source lists.

Additionally, most courts require three to six months to complete the steps needed to create and implement a new master jury list. For courts that routinely do so, NCOA processing is typically one of the last steps in that process.[29] For this reason, the time lag is an important factor to consider when choosing whether to select a Full Service (48-month) or Limited Service (18-month) Provider for NCOA processing. For example, half of the NCOA updated records identified in the New Jersey state income tax list were individuals who moved within 12 months of NCOA processing; 80% moved within 18 months of NCOA processing. An NCOA^{Link} Full Service Provider would have identified only 25% more updated records than an NCOA^{Link} Limited Service Provider. In contrast, only 62% of records on the Tennessee licensed drivers/state ID cardholders list moved within 18 months of NCOA processing. A 48-month NCOA process would return 61% more updated addresses. Depending on the cost differential between the two types of NCOA providers, the additional benefit in terms of deliverable jury summonses (and corresponding decrease in lost printing and postage costs) might outweigh the additional costs.

[28] New Jersey extended the 2020 state income tax filing deadline from mid-April to mid-July 2021 due to the coronavirus pandemic.

[29] The typical process for creating a new master jury list involves (1) obtaining the juror source lists from the source list agencies; (2) cleaning and standardizing records for all source lists; (3) merging the source lists into a single list; (4) identifying and removing duplicate records; (5) assigning a county designation to each record for use in local jury operations; (6) NCOA processing; (7) randomizing records on the master jury list; and (8) distributing the master jury list to local trial courts.

Recommendations for Effective Master Jury Lists

The master jury list is the first step in the juror selection process. An effective master jury list is one that is inclusive of the adult population, geographically and demographically representative of the community, and contains accurate address records for prospective jurors. Maintaining 100% inclusiveness is impractical given the intricacies of duplicate identification algorithms and normal migration rates into, out of, and within communities. To maximize accuracy, the master jury list should be refreshed at least annually.

The quality of juror source lists used to create the master jury list is highly dependent on the source list agency. Each of these agencies have their own statutorily mandated missions and have developed data governance protocols to help them achieve them, including protecting the legitimate privacy interests of individuals to whom the data refer. Although state courts use agency data to create the master jury list, the datasets were not originally developed for that purpose and the source list agencies have no obligation to collect or format data for that purpose. State courts must also recognize that it takes time to engage with state agencies and to develop trust that courts will protect their data, especially when beginning a relationship with an agency that has not previously provided their data for the use in creating the master jury list. Although the present study cannot definitively state the optimal number and types of source lists to use in creating the master jury list, it does provide criteria that state court policymakers can use to conduct their own assessments.

Recommendations for Effective Master Jury Lists

1

Obtain documentation on how source list agencies maintain their databases.

Understanding how source list agencies maintain their databases can alert state courts to potential practices that may facilitate or complicate identification of duplicate or stale records. For example:

a. Does the source list agency employ a data standardization process for record names and addresses? If so, secure a copy of the standardization format.
b. Are records manually entered by agency staff?
c. How frequently do the source list agencies validate records?
d. What indicators does the source list agency employ to flag inaccurate or stale records?
e. For how long are inactive records retained in the database before they are removed?
f. Does the source list agency employ NCOA on its address records? If so, request a copy of a recent NCOA report to assess the general accuracy of the list, especially the proportion of records indicating movers within 12 months and within 18 months of NCOA processing.

2

Request that source list agencies provide data elements that could assist the court with its master jury list creation and maintenance.

Some source list agencies collect personal information that may help the court screen for eligibility and increase the accuracy of duplicate recognition and address validation processes. If available, request the following data elements for each record on the juror source list:

a. First name, middle name or initial, surname, and suffix
b. Street address, city, state, and ZIP+4
c. Mailing address, if different from street address
d. Date record was last updated
e. Date of birth
f. Identification number assigned by source list agency
g. Record status (e.g., active, inactive/suspended)
h. U.S. citizenship status
i. Race, ethnicity, and gender
j. Email address
l. Telephone number

Recommendations for Effective Master Jury Lists

3

Confirm the criteria for the data extract with the source list agency.

Many courts assume that the source list agency will provide all records from the source list. However, the court may be able to specifically request that the source list agency exclude certain records (e.g., under age 18, non-citizen, inactive records). If so, request a count and reason for excluded records with the data extract. Compare the size of the list against the list from the previous data request and seek an explanation if the new list is substantially larger or smaller than can be explained by migration during the same period.

4

Identify and remove internal duplicate records.

Review the source lists individually before merging with other lists to identify and remove duplicate records. Matching criteria should include surname, first name, middle initial, suffix; street address; and date of birth, if available.

5

Before merging source lists, format the source lists to minimize the incidence of unrecognized duplicate records by standardizing the name, address, and dates in data elements.

The U.S. Postal Service has developed the Coding Accuracy Support System (CASS) for standardizing and verifying mailing addresses, including standard formats for directional indicators, apartment indicators, numerical street names, street abbreviations, and ZIP+4 designations. The court should standardize all record addresses according to CASS requirements before merging files.

To minimize unrecognized duplicate names, use matching criteria that control for inconsistent use of punctuation, capitalization, spaces, suffix formats (e.g., Jr. versus 2nd versus II, Third versus 3rd versus III and missing data. Validate date formats in date of birth, date of record updates, etc., including flagging invalid months (greater than 12), days (greater than 31) or years (e.g., before 1922, after 2022).

Recommendations for Effective Master Jury Lists

6

Matching criteria should consist of the following:

a. Surname, first name, middle initial, suffix;

b. Street address only;

c. DOB if the data element is available on all source lists;

d. If most recent date updated is available for all source lists, match on name and DOB, and, in case of a duplicate record, retain the address for most recent update; otherwise, retain the address for the source list with the highest address record accuracy (as indicated by the fewest NCOA updates on recent source list agency reports).

e. Hyphenated surnames can pose unique challenges, especially for individuals who have added their spouse's surname. Source list records may reflect the original (birth) surname, the hyphenated name, the reverse of the hyphenated name, or the spouse's surname only, depending on when the record was first created (before or after the name change) and whether it has been updated. If possible, identify all hyphenated surnames and search for matching records using different combinations of the surname.

7

When matching and removing duplicate records, retain data elements indicating the original source list.

After the list merge/purge is complete, prepare a report showing the number and proportion of master jury list records that are unique to each source list and those that appear in each combination of source lists. In the event of over-inclusiveness on the master jury list, this report may help the court identify source lists that contribute to unrecognized duplicates or stale records.

Recommendations for Effective Master Jury Lists

8

As a quality control measure, spot check for unrecognized duplicates.

After all duplicate records have been removed, sort the master jury list by street address, city, and full name. Randomly select 50 records and extract the record, any records at the same street address that precede the record, and the next 100 records following the record. This will create a dataset consisting of approximately 5,000 records organized by household. Check names living in the same household to identify unrecognized duplicates. Review the cause of the duplicate and develop data matching criteria to identify duplicate records in the future.

9

Use NCOA to update address records.

If the court completes its list merge/purge process within 90 days of receiving source lists, and if NCOA reports indicate that 80% or more updated records moved within 18 months of NCOA processing, it may be more economical to employ an NCOA^{Link} Limited Service Provider; otherwise, use an NCOA^{Link} Full Service Provider for a full 48-month NCOA update. Exclude out-of-state records from the master jury list. Records for out-of-county movers can be transferred to the correct county for inclusion on the correct county's master jury list. In-county records should be retained on the master jury list.

Recommendations for Effective Master Jury Lists

10

Retain all NCOA updated records with the moving date as a separate file for use in future master jury lists. In the same datafile, include self-report address changes from prospective jurors.

The NCOA database includes forwarding addresses for up to 48 months. Records for individuals who moved more than 48 months ago will not be updated through an NCOA^{Link} Full Service Provider. Similarly, records for individuals who moved more than 18 months ago will not be updated through an NCOA^{Link} Limited Service Provider. The court should retain the old and updated records. When reviewing the source lists at each master jury list cycle, use the retained NCOA records to update stale records that persist on the source lists. Over time, the NCOA updated address will allow the court to continue updating records not flagged by most recent NCOA processing.

11

Compare the number of records on the master jury list to the adult population of the jurisdiction.

Inclusiveness is the ratio of the total number of individuals on the master jury list to the total adult population. The resulting master jury list ideally should be at or near 100%. If it is less than 85%, the court should consider adding another source list; if it is greater than 110%, review the list for unrecognized duplicate and stale records.

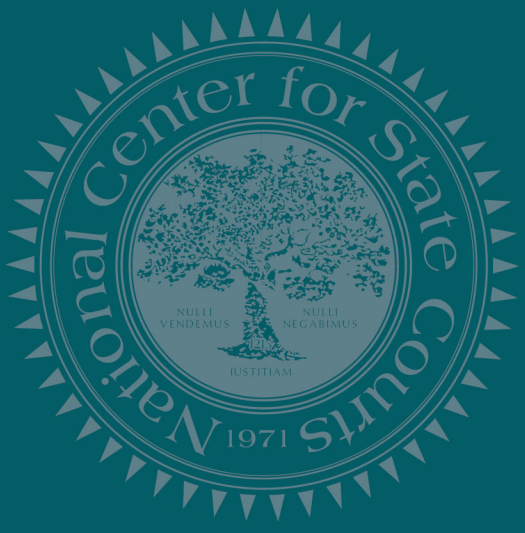
Acknowledgements

The genesis of this research project was renewed state court interest in the diversity of jury pools following widespread racial unrest across the country after the murder of George Floyd in Minneapolis, Minnesota. The project team is grateful to members of the joint CCJ/COSCA Trial Court Management Committee who supported NCSC efforts to secure funding for the project and encouraged its completion.
We also benefitted greatly from the advice and encouragement of project advisory committee members who provided helpful commentary about the scope of the project, the research methods, and the conclusions and recommendations for state courts, including:

- **Glenn Grant**, Administrative Director of the New Jersey Courts
- **Brian McLaughlin**, Manager of Jury Programs, New Jersey Administrative Office of the Courts
- **Jennifer Weatherby-Morris**, Chief of Support Services, New Jersey Administrative Office of the Courts
- **Jessica Lewis-Kelley**, Special Assistant to the Administrative Director, New Jersey Administrative Office of the Courts
- **Judge Danielle Viola**, Chair of the Jury Committee, Superior Court of Arizona in Maricopa County
- **Andrei Rojdestvenski**, Senior Programmer/Analyst, Superior Court of Arizona in Maricopa County
- **Raymond Billotte**, Court Administrator, Superior Court of Arizona in Maricopa County
- **Matthew Martin**, Jury Manager, Superior Court of Arizona in Maricopa County
- **Kathy Lloyd**, Missouri State Courts Administrator
- **Catherine Zacharias**, Legal Counsel, Office of the Missouri State Courts Administrator
- **Tina Jones**, Support Services, Office of the Missouri State Courts Administrator
- **Joanne Martin**, Jury Supervisor, 22nd Circuit Court, City of St. Louis, Missouri
- **Judge Kenneth Garrett**, 16th Circuit Court, Independence, Missouri
- **Deborah Taylor Tate**, former Director of the Tennessee Administrative Office of the Courts
- **Michelle Long**, Director of the Tennessee Administrative Office of the Courts
- **Charlie Baldwin**, Assistant General Counsel, Tennessee Judiciary
- **Rick Pierce**, Judicial Programs Administrator, Administrative Office of the Pennsylvania Courts
- **Ami Levin**, Data Exchange Project Manager, Administrative Office of the Pennsylvania Courts
- **Dr. Cheryl Bailey**, Acting Executive Officer, DC Superior Court
- **Charleston Carter**, Trial Court Administrator, DC Superior Court
- **Bonnie Hoffman**, Director of Public Defense Reform & Training, National Association of Criminal Defense Lawyers
- **Professor Nina Chernoff**, CUNY School of Law

Of course, the project never would have happened without the cooperation of the Administrative Offices of the Courts in Missouri, New Jersey, and Tennessee, which secured permission from their respective juror source list agencies to share the lists with the NCSC for analysis. We learned so much from this project, some of which was suspected, but much of which surprised us and even contradicted long-held beliefs about effective master jury list management practices. We are grateful for their generosity and trust.

At the National Center for State Courts, we are also grateful for support of our colleagues Edwin Bell, Director of Racial Justice, Equity and Inclusion; Judge Gregory E. Mize (ret.), Judicial Fellow and editor of *Jur-E Bulletin*; Keeley Daye, Program Specialist; and Tom Munsterman, founder and continued supporter of the NCSC Center for Jury Studies. Finally, we thank the State Justice Institute for generous financial support for the project (SJI-21-T-030). Views expressed in this report are those of the authors and do not necessarily represent those of State Justice Institute, the National Center for State Courts, or the judicial branches of Missouri, New Jersey, or Tennessee.



ncsc.org

ISBN 978-0-89656-326-1

Suggested citation: Paula Hannaford-Agor, Miriam Hamilton & Erika Bailey, Eliminating Shadows and Ghosts: Findings from a Study of Inclusiveness, Representativeness, and Record Accuracy in Master Jury Lists and Juror Source Lists in Three States.