SUPPORTING STATEMENT (PART B)
Criminal Cases in State Courts (CCSC), 2019

B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS

1. Universe and Respondent Selection

The purpose of the CCSC is to better understand the number and characteristics of felony and serious misdemeanor criminal cases processed in state courts in the 50 states and the District of Columbia. States vary considerably with respect to the types of cases handled in limited jurisdiction or general jurisdiction courts. They also vary with respect to the ability of their courts to provide case-level data on these matters easily and electronically. The CCSC takes advantage of and addresses this variation in data availability by combining a census and sample-based data collection.

The target population for the 2019 CCSC is all original felony and serious misdemeanor cases with a potential punishment of one year or more incarceration, disposed of in state criminal courts of general and limited jurisdiction in 2019, excluding municipal courts.[1] BJS does not have an estimate of the criminal cases filed in every state; however, population is highly correlated with case filings. Using annual reports for 2017 on criminal cases filed in each county for 6 states (Maryland, Maine, Pennsylvania, Colorado, Wisconsin and Montana),[2] and Census estimates of the population from these counties in 2017, the Urban Institute (data collection agent for CCSC) examined the correlation between the total population size and the number of case filings and confirmed that it exceeded 0.90. For that reason, for sampling purposes BJS plans to use the size of the population aged 18 and over as a proxy for the number of cases likely to be charged in criminal courts.[3]

---

[1] This definition excludes violations of probation and all civil cases, including traffic offenses (if charged civilly instead of criminally), municipal ordinance violations, infractions, misdemeanors with potential punishments of less than one year incarceration, fish and game commission charges, and habeas corpus petitions.

[2] The volume of court filings at the county level is often not available. The six states were selected on the basis of data availability and geographic representation.

[3] Five states had a lower age of majority at the end of 2018 (Adult is 16 - Georgia, Michigan, Missouri, Texas, and Wisconsin). Missouri raised the age of majority to 17 in 2018, but the law will not go into effect until 2021. (Juvenile Age of Jurisdiction and Transfer to Adult Court Laws, 1/11/2019, National Conference of State Legislatures, retrieved May 7, 2019 from http://www.ncsl.org/research/civil-and-criminal-justice/juvenile-age-of-jurisdiction-and-transfer-to-adult-court-laws.aspx)

Prior BJS collections of state court cases have focused on felony cases in general jurisdiction courts. The National Judicial Reporting Program (NJRP) collected only felony convictions and included no misdemeanor cases. State Court Processing Statistics (SCPS) collected data on felony case filings and included misdemeanors if the cases were filed as felonies but disposed of as misdemeanors. Both data collections only collected data from general jurisdiction courts.

The National Center for State Courts (NCSC), using data from the Court Statistics Project (CSP), estimated that 64% to 92% of state court criminal filings were misdemeanor cases.[4] Not all state systems receive data from limited jurisdiction courts;[5] BJS estimates that 26 of the 36 statewide systems and the District of Columbia have data from both general jurisdiction and limited jurisdiction courts, based on the data capability interviews from work done under a previous generic clearance. BJS is only collecting limited jurisdiction data if the state or county is able to provide it. BJS is also excluding municipal courts. BJS is also requesting only serious misdemeanors and will not be able to report on the total misdemeanor cases processed in state courts. However, the inclusion of limited jurisdiction courts is critical to understanding the total work of state criminal courts.

Overall Study and Sample Design

The overall CCSC study and sample design will combine a census with a sample-based data collection and estimation approach. In states with centralized, electronic court case-level records systems that cover all courts in the state, all in-scope cases will be processed. In other states, sample-based data collection methods will be appropriate.

For planning purposes, the CCSC study and sample design assumes the following primary classes of states with respect to the availability of centralized, state electronic case-level records systems:

> **Class 1.** States with a centralized records management system that covers most of or the entire state
>
> **Class 2.** States without centralized data

The data collection will be staged as some states may decline to participate. All felony and serious misdemeanor cases disposed in calendar year 2019 will be requested from the state

---

[4] R. LaFountain, R. Schauffler, S. Strickland & K. Holt. (2012) *Examining the Work of State Courts: An Analysis of 2010 State Court Caseloads.* Retrieved May 2, 2019 from http://www.courtstatistics.org/~/media/Microsites/Files/CSP/DATA%20PDF/CSP_DEC.ashx.

[5] Some state court systems are unified and only have courts of general jurisdiction, whereas others are tiered systems and may receive data from both general jurisdiction and limited jurisdiction courts.

systems in Class 1 and from all counties not part of the statewide systems in those states. A sample of counties will be used to represent the balance of states in Class 2. The sample will also include counties from states that are in Class 1 but do not provide statewide data. Once the Class 2 counties are identified, BJS will sample counties from states in Class 2 with probability proportional to size. Any county with a total population of one million residents aged 18 or older or more will be sampled with certainty. The total population aged 18 and over is the target population because these are the individuals likely to be charged in criminal courts.

Table 1 details the expected availability and coverage of centralized state data. Table 2 details the known states without centralized data. Table 3 lists the counties from known Class 2 states with populations of at least one million residents aged eighteen and over. Tables 2 and 3 may be subject to change based on the response of Class 1 states.

Table 1. Class 1 States by Population Aged 18 and Over, 2018

State Name	Population 18+	Additional Counties Needed for Census
New York	15,474,107	
Pennsylvania	10,158,149	
Illinois	9,883,814	Cook and DuPage Counties not on statewide
North Carolina	8,082,975	
Michigan	7,831,247	
New Jersey	6,954,877	
Virginia	6,647,893	Alexandria, Fairfax, and King William Counties not on statewide
Washington	5,872,306	King and Pierce Counties not on statewide
Massachusetts	5,535,291	Pima and Maricopa Counties not on statewide
Arizona	5,528,989	
Missouri	4,749,622	
Maryland	4,702,570	
Wisconsin	4,537,465	
Colorado	4,430,329	Denver County not on statewide
Minnesota	4,308,564	
South Carolina	3,978,182	
Alabama	3,798,031	
Kentucky	3,459,573	
Oregon	3,317,146	
Oklahoma	2,986,593	
Connecticut	2,837,472	
Iowa	2,425,378	
Arkansas	2,310,645	
Utah	2,228,643	
Kansas	2,205,544	
New Mexico	1,613,275	
Nebraska	1,452,427	
Idaho	1,307,236	
Hawaii	1,117,077	
Maine	1,088,000	
Rhode Island	852,102	
South Dakota	664,629	
North Dakota	581,379	
District of Columbia	574,961	
Alaska	553,622	
Vermont	510,326	
Wyoming	442,962	
Total (57% of U.S. 18+ Population)	145,003,401	

Source: U.S. Census Bureau, Population Division. Table 1. Annual Estimates of the Resident Population

for Selected Age Groups by Sex for the United States, States, Counties and Puerto Rico Commonwealth and Municipios: April 1, 2010 to July 1, 2018

Table 2. Class 2 States by Population Aged 18 and Over, 2018

State Name	Population 18+
California	30,567,090
Texas	21,303,746
Florida	17,070,244
Ohio	9,096,117
Georgia	8,013,724
Tennessee	5,263,790
Indiana	5,123,748
Louisiana	3,564,062
Nevada	2,345,395
Mississippi	2,280,389
West Virginia	1,441,672
New Hampshire	1,098,288
Montana	832,871
Delaware	763,555
Total (43% of U.S. 18+ Population)	108,764,691

Source: U.S. Census Bureau, Population Division. Table 1. Annual Estimates of the Resident Population for Selected Age Groups by Sex for the United States, States, Counties and Puerto Rico Commonwealth and Municipios: April 1, 2010 to July 1, 2018

Table 3. Certainty Counties from Class 2 States, by Size of Population Aged 18 and Over, 2018

Certainty County and State	Population 18+
Los Angeles County, CA	7,916,625
Harris County, TX	3,446,935
San Diego County, CA	2,620,956
Orange County, CA	2,487,180
Miami-Dade County, FL	2,203,331
Dallas County, TX	1,948,080
Riverside County, CA	1,834,632
Clark County, NV	1,714,018
San Bernardino County, CA	1,599,325
Broward County, FL	1,538,471
Tarrant County, TX	1,535,868
Santa Clara County, CA	1,513,143
Bexar County, TX	1,478,380
Alameda County, CA	1,324,243
Palm Beach County, FL	1,202,485
Sacramento County, CA	1,177,066
Hillsborough County, FL	1,112,902
Orange County, FL	1,074,728
Franklin County, OH	1,005,657
Total (15.3% of the U.S. Population)	38,734,025

Source: U.S. Census Bureau, Population Division. Table 1. Annual Estimates of the Resident Population for Selected Age Groups by Sex for the United States, States, Counties and Puerto Rico Commonwealth and Municipios: April 1, 2010 to July 1, 2018

BJS is unsure of the availability of data in Puerto Rico. BJS intends to assess the structure and the data reporting capacity of the Puerto Rico court system with the intent of assessing the availability and quality of data. At this time, BJS does not intend to include Puerto Rico in full data collection and reporting for the CCSC.

Cases will be divided into two frames: 1) those that are retrieved from state electronic record systems (Class 1 states, outlined in Table 1) and all counties that will be designated as self-representing (i.e., certainty jurisdictions from Class 2 states, outlined in Table 3), and 2) all other counties from which a sample will be drawn. The self-representing units refer to jurisdictions that provide “census” data within themselves (i.e., Class 1 states and certainty jurisdictions from Class 2 states), and are therefore not subject to sampling error.

National estimates under this design will take the following form:

$$t=\sum_{h=1}^{2} W_h t_h = W_1 t_1 + W_2 t_2$$

where W_h denotes the population weights associated with the self-representing states and counties (h=1) and non-self-representing counties (h=2) with $W_2=(1-W_1)$, and t_h denotes the statistic of interest calculated from frame h.

The corresponding precision of estimates of the combined frames will be sensitive to the percentage of the population covered by each frame. An expression for the variance of a proportion assuming two frames would be as follows:

$$V(t)=\sum_{h=1}^{2} W_h^2 Var(t¿¿h)¿$$

Where:

t	=	statistic of interest
h	=	frame (h=1 denotes self-representing units, i.e., centralized states and certainty jurisdictions; and, h=2 denotes the non-self-representing frame from which a sample of counties will be drawn)
W_h^2	=	the square of the population weight for frame h,

Note that for h=1, there is no sampling variance associated with the statistic t_1 because the data in this frame represent a census of all cases. Thus,

$$Var(t_1)=0$$

and the variance of the national survey statistic t can be simplified to

$$V(t)=\sum_{h=1}^{2} W_h^2 Var(t¿¿h) = W_2^2 Var(t_2)¿$$

For the non-self-representing frame (h=2), presumably, all court data relevant to the CCSC study will be collected for sampled counties. As such, there is no within-county (i.e., within-cluster) variation because all cases are being taken within each sampled county in frame 2. The only source of variation is between sampled counties based on county level statistics aggregated from the court cases within each non-self-representing county. Consequently, the number of sampled non-self-representing counties in frame 2 will primarily determine the statistical precision of the survey estimates.

Indexing the non-self-representing sampled counties by *j* for h=2, the statistic t_2 as a mean (i.e., average, proportion, percentage) takes the following form:

$$t_2 = \sum_{j}^{n_2} \omega_j t_{2_j} / \sum_{j}^{n_2} \omega_j$$

where ω_j denotes an analytic weight for county *j* in the non-self-representing frame 2, and t_{2_j} is the aggregated statistic t from county *j* in frame 2.

The variance of statistic $t_2$simplifies considerably when *t* is an estimated proportion (or a percentage). In this case the variance of $p_2$can be written as:

$$Var(p_2) = ¿ F_w \times p_2(1-p_2)/n_2$$

where p_2 is the estimated proportion from the frame 2 sample (calculated using analytic weights that reflect the unequal probability of each sampled county from probability proportional to size (PPS) selection), and $DEFF_w$ represents the design effect due to weighting (i.e., the weighting effect). The term n_2 denotes the number of non-self-representing counties sampled in frame 2.

The DEFF is due to unequal probability sampling. Differential weighting can therefore vary by sample and can be estimated using the following formula:

$$DEFF_w = \frac{n_2 \sum_{j} \omega_j^2}{¿¿}$$

where the *j* are summed over the n_2 sampled non-self-representing counties in frame 2. Based on simulated samples of n_2=1,155 (somewhat greater than the 1,081 frame 2 counties but a conservative starting value), we estimate $DEFF_w$<1.50 and use 1.50 as a conservative value.

In addition, the maximum variance of an estimated proportion occurs when *p* = 0.50. Thus, the estimated maximum variance of an estimated statistic under this design will take the following form:

$$Var(p) = W_1^2 Var(p_1) + W_2^2 Var(p_2) = W_2^2 Var(p_2)$$

because the variance of p_1 is zero by virtue of it being self-representing.

Using $DEFF_w = 1.50$ and $p_2 = 0.50$, and assuming a sample of $n_2 = 84$ from non-self-representing frame 2, and that the population weight $W_2 = 0.28$ (= 1 – the proportion covered by frame 1 data collection), we expect that the maximum variance of an estimated proportion *p* will be

$$Var(p) = (1.5) \times (0.28)^2 \times \frac{(.5)(.5)}{84} = 0.00035$$

and the maximum margin of error (MOE) at a 95 percent level of confidence will be

$$maximum\ MOE_{.95}(p) = (1.96) \times \sqrt{0.00035} = 0.0366$$

or 3.7 percentage points. This stratified probability sample features a large number of self-represented units. This is a highly efficient approach as the self-represented units contribute zero to the overall standard error.

For the purposes of the overall design, BJS has assumed the following:

Stage 1 – Collection from centralized state data and certainty counties

During work under a generic clearance, BJS contacted state court technology leaders to determine if the state had a statewide electronic case-level records system which can provide all or most of the relevant CCSC data items. The state court leaders indicated that state systems vary in coverage in terms of both geographic location (i.e., counties) and types of courts that contribute data (i.e., general jurisdiction, limited jurisdiction, or both). The state data systems are largely used for case management, and include data elements related to general case information, defendants and attorneys, charges, disposition, and sentencing. Some of these data, primarily related to sentencing, are in text fields or are maintained by separate systems (e.g., probation). Data elements related to pretrial release, diversion, and warrants are also typically maintained separately.

Any state with centralized data (i.e., Class 1 states) will be asked for an electronic file for all criminal cases filed as felonies and serious misdemeanors that were disposed of in 2019. "Disposed of" is defined as a final finding by a judicial officer (typically a judge), and includes dismissal, *nolle prosequi*, placement on an inactive docket (stay of prosecution), placement in a diversion program, guilty, not guilty, acquittal, or other finding. States may provide all such cases in any format, including a formatted data extract, where the data are extracted to meet BJS data coding; unformatted data extract, where the data are extracted from the system as-is and BJS will work with the state to clean and standardize the data; or a system extract ("data dump") of the entire case records system. The state could also allow BJS access to the case management system to help extract the data, or to scrape the data from a website. Based on data submitted in the Survey of Juveniles Charged in Adult Criminal Courts (SJCACC), BJS expects most of the

states to provide unformatted extracts. The data obtained from the state approach is a census of all cases within and across courts (general jurisdiction and limited jurisdiction), excluding municipal courts.

BJS will also collect data in a similar manner from the counties in Class 2 with a total county population of one million or more residents aged eighteen or over.

There may be states in Class 1 that decline to provide data; BJS will ask the state court administrator if BJS can contact counties and move the counties in the state to Class 2. If the state administrator prevents BJS from contacting the counties directly, BJS will move the counties within that state to the sample strata, calculate the number of counties to be drawn from the strata, and then remove the state from the sample. BJS expects that this will be unlikely to occur.

Stage 2 - Sampling of Non-Certainty Counties

States and counties that are not covered in the centralized data and counties not sampled with certainty will be stratified by county population. This will include parts of Class 1 and all of Class 2. Although BJS does not know which states will decline to participate, this section assumes an 80% to 100% response rate in Stage 1, which will provide an estimated 57% to 72% coverage of the U.S. population ages 18 and over (Table 4).

Table 4. Target Population by Stage

Population age 18+	**N**	**Percent**
U.S. Population July 1, 2018, ages 18+	253,768,092	100%
Population ages 18+ covered in Class 1 (where state coverage is likely to be successful)	145,003,401	57%
Population ages 18+ covered in certainty stratum (counties in Class 2 with a total county population of 1,000,000 or more)	38,734,025	15%
Complete Stage 1 coverage (state data and certainty counties)	183,737,426	72%
Successful 80% response in Stage 1		80%
Stage 1 coverage assuming 80% provide data	146,989,941	58%
Balance of population aged 18+ (to be covered by PPS sample)	106,778,151	42%
Estimates of resident population age 18 and over taken from U.S. Census Bureau, Population Division. Table 1. Annual Estimates of the Resident Population for Selected Age Groups by Sex for the United States, States, Counties and Puerto Rico Commonwealth and Municipios: April 1, 2010 to July 1, 2018		

All counties in Class 2 states will be stratified by population size. As discussed, counties with populations exceeding one million will be sampled with certainty, and BJS will collect data from these counties in Stage 1. The remaining counties will be divided into five strata (Table 5):

Table 5. Distribution of Counties and Population in Class 2 by Stratum

Stratum No.	Population	Number of Counties in U.S.	Expected Number of Counties in Class 2	Expected Percent of Class 2 Population
1	500,000-999,999	102	26*	25.6%
2	100,000-499,999	395	151	43.6%
3	20,000-99,999	1,087	398	23.8%
4	5,000-19,999	1,135	389	6.6%
5	<5,000	423	117	0.4%
	Total	3,142	1,081	100.0%

*Does not include the 19 certainty counties (cannot be replaced).
Estimates of resident population age 18 and over taken from U.S. Census Bureau, Population Division. Table 1. Annual Estimates of the Resident Population for Selected Age Groups by Sex for the United States, States, Counties and Puerto Rico Commonwealth and Municipios: April 1, 2010 to July 1, 2018

Alternative Scenarios

The plan described above assumes a high level of success with the response of Stage 1. Because the likely success of the state approach is uncertain, we have considered a few alternative scenarios and have contingency plans built into the overall design. For purposes of illustration, two alternative scenarios are considered. Scenario 2 assumes that the statewide success rate is 25 percent. Table 6 below, which is similar to Table 4 presented above, show the effects of this assumption.

Table 6. Target Population by Census/Sample Approach – Scenario 2

Population	N	%
Total Population aged 18+	253,768,092	100%
Population aged 18+ in Class 1 (where state approach is attempted)	145,003,401	57%
Population aged 18+ in county certainty stratum	38,734,025	15%
Stage 1 coverage fully successful (state and certainty counties)	183,737,426	72%
Stage 1 coverage with low success rate assumed	--	25%
Population covered by 25% successful statewide approach	45,934,357	18%
Balance of population (to be covered by PPS sample)	207,833,735	82%

Scenario 3 assumes that the state approach success rate is 50 percent. Table 7 below shows the effects of this assumption.

Table 7. Target Population by Census/Sample Approach – Scenario 3

Population	N	%
Total Population aged 18+	253,768,092	100%
Population aged 18+ in Class 1 (where state approach is attempted)	145,003,401	57%
Population aged 18+ in county certainty stratum	38,734,025	15%
Stage 1 coverage fully successful (state and certainty counties)	183,737,426	72%
Stage 1 coverage with low success rate assumed	--	50%
Population covered by 50% successful statewide approach	91,868,713	36%
Balance of population (to be covered by PPS sample)	161,899,379	64%

Nonresponse and Substitution

Since the data collection is staged, and nonresponse in Stage 1 moves the counties to Class 2 for sampling, virtually all nonresponse will be in Stage 2, due to the failure of a selected county to provide data. BJS also expects item nonresponse in the context of missing or incomplete data extracts.

Nonresponse adjustment can be used to adjust the weights of responding sampled counties to the weights of all sampled counties. Although unbiased estimates might result, a shortfall in cases and consequently precision might also result. Alternatively, the actual sample could be inflated by the inverse of the expected response rate. This works well when the actual response rate can be estimated accurately in advance. A third possibility is to use release groups of counties (e.g., a reserve sample), with the exact number of counties released to be determined as the sample is implemented and actual response rates manifest themselves. BJS will likely follow this third strategy, with a reserve sample roughly half the size of the expected sample of counties.

If all 36 states, the District of Columbia, and all 18 certainty counties provide data in Stage 1, the total coverage of the U.S. population age 18 and over is 72%. In order to increase coverage of the total population, BJS estimates a sample size for the Class 2 states at approximately 84 counties. **Table 8** below describes the expected draw rate and estimates a reserve sample of roughly 30 additional counties.

Table 8. Expected Sample and Reserve Sample Distribution

Stratum No.	Population	Number of Counties	Expected Number of Counties in Class 2	Expected Percent of Class 2 Population	Expected Number of Sampled Counties	Expected Number of Reserve Counties
1	500,000- 999,999	102	26*	25.6%	22*	8
2	100,000-499,999	395	151	43.6%	37	13
3	20,000-99,999	1,087	398	23.8%	20	7
4	5,000-19,999	1,135	389	6.6%	5	2
5	<5,000	423	117	0.4%	0	0
	Total	3,142	1,081	100.0%	84	30

*Does not include the 19 certainty counties (cannot be replaced).

Estimates of resident population age 18 and over taken from U.S. Census Bureau, Population Division. Table 1. Annual Estimates of the Resident Population for Selected Age Groups by Sex for the United States, States, Counties and Puerto Rico Commonwealth and Municipios: April 1, 2010 to July 1, 2018

Changes to this expected sample size, reserve sample size, and draw rate will be submitted as a material change in circumstance for OMB review once the Class 1 response is completed.

Weighting, Estimation and Variance Estimation

The data collected for estimation purposes will be obtained from a combination of a census and a probability sample. Data weighting, acknowledging and reflecting the probability sampling, will be required in order to enable unbiased estimation. The data weighting proposed will consist of calculating base weights, nonresponse adjustment factors, and post-stratification factors. Base weights will be calculated as the inverse of the overall probability of selection, for all sampled cases, reflecting any variability in the probabilities of selection across all stages. Nonresponse adjustment factors will be calculated reflecting county nonresponse, as the ratio of the sum of base weights for all counties (i.e., at a given stage) to the sum of base weights for all responding counties.

The final resulting weight will be used for estimation. Variance estimation will be supported by providing strata and PPS variables for SUDAAN, as well as a set of replicate weights. The jackknife or balanced repeated replication (JK2 or BRR) approach would be used since variance estimation for medians might be used.

2. Procedures for Collecting Information

Respondents who are able to provide electronic data will be able to submit the files in any format. BJS will assist states or counties unable to provide data as an extract by offering to write a computer program to extract the data (e.g., enterprise custom report), or by offering to scrape data from a public website. A data extraction guide will be provided to all respondents so they

may determine what information will be necessary for those providing electronic data (see **Attachment 2**).

During Stage 1, BJS will mail or email the Office of the State Court Administrator for each state in class 1 a letter introducing the importance and purpose of the collection, the data collection agents (Urban Institute and National Center for State Courts (NCSC)), and invite the court to participate in the collection (**Attachment 5)**. The same letter will be sent to the County Court Administrator in the counties not covered by state data systems identified as having 1,000,000 or more residents.[6]

Once permission to collect data is obtained from the relevant contacts, Urban Institute and NCSC will work with staff who manage the courts' information system to obtain data files **(Attachment 6)**. Urban Institute will process the files, working with the respondent to evaluate data quality and completeness.

As Stage 1 progresses, state systems may elect not to participate. If this occurs, NCSC and Urban Institute will request permission from the state contact to include the counties in the sampling frame. BJS will more fully describe the sampling strategy for Stage 2 once Stage 1 is almost complete (i.e., within 3-6 months of the start of Stage 1) with an updated material change in circumstances memorandum.

The same contact process will be repeated in Stage 2 for county court administrators for the sampled counties.

3. Methods to Maximize Response Rates

Every attempt will be made to collect complete information on criminal cases processed in state and county courts in 2019, excluding the case types noted on page 1. In order to maximize the response rate and minimize nonresponse bias, the data collection agents will work closely with respondents to ensure that they understand the data collection process, including determining the most appropriate format for data submission and obtaining the necessary approvals for providing case level data. It is assumed that BJS will enter into Memoranda of Understanding (MOU) or data use agreements with some or all of the state and county courts (**Attachment 8**). BJS attorneys will review all agreements before accepting.

A team of Urban Institute and NCSC staff members will be assigned to act as the point of contact for each respondent. A toll-free telephone line and email address will be developed to address respondents' technical questions.

4. Testing of Procedures

[6] In some states and counties, NCSC has direct contacts with individuals responsible for court data requests. In those instances, BJS will send the letter to the court administrator and the technology contact, noting that the letter was sent to both parties.

The project team conducted a pilot test with 9 jurisdictions of varying sizes across all regions to ensure that the proposed approach in Stage 2 is feasible. NCSC provided guidance for the selection of a state system that was expected to be more complex. Minnesota was known to have multiple data systems from which the data would need to be drawn. Urban Institute and NCSC then selected four counties that will be included in the certainty stratum, one county expected to be in stratum 1 (500,000-999,999), two counties in stratum 2 (100,000-499,000), and one county in stratum 3 (20,000-99,999). Minnesota was included to explore potential challenges in collection data from states with a complex court structure. No data requests were made to Minnesota.

Table 8. Pilot test sites

Pilot Test Site No	County	State	Region	Population
1	Orange County	California	West	2,487,180
2	Bexar County	Texas	South	1,478,380
3	Palm Beach County	Florida	South	1,202,485
4	New Castle County	Delaware	Northeast	439,108
5	Franklin County	Ohio	Midwest	1,005,657
6	Washoe County	Nevada	West	364,959
7	Penobscot County	Maine	Northeast	123,702
8	Kosciusko County	Indiana	Midwest	60,427
9	Minnesota State	Minnesota	Midwest	4,308,564

Drawing on background research on state court structures and professional contacts, Urban and NCSC identified individuals to reach out to for each of the 9 pilot sites. Urban sent an introduction email to each of these contacts, which outlined the project and requested an opportunity to speak on the phone about their data and the process for making a data request. To minimize burden on those jurisdictions and avoid potentially damaging outcomes for our full data collection effort, the project team did not follow a rigid schedule to complete each data request. How each jurisdiction responded to our request for information also varied considerably. As summarized in Table 9 below, our approach to information gathering yielded varying outcomes.

Table 9: Summary of Pilot Results

Jurisdiction	No Response	Denied	Held Call	Reviewed Data Request	Closed Reason
Orange County, CA			X	X	Denied due to resource constraints
Bexar County, TX	X				No response
Palm Beach County, FL			X	X	Approved
New Castle County, DE*			State Follow-up call	X, referred to State	Referred to state, state drop off (statute, resources)
Franklin County, OH			X	X	Referred to state Ohio Courts Network (OCN) due to lack of resources
Washoe County, NV			X	X	Approved, MOU
Penobscot County, ME*			State Follow-up call	X	Referred to state, state drop off (resource constraint)
Kosciusko County, IN*				X	Referred to state, negotiate MOU
Minnesota				X	Denied, prefer to wait for full collection

* Requests should go through state court office, not the county

Eight of the pilot sites responded to either schedule a call or refer us to another agency. Urban conducted phone calls with the pilot sites that had not been contacted previously, as well as held follow up calls with the state courts previously contacted, to review the availability of their data. In total, three sites referred us to a state court office, three denied our request, and 2 approved our request. Urban and NCSC continued to follow up with states through the end of the pilot period (November 22, 2019) to encourage state participation. One of them participated in a phone call, reviewed our request, and referred us to a state agency. For sites that indicated they could make pilot data available, Urban initiated a data request.

We learned a few important lessons from the pilot test. First, nearly all of the pilot jurisdictions were able to provide the project team with information about the availability of case-level court data within days of the initial outreach. The proposed approach to gathering information would not pose too much of a challenge to most state and local courts. The project team was better positioned to make a successful application for data request by first discussing with the court the availability of case-level data that are easily extracted from the data management systems. Additionally, allowing the court to explain its data management systems was also beneficial in building confidence. For example, one court noted that the quality of the data depended a lot on

the clerk entering the data. This conversational approach is recommended for full data collection when we conduct outreach to courts not previously contacted in either the state data interviews or the pilot test.

Second, the project staff learned that there were a number of state-level data sources for many of the pilot counties (i.e., Maine, Delaware, Indiana, and Ohio). These alternative sources of court data included state administrative office of the courts, state-level data warehouses, information exchange systems, or law enforcement agencies. All of the jurisdictions that referred us to one of those alternative data sources seemed capable of fulfilling our data request as they were also responsible for submitting court records from their county to the state agency or repository. For the full data collection, it would be useful to better understand the circumstances under which the county-level agency may refer us to the state-level agency. However, data elements available from alternative sources are often held in separate management systems, making it difficult to link their records to court data.

Third, data sharing for some sites requires court administration approval, which is not always easy to obtain due to competing priorities and IT resources. Restrictions to data collection include site-specific MOUs, fees, and statutes. BJS is prepared to reimburse courts for any cost of preparing the data extracts. The full data collection will benefit from clear specification on the data elements, and greater emphasis on how the project team can support initial data processing and verification (i.e., documentation, interpretation, and extract modification). The data extraction guide was revised to meet those goals.

5. Contact for Statistical Aspects and Data Collection

The prosecution and courts statistics unit staff at BJS are responsible for the overall design and management of the CCSC data collection, including the development of the data extraction guide and the analysis and publication of the data.

Suzanne Strong, Statistician
Prosecution and Judicial Statistics Unit
Bureau of Justice Statistics
810 7th Street, NW
Washington, D.C. 20531
(202) 616-3666

Attachments

1. 34 USC § 10132
2. Data extraction guide
3. 60 day notice
4. 30 day notice
5. BJS introduction letter
6. Request for data
7. Initial follow-up script
8. Sample generic MOU
9. Second follow-up
10. BJS final follow-up
11. Confirm data script
12. Thank you email
13. Civil Rights Division comments
14. NCAJ comments
15. LAFLA comments
16. BJS response to public comment CCSC