

The author(s) shown below used Federal funds provided by the U.S. Department of Justice and prepared the following final report:

Document Title: Pilot Study Final Report on Statewide Judicial Data for NJRP

Author: Rob Santos, Molly M. Scott

Document No.: 233456

Date Received: February 2011

Award Number: 2008-BJ-CX-K052

This report has not been published by the U.S. Department of Justice. To provide better customer service, NCJRS has made this Federally-funded grant final report available electronically in addition to traditional paper copies.

Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Pilot study final report on statewide judicial data for NJRP

Submitted by:

The Urban Institute

2100 M Street, NW

Washington, DC 20037

(202) 833-7200

Submitted to:

Bureau of Justice Statistics

810 Seventh Street, NW

Washington, DC 20531

November 30, 2010

Final report

This project was supported by Grant No. 2008-BJ-CX-K052 awarded by the Bureau of Justice Statistics. The opinions, findings, and conclusions or recommendations

expressed in this document are those of the authors and do not necessarily reflect the views of the U.S. Department of Justice.

Authorships and Acknowledgments:

This report was written by Rob Santos, Senior Institute Methodologist, and Molly M. Scott, Research Associate.

MEMORANDUM

TO: Thomas Cohen, BJS

FROM: Robert L. Santos and Molly M. Scott, Urban Institute

RE: Pilot Study report on statewide judicial data for NJRP

DATE: November 30, 2010

1. Introduction

The National Judicial Report Program (NJRP) has provided case-level data on felony convictions and sentencing in state courts for over 25 years. Conducted biennially, the NJRP dataset includes information on offenses, convictions, penalties, and demographic characteristics such as age, race and gender. The data can be analyzed both at national and county levels and are used by the Bureau of Justice Statistics (BJS) to report national estimates of conviction and sentencing as well as the characteristics of convicts. Furthermore, federal, state, and local government agencies, as well as researchers and policy analysts, use these valuable data, making NJRP an important national resource.

BJS has initiated a review of the NJRP design in light of societal, demographic, and technological changes since its inception. A prominent design feature under review is a sample design that relies on a two-stage sample of counties (stage 1) and cases within counties (stage 2). Recent iterations of NJRP have found that sentencing and conviction data are increasingly stored on centralized systems at the *state* level, covering all or most of a state. Sizeable gains in both cost and statistical efficiency could be realized by gathering sentencing and conviction data for entire states rather than from samples of counties.

As part of NJRP's redesign efforts, BJS has asked the Urban Institute to assess the feasibility of *statewide* data collection. This memo synthesizes our finding from this effort. Drawing on information gathered from staff in 69 agencies in 29 states that contain over 87% of the U.S. population, we identify candidate states in the best position to provide NJRP data for the entire state. We also include a detailed discussion of trade-offs among data sources and their implications for NJRP sampling.

1.1 *Research focus*

This memo addresses the extent to which state agencies are equipped to generate current *NJRP data*. We define *NJRP data* as case-level information on convictions and sentencing. Specifically, we focus on:

- *expanse* (how many states have these data),
- *location* (which agencies house them),

- coverage (all versus some counties; all versus some types of felony convictions),
- quality (missing data, standardized coding),
- data access policies,
- time required for request approval and processing and
- depth (which items are collected).

Relatively little is known about existing statewide judicial record systems. Consequently, the information gathered from the sample of 29 states represents a substantial contribution to understanding the capacity of state-wide data systems.

2. Methodology

This section provides a detailed account of the process used for selecting the states for the feasibility study as well as for developing the phone interview protocol and contacting agencies within the selected states. We then offer a brief overview of the interviews conducted and the analytical framework applied in order to characterize the quality and accessibility of available NJRP data within individual agencies.

2.1 State selection

Because the feasibility study itself was composed of two phases of research activity, site selection necessarily involved a two-step process. The following sub-sections lay out the criteria used in each distinct phase of the project.

2.1.1 Pilot

We purposively selected nine states for the pilot phase of the feasibility study using several key criteria. Because BJS hypothesized that states with sentencing guidelines in place would be the most likely to house comprehensive statewide NJRP data, we prioritized the selection of such states. We also wanted a good representation of the regions of the country as well as of the national range of technological sophistication. Lastly, we required that selected states contained at least one county sampled in the 2004 State Courts Processing Statistics Program (SCPS).

To begin the state selection process, we first reviewed the report entitled *State Sentencing Guidelines* issued by the National Center for State Courts to identify a subset of states whose state sentencing commissions “regularly report on guideline compliance.”¹ We then performed internet searches on these states to find those that published annual reports using statewide data. From this sub-group, we chose 4 states: Washington, Virginia, Minnesota, and Pennsylvania.

We then examined notes from the implementation of the SCPS survey to identify states without sentencing guidelines that nonetheless maintained some kind of centralized state data.² We expected these states to represent the midpoint of a continuum between those

¹ Kauder NB, Ostrom BJ. *State Sentencing Guidelines: Profiles and Continuum*. July 2008, National Center for State Courts: Washington, D.C.

² For most of these states, only criminal history data was held and maintained by state agencies.

with highly centralized/standardized court data systems at one extreme and those with only decentralized (i.e., county level) data systems at the other. The states we selected were: Arizona, New York, and Texas.

Lastly, we identified 2 states without sentencing guidelines that *did not* have any sampled counties in the SCPS data collection. These were thought to have limited statewide judicial statistics systems. These states were Georgia, and Idaho.

2.1.2 Second phase

For the second phase of data collection, we chose states previously not selected for the pilot phase of the project. We used two factors to rank states: the presence of sentencing guidelines and total state population. Our pilot phase suggested that sentencing guidelines states were most likely to have high quality statewide data systems containing NJRP-type items. But we were also concerned with including/covering the largest sources of sentencing and conviction data in the U.S. – namely, the most populous states. The larger states afford BJS the biggest return on investment (ROI) with regard to NJRP data.

As a result, our second phase sample started with the inclusion of the 11 sentencing guideline states (not previously selected into the pilot) containing more than 1 percent of the total US population. We then added the state of Kansas because of promising practices by its Sentencing Commission despite its relatively small population share. In addition to these 12 states we added the 8 largest states with respect to total population³.

In sum, research activities in the second phase covered 20 states. Combining these with the 9 states used in the initial pilot phase, this study reported on a total of 29 states. These 29 states covered 87% of the total population in the U.S. in 2008.

2.2 Information Collection

2.2.1 Protocol development

We systematically captured information using a semi-structured phone interview protocol. The original pilot study instrument was designed to capture a wide array of information on a broader set of research questions proposed by the Bureau of Justice Statistics, including the availability of not just NJRP items but of criminal history-related variables, misdemeanor cases, and pre-sentencing data. The second stage of data collection was designed to focus exclusively on HJRP data items. In consequence, we revised the interview protocol to both efficiently capture relevant information collection and to maximize the number of states contacted. The revised instrument more directly solicits the information needed to categorize the coverage, completeness, and standardization of each state agencies' NJRP data according to the typologies developed during the pilot stage of the project (see Section 2.3 *Analyses* for more detail). The new protocol streamlined data gathering substantially and allowed more timely analyses of our findings. The revised protocol appears as Appendix A.

³ We used data from the 2008 American Community Survey to obtain estimates for state population.

2.2.2 Processes and respondents

In almost all cases, identifying and obtaining contact information from relevant state agencies required internet searches followed by a series of trial-and-error “cold calls” to identify most knowledgeable staff. Cooperating staff were asked for referrals to other agencies within their state at the end of each interview.

In many cases, we needed to speak directly to *several staff within an agency* to provide the information requested. For instance, the person with best global knowledge of the data and its day-to-day use (e.g., Research Analyst or Criminal Records Supervisor) was not typically the best person to ask about more technical issues such data access protocols or the availability of a codebook or comprehensive list of data fields. In this report we do not count these multiple contacts as separate “interviews,” but rather as components of a single interview meant to capture information about a unique potential data source.

2.2.3 Overview of interviews and agencies

In the pilot stage of this feasibility study, the agencies we contacted (i.e., state court administration, criminal records depositories, and sentencing commissions) were independent of each other. However, during the second stage of data collection, we found that these roles were sometimes vested within the same agencies. For example, in Alabama, Massachusetts, and North Carolina, sentencing commissions were located within the each state’s administrative office of the courts. Similarly, we discovered that in Kentucky, the administrative office of the courts also served as the criminal records depository. In all these cases, we completed interviews with staff in each separate department.

In the second round of data collection, we interviewed staff at the department of corrections in a total of 4 states. In three of these states (Missouri, Ohio, and Louisiana), we engaged this agency because it provided data and analyses to the state’s sentencing commissions and held exclusive jurisdiction over the available data. In the remaining state, Tennessee, we spoke with department of corrections staff because our contacts at the Tennessee Administrative Office of the Courts and the Tennessee Bureau of Investigation reported that corrections was the only agency in the state likely to have NJRP data.

During the data collection process, we also ascertained that two states chosen for their use of sentencing guidelines in the first stratum of the second round of data collection *did not have a sentencing commission in place at the time of the interview.*⁴ The Wisconsin Sentencing Commission was abolished in 2007 and the Michigan Sentencing Commission was disbanded following the enactment of the statutory guidelines in 1998.

In total, we conducted 73 interviews by phone for the NJRP feasibility study. During the pilot, we performed interviews with 21 different agencies in 9 states between April 5th and May 28th. During the second phase of data collection between August 31st and

⁴ These states do indeed have sentencing guidelines at the state level, but have no active commission to monitor compliance with these guidelines.

October 22nd, we added 52 interviews in a total of 48 agencies and 20 states. There were only two agencies during the study—the Pennsylvania State Police in the pilot period and the Massachusetts Criminal History Board in the second round of data collection-- that we were unable to interview. Table 1 presents the total number of interviews conducted by type of agency, classifying states by the presence of a sentencing commission. We interviewed between 2 and 4 agencies in each state.

State	Type of Agency				Total
	Stand-alone Sentencing Commission	Court Administration	Criminal Records Depository	Department of Corrections	
AL*		2	1		3
KS	1	1	1		3
LA	1	1	1	1	4
MA*		2			2
MD	1	1	1		3
MN	1	1	1		3
MO	1	1	1	1	4
NC*		2	1		3
OH	1	1	1	1	4
OR	1	1	1		3
PA	1	1			2
VA	1	1	1		3
WA	1	1	1		3
AZ		1	1		2
CA		1	1		2
CO		1	1		2
FL		1	1		2
GA		1	1		2
ID		1	1		2
IL		1	1		2
IN		1	1		2
KY**		2			2
MI		1	1		2
NJ		1	1		2
NY		1	1		2
SC		1	1		2
TN		1	1	1	3
TX		1	1		2
WI		1	1		2
TOTAL	10	33	26	4	73
Note: States with Sentencing Commissions are yellow; others are green					
* Administrative Office of the Courts contains a department in which the state's sentencing commission is located.					
** Administrative Office of the Courts is also responsible for the state's criminal records system					

2.3 Analyses

Before delving into our analyses, it is important to provide some preliminary remarks that explain the exposition of findings below. First, during the interview process we found that sentencing *guideline* states did not uniformly have operating sentencing *commissions*. This prompted us to re-think how to classify states for analyses. Because it is vital for an institution to actually exist before one can solicit information from it, we distinguish between states with and without an *active sentencing commission*. Consequently, even though Wisconsin and Michigan were originally selected because of their use of sentencing guidelines, we grouped them with the other (non-commission) states for analyses.

We also chose to examine the data at the agency level when state sentencing commissions or criminal records depositories were located within the administrative offices of the courts. This approach is the most appropriate and parsimonious given that the data sources themselves were identical.

We now present a detailed explanation of the analytical framework used to rate the quality of available NJRP data at individual agencies as well as a discussion of how we classified ease of data access.

2.3.1 Data quality analytic framework

In this second stage of the feasibility study, we retained the same analytical framework developed for the pilot to report the quality of potential NJRP data sources. The three dimensions for which we assess quality are coverage, completeness, and standardization, and the typologies are defined as follows:

1. *Coverage: Reporting compliance among jurisdictions*

High: All jurisdictions report

Medium: A few large volume jurisdictions are excluded or particular types of sub-agencies do not report

Low: At least half of jurisdictions do not report

2. *Completeness of NJRP data items*

High: Data on both conviction and sentencing are complete

Medium: Some missing data mostly because of matching issues or types of felony offenses that are systematically not included

Low: A lot of missing data due to lack of agreement on required fields

3. *Standardization of data items*

High: All data are cleaned and standardized and put into analytic datasets

Medium: Administrative data exist with standardized fields and coding systems, but coding practice standardization is not verified

Low: No standardized fields or codes (e.g., each jurisdiction employs their own schemes to code convictions and sentencing records)

We classified all agencies' data along these three dimensions. We also combined the three dimension ratings to develop an *overall quality index* in order to identify the best candidates for NJRP data collection. The overall quality groupings we identified were:

- **high quality:** this group of agencies appears best positioned to provide statewide judiciary records with some/all NJRP items; agencies in this group all display high quality ratings on all three dimensions;
- **medium quality:** These agencies appear to have the data but there may be some “nonfatal” issues of geographic coverage (e.g., most but not all counties included), completeness (e.g., some notable issues with missing data), and/or standardization (e.g., requiring coding/processing); agencies in this classification display a mixture of “high” and “medium” ratings across the three dimensions;
- **low quality:** agencies in this group have nontrivial limitations/gaps, and accordingly earned a “low” quality rating on one or more dimensions;
- **no NJRP data:** these agencies either fail to provide key NJRP items like sentencing or do not have centralized state-wide data systems in place that capture the micro data necessary to generate NJRP data.

For the purposes of this report, we focus on “medium” and “high quality” agencies which we consider to meet the minimum standards for consideration as candidates for state-wide NJRP data collection.⁵

2.3.2 Data access classifications

Another key element that BJS must take into consideration is the time necessary to acquire clean and processed NJRP data from individual agencies. In order to simplify respondents’ characterizations of their data access processes, we chose to consider estimated *approval* time and *processing* time separately. We also decided to classify each of these parts of the process as “fast” or “slow” so that BJS could more easily identify which agencies might be most responsive. In terms of *approval* time, we considered any estimate of less than a week to be “fast.” For *processing* time, we flagged as “fast” all agencies that expected to produce NJRP data in a month or less.

3. Findings

We now consider each research question in turn and use our qualitative assessments to give insight into statewide NJRP data collection.

3.1 Availability of NJRP data at the state level

Our analysis finds that “high quality” NJRP data are available in a third of the states that we examined (10 of 29). Eight of these ten states have sentencing commissions in place, but agencies in two other states without sentencing commissions not only house highly standardized data that covers the whole state but regularly produce clean analytic datasets for research purposes. *Collectively, these 10 “high quality” states account for about a quarter of the U.S. population.* Table 2 exhibits this information in more detail.

⁵ We acknowledge that this three dimensional qualitative rating (high, medium, low) is a subjective, quasi-qualitative assessment, but consider it to be the most feasible and informative option for this exploratory study. Time and resource limitations prevented a more concise characterization of data quality using test data files to crosswalk to individual NJRP items and verify rates of missing data, etc.

Table 2. State availability of NJPR data by quality and coverage of the US population			
State	High Quality	High or Medium Quality	Cumulative % of US Population
AL	1	1	1.5
KS	1	1	2.5
LA	1	1	3.9
MA	1	1	6.0
MN	1	1	7.8
NC	1	1	10.8
VA	1	1	13.3
WA	1	1	15.5
NY	1	1	21.9
WI	1	1	23.8
MD	0	1	25.6
MO	0	1	27.6
OR	0	1	28.8
PA	0	1	33.0
CA	0	1	45.0
CO	0	1	46.6
ID	0	1	47.1
IL	0	1	51.4
FL	0	1	57.4
GA	0	1	60.6
KY	0	1	62.0
NJ	0	1	64.9
TX	0	1	72.8
OH	0	0	76.6
AZ	0	0	78.7
IN	0	0	80.8
MI	0	0	84.1
SC	0	0	85.6
TN	0	0	87.6
Total	10	23	87.6
Note: States with Sentencing Commissions are yellow; others are green.			

Another 13 states have agencies with *potentially usable* NJRP data. However, BJS would need to consider carefully the time and resources needed to produce a research-ready dataset from these states. If BJS were amenable to working through issues of data-cleaning, geographic coverage, and/or missing data, it could obtain NJRP data from a total of 23 states, home to roughly $\frac{3}{4}$ of the U.S. population.

Note that the two largest states appear in the “medium quality” group: California and Texas together account for 20 percent of the U.S. population. Moreover, collectively these 13 states represent about half of the U.S. population. As discussed later in this

report, these findings can be used to influence the development of a transition strategy from county based to state based NJRP data collection.

Table 2 also shows that six states are currently not ready for statewide NJRP data collection. Surprising, one of these, Ohio, is a *sentencing commission state* but has no apparent capacity for statewide case-level data collection and analysis.

3.2 Key agencies and quality trade-offs in states with available NJRP data

Table 3 confines attention to states with “high quality” NJRP data and provides a more detailed listing of the key agencies to contact for data acquisition. While the great majority of individual agencies are sentencing commissions, we identify some notable exceptions. In Louisiana, a sentencing commission state, it is not the commission but rather the department of corrections that houses micro data on felony offenders. This agency not only produces and analyzes all case-level data for the state’s sentencing commission, but has also worked directly with BJS in the past to produce data for the NJRP. New York’s Division of Criminal Justice Services, the state’s criminal records depository, also stands out for the “high quality” analytic datasets it releases annually through its close collaboration with the administrative office of the courts. Wisconsin stands out as the only state in which the court system itself, absent a sentencing commission, promises to provide “high quality” NJRP items.

State	High Quality	Medium Quality
AL	Admin Office of the Courts*	-
KS	Sentencing Commission	Bureau of Investigation
LA	Dept of Corrections	-
MA	Admin Office of the Courts*	-
MN	Sentencing Commission	Admin Office of the Courts Bureau of Criminal Apprehension
NC	Admin Office of the Courts*	-
VA	Sentencing Commission	Supreme Court
WA	Sentencing Commission	Admin Office of the Courts
NY	Division of Criminal Justice Services	Office of Court Admin
WI	State Courts	Crime Information Bureau

Note: States with Sentencing Commissions are yellow; others are green.
 * Administrative Office of the Courts contains a department in which the state's sentencing commission is located.

In four of the 10 “high quality” states, there is only one viable agency source for the NJRP. Even so, in three of these states, the sentencing commission’s annual dataset is produced within the administrative office of the courts. As a result, BJS may have relatively seamless access to any secondary NJRP data items should they be missing. In the remaining 6 “high quality” states, BJS has other “medium quality” options available if data access proves to be too time-consuming or there are other unforeseen difficulties.

In order to expand NJRP statewide data collection beyond these 10 states, BJS may also want to consider approaching the states whose agencies offer “medium quality” data. Each of these data sources has their limitations; however, there may be trade-offs that

BJS finds advantageous. In Table 4, we offer a list of these states and key agencies along with individual ratings on coverage, completeness, and standardization to preface a brief discussion of these trade-offs.

Table 4. Key agencies and trade-offs in states with only "medium quality" NJRP data				
State	Agency	Quality Dimensions		
		Coverage	Completeness	Standardization
MD	Commission on Criminal Sentencing Policy	+	+	+
	CJIS Central Repository	+	+	+
	Maryland Judiciary	+	+	+
MO	State Highway Patrol	+	+	+
	Dept of Corrections	+	+	+
	Admin Office of the Courts	+	+	+
OR	Sentencing Commission	+	+	+
	Admin Office of the Courts	+	+	+
PA	Sentencing Commission	+	+	+
	Admin Office of the Courts	+	+	+
CA	Div of Justice Information Services	+	+	+
CO	Admin Office of the Courts	+	+	+
ID	Idaho State Judiciary	+	+	+
IL	State Police	+	+	+
FL	Dept of Law Enforcement	+	+	+
GA	Bureau of Investigation	+	+	+
KY	Admin Office of the Courts**	+	+	+
NJ	Admin Office of the Courts	+	+	+
TX	Bureau of Investigation	+	+	+

Note: States with Sentencing Commissions are yellow; others are green.
 ** Administrative Office of the Courts also serves as state's criminal records depository
 + indicates a high rating; + a medium rating

BJS should be able to prioritize the “medium quality” states by differentially weighing quality assessments on the dimensions of coverage, completeness, and standardization. For example, if BJS’ primarily values the acquisition of *clean, standardized data for an entire state*, then it might choose to approach the populous states of California and Florida first. Their best source of NJRP data comes from criminal records depositories which, as such, suffer from some issues with lag-time and matching associated with putting together arrest and court records and court disposition records. Nevertheless, both states produce regular analytic datasets for policy analysis that would be relatively easy to translate into NJRP data. The downside is that not all NJRP data items are likely to be included for CA and FL, leaving possibly some important items completely missing.

In contrast, BJS might decide to favor states that don’t produce analytic datasets but that house the most *comprehensive and complete data*. In this light, the administrative offices of the courts in states like Kentucky, New Jersey, Idaho, and others look more attractive than criminal records depositories in other states.

Similar trade-offs also affect BJS decisions about which agency to approach *within a state*. In Maryland, the Commission on Criminal Sentencing Policy produces the most clean analytic datasets in the state. However, these data exclude cases tried in district court, death penalty cases, and felony cases for which incarceration is not an option. If BJS wants to ensure that Maryland's NJRP data encompass these cases, it may choose to work with the CJIS Central Repository, with the caveat that many arrests will not have a corresponding disposition or sentence and the data are likely to require some cleaning.

3.3 Data access issues

Another important consideration in prioritizing states and individual agencies for NJRP data collection is the ease with which BJS may access the data. In Table 5 (see next page) we display the request procedure required by each individual agency as well as information on how quickly data requests might be approved and processed. We observe that one agency, the Idaho State Judiciary, has no policy of releasing data to outside users, and another that is unsure of what the proper data access procedure would be. However, most of the 35 agencies with "medium" or "high quality" NJRP data have formal protocols for accessing their micro data, 13 requiring that requests be submitted using a standardized online or hard-copy form, and another 8 preferring a formal letter laying out the particulars of the request. One of these agencies, the Oregon Sentencing Commission, also stresses that the request form has to be directed to the state's Administrative Office of the Courts, not the commission itself since it has no authority over the data. The remaining 13 agencies make the initial data request as easy as sending an email or telephoning.

In Table 5, we also identify those agencies able to expedite the approval of BJS data requests. The great majority (21 of 35) are likely to approve an initial request within a week. Those agencies that estimate their approval time to exceed a week often describe detailed review processes and schedules that impede a timely decision. For example, all requests made of the Missouri Administrative Office of the Courts have to be reviewed by the State Judicial Records Committee which only meets 4 times a year. Respondents at other agencies like the Massachusetts Sentencing Commission/AOC, the Minnesota Bureau of Criminal Apprehension, and the New York Office of Court Administration explain that data requests often take lengthy periods of time to approve because data access is governed by state statute.

As many agencies reported being able to expedite the processing of a data request as reported the capacity to expedite approval, though they were not necessarily the same agencies. There are several principal reasons for delays in data processing. First of all, many offices of court administration experience a great deal of seasonal variation in workload flows; they have to prioritize requests made by the state legislatures which leads to a longer turnaround time for all other lower priority data requests, including those made by federal agencies. Second, many agencies are in the process of implementing new data systems to which they have not fully migrated. Third, some agencies have limited staff or have experienced recent cut-backs that have diminished

Table 5. Ease of access to available NJRP data by state and agency								
State	Agency	Quality	Request Procedure			Ease of Access		
			Form	Call/email	Formal letter	Fast approval	Fast processing	
AL	Admin Office of the Courts*	+	✓			✓		
KS	Sentencing Commission	+	✓			✓	✓	
	Bureau of Investigation	+		✓		✓	✓	
LA	Dept of Corrections	+			✓	✓	✓	
MA	Admin Office of the Courts*	+			✓			
MN	Sentencing Commission	+	✓			✓	✓	
	Admin Office of the Courts	+		✓				
	Bureau of Crim Apprehension	+	✓					
NC	Admin Office of the Courts*	+	✓			✓		
VA	Sentencing Commission	+		✓		✓		
	Supreme Court	+		✓		✓	✓	
WA	Sentencing Commission	+		✓		✓	✓	
	Admin Office of the Courts	+	✓			✓	✓	
NY	Div of Crim Justice Services	+	✓			✓	✓	
	Office of Court Admin	+	✓					
WI	State Courts	+			✓	✓		
	Crime Information Bureau	+			✓			
MD	Comm Crim Sentencing Policy	+	✓			✓	✓	
	CJIS Central Repository	+		✓		-	-	
	Maryland Judiciary	+		✓			✓	
MO	State Highway Patrol	+			✓	✓		
	Dept of Corrections	+			✓		✓	
	Admin Office of the Courts	+	✓				✓	
OR	Sentencing Commission	+		✓		✓	✓	
	Admin Office of the Courts	+		✓		✓		
PA	Sentencing Commission	+	✓			✓	✓	
	Admin Office of the Courts	+	✓			✓	✓	
CA	Div of Justice Info Services	+		✓		✓	✓	
CO	Admin Office of the Courts	+			✓			
ID	Idaho State Judiciary	+	--no access--					
IL	State Police	+		✓			✓	
FL	Dept of Law Enforcement	+	-	-	-		✓	
GA	Bureau of Investigation	+		✓		✓	✓	
KY	Admin Office of the Courts**	+	✓				✓	
NJ	Admin Office of the Courts	+			✓			
TX	Bureau of Investigation	+		✓		✓	✓	

Note: States with Sentencing Commissions are yellow; others are green.
* Administrative Office of the Courts contains a department in which the state's sentencing commission is located.
** Administrative Office of the Courts also serves as state's criminal records depository
+ indicates a high rating; + a medium rating

their capacity to respond to data requests. Lastly, some agencies have outsourced their data processing to an external provider, which adds a layer of bureaucracy to the process.

Like the quality trade-offs we presented in an earlier section, it may be useful for BJS to use this information on data access to plan data collection timelines or prioritize states or agencies within states for NJRP data collection. For example, among “high quality” states, BJS may consider approaching the Louisiana Department of Corrections before the Massachusetts Administrative Office of the Courts/Sentencing Commission because of the relative ease of data access. For the same reason, when choosing among “medium quality” states, BJS may decide to solicit NJRP data from the Texas Bureau of Investigation before submitting a data request to the Colorado Administrative Office of the Courts if timing is particularly sensitive.

3.4 Individual NJRP data items and documentation

As we found during the pilot phase of this project, obtaining documentation of available data proves to be challenging. Table 6 (see next page) illustrates the type of documentation furnished by individual agencies and summarizes the clarity it provides about the fields and codes employed.

Not surprisingly, all 10 agencies with “high quality” data can provide some sort of documentation, although only one has a traditional codebook with full explanations of all fields and codes. Of the 26 “medium quality” agencies, 8 can furnish some type of explanation for the data they house. Overall, the most common form of documentation is a printout of database fields from a statistical package. Other types of documentation include sample dockets, database diagrams, traditional codebooks, and even annual reports.

A review of the documents supplied by agencies reveals that the documentation is often incomplete. Only 12 of the agencies were able to produce materials that documented both the fields and the codes that they use. For administrative offices of the courts, which generally have massive case management databases with fields for every possible transaction in a court process from the filing of charges to the paying of any final fees, producing a codebook is particularly burdensome. Many point us to sample dockets and database diagrams to get an idea of the specific data items available, but lament that the only way to understand how each item is coded is to interact directly with the system on-screen.

That said, there are some things we can discern from the documentation we amassed. For example, certain types of agencies are more likely to collect particular NJRP data items than others. While all criminal records depositories will reliably have the date of arrest because of direct transmissions of these data from law enforcement, sentencing commissions and administrative offices of the courts do not routinely capture this field. Sentencing commissions usually track the incident date, while court systems’ first data point is the date charges are filed which may or may not correspond to the date of arrest.

State	Agency	Quality	Type of Documentation					Items Documented	
			Codebook	Statistical printout	Sample Docket	Diagram	Manual/guide	Other	Fields
AL	Admin Office of the Courts*	+	✓					✓	
KS	Sentencing Commission	+	✓					✓	
	Bureau of Investigation	+	not provided						
LA	Dept of Corrections	+					✓	✓	✓
MA	Admin Office of the Courts*	+					✓	✓	
MN	Sentencing Commission	+	✓					✓	✓
	Admin Office of the Courts	+		✓				✓	
	Bureau of Crim Apprehension	+				✓		✓	
NC	Admin Office of the Courts*	+	✓					✓	✓
VA	Sentencing Commission	+	✓				✓	✓	✓
	Supreme Court	+				✓		✓	✓
WA	Sentencing Commission	+	✓					✓	✓
	Admin Office of the Courts	+	not provided						
NY	Div of Crim Justice Services	+	✓					✓	✓
	Office of Court Admin	+	not provided						
WI	State Courts	+		✓				✓	
	Crime Information Bureau	+			✓			✓	
MD	Comm Crim Sentencing Policy	+	✓					✓	✓
	CJIS Central Repository	+	not provided						
	Maryland Judiciary	+		✓				✓	
MO	State Highway Patrol	+			✓			✓	
	Dept of Corrections	+	✓					✓	
	Admin Office of the Courts	+	not provided						
OR	Sentencing Commission	+	✓					✓	
	Admin Office of the Courts	+	not provided						
PA	Sentencing Commission	+	✓					✓	✓
	Admin Office of the Courts	+		✓				✓	
CA	Div of Justice Info Services	+	✓					✓	✓
CO	Admin Office of the Courts	+	not provided						
ID	Idaho State Judiciary	+	not provided						
IL	State Police	+	✓					✓	✓
FL	Dept of Law Enforcement	+	✓					✓	✓
GA	Bureau of Investigation	+			✓			✓	
KY	Admin Office of the Courts**	+	✓					✓	
NJ	Admin Office of the Courts	+	not provided						
TX	Bureau of Investigation	+					✓	✓	

Note: States with Sentencing Commissions are yellow; others are green.
 ** Administrative Office of the Courts also serves as state's criminal records depository
 + indicates a high rating; + a medium rating

Similarly, an NJRP data item like the nature of a conviction is not equally likely to be present in all types of data sources. Administrative offices of the courts regularly collect this information as part of their case management database and sentencing commissions often include it for their analyses of compliance with sentencing guidelines. However, the

nature of conviction is immaterial for criminal records depositories which are mainly concerned with recording guilt or innocence and the subsequent sentence. As a result, very few agencies of this type are able to furnish BJS with these data—only those that work closely and collaboratively with their administrative office of the courts, like the Division of Criminal Justice Services in New York.

Other NJRP data items are more universally problematic. For example, while almost all agencies include in their data bases a variable reflecting the convicted individual's race and ethnicity, relatively few break this information down into separate variables like those used in NJRP. The end result of this is that, in many states, it would be impossible to make racial distinctions among Hispanics. In addition, when asked about which fields are most likely to have missing or incomplete data, many respondents named the ones capturing race, ethnicity, and even gender. Our discussions with state staff revealed that criminal records depositories regularly receive complaints about incorrect demographic data from individuals when they review their criminal records. Many courts also do not require these characteristics from all defendants.

Turning to NJRP *sentencing items*, the biggest challenge will be using case data to calculate the minimum and maximum incarceration sentence, as well as the maximum incarceration term for the most serious offense. Very few agencies store variables with this information; most simply have a field for incarceration term without a range. Other areas where the NJRP sentencing data are tenuous include sentences to community service or treatment as well as flags for consecutive and suspended/deferred sentences. Such deficiencies may or may not prove to be an obstacle for BJS depending on NJRP editing/coding conventions when such limitations are encountered.

4. Discussion

4.1 *Methodological considerations under a hybrid sampling approach*

It is clear from the findings of this study that an *immediate* move to an all-inclusive, statewide data collection design for NJRP is neither realistic nor feasible. However, it is also clear that a statewide data collection is on the horizon; and it is appropriate to plan for the transition from the current county-based design to an eventual statewide design. A majority of states may not have statewide NJRP data available for a decade or more, especially given the ongoing economic crisis. Consequently, BJS may find it advantageous to adopt a *hybrid design strategy* to navigate from the current county-based design to a design relying on statewide data collection.

A *hybrid design* strategy is one that exploits the availability of statewide sentencing and convictions micro data. For states where it is easily available, NJRP data are harvested for the entire state; and for all other states, the usual county based design would be used to represent the balance of the U.S. From a sample design perspective, states providing statewide data *become self-representing strata*. The balance of the U.S. becomes a separate sampling task where strata are formed and counties are sampled via stratified sampling with probabilities proportional to size (similar to the current NJRP design).

Over consecutive iterations of NJRP, a hybrid design strategy allows for increasing numbers of states to participate on a statewide basis. It is possible that some states may never have the resources to adopt a centralized data system. However, if eventually 90+ percent of the U.S. population is covered by statewide NJRP data collection, then it can still be considered a successful statewide data collection model since national statistical estimates are not sensitive to low amounts of noncoverage.

Two more issues are worthy of discussion in the development of a hybrid sample design:

- states with easily accessible data but having only partial coverage
- stratification of the balance of the U.S.

Partial Coverage. Some states (e.g., PA) have centralized NJPR data for some but not all jurisdictions inside their borders. If the population in the missing areas is small (e.g., less than 5 percent), then the missing areas can just be ignored. On the other hand, if the population is substantial (as in Pennsylvania where Philadelphia is not included) then the appropriate solution would be to include the missing areas in the “balance of the US” sampling frame, where “balance of the U.S.” refers to the collection of states for which statewide data are not yet available.

Stratifying the balance of the U.S. Ideally, one first identifies the states where statewide data collection will occur and then constructs the “balance of the U.S.” sampling frame. Unfortunately, the first of these steps may not be possible far enough in advance of the NJRP field period. . . Moreover, even after the statewide data collection states are identified, problems may emerge that render it impossible to gather statewide data. . . This is particularly problematic if, by that point in time, the “balance of the U.S.” sample of counties has already been defined and selected. Any state that was intended for statewide data collection but fails to provide it poses a problem to the hybrid design.

The easiest solution is to identify “statewide states” far in advance of the NJRP sample design milestone, then stratify and select a sample of counties for the “balance of the U.S.” During the latter process, , separate sample designs should be developed for sampling counties independently within each “statewide state.” The county samples from a given statewide state are held and only used in the event that the “statewide state” fails to deliver statewide data. This strategy protects against the risk of a statewide state disaster (i.e., failure to provide statewide data).

The proposed solution does not address the issue of how far in advance the statewide states can be identified. From a pragmatic perspective it seems logical to start the transition process “small,” selecting only 2-3 statewide states for the first “NJRP Hybrid” sample. This first data collection period could be used not only to collect statewide data from 2-3 states but also to identify and set-up the next set of statewide states (e.g., 5-8) for the second iteration of the hybrid design. That way the statewide states are known far in advance of the second iteration NJRP sample design task. This process can proceed

and be enhanced over consecutive iterations of the NJRP until the full statewide goal is attained.

We now discuss approaches to prioritizing states for statewide data collection during the *first iteration* of an NJRP hybrid design.

4.2 Prioritization of states for NJRP sampling

In deciding how to prioritize states with available NJRP data for NJRP, BJS should identify where the statewide approach might yield the greatest benefits relative to the current county-based approach. Table 7 presents states with “high” and “medium quality” NJRP data by percent of the U.S. population and cost factor⁶.

Statewide NJRP collection has the potential to dramatically reduce costs in states with elevated cost factors like North Carolina, Texas, California, Illinois, or Wisconsin. In a hybrid sampling design, the prioritization of states for statewide data collection could incorporate the size of the state population. The state of California accounts for more than 12 percent of the total U.S. population. And Texas represents 9 percent of the U.S. population. Being able to convert just these two states earlier rather than later in the transition process would represent great progress.

4.3 Considerations for agency selection

When states have centralized NJRP data but do not include all jurisdictions, caution should be exercised. On the one hand, if 90 percent or more of all eligible sentencing and convictions are included, then noncoverage bias can be considered negligible. When noncoverage is more substantial, then options should be weighed carefully. Pennsylvania provides a good case in point. Its Sentencing Commission maintains complete, highly standardized data but excludes the state’s most populous city, Philadelphia. In this case, an immediate solution would be to consider Philadelphia a “pseudo- state” and exploit the centralized availability of NJRP data from PA state agencies. Philadelphia City could be included with certainty (making it self-representing) or it could be placed in the “balance of the U.S.” sampling frame and given a chance of selection. In either case it is represented in the NJRP. However, there is some advantage to being able to declare PA as a statewide data collection state. So if Philadelphia City has low cost NJRP data availability we would recommend that it be included with certainty whenever PA is transitioned to statewide NJRP data collection.

Other state noncoverage situations may not be as easy to resolve. The alternatives are to: (1) hold off on statewide data collection for those states, or (2) exploit whatever centralized data is available at the (partial) state level, then sample the “noncovered” counties. We believe the latter approach is most advantageous to NJRP when statewide

⁶ To select the 2004 NJRP sample, BJS assigned each state a cost factor of 1, 3, or 5. A cost factor of 1 means that the data collection and processing costs are low while maintaining good data quality. Cost factors of 3 and 5 indicate that it is 3 or 5 times as expensive to collect the data and the data quality is not as good in these states’ jurisdictions.

Table 7. State availability of NJRP data by cost factor, counties sampled in 2004 NJRP, and % of US population

State	High or Medium Quality	Cost Factor	% of US Population
OR	+	1	1.2
KY	+	1	1.4
MD	+	1	1.9
WA	+	1	2.1
KS	+	1	0.9
LA	+	1	1.4
MN	+	1	1.7
GA	+	1	3.2
NJ	+	1	2.9
VA	+	1	2.6
NY	+	1	6.4
PA	+	1	4.1
ID	+	3	0.5
CO	+	3	1.6
AL	+	3	1.5
MA	+	3	2.1
MO	+	3	2.0
FL	+	3	6.0
WI	+	5	1.9
NC	+	5	3.0
IL	+	5	4.3
CA	+	5	12.1
TX	+	5	7.9

Note: States with Sentencing Commissions are yellow; others are green.
 + indicates a high rating; + a medium rating

data are available for substantial portions of the state – say, for jurisdictions representing 60+ percent of the state’s population. To achieve full coverage of the U.S., the “noncovered” counties/jurisdictions would be added to the “balance of the U.S.” sampling frame.

We do not believe that the availability of individual **NJRP data items** should be a primary factor guiding the prioritization of states and agencies. The county level data gathered in the last round of NJRP suffers from many of the same deficiencies discussed in more detail in an earlier section of this memo. To illustrate this point, of the 299 counties sampled in 2004, 137 are lacking the date of arrest and 138 do not have a variable for the ethnicity of the convicted individual. In this context, there is no clear disadvantage to using state level data that may be similarly flawed.

Regardless of how state agencies are selected for statewide NJRP data collection, BJS should request data from agencies well in advance. This is especially true in states like Colorado where the only potential NJRP data source estimates prolonged wait-times for both the request approval and processing. BJS may also want to allow extra time in states where the primary sources of NJRP data are likely to be the *criminal record depositories*, like Florida and Georgia. Most of these systems are *offender* based since they are used mostly as look-up systems of individuals for law enforcement and employers. As a result, additional processing may be necessary to transform the data into a *case-level format* that would be usable for NJRP analytical purposes.

5. Conclusion

The results of this project are encouraging for replacing county-based collection of NJRP data with statewide level collection. Twenty-three of the 29 states we purposively selected were assessed to have high or medium quality NJRP data available at for the state. Collectively these states account for almost $\frac{3}{4}$ of the U.S. population. Further, ten of these states were assessed to have *high quality* NJRP data available; these account for a quarter of the U.S. population.

We recommend that a transition to statewide data collect be conducted over a transition period covering several iterations of NJRP. The total number may span a period of ten years or more depending on the ability of states to develop a centralized, integrated data system for felony sentencing and convictions. As discussed earlier, a conservative approach would be to commence statewide data collection in a few states. BJS could use this experience as well as contacts with other states to develop and refine the protocol for state selection and to identify candidate states to add in subsequent iterations of NJRP. In this fashion, each implementation of NJRP would be used to develop and expand statewide data collection for subsequent iterations over the course of the transition periods. The design could be developed to reduce or eliminate the risk of state level dropouts (via a backup sample of counties for each self-representing state).

We reiterate two of our recommendations from the pilot study: First, state agencies should be identified and contacted a year or more in advance of NJRP data collection to identify and overcome access barriers and more generally to establish a solid working relationship. This will facilitate the data acquisition process once the NJRP is launched. Secondly, we recommend that the NJRP program consider an outreach strategy to engage state agencies and their staff, possibly by providing technical assistance, with the objective of promoting standardization, full geographic coverage, completeness of data, etc. in state agency judicial data systems for subsequent NJRP-type data extract requests. Moreover, BJS should instill and promote the value of the NJRP to states so that state agencies become stakeholders rather than entities upon which the burden of NJRP data collection is placed. Venues for such outreach activities would include professional conferences, workshops, etc. The objective would be to conduct outreach on a *continuous, ongoing basis* as part of the NJRP research program.

Limitation. One important factor aspect we did not consider in this report was the volume of data to be processed and collected at the state level. There may be non-trivial capacity issues with the volume of data being collected at the state level for multiple states. This may impede processing times and tax the systems being used to store the data.

Appendix A.

Screenener

Hi, I'm calling you from the Urban Institute in Washington, D.C. We're currently working with the Bureau of Justice Statistics to think about strategies for making data collection more efficient.

Currently, NJRP (National Judicial Reporting Program) collects data on felony convictions and sentencing. Data is gathered from selected counties within states, but BJS wants to know how much of these data states like yours might be collecting already.

- **What is your role at _____ (name of agency)?**
- **How long have you been in this position?**

We really want to talk to both someone who can give talk to us about:

- The judicial data available at your office
- what areas of the state are included
- what judicial items are included and how they are coded and recorded
- how to request and access the data
- **Would you be the right person to speak to? Is there someone else who might want to join us?**
- **Do you have time to talk right now? If not, when would be a convenient time to schedule an interview?**

Overview

To start out, it would be wonderful if you could give us a brief overview of the data your agency maintains on judicial proceedings for felony offenses?

- **How are the data stored -- by offender, court case or some other way?? (offense, offender, aggregate only)**
- **Which points in the judicial process do these data cover?**
 - **Arrest**
 - **Disposition**
 - **Sentencing**
 - **Can you send me documentation or description of the fields and codes?**
- **Where does your data on felonies come from?**
 - **[If they get their data from courts] Which type of courts?**
 - **[If they get their data from law enforcement] Which type of agencies?**

Coverage

- **Do all of these courts/agencies report their data to you?**
 - **[If not] Which ones do not?**
 - **What percentage of cases statewide are not covered?**

Data Completeness

- **What percentage of their felonies are reported to you?**
 - Is there general agreement about which fields are mandatory?

- [For criminal records depositories]
 - **What percentage of arrests are missing a disposition?**
 - **What percentage of dispositions are missing an arrest?**

- [for criminal records depositories and sentencing commissions]
 - **Are there any types of felonies that are NOT reported to you?**

Data Standardization

- **Do you produce analytic datasets?**

- **Does your database use standardized fields? Standardized codes?**

- **In practice, to what degree do agencies code things the same way or do they have local coding practices? How do you know?**

Data Requests and Access

- **What steps would be needed for BJS to access these data?**
[Probe: request form, review process, fees if applicable]

- **How long does it typically take to get a request approved? (ballpark is fine—i.e. days, weeks, months)**
To process a request?

Closing

Those are all the questions I have today. I may need to follow-up with you at a later time.

- **What your email address?**

We are also trying to talk to your counterparts at _____.

Could you refer me to someone at that agency?

Thanks so much for taking time to talk to me today.