Data Quality Assessment NCIC Criminal Histories for Scoring the Arnold Public Safety Risk Assessment

Background

For our retrospective study to evaluate the Arnold Public Safety Assessment (PSA) tool, we needed out of state criminal histories. Three of Vermont's population centers are on the border with another state, and it takes one wrong turn to cross state lines. The proximity of other states, in addition to migration patterns, make out of state criminal histories necessary when evaluating criminal justice policy. Forty-two percent of the total cohort of people had a rap sheet from another jurisdiction. Not all contained convictions of interest, but it illustrates why out of state criminal histories are needed to evaluate a tool that is based on criminal histories.

Data Permissions

CRG filed an application with the Federal Bureau of Investigation for Institutional Review Board (IRB) approval for the research design. This was necessary to gain access to the out of state criminal histories. After the IRB was approved, the FBI Criminal Justice Information Services approved the design. An agency number (ORI) was issued for this project, allowing the International Public Safety and Justice Network (NLETS) to parse the data into an analyzable Extensible Markup Language (XML) and transfer the data to Vermont Crime Information Center. This process took over four years.

CRG has a data sharing agreement with Vermont Crime Information Center to access Vermont criminal histories.

Data Acquisition

After identifying the cohort in CRG's Court Adjudication Database, we submitted the names and dates of birth to Vermont Crime Information Center to obtain the Vermont criminal histories. We extracted the FBI numbers from the criminal histories and transferred the information to NLETS.

NLETS then acquired the rap sheets and used its parsing service to deliver them to CRG in XML to be manipulated for data analysis. All 50 states and the federal government returned rap sheets on the cohort.

NLETS Parsing Service for Data Analysis

For the Arnold PSA we were only interested in convictions and sentences, and this section reflects our views on what work needs to happen to extract that data from the NLETS parsing service.

The NLETS parsing service returned the rap sheets with XML tags that would allow the extraction of elements needed for analysis. The following states did not parse: Massachusetts, Oklahoma, New Jersey, Virginia, Montana, Alaska, and Nebraska. The information from those states would require hand scoring, or perhaps some other type of programmed text recognition.

Other states have the potential, however, to be read in using R or other programming language. We were unable to do so for this project because the data came to us two weeks before the end of the grant. The following would need to happen for the parsing service to be usable for data analysis:

1. Map all the data structures and elements for each state.
    a. Each state has its own way of reporting basic information, such as the charge convicted of, sentence, and other events related to a case. For example, Vermont records parole violations as part of a cycle that includes the arrest offense, convicted offense, and then the parole violation. Other states report parole violations as a separate, unconnected event to the offense for which the person is serving parole. These need to be mapped for each state for data importation for analysis.
2. A data dictionary for all states that includes disposition types, sentence types, offense levels, etc.
    a. For scoring the Arnold PSA, we needed to count the number of misdemeanor convictions and number of felony convictions. Many states have a disposition that would not be considered a conviction, such as Continued Without a Finding, or Good Behavior, or other intermediate sanctions before a

conviction. Some states also reported levels of crimes that were infractions and distinct from felonies or misdemeanors.

3. Programming to turn words into meaningful units of analysis.

    a. String manipulation programming is needed to extract sentence information, especially durations of sentences; this would happen after the creation of a data dictionary, which is needed to catalog all the ways states record sentences. For example, sometimes sentences were recorded "2Y6M" or spelled out as "Two Years and Six Months." Each phrase has a programming answer, but it needs to be programmed in.

Data Quality for the PSA:

The PSA did not work for Vermont. However, the out of state criminal histories did change the scores of many in the cohort. The accuracy of the records would affect the scores, and this may have a differential impact on defendants. In addition to convictions, the PSA also scores prior failures to appear and pending charges. Whether these are accurately recorded in the criminal histories is a concern.

For example, Vermont lists failure to appear as an event that happens on a case. A case is called a cycle in the rap sheet and the cycle contains all the charges and events that happen until the end of the case. Other states, however, may not be reporting them at all. Of the 2,521 out of state rap sheets we scored, only 53 had an obvious failure to appear on their rap sheets. This seems an improbably low 2.1% rate of a history of failure to appear.

For many states, it was impossible to determine what was a "pending case" or simply wasn't prosecuted. Massachusetts, for example, reported many prior arrests for the 196 people in our cohort, but only 20 people had a reported a prior conviction. In several states, there were cases that appeared to have been filed but hadn't been disposed of in 20 years. There was no way of knowing if just the disposition was not reported, or if, in fact, the case was pending for 20 years.

Conclusion:

We were excited to again work with out of state criminal histories. The NLETS procedure has great promise, and relevant institutions such as the Bureau of Justice Statistics, Justice Research and Statistics Association, SEARCH, and the SACS should put a concerted

effort into programming and data dictionaries. A one-time push, with a continual data governance structure for researchers, could turn this really important service from NLETs into a very viable resource for researchers desiring to use criminal histories in their research.