This report was prepared by Abt Associates using federal funding provided by the Bureau of Justice Statistics.

Document Title: Event- and Offender-Based Recidivism Methodology Using the National Corrections Reporting Program

Authors: William Rhodes, Abt Associates
Garald Gaes, Abt Associates
Christopher Cutler, Abt Associates
Ryan King, Abt Associates
Jeremy Luallen, Abt Associates
Tom Rich, Abt Associates
Michael Shively, Abt Associates
Joshua Markman, Bureau of Justice Statistics

Document No.: 250749

Publication Date: February 14, 2019

Award No.: This project was supported by award numbers 2010-BJ-CX-K067 and 2015-R2-CX-K135

Abstract:

This report presents a new methodology using National Corrections Reporting Program (NCRP) data to produce return-to-prison recidivism estimates. This offender-based methodology uses longitudinal data to estimate recidivism and produces estimates of offenders released over long periods rather than specific release cohorts. Though each type of recidivism methodology is suited to answer different criminal justice questions, scholars formerly relied on the event-based approach alone. This report discusses how to produce offender-based recidivism estimates and the circumstances in which offender-based estimates are more appropriate, based on the type of policy or research question.

This page intentionally left blank.

Event- and Offender-Based Recidivism Methodology Using the National Corrections Reporting Program

William Rhodes
Garald Gaes
Christopher Cutler
Ryan King
Jeremy Luallen
Tom Rich
Michael Shively

Abt Associates
10 Fawcett Street
Cambridge, MA 02138

Joshua Markman

Bureau of Justice Statistics
810 Seventh Street, NW
Washington, DC 20531

Executive Summary

This Technical Report describes methods that the Bureau of Justice Statistics (BJS) uses to estimate recidivism, defined in this report as a return to prison, using data from the National Corrections Reporting Program (NCRP). NCRP is a compilation of administrative records provided by state departments of correction and community supervision. The states submit annual files of offender-level data on entries and exits from prison and post-confinement community supervision. In addition to using the data to report on the characteristics of correctional populations in the United States, BJS reassembles these data into "term records," representing a single continuous period of time that a particular offender spends in prison or on post-confinement community supervision. The term record also includes unique offender identifiers, enabling analysts to easily construct prison and post-confinement community supervision histories for individual offenders. An offender's history may include multiple terms for either prison or post-confinement community supervision, and for many states, the data goes back to 2000.

This report describes two ways to calculate recidivism: event- and offender-based. The traditional recidivism calculation used by most state departments of correction is event-based. This type of calculation examines a cohort of n prisoners released in a given year and then determines the percentage of offenders who return to prison within some specified time frame, such as 3 years. Another approach for estimating recidivism is called "offender-based." The offender-based approach recognizes that during a long period (2000–2014 in this report), N distinct offenders will experience prison terms. Using the offender-based approach, an analyst estimates an average rate of recidivism for all N offenders observed during that long observation window. Estimating recidivism using the offender-based approach requires datasets such as the NCRP that collect prison admissions and releases during a long period, while event-based estimates require no such lengthy data collection.

A key difference between the two identified recidivism estimation approaches, offender- and event-based, is that each estimation pertains to a different population. The event-based approach makes inferences about a population of n offenders defined as everyone released from prison in a given year, while the offender-based approach makes inferences about a population of N offenders defined as everyone released from prison during a long period. That is, the two approaches make inferences about two different populations.

This Report makes several important points about these two approaches. First, both are firmly grounded in criminological research traditions: criminal career and life course research for the offender-based approach, and developing predictive instruments, correctional program evaluation, and correlates of recidivism outcomes for the event-based approach. Thus, the event- and offender-based approaches are not in competition; the best approach depends on the research question.

Second, offender-based estimates of recidivism are lower than event-based estimates because there is a higher proportion of high-risk offenders, who are more likely to recidivate, in the pool of n offenders than there are in the pool of N offenders. The criminological literature has documented that high-risk offenders cycle in-and-out of prison, Beck/Shipley and Langan/Levin provide the percent of prisoners who returned to prison, but do not provide information those who cycle in-and-out of prison multiple times.(Beck & Shipley, 1989; Hunt & Dumville, 2016; Langan & Levin, 2002; National Research Council, 2008). Thus, high-risk offenders have a higher concentration in yearly release cohorts of size n than among the N

offenders who are incarcerated sometime during the long time period. For example, using NCRP data in 16 states and the offender-based approach, an estimated 65% of the population of *N* offenders never returned to prison between 2000 and 2014. In comparison, according to the event-based approach, 43% of *n* offenders released in 2000 from the same 16 states never returned to prison by 2014. This percentage is comparable for other release cohorts.

Third, the two approaches address different policy questions regarding recidivism. Event-based estimates, for example, help policymakers understand the risk of reoffense posed by a particular release cohort. Offender-based estimates, in comparison, assess the "performance" of the criminal justice system over time. This report provides guidance on which approach to use to answer a particular policy or research question.

This Technical Report includes examples that show why offender-based estimates are lower than event-based estimates, and data tables that illustrate the findings for both approaches. BJS has indicated that it intends to release future BJS publications and web tools that will include additional tables and charts that rely on both event- and offender-based recidivism approaches.

1. Introduction and Background

Criminal recidivism, the act of returning to crime after some form of adjudication for a prior crime, is a common benchmark appearing in criminal justice policy analysis. Michael Maltz's (1984, 2001) seminal work on recidivism articulates two fundamental questions about its measurement: (1) For what purpose is the statistic used? (2) How do you calculate a proper statistic? This Technical Report addresses both of these questions distinguishing between offender- and event-based recidivism. After providing the definition of offender- and event-based approaches used for this report, the report discusses data from BJS's NCRP and calculates offender- and event-based recidivism estimates. The lengthy time-series the NCRP provides is essential for offender-based recidivism estimates and leads to estimates of incarceration desistance. This report provides further intuition on the meaning and policy implications of event- and offender-based recidivism, and offers guidance on when to use the different measures.

Terminology and Definitions

- ***At risk.*** Once released from prison, offenders are at risk of returning to prison. The at-risk period is always of fixed length but a variable length risk period could be used instead.

- ***Custody.*** Prisoners held in the physical custody of state or federal prisons or local jails, regardless of sentence length or the authority having jurisdiction.

- ***Event-Based Population.*** All offenders who were released from prison during a short time, typically 1 year, comprise an event-based population. Each offender is counted as many times as he or she is observed.

- ***High-risk offender.*** Offenders who are released and reenter prison repeatedly during a long time.

- ***Low-risk offender.*** Offenders who are released and reenter prison infrequently during a long time.

- ***Jurisdiction.*** The legal authority of state or federal correctional officials over a prisoner, regardless of where the prisoner is held.

- ***Observation window.*** An observation window is the length of time that a state has reported corrections data to the NCRP. All states are reporting to the NCRP. For most of this Report, data are used from 16 states for which admissions and releases have been linked from 2000 to 2014. Other states have reported for shorter observation windows, and their data are used where appropriate.

- ***Offender-Based Population.*** All offenders who were released from prison during a long time comprise an offender-based population. Each offender is defined uniquely.

- ***Recidivism.*** A return to prison within the state from which the offender was released. The return may be for a new commitment, technical violation, or any other reason as long as the offender stays for more than 30 days.

- ***Release cohort.*** A release cohort comprises individuals who were released during a specified year, for example, the 2000 release cohort. When using an event-based sample, every release is given a weight of 1.

The definitions provided on this page (e.g,, high-risk offender and recidivism) are for this particular study and not universal definitions.

1.1 Offender- and Event-Based Populations

Offender-based analysis has a rich tradition in criminology and underscores the importance of longitudinal research, that is, individual behavior during a long period. Piquero (2008) identifies scholarship beginning in the late 19th century evolving into modern paradigms of longitudinal analysis. Two of these paradigms are criminal career and life course research. The essential idea underlying this research is the longitudinal sequence of an offender's crimes (Blumstein, Cohen, Roth, & and Visher, 1986; Piquero, Farrington, & Blumstein, 2003). Within this tradition, criminologists study the factors that affect the onset, termination (desistance), patterning, and rate of crime during the period of study (Bushway, Piquero, Broidy, Cauffmann, & Mazerolle, 2001; Brame, Bushway, & Paternoster, 2003; Maruna, 2001; Laub & Sampson, 2001). Life course research embeds the criminal career paradigm within a developmental framework (Liberman, 2008; Piquero A. , 2008).The arc of criminality is examined from its earliest developmental period in conjunction with other life events that modify or mediate different trajectories. Although a great deal of this longitudinal research focuses on early childhood aggression and delinquency (Nagin & Tremblay (1999); Nagin & Tremblay (2001)), there are longitudinal databases that follow individuals into adulthood and some even into middle age and beyond (Laub & Sampson, 2006; Sampson & Laub, 1993).

These offender-focused, rather than event-focused, criminal justice sequences require methods such as panel and multilevel models that allow the analyst to analyze offender-specific events over time. Especially noteworthy is the seminal group-based trajectory modeling approach of Nagin and Land (1993) and Nagin (2005) that attempts to identify latent groups of individuals distinguished by their longitudinal patterns of criminal events. The trajectory of events that is modeled may be based on self-reported delinquency or crime, independent judgments of aggression, or administrative records of criminal justice contacts such as arrests and commitments. The person is the unit of analysis and the criminal events are repetitions of behavior over time, acknowledging that individuals who do not commit crimes will have no instance of an event in the study period.

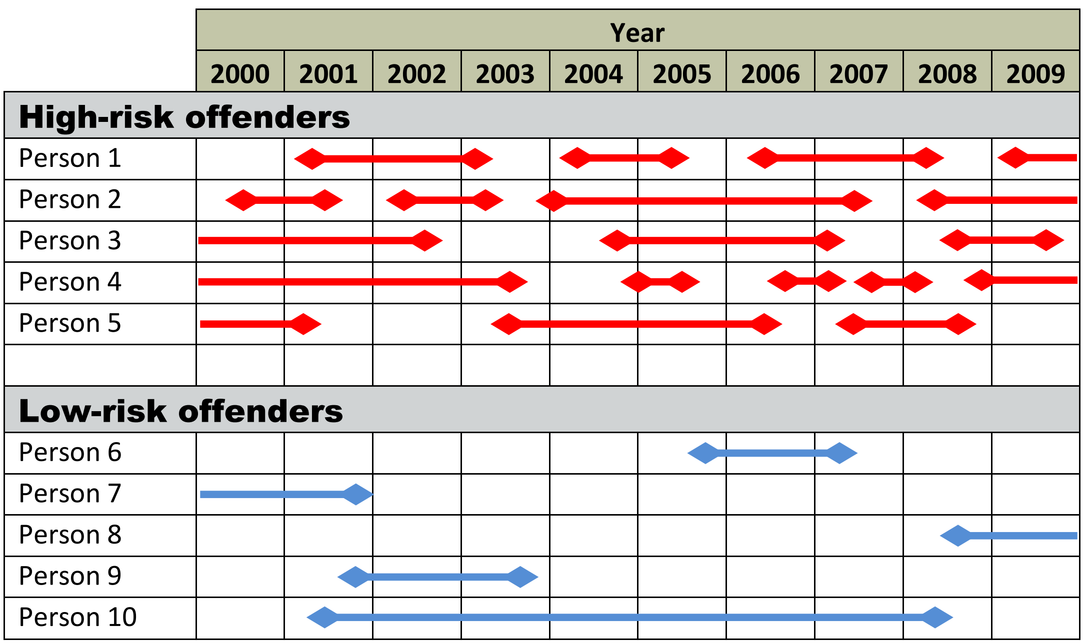
A second tradition may be viewed from a cross-sectional analysis perspective. There is a rich literature that uses alternative definitions of recidivism—self-reported crime, and administrative measures of arrest, conviction, or imprisonment. This literature covers such diverse topics as developing prediction instruments (Gottfredson & Moriarty, 2006; Berk, 2012); correctional program evaluation (Andrews, et al., 1990; Aos, Phipps, Barnoski, & Lieb, 2001; MacKenzie, 2006); or correlates of recidivism outcomes (Gendreau, Little, & Goggin, 1996; Hanson & Bussiere, 1998; Bonta, Law, & Hanson, 1998; Cottle, Lee, & Heilbrun, 2001). These research domains provide guidance about predicting recidivism outcomes, targeting who should qualify for increased levels of supervision, evaluating the effectiveness of interventions, and measuring potential covariates that will differentiate between those who do and do not recidivate.

Rhodes, et al. (2014), formally introduced the distinction between offender- and event-focused sequences, characterizing the differences in populations of interest, providing calculations, and discussing policy implications. Note that these two lines of research inquiry, offender- and event-based, typically make inferences about different populations. The criminal careers literature tends to make inferences about a general population of offenders or potential offenders, analogous to the *N* offenders that the offender-based approach associates with a long time period. The cross-sectional approach makes inferences about a specific population of offenders released from prison during a relatively short period of 1 year or so, analogous to the population of *n* offenders studied by the event-based approach. Both lines of inquiry are valuable to criminological research. However, the applicability of employing one approach over the other depends on the posed research question; the two approaches are not interchangeable and may lead to faulty conclusions if applied incorrectly.

1.2 Introductory Explanation of the Event- and Offender-Based Approaches

A simplified diagram using a population of five high-risk and five low-risk offenders (a population of N=10) demonstrates the difference between the offender- and event- based approaches (Appendix Figure 1). The high-risk offenders, those especially likely to return to prison, return to prison more often than the low-risk offenders during the observation window of 2000–2009. Person 4, one of the high-risk offenders, has five prison terms in this window; person 7, one of the low-risk offenders, has one. Each term is represented by a beginning and an end (using diamonds), unless the term starts before the observation window and/or ends after the observation window.

Appendix Figure 1: Illustrative depiction of prison terms for high- and low-risk offenders



Consider the event-based approach. With this small population of offenders, a release cohort in 2001 would identify high-risk offenders 2 and 5, and low-risk offender 7 in the sample of releases, and then follow them for some period of time to observe whether they return to prison. While there are five high-risk and five low-risk offenders in the population of N=10 offenders who are in prison at some point from 2000 to 2009, an event-based sample of n=3 drawn in 2001 will contain two high-risk and one low-risk offender. The year 2001 is not unusual. The year 2003 event-based release cohort has three high-risk offenders and two low-risk offenders; the year 2005 event-based release cohort has two high-risk offender and no low-risk offenders. In the real world, where *n* and *N* are large numbers, an event-based sample will always include a higher proportion of high-risk offenders than will an offender-based sample. Comparing the number of red and blue lines in Figure 1, the ratio of high-risk to low-risk offenders in an event-based release cohort will be about four times as large (19 prison terms for high-risk offenders divided by 5 prison terms for low-risk offenders) as the actual ratio of high-risk to low-risk offenders, a ratio of 1.

Now consider the offender-based approach. Seeking to estimate the rate of recidivism for the N=10 offenders who comprise the population of offenders who experienced prison during the “long” 10-year period, all five of the high-risk offenders recidivated at some time during the 10-year period none of the low-risk offenders recidivated. On average, the rate of recidivism is 0.5 for this population of N=10 offenders. (For this illustration, calculations ignore censoring.) In comparison, using the event-based approach and defining the population of all offenders released in 2001, two of three offenders were readmitted to prison, resulting in a recidivism rate of 0.66. Defining the population as all offenders released in 2003 again leads to a higher recidivism rate: three of four offenders were readmitted to prison, resulting in a recidivism rate of 0.75. Event-based recidivism rates are higher than offender-based recidivism rates. Section 3 provides a more formal explanation for why event- and offender-based estimates differ, but before turning to that formal explanation, the next section explains how offender- and event-based estimates are computed.

1.2.1 Computing Offender-Based Estimates

Turning to estimation, inverse probability weights applied to a defined release cohort result in offender-based estimates. Continuing with the Figure 1 example, a release cohort from any year (such as 2000) is a sample from the population of N=10 offenders. The sampling probability is proportional to the number of releases during the 10-year period, so for person 1, the sampling probability is proportional to 3 and for person 6 the sampling probability is proportional to 1. Inverse probability weights applied to the 2000 release cohort will then produce offender-based statistical estimates. An approximately equivalent way to compute the offender-based estimates is to randomly select one release for every offender during the 10-year period, but operationally, the first method avoids right-hand censoring. As in survey research, good estimates require understanding the sampling procedure and using inverse probability weights to correctly estimate parameters for the population of interest. Computationally, the difference between event- and offender-based estimates is due to different weighting procedures. Thus, an offender-based study is only possible using a lengthy window, such as that provided by the NCRP, because the long window provides the means to correctly apply weights to the data. Event-based estimates do not require a long window. The rationale for this requirement is justified in the following section on methodology.

The inverse weighting required to produce offender-based estimates of recidivism is similar to other applications called choice-based or endogenous sampling. Manski and Lerman’s (1977) paper describes studies of commuter behavior in which analysts want to assess factors that determine whether a person will choose between a taking a commuter train or driving. Rather than sample households, sometimes it is more convenient for the analyst to gather data by surveying transit users at a station and auto users in a parking lot. Manski and Lerman show that if the resulting sample overrepresents one mode and underrepresents the other relative to the population, an analysis of predictors of the mode of transportation will result in inconsistent parameter estimation. Englin and Shonkwiler (1995) provide a second example analogous to prison releases, where surveying people who exit the park is analogous to a release cohort of prisoners (see the *Studying the Use of State Parks* text box). Rhodes et al. (2007), provide a criminology application of endogenous-based sampling. The next section of this Report formalizes the mathematics of offender- and event-based recidivism.

One method is not inherently better than the other. The research question dictates which method of recidivism estimation is appropriate. If interactions that offenders (e.g., the 10 in Figure 1) have with prisons is the primary motivation for the study, such as “how frequently do offenders cycle in and out of prison?” then offender-based estimators are better suited to answer the question. On the other hand, studying the effect of a prison intervention on recidivism, such as the effect of in-prison vocational training on post-release recidivism, would require the use of event-based estimators. Event-based estimation would also be used by prison administrators to forecast how many beds would be needed for repeat offenders in the future.

Studying the Use of State Parks

Consider a survey researcher who is studying the use of state parks (Englin & Shonkwiler, 1995). The research question is: How many unique individuals use the park and how often? The population of interest is park users and the unit of analysis is also park users. When answering such a question about park usage, survey researchers often use an intercept survey that selects respondents at the time that they enter the park. The survey is administered during a long period to capture infrequent park users, but that extended survey period also means that frequent park users are likely to answer the survey multiple times. When the general population of park users is the population of interest, survey respondents over represent frequent users and consequently a naïve analysis of survey data will produce biased estimates of the frequency at which campers use parks and it will understate the average elapsed time between one visit and the next.

Criminal offenders are like park visitors in that some are incarcerated frequently and others are incarcerated infrequently. Rather than being questioned about their prison usage at the time they enter prison, offenders are identified at the time that they leave prison, and rather than being asked to self-report returning to prison, their returning to prison is monitored. Researchers studying park usage know how to weight their sample to represent the population of interest for their study. The same is true for researchers studying recidivism. This Technical Report discusses one method for weighting prison release cohorts (the population of n) to represent the broader population of N offenders.

2. Methodology

To this point, the argument has been intuitive; at this point, the argument is formalized using a stylized illustration. When considering the mathematics underlying the offender- and event-based approaches, it is important to continue to think of two distinct populations: N individuals who are released at some time during an infinite observation window (approximated by the NCRP observation window) and the n individuals who are released (possibly multiple times) and comprise a release cohort. The offender-based approach estimates population parameters for the population of N; the event-based approach estimates population parameters for the population of n. One population parameter of interest is the rate of recidivism, and on average, this is smaller in the population of N because high-risk offenders accumulate in the population of n.

Turning to estimation, it is convenient to think of the population of n as being a sample from the population of N. Then, with a known sampling probability, standard survey procedures are available to weight the sample of n when estimating population parameters for the population of N. To substitute for unknown sampling probabilities, this approach estimates them as being proportional to the number of times that an offender is released from prison during the NCRP window period.

Though both the event- and offender-based approaches compute estimates from the sample of n offenders, they result in different estimates. One uses no weights (event-based) because it attempts to estimate parameters for the sample of n offenders. The other (offender-based) does apply weights because it is attempting to estimate parameters for the population of N offenders using the sample of n offenders. Event-based estimates need no weighting, or equivalently, they use a weight of 1 to every offender/release event. In comparison, offender-based estimates receive weights that are inversely proportional to the probability of selection into the sample. The long window period provided by the NCRP is essential for deriving the weights.

This weighting is more complicated than simply retaining a single record per offender from the release cohort. Because offenders infrequently experience two prison releases during the same year, such an approach will not yield estimates that differ much from the traditional event-based approach. Moreover, when an offender has two or more records of release during the same year, which record to retain is arguable.

2.1 What Is Being Estimated by the Offender-Based Estimator?

The more formal argument that follows provides clarification to the explaination above at the expense of adopting a sytlized mathematical model. For this model, assume a hypothetical cohort of N individuals comprising offenders who begin criminal careers on the same day and experience criminal careers of the same infinite duration. Assume these offenders may be partitioned by risk of recidivating such that P_L is the proportion of low-risk offenders in the population, P_H is the proportion of high-risk offenders, and $P_N = (1 - P_L - P_H)$ is the proportion of zero-risk offenders. Assume further that during their criminal careers, offenders generate events (zero-length prison terms) according to a Poisson process so that an event may never occur or may occur multiple times. The Poisson rate parameter for this event-generating process differs across strata such that λ_L is the rate of prison entry for low-risk offenders, λ_H is the rate of prison entry for high-risk offenders, and $\lambda_N = 0$ is the rate of prison entry for zero-risk offenders. Let λ_H be defined such that $\lambda_H > \lambda_L$. Finally, assume an observation window of length W that begins on the first day that offenders begin their criminal careers and extends as long as researchers may observe events.

Using the notation introduced above, the population parameters are P_H, P_L, λ_H, and λ_L. If P_H is much larger than P_L, then most offenders who enter prison will be at high-risk of recidivism. Estimating these parameters is fundamental to thinking about how prisons operate. Suppose that the objective is to estimate these four parameters. An ideal approach would be to use equal probability sampling from the general population survey of offenders and nonoffenders. Within that sample, the three strata are latent and the difficult inferential problem is to estimate the population parameters of interest, but the sample is not biased in favor of one group over the other. After applying appropriate estimation procedures, this sample should lead to unbiased (or at least consistent) estimates of the population parameters P_H, P_L, λ_H, and λ_L (Nagin, 2005). This is the ideal sampling procedure when estimating population parameters P_H, P_L, λ_H, and λ_L, but typically the ideal is unavailable.

Given the unavailability of the ideal sampling procedure, selecting offenders when they leave prison is often the next best and readily available alternative. As noted earlier, this is known as choice-based sampling or endogenous stratification. To be included in the sampling frame using this method, an offender must have been released at least once. However, some offenders appear in the sample multiple times because they enter and are released from prison multiple times. The Poisson model is helpful for understanding how choice-based sampling complicates estimation about populations of interest.

To explain, let an observation window of length W be the duration of time available for observing prison admissions and releases. To enter the study, an offender must be released from prison at least once during the observation window. Given a Poisson process, the probability that a low-risk offender will generate at least one event during the observation window of length W is $1-e^{-\lambda_L W}$. Likewise, the probability that a high-risk offender will generate at least one event during that same window is $1-e^{-\lambda_H W}$. Thus, the ratio (R) of high-risk to low-risk offenders who appear among the N offenders is a function of the length of the observation window:

$$R^N = \frac{P_H}{P_L}\left(\frac{1-e^{-\lambda_H W}}{1-e^{-\lambda_L W}}\right) \qquad (1)$$

As the length of the observation window approaches zero, the term in parentheses approaches λ_H / λ_L, and the N offenders observed during the window period will contain a disproportionately large representation of high-risk offenders compared to the actual rate of high-risk to low-risk offenders in the population of N offenders. As the window approaches infinity, the term in parentheses approaches 1 and the bias disappears: the proportion of high-risk to low-risk offenders among the N offenders observed during the window period will equal the proportion in the population, and the choice-based sample will be as informative as the ideal sample. Given that observations windows are always finite, the N observed offenders has a persistent bias toward representing high-risk offenders, but this bias is lessened by using a lengthy observation window W , which is available in the NCRP. Importantly, this discussion demonstrates the effect of the length of the observation window in estimating the true ratio of high-risk to low-risk offenders among the *N* offenders.

While the *N* offenders comprise unique individuals who are released from prison during the window period, the population of *n* offenders comprises individuals whenever they are released from prison, so that same individual may be selected multiple times. Assuming events occur according to a Poisson process, the ratio of high-risk to low-risk offenders in an event-based population is:

$$R^n = \frac{P_H}{P_L}\left(\frac{\lambda_H}{\lambda_L}\right) \qquad (2)$$

In this stylized illustration, it must be true that $R^n > R^N$.

The stylized illustration makes the point that an event-based population will have a higher concentration of high-risk offenders than appear in the general population of *N* offenders and that concentration does not change with the length of the observation window. To be clear, the proportion of high-risk to low-risk offenders in a release cohort is larger than the proportion of high-risk to low-risk offenders in the general population.

Granted, the use of the Poisson is entirely heuristic, and the comparative concentration of high-risk offenders in n relative to *N* is unclear when accounting for more realistic stochastic processes. However, unless more realistic stochastic processes cause $\lambda_H = \lambda_L$, that is unless more realistic stochastic processes cause low-risk and high-risk offenders to generate the same rate of prison releases or cause low-risk offender rates to exceed high-risk offender rates, the implications of the heuristic Poisson model hold. The accompanying text box illustrates how high rate offenders are more heavily concentrated in release cohorts than in the general population of offenders.

High-Risk Offenders Are Concentrated in Release Cohorts

Using a stylized mathematical model, this Report asserts that high-risk offenders are concentrated in a release cohort. An analysis of NCRP data from Georgia and Illinois, two states with long observation windows, demonstrates the veracity of this assertion.

Georgia has reported prison terms retroactively to 1971, and this example uses all prison terms for age cohorts born from 1965 to 1970. To be included in the population of *N* offenders who have ever been released from a Georgia prison, members of this age cohort must have been released from prison at least once: 49,888 offenders meet this criterion. Table 1 in this box shows the distribution of offenders by percentage who never returned to prison (54.1%), who returned once (20.4%), who returned twice (10.8%) and so on. Those who return zero times are at the lowest risk level and those who return 11 times are at the highest risk level and others are intermediary.

Table 1. Distribution of offenders by risk level among N offenders born 1965–1970 (Georgia)

Risk level	Percentage
0	54.1%
1	20.4%
2	10.8%
3	6.2%
4	4.0%
5	2.2%
6	1.2%
7	0.6%
8	0.3%
9	0.1%
10	0.0%
11	0.0%

What is the distribution of these risk levels (using the same age cohorts) in release cohorts? Table 2 in this box shows the distribution for the *n* members of the 2000–05 release cohorts, where *n* is the number released per year. In the population of *N* offenders (see Table 1), 54.1% were in the lowest risk category. In the population of *n* offenders comprising a release cohort, an average of only 26.1% of the offenders are in the lowest risk category. In Table 1, 6.2% of the *N* offenders are in the risk category "return 3 times;" in Table 2, 12.6% are in this risk category. Likewise, in Table 1, 4.0% of the *N* offenders are in the risk category "return 4 times;" in Table 2, 10.2% are in this risk category.

Table 2. Distribution of n offenders by risk category in six release cohorts based on birth cohorts, 1965–1970 (Georgia)

Number of returns	Release cohort year						Average
	2000	2001	2002	2003	2004	2005	
0	27.1%	23.9%	26.5%	25.6%	25.9%	27.5%	26.1%
1	20.6%	19.9%	20.4%	19.8%	20.6%	21.1%	20.4%
2	15.0%	17.6%	15.4%	16.1%	16.5%	15.6%	16.0%
3	13.2%	12.5%	12.6%	13.3%	12.4%	11.8%	12.6%
4	10.0%	10.5%	10.4%	9.9%	10.1%	10.0%	10.2%
5	6.2%	7.0%	7.0%	6.3%	6.0%	5.7%	6.4%
6	4.1%	4.0%	3.6%	4.7%	4.4%	4.1%	4.2%
7	2.3%	2.6%	2.5%	2.3%	2.1%	2.5%	2.4%
8	0.8%	1.4%	0.8%	1.2%	1.1%	0.8%	1.0%

Number of returns	Release cohort year						Average
	2000	2001	2002	2003	2004	2005	
9	0.6%	0.4%	0.5%	0.5%	0.6%	0.8%	0.6%
10	0.1%	0.2%	0.1%	0.2%	0.1%	0.1%	0.1%
11	0.1%	0.1%	0.1%	0.1%	0.0%	0.1%	0.1%
Offenders	3,245	3,107	3,193	3,269	3,479	3,181	3,246

Clearly, the population of *n* offenders who appear in the typical release cohort has a higher proportion of high-risk offenders than do the *N* offenders who appear in the population of all offenders who were released at some time during observation window. Although percentages will vary across states because states use their prisons in different ways, this concentration of high-risk offenders in a release cohort is not a state-specific phenomenon. Illinois has reported NCRP retroactively to 1989. Analyses of Illinois NCRP data, using the age cohorts born between 1972 and 1977 (because the data contain most of their adult correctional experiences), show that: 47.2% of the *N* members of the 1972–1977 age cohorts (*N*=53,940) are in the risk category "do not return to prison;" 22.0% are in the risk category "return once;" 12.0% are in the risk category "return twice;" and 7.3% are in the category "return more than twice." In comparison, 16.6% of the *n* members of a release cohort (*n*=7,090 averaged across 2000 and 2005) are in the risk category "never return;" 18.1% are in the risk category "return once;" 16.4% are in this risk category "return twice;" and 48.9% are in the risk category "return more than twice." Again, high-risk offenders are more heavily concentrated in the *n* offenders comprising release cohorts.

These calculations show why event-based estimates show higher rates of recidivism than do offender-based estimates. Recidivism may be viewed as a weighted average based on the relative contribution of high, intermediate, and low rate offenders in any given release cohort. As Tables 1 and 2 show, a sample of *n* offenders in a release cohort over represent higher rate offenders relative to the number of offenders observed in the entire observation window. That is why the event-based recidivism calculation is higher than the one based on the offender-based approach.

2.2 Weighting to Derive the Offender-Based Estimates

The methodology for computing offender-based estimates relies on the assertion that if the length of the observation window is long, then low-risk and high-risk offenders have proportional representation in the population of unique offenders collected during that period. The assertion rests on a stylized Poisson model but an empirical demonstration from two states (see the previous text box) showed that the stylized illustration is not misleading. This section explains estimation procedures leading to parameter estimates for the population of *N* offenders and the population of *n* offenders, where the stylized Poisson model is no longer necessary.

Suppose that from 2000 to 2014, *N* unique offenders were released from prison at least once, that the study motivation is to estimate some valid descriptive statistics for these *N* active offenders, and these *N* offenders are representative of the population of offenders who have experienced prison. One simple statistic of interest is the rate at which these *N* active offenders will recidivate on average when released from prison.

For each of these active offenders, it is possible to count the number of times that they were released from prison from 2000 to 2014. Call this *R*. As an approximation, assume that these releases are randomly distributed across the 15-year period and that two releases cannot occur during the same year. Then, the probability that an offender with *N*=1 prison releases during the 15-year period will be released during 2000 equals 1/15. The probability that an offender with *N*=2 prison releases during the 15-year period will be released during 2000 equals 2/15 and so on. Now, think of the 2000 release cohort as being a sample from the *N* active offenders who were released from prison from 2000 to 2014. This notion of the 2000 release cohort (or any other release cohort) as being an approximately random sample from the general population of offenders is key to the estimation methodology.

For every offender who was released in 2000, assign a weight that is inversely proportional to the estimated sampling probability. Given that weight, compute statistics using whatever statistical procedure is desirable. To derive offender-based estimates, this report estimates the weighted proportion of the 2000 release cohort who returned to prison within 1 year, 2 years, 3 years, and so on. This approach is not unique to the year 2000 release cohorts. It is as easy to weight every offender who was released in 2005, although that would have caused a complication eliminating the possibility of following recidivism for the full 15 years. Using survival analysis, the complication is surmountable, but it is simpler to use the 2000 release cohort and to define recidivism as occuring within 15 years.

Applying weights equal to one instead of applying weights that were inversely proportional to the sampling probability produces event-based recidivism statitics. The difference between event- and offender-based statistics is explained by the intent to define the population as “everyone released from prison during 2000” for the event-based approach and “everyone released from prison from 2000 to 2014” for the offender-based approach. Operationally, then, when estimating event-based recidivism statistics every member of the 2000 release cohort receives a weight of one. When estimating offender-based recidivism statistics, every member of the 2000 release cohort receives a weight that is inversely proportional to the sampling probability. When thinking of probability-based samples as the means to estimate the parameters of a clearly defined population, the estimation procedure makes sense. Note again that this approach makes no strong assumptions about the process generating prison terms. Assuming a Poisson process is unnecessary.

When considering the 15-year recidivism described above, users of offender-based estimates should consider them to be approximations for several reasons:

1. As implied by the Poisson illustration, an infinite observation window is necessary to determine the actual number N. However, a finite observation window does not seem to be especially binding because if recidivism has not occurred within 7 years, based on data shown later, it is unlikely to occur. Thus, a long but finite window period is acceptable. Of course, sensitivity testing is practical because users may always artificially restrict the window period to see how estimates change at the margin when the window period is shortened. Though this report does not contain any sensitivity tests, tests conducted outside the scope of this report support the assertion that the bias is small or nonexistent.
2. *N* may be miscalculated because some offenders turn 18 after 2000, so their time-at-risk is not really 15 years. It is easy to deal with this problem by requiring offenders to be at least 18 as of 2000 to be included in the calculations. Selecting a 2005 release cohort as the sample would require a similar

rule that the offender must have been 18 as of 2005. Thus, the method does not require that any specific release cohort serve as the sample. Most studies show that the risk of recidivism decreases with age (Beck & Shipley, 1989; Blumstein A. , 1995; Fox & Zawitz, 2007; Langan & Levin, 2002; Nagin, Cullen, & Jonson, 2009; Hindelang, 1981), so the assumpton that prison releases are randomly distributed over time is wrong. It is therefore likely that prison admissions and releases are more concentrated early in the offenders' criminal careers when the offenders are more recidivistic. If true, the offender-based estimates using the 2000 release cohort would be biased downward. This bias is unlikley to be large because an offender's rate of recidivism changes gradually during a relatively short 15-year period. Moreover, as a sensitivity test, it is possible to alternatively use the 2000 release cohort, the 2005 release cohort, and the 2010 release cohort as the sample. The bias shifts the other way as the sample comes from more recent release cohorts. Though this report does not contain any sensitivity tests, tests conducted outside the scope of this report support the assertion that the bias is small or nonexistent.

3. Assuming that offenders are active (e.g., subject to reincarceration) from 2000 to 2015 does not allow for offenders to cease criminal careers, but practically this is a manifestation of recidivism falling with age.

4. The 2000 release cohort is convenient for making estimates, but as noted, another release cohort could be used. For example, if the 2005 release cohort was used, the weights would not change materially but young offenders cause some complications. Consider an offender who turned 18 in 2005. Assuming that offenders cannot serve adult prison terms until they are 18, it is not possible to observe any meaningful correctional history before 2005. The solution is to prorate such an offender's value of N as:

$$N^* = N\left[\frac{15}{2014-2005}\right]$$

- Offenders aged 19–22 have the number of events prorated similarly. Sensitivity tests conducted outside the scope of this report show that prorating in this way makes little difference to estimates. An advantage of using more recent cohorts is that they are more timely, although the weighting procedure tends to homogenize the estimates because the population of N offenders does not change greatly during the 15 years. A disadvantage of using a more recent release cohort is that the follow-up period is abbreviated and censoring must be taken into account.

The primary factor that complicates offender-based estimation is the introduction of weighting. Otherwise, estimated recidivism is based on tabulation. However, tabulation forces analysts to discard some data. Data are discarded depending on the specific date an offender appearing in the release cohort exits prison. Survival models have been developed to address this issue. The Kaplan-Meier Survival Example (see text box) shows this by comparison a traditional Kaplan-Meier survival estimate to one produced after applying inverse weights to produce an estimate in offender-based form.

Kaplan-Meier Survival Example

The table uses contrived data to demonstrate the Kaplan-Meier (KM) cumulative probability of survival. The top half of the table shows an event-based estimate and the bottom half an offender-based estimate.

Event-based survival

Time to return in days (t)	Number returned (d)	Number not returned at start of day (n)	Number censored	Estimated probability		Cumulative prob. of survival
				Return (d/n)	Survival (1-d/n)	
6	1	23	0	0.04	0.96	0.96
12	1	22	0	0.05	0.95	0.91
21	1	21	0	0.05	0.95	0.87
27	1	20	0	0.05	0.95	0.83
32	1	19	0	0.05	0.95	0.78
39	1	18	0	0.06	0.94	0.74
43	2	17	1	0.12	0.88	0.65
89	1	14	5	0.07	0.93	0.61
261	1	8	0	0.13	0.88	0.53
263	1	7	0	0.14	0.86	0.45
270	1	6	1	0.17	0.83	0.38
311	1	4	3	0.25	0.75	0.28
Total	13		10			

Offender-based survival

Time to return in days (t)	Number returned (d)	Number not returned at start of day (n)	Number censored	Estimated probability		Cumulative prob. of survival
				Return (d/n)	Survival (1-d/n)	
6	1.00	23.00	0	0.04	0.96	0.96
12	1.00	22.00	0	0.05	0.95	0.91
21	0.20	21.80	0	0.01	0.99	0.90
27	1.00	20.80	0	0.05	0.95	0.86
32	0.20	20.60	0	0.01	0.99	0.85
39	1.00	19.60	0	0.05	0.95	0.81
43	2.00	17.60	1	0.11	0.89	0.72
89	1.00	16.60	5	0.06	0.94	0.67
261	0.25	16.35	0	0.02	0.98	0.66
263	1.00	15.35	0	0.07	0.93	0.62
270	0.50	14.85	1	0.03	0.97	0.60
311	1.00	13.85	3	0.07	0.93	0.56
Total	10.15		10			

This example is based on Allison’s (2010) explanation of KM estimators.

The Kaplan-Meier estimator is defined as:

$$\hat{S}(t) = \prod_{j:t_j \le t} \left(1 - \frac{d_j}{n_j}\right)$$

Where t_j are distinct event times such as days. At each time t_j, there are n_j individuals who are at risk to recidivate and d_j is the number of offenders who recidivate at time t_j. People at risk are those who have not recidivated prior to this time, nor have they been censored prior to time t_j. For example, an offender who died would not be at risk to recidivate, and would no longer be part of the at risk pool.

The columns of the table walk through the necessary steps to perform the KM calculation and the last column shows the cumulative survival at each point in time as offenders recidivate and are returned to prison. The column (t) is the number of days up to the time recidivism occurs and (d) is the number recidivating on a particular day. The column (n) is the number of offenders at risk to recidivate and the column labeled "number censored" shows how many offenders are censored on a given day. The remaining columns demonstrate the KM calculation and the last calculation performs the multiplication indicated by the $\prod operator$ which is the cumulative probability of survival over time. The difference between the top and bottom half of this table is that the bottom half shows the application of inverse probability (sampling) weights in the (d) and (n) columns to produce offender-based estimates. The weights are the inverse of the number of offender terms. In the top half of the table, all offenders have a weight of 1.0. In the bottom half of the table there are four offenders who have multiple terms and inverse weights. These occur on days 21, 32, 261, and 270. The respective inverse weights on these days are $\frac{1}{5}$, $\frac{1}{5}$, $\frac{1}{4}$, and $\frac{1}{2}$ indicating that *k* for these four offenders was 5, 5, 4, and 2. All other offenders in the bottom half of the table have a weight of 1.0.

The offender-based estimate of cumulative survival in this contrived example is 0.56 (56% survive), while the event-based estimate is 0.28 (28% survive). Different weights would produce different estimates demonstrating how offender based estimates depend on the ratio between high and low rate offenders.

The next section demonstrates how event- and offender-based estimates differ using actual data from the NCRP. Note that there are no tests of statistical significance on recidivism estimates because the estimates utilize entire populations of offenders rather than samples.

2.3 NCRP Data

The NCRP provides a unique data source for studying returns to state prison. Working with state correctional agencies in all 50 states, BJS assembles offender-based records of adult prison admissions and releases that occurred within most states during calendar years. (The NCRP data exclude juvenile detention facilities.) The records identify offenders for whom the state has custody in public or private prison facilities. BJS reassembles the admissions and releases into prison terms identifying when offenders entered prison and when they were released, unless the offender remained in prison at the end of the reporting period, in which case the term is censored. The NCRP uses unique offender identifiers to merge multiple terms for the same offender into offender-based prison histories. The number of years of prison history data available in the NCRP varies by state, but the most common starting point is 2000. Thus, this is the year selected to serve as the beginning of the observation window's analyses. Methods for assembling data and assuring quality are discussed elsewhere (Luallen, et al., 2014). These prison history profiles are essential for studying criminal recidivism, defined as a readmission to state prison for a term exceeding 30 days. The 30-day rule excludes some ambiguous, short terms that may be for court-ordered clinical observation or for disciplinary responses to minor infractions.

To be included in the analysis, an offender must have been released from prison sometime during the observation window (and might have served multiple terms). All 50 states currently report to the NCRP, but not all report with sufficient detail to link prison terms across admissions for the same offender. Further, some states have gaps in their reporting, so those states' data are not suitable for use in this analysis. Of those states used in the analysis, 16 have an observation window from 2000 to 2014. As of yearend 2014, offenders in these 16 states—Arizona, California, Colorado, Florida, Georgia, Kentucky, Minnesota, Missouri, New York, North Carolina, Oklahoma, South Carolina, Tennessee, Utah, Washington, and Wisconsin—comprise 48% of all offenders in the 50 state prison systems, according to BJS (Carson, 2015). Other states provide data for shorter windows.

3. Event- and Offender-Based Recidivism

The following sections compare event and offender-based recidivism estimates.

3.1 Event-Based Recidivism Estimates by Release Year and Years at Risk

Table 1 shows the recidivism rates for each annual release cohort from 2000 to 2013 for varying numbers of years at risk. This table is based on an event-based estimation procedure. For example, for the 2009 release cohort, the table shows the 1-year (20%), 2-year (31%), 3-year (38%), 4-year (42%), and 5-year (45%) return rates. The return rate is cumulative: 57% of the 2000 release cohort recidivated at some point from 2000 to 2014. The number of states included in the analysis varies by release cohort. Recidivism figures for the 2000 release cohort (leftmost column), for example, include the 16 states with an observation window from 2000 to 2014 in the NCRP. States are weighted equally in Table 1 (i.e., the rates are the arithmetic mean of the rates for each of the states in the calculation).

Appendix Table 1: Return to prison rates by year of release and years at risk, using the event-based approach

Years at risk	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
1	22%	21%	21%	21%	22%	23%	22%	21%	20%	20%	20%	20%	20%	20%
2	33%	34%	33%	34%	35%	36%	35%	33%	32%	31%	32%	31%	31%	x
3	40%	40%	40%	41%	42%	43%	41%	40%	38%	38%	38%	38%	x	x
4	44%	44%	44%	45%	46%	47%	45%	44%	42%	42%	42%	x	x	x
5	47%	47%	47%	48%	49%	50%	48%	47%	45%	45%	x	x	x	x
6	50%	50%	50%	50%	51%	52%	50%	49%	47%	x	x	x	x	x
7	51%	51%	51%	52%	52%	53%	52%	50%	x	x	x	x	x	x
8	53%	53%	53%	53%	54%	54%	53%	x	x	x	x	x	x	x
9	54%	54%	53%	54%	55%	55%	x	x	x	x	x	x	x	x
10	55%	54%	54%	55%	55%	x	x	x	x	x	x	x	x	x
11	55%	55%	55%	56%	x	x	x	x	x	x	x	x	x	x
12	56%	55%	55%	x	x	x	x	x	x	x	x	x	x	x
13	56%	56%	x	x	x	x	x	x	x	x	x	x	x	x
14	57%	x	x	x	x	x	x	x	x	x	x	x	x	x
# of states	16	17	19	19	21	22	25	26	27	28	30	32	33	33

An advantage of this table is that the reader may examine the extent to which a cumulative recidivism period of, for example, 3 years, varies by release cohort. Table 1 shows that the 3-year recidivism percentage varies from 38% to 43% among the 2000 to 2011 release cohorts. Although the changing number of states seems to obscure the comparison, the recidivism rates for the post-2000 release cohort years (not shown) are fairly insensitive to the expansion of states beyond the original 16. Future publications will provide greater detail, including statistics that hold the number of states constant across the years.

Table 1 also highlights an advantage of using the NCRP to estimate recidivism. NCRP is an ongoing BJS data collection, allowing for estimates for each release cohort and for varying numbers of years at-risk. With each successive year of data collection, the observation window for each release cohort expands by one year. Currently, 2016 data are being collected. States have also been submitting older data to eliminate reporting gaps, so the number of state with long reporting periods is expanding.

3.2 Event-Based Recidivism Estimates for the 2000 Release Cohort

Table 1 masks state-to-state variation in recidivism rates. Figure 2, in comparison, shows that variation in a series of box plots (i.e., box and whisker diagram) derived from the event-based estimation procedure and based on prison returns for the 2000 release cohort. The boxplots show the lowest state recidivism rate (bottom whisker); the 25th (bottom of the box), 50th (middle of the box), and 75th percentile rates (top of the box); the highest rate (top whisker); and the average (dashed line) across all 16 states for the 2000 release year cohort. As with Table 1, the average assigns every state's population a weight of one rather than a weight proportional to state size. Typical of long term recidivism results, the earlier years show the largest average growth in recidivism and the latter years show less growth. That is, the dashed line representing the average rate of recidivism increases at a decreasing rate.

Appendix Figure 2: Return to prison rate for the 2000 release cohort, using the event-based approach

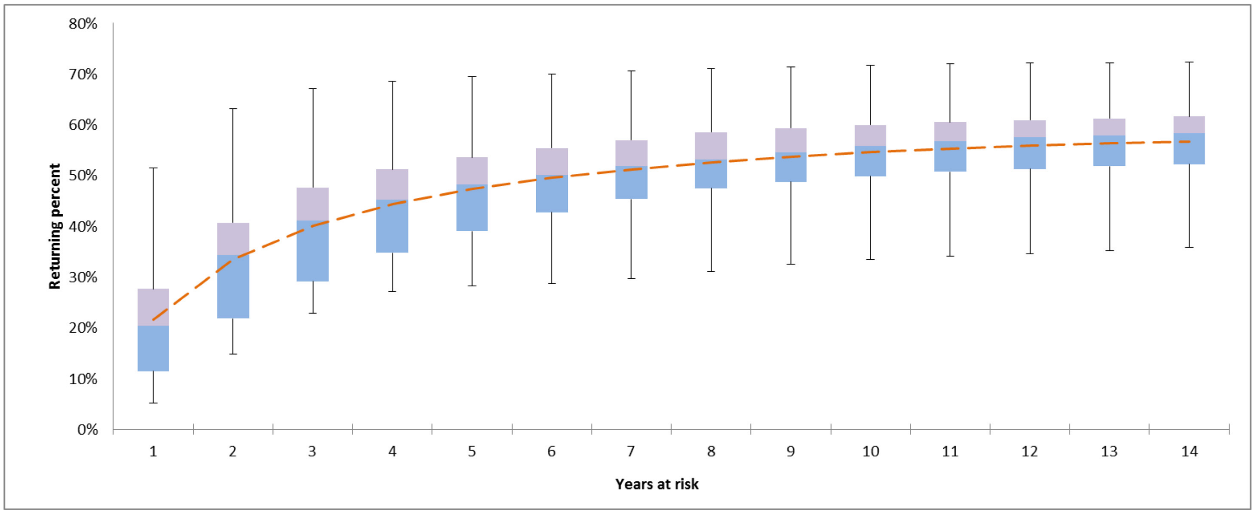


Figure 2 shows considerable variation across the 16 states—fewer than 11% of offenders in the 2000 release cohort returned to prison within 1 year in a quarter of the states and more than 28% of offenders returned to prison within 1 year in another a quarter of the states. Averaged across the 16 states, 22% of offenders returned to prison within 1 year. Some states may in fact have lower recidivism rates than do other states, but there are at least two other possible explanations. First, the NCRP currently collects prison admission and release data from state correctional agencies. Local and county jails do not currently participate in the NCRP, unless the state has a unified prison/jail system. This results in widely varying offender profiles for states in the NCRP, because state policies vary widely on which offenders serve time in state prisons and which serve time in local and county jails. At one extreme are the six states (Alaska, Connecticut, Delaware, Hawaii, Rhode Island, and Vermont) that have unified prison/ jail systems. In these states, NCRP receives data on all offenders, regardless of sentence length. In comparison, in Massachusetts, offenders who receive a minimum sentence of 2.5 years serve time in state prisons (and hence, appear in the NCRP), while offenders with shorter sentences serve time in county jails. Many states have a 1-year cut off between a jail and prison sentence, the traditional distinction between a misdemeanor and felony. As a result, offender profiles differ and the level of criminality required to recidivate (i.e., reappear in the NCRP) varies by state. However, even if a group of states has the same prison/jail cut off, those states may have different compositions of offenders due to varied state sentencing policies by type of crime, so these offenders will, have different risks of recidivating.

A second reason recidivism rates vary by state are differences in policies regarding post-confinement community supervision. A state that mandates community supervision following a prison sentence for all released offenders will have a higher recidivism rate than a state where few offenders receive community supervision following release from prison. States also vary regarding revocation practices: some states are reluctant to revoke offenders for technical violations while other states consider revocations to be good correctional practice.

Whatever the reasons for the intra-state differences, they narrow over time (see Figure 2). For offenders who were potentially at risk for 14 years, a maximum of 52% returned to prison during the 14-year period in a quarter of the states. A minimum of 62% returned to prison during a 12-year period in another quarter of the states. Overall, 57% of offenders across 16 states had returned to prison within 14 years using the event-based approach.

3.3 Offender-Based Recidivism Estimates for the 2000 Release Cohort

Figure 3 has the same structure as Figure 2, but the statistics are derived using the offender-based estimation approach. Figures 2 and 3 highlight how the event- and offender-based approaches yield different estimates of recidivism. Whereas 57% of the population of n offenders released from prison in 2000 return to prison within 14 years using the event-based approach (Figure 2), an estimated 35% of the *N* offenders return to prison within 14 years using the offender-based approach (Figure 3). That is, after weighting to compute offender-based estimates, two of every three members of the 2000 release cohort did not return to prison. Given the shape of the “survival curve” it appears that nationally, the percentage returning to prison is unlikely to greatly exceed a third even if the time-series could be extended. This suggests that two of every three offenders desist from crime of sufficient seriousness to return them to prison.

Appendix Figure 3: Return to prison rate for the 2000 release cohort, using the offender-based approach

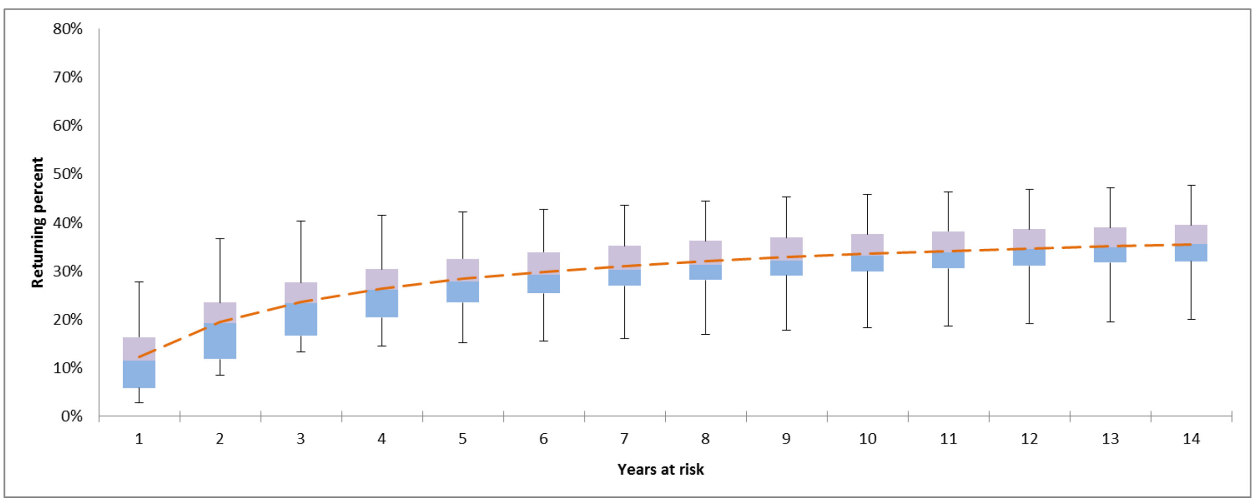
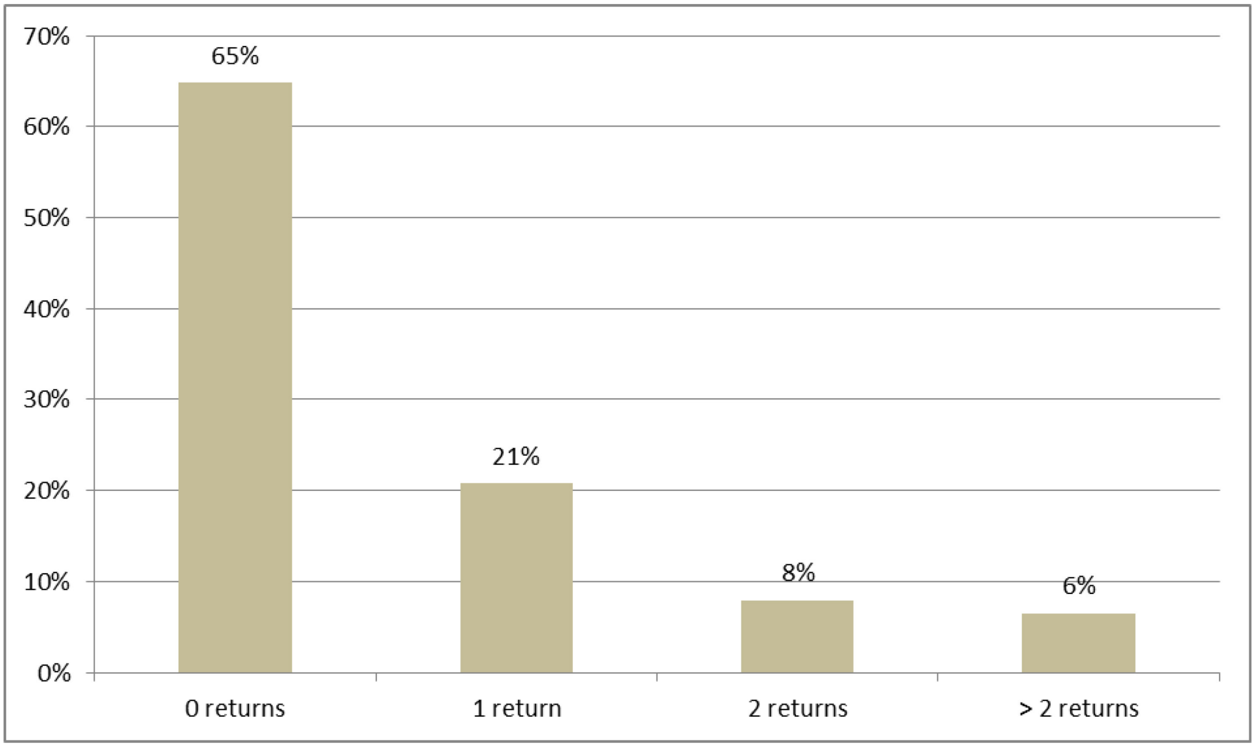


Figure 3 also shows that, after applying the offender-based weighting to the 2000 release cohort, for a quarter of the states, 6% or fewer of offenders returned to prison within 1 year. For another a quarter of the states, 16% or more of offenders returned to prison within 1 year. The differences across states narrow over time so that after 14 years, 32% or fewer of offenders returned to prison within 14 years for a quarter of the states. For another a quarter of the states, 39% or more of offenders returned to prison within 14 years.

3.4 Number of Returns to Prison

In addition to determining whether an offender returned to prison, NCRP data may be used to determine the number of times an offender returned to prison. Considering the *N* offenders who were released from prison sometime from 2000 to 2014, the offender-based procedure may estimate the proportion who never returned to prison, who returned once, who returned twice and so on. Those proportions are shown in Figure 4.

Appendix Figure 4: Number of returns to prison, for the offender-based approach (offenders released from 2000 to 2014)



3.5 Other Possible Analyses

Other refinements are possible, and Table 1 and Figures 2, 3, and 4 could be replicated to distinguish across sex, race, age groups, or offense types or include interactions between race, Hispanic origin, and sex. Eventually, the breakdowns become too complex to represent with tables and figures, and the analyses would have to switch to multivariate models such as survival models. The statistics demonstrated here give equal weight to every state reporting to the NCRP but an alternative approach would give equal weight to offenders or events (e.g., number of releases in a given year). This approach would give much greater weight to states such as California with large inmate populations.

Finally, the presented tables vary the number of states used to calculate the recidivism percentages. The 2000 release cohort uses 16 states. More recent release cohorts involve as many as 33 states (see Table 1). An alternative would be to report statistics for a consistently defined set of states and for some purposes that may be the preferred approach. Specifically, this would be the preferred approach if the interest was to study how recidivism rates change over time.

4. Conclusions

Though surveying the entire recidivism literature was not practical for the purposes of this report, the great majority of identified studies of recidivism have used event-based estimation procedures. In many contexts, these have been used to draw policy implications that might logically be based on offender-based estimation procedures instead. The following passage demonstrates how one publication portrays their calculations.

> "…45.4 percent of people released from prison in 1999 and 43.3 percent of those sent home in 2004 were reincarcerated within 3 years, either for committing a new crime or for violating conditions governing their release." (Pew Center on the States, 2011, p. 2)

These estimates are not wrong and, in fact, they are close to the estimates derived from the NCRP using the event-based estimation procedure.[1]

However, taken at face value, event-based estimates of recidivism define what the public may think of as corrections failures to be on the order of 45–55% during a 3- to 5-year period following release from prison. As was already discussed, these statistics are higher than the offender-based estimates because there is a higher proportion of high-risk offenders in yearly release cohorts. The offender-based estimate would tell a more salutary story of correctional success than the "revolving door" metaphor implies. An analyst may, and probably should, report estimates both ways.

This report demonstrates differences between event- and offender-based estimates of recidivism. The contrast should cause scholars and practitioners to rethink which counting rule is appropriate for their particular criminal justice application. Widely cited reports of recidivism may unintentionally promote this misunderstanding.

In fairness, the offender-based estimates may also be misused. For example, many criminal justice reforms focus on revocations as explaining mass incarceration. Using the offender-based approach, revocations would seem to be a minor problem, because if most offenders avoid returning to prisons, even more must avoid revocations. Of course, this conclusion misses the point when revocation policies and practices pertain to specific release cohorts whose members either fail or succeed. A policy analyst should use the event-based approach to understand revocation policy and practice.

The event- and offender-based methodologies are not competing; there is no best methodology, per se. Rather, the point is that the research question should guide the methodology and choosing the wrong methodology for a given research question will lead to misleading answers. The following provides guidance on the appropriateness of the offender- and event-based approaches:

---

[1]The estimates in this report give every state a weight of 1. The Pew study gives each state a weight that is proportional to the prison release population from that state. California drives estimates that are proportional to prison release populations so the estimates cited in these passages are higher than the cited estimates. Repeating this report's estimates using weights that are proportional to the size of the release populations causes the estimates to conform.

- If the practitioner wants to understand the number of offenders who are cycling through their system, the offender-based estimate is more appropriate. Legislators, prison administrators and other executives often want to track the "performance" of their prison system over time. By performance, they want to know if the prison system has some impact on offender recidivism. Offender-based recidivism is more appropriate for this inquiry.

- If the practitioner wants to understand the risk posed by members of a release cohort, then he or she should use the event-based estimate. Practitioners who want to allocate rehabilitation and supervision resources for a release cohort should use event-based estimates of recidivism.

- If a trend in recidivism is the main goal of the inquiry, then the practitioner will have to ask "Am I interested in the trend in offenders over time (offender-based), or am I interested in the distribution of risk of release cohorts over time (event-based)?"

This guidance does not exhaust the different usages, but it should provide insight on when each approach is most appropriate in answering policy questions.

4.1 Limitations

There are some limitations to this study. This study primarily uses data from 16 states, but initial analysis so far shows that the inclusion of more states, whose data are available for shorter observation windows, does not change the story much. The NCRP continues to assemble data, including requests for data reported retrospectively for those states that do not report for all years from 2000 to 2014, so current limitations will be reduced. Over time, with the expansion of observation windows by adding additional years of data as they occur and by including historic data from past years, it is likely that the NCRP will become an increasingly important platform for examining recidivism.

Another limitation is that offenders sometimes cross state lines and are arrested in other states where they may be incarcerated. Because NCRP records are not currently linked across states, reincarceration that occurs outside the state used in the calculations will not be included in the statistics reported above. Although that is a problem causing a downward bias in recidivism estimates, the bias is probably small. Only 10.9% of rearrests that occur within five years of release occur in another state (Durose, Cooper, & Snyder, 2014), and only 28.2% of arrests result in a prison stay (Durose, Cooper, & Snyder, 2014). Moreover, if an arrest results in a revocation, the revocation would likely result in the offender's return to the state from which he or she was originally released, so the reincarceration would appear in the NCRP. Reinforcing the conclusion that the bias is small, event-based statistics reported in this study are about the same as event-based statistics from a BJS study (Durose, Cooper, & Snyder, 2014), which accounted for cross-jurisdictional recidivism.

Another bias is that short observation windows bias statistics toward overrepresentation of high-risk offenders even when using the offender-based methodology, but this bias does not appear to be severe. Table 1 and Figure 2 show that most recidivism will occur within 5 to 6 years, so a window as short as 5 or 6 years is probably adequate to derive decent estimates. Table 1 also suggests that windows even shorter than 5 years do not seriously distort findings. It might be advantageous to be cautious about inferences for the youngest offenders, especially those released toward the end of the observation window, for whom the window is truncated, because they have not had the opportunity to acquire extensive adult criminal records. Any resultant bias will move the offender-based estimates closer to the event-based estimates.

These limitations do not affect the main conclusions of this paper. Offender-based estimates of recidivism are rarely estimated. Furthermore, both practitioners and scholars must carefully consider when it is appropriate to use offender- or event-based recidivism in their public policy deliberations.

5. References

Allison, P. (2010). *Survival Analysis Using SAS: A Practical Guide, 2nd Edition.* Cary, North Carolina: SAS Institute, Inc.

Andrews, D. A., Bonta, J., & Hoge, R. (1990). Classification for effective rehabilitation. *Criminal Justice and Behavior*, 17,19-52.

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does Correctional Treatment Work? A Clinically Relevant and Psychologically Informed Meta-analysis. *Criminology*, 369-404.

Aos, S., Phipps, P., Barnoski, R., & Lieb, R. (2001). *The Comparative Costs and Benefits of Programs to Reduce Crime, Version 4.0.* Olympia, WA: Washington State Institute for Public Policy.

Beck, A. J., & Shipley, B. E. (1989). *Recidivism of Prisoners Released in 1983.* Washington, DC: Bureau of Justice Statistics.

Berk, R. (2012). *Criminal Justice Forecasts of Risk: A Machine Learning Approach.* New York: Springer.

Bhati, A., & Piquero, A. (2008). Estimating the Impact of Incarceration on Subsequent Offending Trajectories: Deterrent, Criminogenic, or Null Effects? *The Journal of Criminal Law and Criminology Vol. 98, no. 1*, 207-253.

Blumstein, A. (1995). Youth Violence, Guns, and the Illicit-Drug Industry. *Journal of Criminal Law and Criminology*, 86(1): 10-36.

Blumstein, A., Cohen, J., Roth, J., & and Visher, C. E. (1986). *Criminal Careers and Career Criminals.* Washington, DC: National Academy Press.

Bonczar, T., & Beck, A. (1997). *Lifetime Likelihood of Going to State or Federal Prison.* Washington, DC: Bureau of Justice Statistics Special Report.

Bonta, J., Law, M., & Hanson, K. (1998). The Prediction of Criminal and Violent Recidivism among Mentally Disoriented Offenders: A Meta-Analysis. *Psychological Bulletin, vol 123 no 2*, 123-142.

Brame, R., Bushway, S. D., & Paternoster, R. (2003). Examining the prevaslence of criminal desistance. *Criminology*, 41(2), 423-448.

Bushway, S., Piquero, A., Broidy, L., Cauffmann, E., & Mazerolle, P. (2001). An empirical framework for studying desistance as a process. *Criminology*, 39, 491-513.

Carson, E. A. (2015). *Prisoners in 2014.* Washington, DC: Bureau of Justice Statistics.

Cosslett, S. (1993). Estimation from Endogenously Stratified Samples. In G. Maddala, C. Rao, & H. Vinod, *Handbook of Statistics, vol. 11.* Elsevier Science Publishing.

Cottle, C., Lee, R., & Heilbrun, K. (2001). The Prediction of Criminal Recidivism in Juveniles: A Meta-Analysis. *Criminal Justice and Behavior, vol 28 no. 3*, 367-394.

DeFina, R., & Hannon, L. (2009). The Impact of Mass Incarceration on Poverty. *Crime & Delinquency*.

Durose, M., Cooper, A., & Snyder, H. (2014). *Recidivism of Prisoners Released in 30 States in 2005: Patterns from 2005 to 2010.* Washington, DC: Bureau of Justice Statistics.

Englin, J., & Shonkwiler, J. (1995). Estimating Social Welfare using Count Data Models: An Application to Long-Run Recreational Demand under Conditions of Endogenous Stratification and Truncation. *The Review of Ecoonomics and Statistics*, 104-112.

Fox, J. A., & Zawitz, M. W. (2007). *Homicide Trends in the United States.* Washington, DC: Bureau of Justice Statistics.

Gendreau, P., Little, T., & Goggin, C. (1996). The Meta-Analysis of the Predictors of Adult Offender Recidivism: What Works? *Criminology vol. 34*, 575-607.

Gottfredson, S., & Moriarty, L. (2006). Statistical Risk Assessment: Old Problems and New Applications. *Crime and Delinquency, 52*, 178-200.

Hanson, R., & Bussiere, M. (1998). Predicting Relapse: A Meta-Analysis of Sexual Offender Recidivism Studies. *Journal of Consulting and Clinical Psychology, vol(66) no 2*, 348-366.

Hindelang, M. J. (1981). Variations in Sex-Race-Age Incidence Rates of Offending. *American Sociological Review*, 46(4): 461-74.

Hunt, K. S., & Dumville, R. (2016). *Recidivism Among Federal Offienders: A Comprehensive Review.* Washington, DC: United States Sentencing Commission.

Killettaz, P., Villattaz, P., & Zoder, I. (2006). *The Effects of Custodial vs. Non-Custodial Sentences on Re-Offending: A Systematic Review of the State of Knowledge.* The Campbell Collaboration.

Kling, J. (2006). Incarceration Length, Employment, and Earnings. *The American Economic Review vol. 96, No. 3*, 863-876.

Langan, P., & Levin, D. (2002). *Recidivism of Prisoners Released in 1994.* Washington, DC: Bureau of Justice Statistics.

Laub, J. H., & Sampson, R. J. (2001). Understanding desistance from crime. In M. Tonry, *Crime and Justice: A Review of Research Vol. 28* (pp. 1-69). Chicago: University of Chicago Press.

Laub, J. H., & Sampson, R. J. (2006). *Shared Beginnings, Divergent Lives: Delinquent Boys to Age 70.* Cambridge, MA: Harvard University Press.

LaVigne, N., Bieler, S., Cramer, L., Ho, H., Kotonias, C., Mayer, D., . . . Samuels, J. (2014). *Justice Reinvestment Initiative State Assessment Report.* Washington, DC: Urban Institute.

Liberman, A. (2008). Synthesizing longitudinal findings. In A. Liberman, *The Long View of Crime: A Synthesis of Longitudinal Research* (pp. 3-22). New York: Springer.

Listokin, Y. (2003). Does More Crime mean More Prisoners? An Instrumental Variables Approach. *Journal of Law and Economics*.

Lowenkamp, C. T., Latessa, E. J., & Holsinger, A. M. (2006). The risk principle in action: What have we learned from 13,676 offrnders and 97 correctional programs? *Crime and Delinquency*, 52(1):77-93.

Luallen, J., Neary, K., Kling, R., Rhodes, B., Gaes, G., & Rich, T. (2014). *A Description of Computing Code Used to Identify Correctional Terms and Histories.* Abt Associates Inc. NCRP White Paper #3.

MacKenzie, D. L. (2006). *What Works in Corrections? Reducing the Criminal Activities of Offenders and Delinquents.* Cambridge, UK: Cambridge Press.

Maltz, M. D. (1984, 2001). *Recidivism.* Orlando, FL: Academic Press.

Manski, C., & Lerman, S. (1977). The Estimation of Choice Probabilities from Choice Based Sample. *Econometrica, vol 45 (8)*, 1977-1988.

Maruna, S. (2001). *Making Good: How Ex-Convicts Reform and Rebuild Their Lives.* Washington, DC: American Psychological Association.

Nagin, D. (2001). Test. *Archives of General Psychiatry*, 389-394.

Nagin, D. (2005). *Group-Based Modeling of Development.* Cambridge, MA: Harvard University Press.

Nagin, D., & Land, K. (1993). Age, criminal careers, and population heterogeneity -- specification and estimation of a non-parametric mixed Poisson model. *Criminology*, 31, 327-362.

Nagin, D., & Tremblay, R. (1999). Trajectories of Boys' Physical Aggression Opposition, and Hyperactivity on the Path to Physically Violent and Non-Violent Juvenile Delinquency. *Child Development*, 1181-1196.

Nagin, D., & Tremblay, R. (2001). Parental and Early Childhood Predictors of Persistent Physical Aggression in Boys from Kindergarten to High School. *Archives of General Psychiatry*, 389-394.

Nagin, D., Cullen, F., & Jonson, C. (2009). Imprisonment and Reoffending. *Crime and Justice: A Review of Research (M. Tonry, ed)*, Vol 38, pp.115-200.

National Research Council. (2008). *Parole, Desistance from Crime, and Community Integration.* Washington, DC: National Academies Press.

Pettit, B., & Western, B. (2004). Mass Imprisonment ad the Life Course: Race and Class Inequality in U.S. Incarceration. *American Sociological Review, Vol. 26*, 151-165.

Pew Center on the States. (2011). *State of Recidivism: The Revolving Door of America's Prisons.* Washington, DC: The Pew Charitable Trusts.

Piquero, A. (2008). Taking stock of developmental trajectories of criminal activity over the life course. In A. Liberman, *The Long View of Crime: A Synthesis of Longitudinal Research* (pp. 23-78). New York: Springer.

Piquero, A., Farrington, D., & Blumstein, A. (2003). The Criminal Career Paradigm. *Crime and Justice*, 359-506.

Rhodes, E., Gaes, G., Luallen, J., Kling, R., Rich, T., & Shively, M. (2014). Following Incarceration, Most Released Offenders Never Return to Prison. *Crime & Delinquency*.

Rhodes, W., Kling, R., & Johnson, P. (2007). Using Booking Data to Model Drug Use Arrest Rates: A Preliminary to Estimating the Prevalence of Chronic Drug Users. *Journal of Quantitative Criminology, vol 23 no 1*.

Sampson, R., & Laub, J. (1993). *Crime in the Making: Pathways and Turning Points Through Life.* Cambridge, MA: Harvard University Press.

Spelman, W. (2008). Specifying the Relatinship between Crime and Prisons. *Journal of Quantitative Criminology, vol. 24*, 149-178.

Western, B., Kleykamp, M., & Rosenfeld, S. (2004). Crime, Punishment, and American Inequality. In K. Necherman, *Social Inequality.* New York: Russel Sage Foundation.

Wheelock, D., & Uggen, C. (2010). Race, Poverty and Punishment: The Impact of Criminal Sanctions on Racial, Ethinc, and Socioeconomic Inequality. In A. Lin, & D. Harris, *The Color of Poverty: Why Racial amd Ethnic Disparity Persists* (pp. 261-292). New York: Russell Sage Foundation.

Wildeman, C. (2009). Parental Imprisonment, the Prison Boom, and the Concentration of Chidhood Disadvantage. *Demography, vol. 46, no. 2*, 265-280.

6. Appendix Tables and Figures

Appendix Table 1: Return to prison rates by year of release and years at risk, using the event-based approach

Years at riisk	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
1	22%	21%	21%	21%	22%	23%	22%	21%	20%	20%	20%	20%	20%	20%
2	33%	34%	33%	34%	35%	36%	35%	33%	32%	31%	32%	31%	31%	x
3	40%	40%	40%	41%	42%	43%	41%	40%	38%	38%	38%	38%	x	x
4	44%	44%	44%	45%	46%	47%	45%	44%	42%	42%	42%	x	x	x
5	47%	47%	47%	48%	49%	50%	48%	47%	45%	45%	x	x	x	x
6	50%	50%	50%	50%	51%	52%	50%	49%	47%	x	x	x	x	x
7	51%	51%	51%	52%	52%	53%	52%	50%	x	x	x	x	x	x
8	53%	53%	53%	53%	54%	54%	53%	x	x	x	x	x	x	x
9	54%	54%	53%	54%	55%	55%	x	x	x	x	x	x	x	x
10	55%	54%	54%	55%	55%	x	x	x	x	x	x	x	x	x
11	55%	55%	55%	56%	x	x	x	x	x	x	x	x	x	x
12	56%	55%	55%	x	x	x	x	x	x	x	x	x	x	x
13	56%	56%	x	x	x	x	x	x	x	x	x	x	x	x
14	57%	x	x	x	x	x	x	x	x	x	x	x	x	x
# of states	16	17	19	19	21	22	25	26	27	28	30	32	33	33