Providing Another Chance

Resetting Recidivism Risk in Criminal Background Checks

SHAWN D. BUSHWAY, BRIAN G. VEGETABLE, NIDHI KALRA, LEE REMI, GREG BAUMANN

Sponsored by Arnold Ventures
Criminal background checks are regularly conducted in the United States to screen candidates for employment, volunteer, and housing opportunities. Individuals with past convictions are routinely subject to evaluations amid concerns that, because of their criminal history or concerns about recidivism risk, they pose a risk of creating avoidable costs. Current methods of assessing recidivism risk estimate risk at the time of a person’s last conviction. As a result, they do not fully reflect important information about the time since that conviction, during which the individual has lived in the community without a new conviction. Given that recidivism risk declines the longer a person lives in a community without a new conviction, risk assessments that omit this information may make people look riskier than they are. As a result, people may be denied opportunities for which they might otherwise be qualified.

This report describes a principle for developing recidivism risk models that are anchored at the time of a criminal background check. We refer to this as the reset principle. Risk-prediction models that use this principle reset the assessment of risk to account for the amount of time a person spends in the community without a new conviction. Although the report focuses on employment background checks, the reset principle also can be applied in other settings, such as in the vetting of volunteers or potential renters.

We demonstrate the viability of a particular version of a recidivism risk model in this report using an extensive criminal history data set from North Carolina. In the process, we add to the nascent literature demonstrating that most people with a conviction do not have a subsequent one and that their risk of recidivating declines dramatically over the period from their last major interaction with the criminal justice system. We also offer considerations for developing recidivism risk models that conform with best practices to estimate recidivism risk in criminal background checks. This report was shaped by a RAND Corporation workshop in June 2021.

This report should interest policy- and decisionmakers who want to develop more-accurate risk prediction models than what are currently available. Researchers interested in recidivism risk prediction can use this report as a starting point to develop models and tools that more accurately predict recidivism risk.

This research was sponsored by Arnold Ventures through Grant No. 20-03131.

Justice Policy Program

RAND Social and Economic Well-Being is a division of the RAND Corporation that seeks to actively improve the health and social and economic well-being of populations and communities throughout the world. This research was conducted in the Justice Policy Program within RAND Social and Economic Well-Being. The program focuses on such topics as access to justice,
policing, corrections, drug policy, and court system reform, as well as other policy concerns pertaining to public safety and criminal and civil justice. For more information, email justicepolicy@rand.org.

Acknowledgments

We acknowledge the support of the team at Arnold Ventures, in particular Jocelyn Fontaine, for seeing the promise in the original idea and working with us to bring the idea to life. We also acknowledge the benefit of conversations with David Roberts and his team at SEARCH, which has been a partner on the larger project about the strengths and weaknesses of various kinds of criminal history record searches. We thank reviewers Greg Ridgeway, Matt Cefalu, and James Anderson for suggestions that greatly improved the work. We are enormously grateful to all the workshop speakers and attendees for their unique insights and perspectives; their insights have shaped how we think about and communicate about this work.
Criminal background checks are commonly used in the United States to screen people who are applying for jobs, loans, housing, volunteer activities, and other opportunities. The evaluations operate on a broadly accepted assumption that past behavior is a good predictor of future behavior. They also assume that people who have committed crimes in the past may do so again and may create costs for the other party. For those with criminal histories—about 30 percent of all adults in the United States (according to the 1997 National Longitudinal Survey of Youth) (Brame et al., 2014)—background checks can limit life opportunities.

Contrary to popular belief, however, most people who enter the criminal justice system ultimately desist from crime (Rhodes et al., 2014; Sampson and Laub, 2003). The risk of recidivism declines the longer a person is in the community and does not commit a crime. Eventually, a past criminal record is no longer predictive of future convictions.

However, most background checks do not adequately take into account the amount of time someone with a conviction spends in the community without a new conviction. The evaluations miss information about a person’s propensity (or lack thereof) to reoffend, which may be revealed by examining their time in the community. Not having this information may skew the risk assessment to make people look riskier than they are, resulting in denied opportunities. Lost opportunities impose costs on people with convictions (e.g., lost wages), employers (e.g., not having the most suitable candidate), and society as a whole (e.g., underemployment). The predictive shortcomings of current recidivism risk assessments also contribute to the use of set-time exclusions, which keep people convicted of certain crimes from getting a job for a defined period, regardless of the actual level of recidivism risk.

The Reset Principle

Background checks are designed, in part, to identify people who might recidivate. We describe how background checks can be improved to include information about time spent in the community, thereby more accurately reflecting risk of recidivism. In particular, we suggest that any viable risk-assessment instrument used in background checks should reset the assessment of recidivism risk to the time of the background check and not the time of conviction, as current methods derived from the criminal justice context do. We call this the reset principle. If models are developed into tools that employers and others can use to assess recidivism risk, they may offer a more accurate way to distinguish candidates’ risk of recidivism. Thus, they may offer many with criminal histories a way to demonstrate that they should be offered another chance.

This reset principle is easy to explain but difficult to implement in decision tools for background checks. We demonstrate one approach to meeting the reset principle with a risk-
prediction model that satisfies the principle using survival analysis (a branch of statistics) or, more generally, time-to-event analyses. In this report, we develop a model for the employment context using state-of-the-art statistical and machine-learning techniques applied to a large data set. The data set comes from the North Carolina Department of Public Safety (NCDPS) Offender Public Information Search/Inmate Locator and represents public information related to convictions, sentencing events, and prison incarceration spells dating back to 1972. As of April 2021, the data set contained more than 4 million sentencing records, representing more than 1 million individuals with unique IDs in the database.

Considerations for Creating Models to Satisfy the Reset Principle

We created a model that is statistically validated and well calibrated for North Carolina data. In doing so, we articulated five considerations that should guide the creation of models that satisfy the reset principle:

1. Accurate estimation requires properly defining the relevant population for prediction—in this case, people with criminal histories and not currently incarcerated.
2. Estimation of the model should be performed using data based on convictions, not arrests.
3. Data sets should have at least ten years of conviction records to accurately estimate recidivism risk over extended periods.
4. Model estimates should be calibrated and reflect true rates of recidivating in the population.
5. Model estimates must be validated on an independent test sample.

These considerations represent good research practices for creating a useful model. This report contributes to a discussion of models in the specific policy context of a criminal background check. Additionally, although our model uses survival functions to estimate risk, other methodologies also may be used to implement the reset principle.

Observations About Recidivism Risk from the North Carolina Data

We made the following observations about recidivism risk based on the North Carolina data:

- In the process of developing the model described in this report, we applied a statistical method called Kaplan-Meier estimation to the North Carolina data and made several observations that substantiate past research findings: The majority of individuals with a conviction do not have a subsequent conviction.
- A person’s likelihood of reoffending declines rapidly as more time passes without a conviction.
- After a sufficient period without a new conviction, even people initially deemed to be at highest risk for reoffending (such as those with a more extensive criminal background) transition to risk levels that appear similar to those initially at the lowest risk.
Implications for Employers, Policymakers, and Researchers

As part of this research project, we sponsored a workshop to gather feedback from stakeholders involved in background checks for employers. This workshop helped identify the implications of our work so they can be considered by people who are interested in advancing the reset principle in practice.

Implication: The Majority of People Who Have Been Convicted Do Not Get Reconvicted

The North Carolina data confirm what Rhodes and colleagues showed in a sample of people who have been sent to and released from prison: Most people who have had that experience do not recidivate (Rhodes et al., 2014). This observation adds new weight to previous research about the subject and should move practitioners away from the oft-cited view, based on misunderstandings of Bureau of Justice Statistics recidivism analyses of prison release cohorts, that (nearly) everyone who has ever been released from prison recidivates.

Implication: Models Can Be Used to Predict Risk of a New Conviction

U.S. Equal Employment Opportunity Commission (EEOC) guidelines allow for validated employment screenings that give employers a more certain defense to challenges to their employment decisions. The EEOC has noted, however, that research and evidence supporting predictive models that qualify for validation did not exist when the guidelines were issued. This report demonstrates that it may be possible to create a validated model under the Uniform Guidelines on Employee Selection Procedures. Model calibration and validation are well-developed methods in statistics. There is no need to reinvent the wheel, but the Uniform Guidelines do need to be formally updated to recognize this new class of models (Maurer, 2020). Unlike other current decision rules, decision rules based on models that satisfy the reset principle could allow employers to consider a candidate’s level of risk in relation to a threshold the employer defines as acceptable.

Implication: Risk Prediction Models Do Not Define an Absolute Threshold for Acceptable Risk; Setting Acceptable Risk Thresholds Requires Policy Discussion and Decisions

The models presented in this report allow for the comparison of the recidivism risk of different individuals. However, such models do not assess whether any individual’s risk is low enough for a particular job. That threshold must be defined by decisionmakers based on specific circumstances. The threshold decisions drive the final evaluation of equity and fairness. The tools simply rank order people in terms of risk—the thresholds determine who actually get hired. Evaluating the hiring threshold also helps anchor the decision relative to current practices—an important feature of any evaluation.
Implication: Data Quality Can Limit the Development of Successful Recidivism Risk Models; Data Infrastructure Should Be a Primary Concern for Equitable Application of Tools Drawing from These Models

Currently, decision tools for background checks are based on expert assessments of risk and do not rely directly on models estimated with primary data sources. As a result, there is no data infrastructure designed to predict the risk to recidivate. Policymakers should be aware that cleaning data from secondary data sources to create new models is a complex task. Furthermore, even high-quality data sets, such as the NCDPS data, may be missing important covariates for predicting risk of a future conviction (NCDPS data, for example, do not include incarceration in county jails). Policymakers should consider the costs and benefits of creating purpose-built data sets for assessing recidivism risk.

Implication: Prediction-Based Models Are Associational, Not Causal; Reflect Historical Trends on Recidivism to Predict Future Outcomes; and Are Grounded in a Specific Place and Time

Prediction models that work in a specific context are not guaranteed to work in other contexts. Therefore, it is bad practice to take a model built using the North Carolina data (or other data, for that matter) and simply use it other settings. The model must be validated in these new settings. This is standard practice in the criminal justice system (Slobogin, 2021). We expect that validation might look different in the background-check context. Practitioners should be aware that there is no one-size-fits-all solution. However, it is possible that one tool might work well in various settings and that the necessary modifications to that tool will be minor. This is particularly true if the model represents a significant improvement over current practice. The key is that this application in other settings must be done carefully and ultimately must be validated.

Implication: The Use of a Model in Employment or Other Contexts Could Alter Its Performance

New models (reset or otherwise) predict outcomes in a system that they currently do not influence. However, if the model we developed, based on the North Carolina data, were to be developed into a tool used to inform hiring in North Carolina, the predictive ability of the model could change (Bushway and Smith, 2007; Kleiman, Ostrom, and Cheesman, 2007). That would be because employment affects recidivism, and this model could ultimately affect who gets hired. As a result, it would be neither surprising nor problematic that the risk-prediction model would need to be periodically recalibrated to maintain accurate predictions moving forward. This is not a bug but rather a feature of signaling systems that use information to systematically make inferences about an unobservable characteristic, such as recidivism risk or desistance (Bushway and Apel, 2012).
Implication: Exploring and Stressing Models Under the Reset Principle for Bias Will Be Crucial

The reset principle should begin a conversation about developing methods that mathematically demonstrate how peoples’ risks of recidivism changes. Necessarily, however, predictions will reflect the current criminal justice system, which includes many unfair systemic biases—notably many racial biases (Hetey and Eberhardt, 2018; Hinton, Henderson, and Reed, 2018; Mauer, 2010; The Sentencing Project, 2018).

Therefore, models based on the current system may not satisfy requirements for creating fair predictions and risk assessments, even if they are well calibrated and adequately assess empirical risk. Therefore, such tools should be developed judiciously and after carefully considering the many systemic factors regarding fairness. However, the search for a perfect tool should not ignore the rudimentary and subjective nature of many screens currently in use in background-check processes. An adequate assessment of bias should also include a comparison of the reset tool relative to the current practice it is intended to replace. Even an imperfect tool could provide more opportunities for candidates against whom the current system is biased, notably people of color, than the current tools that are not anchored on recidivism risk (Siwach, Bushway, and Kurlachek, 2017).

Implication: Other Concerns Besides Recidivism Risk That Employers Consider When Setting Decision Rules of a Criminal Background Check Need to Be Integrated into an Approach That Considers the Reset Principle

Employers consider other things besides the risk of an individual committing a new crime, including reputational risk for hiring someone who has damaged the community, statutory or licensing restrictions, and negligent hiring concerns. Although the last two are ostensibly tied to recidivism risk, they often are not informed by recidivism research and therefore may have features that clash with the recommendations of an approach based on the reset principle. An employer might respond to these conflicts by creating a multistep hiring-decision process in which they first consider recidivism risk and then review legal concerns about brand value, regulatory compliance, and liability. Distinguishing among the reasons behind hiring decisions is only a first step; future research is needed to understand how these competing concerns can be effectively integrated into a coherent decision framework.

For researchers, the reset principle and the models that we present should be viewed as starting points in the development of these improved recidivism risk models. Further refinement of techniques and data sources is required before models can be integrated into tools actually used in any given context. For example, we developed our model for the employment context, and therefore designed a model that can be validated under the EEOC’s Uniform Guidelines on Employee Selection Procedures. That validation may give employers a more certain way to
defend their employment decisions than the targeted employment screens that are currently in use, which lack the predictive qualities of statistical models.
Appendix D. Additional Model Calibration Curves.................................................................75
Abbreviations .......................................................................................................................77
References ..............................................................................................................................78
Figures

Figure 3.1. Example of a Single Survival Function ...................................................................... 18
Figure 3.2. How to Interpret a Single Survival Function at a Given Point in Time ............... 18
Figure 3.3. Comparing Risk from Multiple Survival Functions ............................................. 19
Figure 3.4. Demonstration on Updating Survival Functions ................................................... 20
Figure 3.5. Kaplan-Meier Estimated Survival Functions of NCDPS Data, Grouping Risk Estimates on Prior Numbers of Convictions ................................................................. 22
Figure 3.6. Demonstrating the Updating of Survival Functions in the NCDPS Data Set ............. 23
Figure 4.1. Five Considerations for Developing Recidivism Prediction Models for Criminal Background Checks ........................................................................................................ 25
Figure 4.2. Hypothetical Random Sample of Individuals with Criminal Histories All Reentering Society on the Same Day ................................................................. 26
Figure 4.3. Following Sample Through Time to an Individual’s Background Check and Assess Rates of Reoffending ....................................................................................... 27
Figure 4.4. Distribution of Number of Dates with Conviction Events Among Individuals ...... 28
Figure 4.5. Demonstration of the Importance of Considering Individual Versus Event-Based Risk .......................................................................................................................... 29
Figure 4.6. Demonstration of the Necessity of Incarceration Time on Estimated Survival Functions, Part 1 ............................................................................................................. 31
Figure 4.7. Demonstration of the Necessity of Incarceration Time on Estimated Survival Functions, Part 2 ............................................................................................................. 32
Figure 5.1. Heuristic Example to Illustrate Averaging Experts to Obtain Flexible Estimation .... 36
Figure 5.2. Model Calibration at One Time Point: Assessing Calibration of Estimated Five-Year Survival Probabilities Against Observed Kaplan-Meier Estimates Within Bins of Estimated Probabilities ........................................ 40
Figure A.1. Example 1 of an Individual’s Sentencing History, Including a Subset of Covariate Information .................................................................................................................. 62
Figure A.2. Additional Information on Example 1’s Criminal Demonstrating Offense Codes That Occur Through Time .................................................................................................. 62
Figure A.3. Example 2 of an Individual’s Sentencing History, Including a Subset of Covariate Information .................................................................................................................. 63
Figure A.4. Additional Information on Example 2’s Criminal History Demonstrating Offense Codes That Occur Through Time .................................................................................. 64
Figure A.5. Possible Transitions Among Events and States That Are Considered in These Analyses and How We Track Time Between Events .................................................................. 65
Figure A.6. Changes in DWI/DUI Names over Time in NCDPS Data Sets ............................... 67
### Tables

Table 1.1. Description of NCDPS Data Files ................................................................. 6  
Table 4.1. Importance of the Distinction Between People Versus Events ...................... 28  
Table 5.1. Concordance Index Estimates of the Model Across Time ................................ 39  
Table 5.2. Variable Importance from the Mixture-Model ............................................. 41  
Table 6.1. Workshop Agenda ...................................................................................... 50  
Table C.1. Detailed Description of Variables Included in Model .................................... 73  
Table C.2. List of Offense Types .................................................................................. 74
1. Introduction

Criminal background checks are commonly used in the United States to screen people who are applying for jobs, loans, housing, volunteer activities, and other opportunities. The evaluations operate on a broadly accepted assumption that past behavior is a good predictor of future behavior. But such an assumption must be balanced with the fact that the vast majority of people who enter the criminal justice system ultimately desist from crime (Rhodes et al., 2014; Sampson and Laub, 2003). Eventually, a past criminal record is no longer predictive of future convictions.

Current methods used to evaluate the recidivism risk posed by people with convictions do not sufficiently take into account the time that has passed since their last conviction. In the context of employment background checks, these approaches can ignore valuable information that a person has signaled about their propensity, or lack thereof, to reoffend between the time of their last major interaction with the criminal justice system and when they apply for a job. The predictive shortcomings of current recidivism risk assessments also contribute to the use of set-time exclusions that keep people convicted of some crimes from getting a job regardless of their actual level of recidivism risk.

In this report, we propose the reset principle, which requires that risk-assessment instruments (RAIs) be anchored at the time of a background check. This approach more fully accounts for an individual’s history since their last conviction or release from prison. Developing models and risk assessment tools using this principle may allow employers to better distinguish between candidates with different levels of recidivism risk. This could help employers hire lower-risk candidates and avoid higher-risk people who cannot be identified using current models. We describe characteristics that such models must have and demonstrate the viability of one such recidivism risk model. We also discuss policy implications that the reset principle could have on background checks not just in the employment context but across other domains, such as volunteer opportunities or housing.

More-accurate recidivism RAIs are useful given the many people who are involved in the U.S. criminal justice system. An estimated 30 percent of all adults in the United States had at least one arrest for a non-traffic offense by the age of 23, according to data from the National Longitudinal Survey of Youth, 1997 (Brame et al., 2014). An article in Demography estimated that 8 percent of all U.S. adults had a felony conviction as of 2010, a number that increases to 33 percent for Black males (Shannon et al., 2017).

Convictions, particularly for felonies, can block people from many opportunities. For example, the National Inventory of Collateral Consequences of Conviction reports 7,562 unique local and state mandatory exclusions applied to people with felony convictions across various activities: business licensure and participation, property rights, employment and volunteering,
occupational or professional licensure and certification, education, family and domestic rights, government benefits, government programs, government grants and loans, political participation, housing, and recreational licensing.¹ Even when opportunities are not mandatorily restricted, applicants who were convicted of felonies still may be restricted through discretionary decisions.

Employment restrictions have been the focus of much of the existing academic research on the effect of background checks. Research has clearly documented that people with felony convictions experience fewer callbacks for jobs that do not have mandatory felony restrictions than individuals who do not have records (Pager, 2003; Pager, 2007). Most employers report that they would not knowingly hire applicants with criminal records (Pager, 2007). Employers also worry about legal liability if the job candidate goes on to commit a crime (Lageson, Vuolo, and Uggen, 2015), a reality that builds on a general desire to create a safe work environment (Society for Human Resource Management, 2012). Although the degree to which criminal history information is used to restrict employment varies among employers that do background checks (Stoll and Bushway, 2008), most employers do some type of criminal background check prior to employment.²

Background checks can be challenging because they evaluate people who are not necessarily in the middle of an active criminal career. The vast majority of people who enter the criminal justice system ultimately desist from crime (Bersani and Doherty, 2018; Rhodes et al., 2014; Sampson and Laub, 2003). Research also has shown that individuals with records can be good employees in terms of common measures of job performance (e.g., Lundquist, Pager, and Strader, 2018). Moreover, another study provides evidence that, for some individuals with a record, the opportunity to work decreases recidivism (e.g., Denver, Siwach, and Bushway, 2017). If an employer’s inaccurate risk assessment stops them from hiring a low-risk person with a conviction, and the continued unemployment contributes to the candidate’s likelihood to reoffend, both the individual and society bear the costs of subsequent arrests and convictions stemming from that error. Even if someone does not recidivate, that person has lost a chance to work in a preferred job (perhaps with better pay, benefits, or promotion opportunities), and the employer has missed a chance to hire a good employee—a real loss, especially in tight labor markets. The need for accurate risk assessments is heightened by the fact that exclusion from employment because of convictions has been demonstrated to have disparate effects on historically marginalized populations (EEOC, 2012).

¹ Search done on April 15, 2021 (National Reentry Resource Center, undated). There are more restrictions that focus narrowly on specific felony offenses.

² A survey cited by the U.S. Equal Employment Opportunity Commission (EEOC) noted that 93 percent of employer respondents used criminal background checks. Of a representative sample of recent job seekers, 71 percent reported experiencing a criminal background check with their most recent job application (Denver, Pickett, and Bushway, 2018).
The legal discussion about criminal background screens has focused on these two questions:

1. Are people who have been convicted of a crime always at higher risk of committing subsequent crimes than those without criminal records?
2. Can criminal history record information be used to differentiate between applicants with records in terms of their levels of recidivism risk?

The answer to the first question is undoubtedly no (Bushway, Nieuwbeerta, and Blokland, 2011; EEOC, 2012). The answer to the second question is an equally enthusiastic yes (EEOC, 2012). These conclusions create the need for new tools that accurately identify the different levels of recidivism risk of people with different criminal history records. This need has evolved most directly in the context of employment screens and is central to accurate background checks in various contexts.

In this report, we focus primarily on one variable that fundamentally shapes the challenge of risk prediction in this context: the time since an individual’s last major interaction with the criminal justice system, specifically the date of their last conviction or release from prison. The estimate of risk of recidivism changes as time passes from an individual’s last major interaction with the criminal justice system. The highest-risk people fail and exit the pool of potential applicants, leaving a pool of lower-risk applicants. At the same time, those who continue to survive without a new conviction may be less likely to commit a crime as they become more integrated in their new life. Employers and others who conduct background checks should take advantage of knowing that a certain amount of time has passed during which someone has not incurred a new conviction and use this information to reset the risk assessment. In this report, we suggest that an RAI used in background checks should reset the assessment of recidivism risk to the time of the background check—not the time of their last criminal justice event, as current instruments do. If models are used to develop accurate tools that employers and others can use to assess recidivism risk and can be reset at the time of the background check, these tools may offer a more accurate way to screen candidates. These tools may offer many people with criminal histories a way to demonstrate that they should be offered another chance sooner than many set-time exclusions require.

In a workshop that was also part of this research project (described in Chapter 6), stakeholders involved in employment background checks recognized that this reset principle could fundamentally change policy in this area. However, these stakeholders also noted that recidivism risk is not the only consideration that informs background checks: Employers have legitimate concerns about brand reputations, legal restrictions, and liability for negligent hiring that may not be informed directly by recidivism risk. This report is designed to help better articulate how concerns about recidivism risk can be addressed accurately. This clearer articulation of recidivism risk should help clarify and distinguish that concern from other valid ones that motivate employer decisionmaking.

To help readers understand the potential effect of our ideas on how background checks should be conducted, we introduce a hypothetical employment scenario grounded in actual
reentry policy. Suppose we have two people, Anthony and Bob, who have been hired to work as janitors at a long-term care facility in Nevada and now need to pass a criminal background check. Both Anthony and Bob have criminal records, and their most recent convictions were for misdemeanor theft. Anthony was 22 years old at the time of his most recent conviction, and he had three previous convictions for misdemeanor theft. Bob was 40 years old at the time of his most recent conviction, and he had no prior convictions. Like many states, Nevada’s rules for workers at long-term care facilities require a waiting period before people with certain types of convictions can be hired to work in those facilities (State of Nevada, 2015). This policy gives equal treatment to both Anthony and Bob, and each face a seven-year waiting period from the time of conviction. This means that Anthony cannot clear his background check until he is 29, and Bob cannot clear his background check until he is 47.

This blunt approach to considering the time since the last offense fails to distinguish the differences in circumstances. Anthony might have a much higher risk of recidivating soon after his conviction than Bob, whose older age and single conviction imply a much lower risk (Bushway, Nieuwbeerta, and Blokland, 2011; Siwach, Bushway, and Kurlychek, 2017). Imposing a seven-year ban on both individuals ignores this difference to Bob’s detriment (DeWitt et al., 2017). The modeling we describe in this report supports the development of tools that an employer could use to determine whether Bob’s level of risk may in fact be acceptable long before the seven years expire, providing Bob another chance to work—a reset—much sooner than the fixed exclusion period required by Nevada policy.

In addition, models might show that Anthony, who was at very high risk for recidivism in the first year or two after his most recent conviction, may have a relatively low risk for recidivism when applying for a job after several conviction-free years. Indeed, after five years, the model might show that Anthony may be a lower risk than Bob for recidivating after two years, despite the fact that Anthony posed a much higher risk than Bob at the time of conviction. Predicting risk at the time of conviction or release from prison is a very different exercise than predicting risk many years later. Resetting the risk calculation to take into account the time spent without recidivating is particularly useful for evaluating risk among people who are at different stages of the desistance process (Bersani and Doherty, 2018).

We acknowledge the potential dangers that accompany the use of predictive risk models in this context. A statistical tool used to eliminate some subjective human biases could unintentionally perpetuate other unfair practices and biases. Even algorithms designed to avoid unfairness can result in inequality, although it is important to recognize that different definitions of fairness exist and can present competing or conflicting requirements (Corbett-Davies et al., 2016; Harris, Goss, and Gumbs, 2019). It is helpful to remember that it is difficult to create perfect systems. Reducing inequalities in the current system can yield net benefits, even if the new approach is not perfect. Research suggests that current methods that rely on simple exclusionary periods based only on the type of crime, such the above-cited example from Nevada, can create inequality without effectively identifying low-risk people (Siwach, Bushway,
and Kurlychek, 2017). Formal risk models create the promise of both fairer and more-efficient outcomes, along with opportunities to interrogate the formal decision rules that might replace more arbitrary or subjective rules.

**Sentencing and Conviction Data from North Carolina**

Throughout this report, we illustrate the importance of the principle that risk must be reset at the time of the background check using a data set on criminal sentencing events from North Carolina. The data set is from the North Carolina Department of Public Safety (NCDPS) Offender Public Information Search/Inmate Locator and represents public information related to convictions, sentencing events, and prison incarceration spells dating back to 1972. One obstacle to creating successful models that estimate long-term risk for background checks is acquiring high-quality data that represent many instances of individuals’ conviction records over long periods. That is, to understand how risk changes ten years after a conviction, we need data sets with at least ten years of data to understand rates of reoffending over those time periods. To that end, the North Carolina data include long-enough periods and can be used to create accurate models for recidivism risk for criminal background checks.

In this report, we focus on data from NCDPS from 1995 to the present—about 25 years of records. We start at 1995 because that year saw major structural change in sentencing practices (Wright, 1998), there was a change in NCDPS’s data-collection process, and data on prison time served becomes more comprehensive from 1995 on (Shen et al., 2020). As of April 2021, the data set contained more than 4 million sentencing records representing more than 1 million individuals with unique IDs in the database. For our analyses in this report, the data were restricted to events occurring after 1995, events for individuals who are between 18 and 70 years of age, individuals born in North Carolina, and records through April 8, 2021 (when we downloaded the data). Table 1.1 shows an overview of the files used and their relevant information.

---

3 See NCDPS, undated-a. This portal provides access to the data files, which can be downloaded for research purposes:

The files . . . are provided to facilitate users interested in data analysis of NC DPS Division of Adult Correction offender related data. Users may download these database files which can then be imported into their own database software. The files contain all public information on all NC DPS Division of Adult Correction offenders convicted since 1972. . . . The North Carolina Department of Public Safety does not offer any technical support or systems analysis assistance to users wishing to work with these files. (NCDPS, undated-b).

See their help page for more information.

4 Of interest to the analyses in this chapter are the files listed at the NCPDS, undated-b, pertaining to the “Offender Profile,” “Sentence Component,” and “Sentence Computations.”
Table 1.1. Description of NCDPS Data Files

<table>
<thead>
<tr>
<th>Table</th>
<th>Information Contained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Offender profile</td>
<td>Demographic information, including birth date, race/ethnicity characteristics, and birth/residency information</td>
</tr>
<tr>
<td>Sentence component</td>
<td>Individual sentence components for each conviction event that an individual experienced; contains information on conviction date, connections to other sentences, penalty information, offense codes, and other important criminal history information</td>
</tr>
<tr>
<td>Sentence computations</td>
<td>Information on incarceration dates, including incarceration start date, projected/actual release dates, and parole information (not used here)</td>
</tr>
</tbody>
</table>

Appendix A contains further detail on how the data were structured for the analysis conducted in this report.

Roadmap of This Report

We begin our exploration in Chapter 2, where we examine the role of risk-prediction techniques in the criminal justice context, trace their adaptation into the employment context, and examine their limitations. In Chapter 3, we posit the reset principle that can be used to guide the development of empirical models for criminal background checks across domains. We follow in Chapter 4 with an exploratory empirical assessment of considerations that can be used to guide model development. In Chapter 5, we apply advanced statistical and machine-learning techniques for a data set of people in North Carolina with convictions to demonstrate the viability of creating models for background checks that could be updated through time. Chapters 4 and 5 are primarily directed at researchers interested in working with policymakers to develop applicable models and can be skipped by readers who are less interested in the technical details. Chapter 6 describes the workshop held in June 2021, which helped sharpen and clarify our ideas in the context of employment background checks. This workshop featured a speech by EEOC chair Charlotte Burrows and a series of discussions about this report. Chapter 7 discusses the implications of this report and points for policymakers to consider as they move forward with new policies on background checks. Appendix A provides information about issues that we encountered that could be useful for researchers constructing data sets. In Appendix B, we describe survival analysis and its relation to the reset principle. Appendix C outlines the data that were used to estimate the model in Chapter 5. Finally, in Appendix D, we offer additional model calibration curves.
Given our goal to improve models used to assess risk during criminal background checks, it is important to understand how those risk models evolved. In this chapter, we briefly discuss the origins of risk models in the criminal justice system. We then discuss their adaptation for the employment context, regulatory responses, court decisions, and research that has informed background checks in hiring. This discussion grounds a more detailed examination of the reset principle that should guide the creation of recidivism risk models for background checks in Chapter 3. We focus on employment and not other contexts, such as housing, because the jurisprudence around criminal background checks appears most developed around employment. The use of a specific context also helps make the problem and its solutions more concrete.

Risk Assessment in the Criminal Justice System

Fifty years ago, decisions about how to treat different people in the criminal justice system were made subjectively, using the judgment and experience of administrators. Starting in the 1970s, however, correctional institutions faced overcrowding lawsuits and pressure to release inmates (Clements, 1996). As a solution, courts pressed correctional systems to create transparent and testable systems that could classify prisoners and allocate resources fairly, consistently, and effectively across the prison population to advance the goals of rehabilitation and good management. Later, this pressure from the courts to classify prisoners according to risk would expand to parole and probation systems, which faced lawsuits for releasing individuals who subsequently harmed people in the community (Clements, 1996). Today, risk-classification systems are widely used throughout the criminal justice system, from pretrial detention to parole board hearings.

There are two basic classes of risk-assessment tools in criminology and criminal justice (Gottfredson and Moriarty, 2006). The first is an expert assessment, in which human decisionmakers judge risk using their expertise and experience, sometimes with a structured tool that formalizes the factors that are used to make the decision and/or that provides a rudimentary ranking of risk among people with different risk factors. The most formal form of this approach is a risk score known as the Burgess tool (Burgess, 1928). The approach makes no claim that the given factors are strictly predictive in any particular case, but the factors have been correlated with recidivism in other contexts. A slight modification of this technique makes use of a weighted sum of the different risk factors. The weighted Burgess method uses expert judgment to assign weights that correspond to the perceived power of these factors, drawing from the research literature available at that time. Both methods are empirically guided risk tools and
occupy the middle ground between purely subjective techniques and statistical risk tools, which are the second major class of risk-assessment tools.¹

Statistical risk-assessment tools (sometimes called RAIs, see Slobogin, 2021) take into account factors that influence recidivism and estimate a formal statistical model (i.e., regression) to predict the risk of recidivism. The coefficients of the various factors become the weights used to create the risk scale. These models can be quite complicated and are often considered less transparent than Burgess scales. These approaches are sometimes called actuarial or algorithmic models of risk estimation. There is general consensus in the field that risk-assessment tools outperform decisions made using expert assessment (Gottfredson and Moriarty, 2006), although the difference between the decisions can be small, especially among Burgess scores, tools, and RAIs (Gottfredson and Snyder, 2005; Pennsylvania Commission on Sentencing, 2013).

The main difference between a statistical scale and the Burgess scale is that the former predicts the actual probabilistic risk of a particular event (e.g., conviction or arrest) in a given time window, whereas the Burgess scale produces a score that does not necessarily have a probabilistic interpretation (it is only used to rank people with different histories). In the statistical approach, the predicted risk can be used both to compare and rank individuals and to evaluate individuals’ risks relative to a benchmark or threshold for acceptable risk. For example, a parole board might decide it only wants to release people who have a 10-percent chance or less of being arrested in the first year after release. Unlike a typical Burgess scale, a statistical scale built around the risk of a new arrest in a year can identify individuals who are below this level of risk.² If the models are simply used to identify the lowest-risk individuals, such as the lowest 25 percent, there may be very little practical difference between Burgess tools and formal RAIs (Kleiman, Ostrom, and Cheesman, 2007; Pennsylvania Commission on Sentencing, 2013).

Risk Assessment in the Employment Context

The history of employers using criminal history information to predict risk is framed by the federal regulations arising from the Civil Rights Act of 1964. Title VII of this act prohibits employment practices that have a disparate effect on minorities, unless the employer can show that the practice is related to the job in question and is consistent with business needs. The

¹ Burgess risk tools cross over to become statistical risk-assessment tools when they are empirically validated on real data; the scores from the tools have proven to be reliably correlated with recidivism. Validation, in this context, means that the instrument is correlated with a desired outcome.

² At least two studies compare the performance of these two basic approaches in the criminal justice system (Gottfredson and Snyder, 2005; Pennsylvania Commission on Sentencing, 2013). Individuals are rated on competing scales, and then they are followed to measure their level of recidivism. For example, the Pennsylvania Commission on Sentencing looked at recidivism outcomes after assigning people to risk groups at sentencing. Interestingly, in the two cited cases, the simple Burgess method was noted as the best approach based on both simplicity and validity. In other words, the method using arbitrary weights does nearly as well in predicting outcomes as the more sophisticated approaches using empirically derived weights. However, a multiple linear regression still may be preferred because it is less likely to be influenced by human judgment and biases in the decisionmaking process.
EEOC, which is responsible for creating regulations that enforce Title VII, has consistently acknowledged that businesses have a legitimate need to conduct background checks to reduce theft or fraud in businesses and to guard against negligent-hiring lawsuits. Nonetheless, the EEOC has also been concerned that employers may impose overly broad restrictions that unnecessarily limit employment opportunities. The EEOC has issued two major guidance statements—one in 1987 and an update in 2012—about the use of criminal history records by employers.

In 1987, the EEOC identified the three factors that employers can use to consider risk: (1) the nature of the job, (2) the type of crime committed, and (3) the time since the individual’s last conviction. This three-part test was called the Green test, which is based on the ruling in an early Title VII lawsuit, Green v. Missouri Pacific Railroad (Buck Green et al., 1975).

Although there is little rigorous research about employer background-check practices, anecdotal evidence suggests that most employers have taken the Green factors into account (Connor and White, 2013). Current Green-based approaches applied by employers are very basic, empirically guided Burgess scales. For example, the type of conviction and the time since conviction can be combined in a matrix to produce a recommendation about whether to hire someone. This recommendation can often be adjusted by an expert within the hiring company who is tasked with making the final decision. In general, this approach meets many of the objectives of a good risk assessment tool: It can be applied to most applicants, is fairly unambiguous, can be consistently applied, and can be applied efficiently to large numbers of applicants (Megargee, 1984). As part of these rules, employers often had very long exclusionary periods, including lifetime bans, for very serious offenses.

These long bans led to a prominent Title VII lawsuit known as Douglas El v. Southwestern Pennsylvania Transportation Authority (SEPTA) (Douglas El v. Southwestern Pennsylvania Transportation Authority, 2005). In this lawsuit, Douglas El, a former employee, sued Philadelphia’s mass-transit organization for discrimination after he was fired because of a 40-year-old murder conviction that was discovered as part of a newly implemented criminal background check. SEPTA’s experts claimed that (a) everyone with a criminal record is always higher risk than anyone without a criminal record and that (b) it is not possible to differentiate between those with records. The district court dismissed El’s case in a summary judgment, which he appealed (El v. Southeastern Pennsylvania Transportation Authority, 2007). The U.S. Court of Appeals for the Third Circuit upheld the summary judgment but created room for future change when it stated that “[a]lthough we have reservations about such a policy in the abstract, we affirm here because El did not present any evidence to rebut SEPTA’s expert testimony” (El v. Southeastern Pennsylvania Transportation Authority, 2007). Criminologists responded to this expressed skepticism with a flurry of research that clearly refuted both of the claims of the defense experts in the original case (Blumstein and Nakamura, 2009; Bushway, Nieuwbeerta, and Blokland, 2011; Kurlychek, Brame, and Bushway, 2006; Kurlychek, Brame, and Bushway, 2007; Soothill and Francis, 2009).
Satisfying New EEOC Guidance

One result of this litigation and research was new guidance from the EEOC in 2012. The agency advised strongly against blanket bans on hiring people with criminal records and required nuanced, case-by-case consideration about whether a particular employment policy or action satisfies Title VII’s business necessity test. The EEOC also identified additional factors for employers to review for individualized assessment, over and above the three Green factors. These were based largely on the research on risk prediction in the criminal justice system. These include the following:

1. circumstances surrounding the offense
2. number of prior offenses
3. age at the time of conviction or release from prison
4. prior employment history, including whether a similar job was performed without incident following the most recent conviction
5. evidence of rehabilitation and references
6. participation in a government bonding program.

The EEOC’s new guidance further notes that, where a plaintiff in litigation establishes the disparate effect of an employer’s use of criminal background checks, the burden shifts to the employer to “demonstrate that the challenged practice is job related for the position in question and consistent with business necessity” (see 42 U.S.C. § 2000e-2(k)(1)(A)(i); Griggs v. Duke Power Co., 1971). The EEOC guidance indicated that an employer could establish that its criminal-history exclusions were “job related and consistent with business necessity” under two distinct circumstances, the first being that

[t]he employer validates the criminal conduct exclusion for the position in question in light of the Uniform Guidelines on Employee Selection Procedures (if there is data or analysis about criminal conduct as related to subsequent work performance or behaviors).

The second circumstance is that

[t]he employer develops a targeted screen considering at least the nature of the crime, the time elapsed, and the nature of the job (the three factors identified by the court in Green v. Missouri Pacific Railroad, 549 F.2d 1158 (8th Cir. 1977)). The employer’s policy then provides an opportunity for an individualized assessment for those people identified by the screen, to determine if the policy as applied is job related and consistent with business necessity. (EEOC, 2012)

---

3 42 U.S.C. § 2000e(m) defines *demonstrates* to mean “meets the burdens of production and persuasion.”
The EEOC almost immediately acknowledged that the first circumstance could not be met at the time because of a lack of data and research (EEOC, 2012).\textsuperscript{4} The rest of the EEOC guidelines outline how firms could satisfy the second circumstance and recommend consultation with experts who could craft new policy.\textsuperscript{5}

\textit{Using Targeted Screens Per the Second Circumstance}

The second circumstance corresponds to the first class of approaches used in criminal justice that is based on expert opinion and empirical evidence. The “targeted screen” described by the EEOC could be developed as a Burgess scale that considers time since last conviction, type of crime, age at last conviction, and number of past convictions.\textsuperscript{6} There would be no need to wait until the individualized assessment stage to consider these last two factors because these pieces of information are in the formal criminal record obtained as part of the criminal record check. Empirical evidence from the criminology literature, along with the insights from the firm’s loss-prevention and liability experts, could be used to assign the weights.

\textsuperscript{4} In one paragraph of Section 5, the EEOC, 2012, concludes that

[\textit{a}]\textit{though there may be social science studies that assess whether convictions are linked to future behaviors, traits, or conduct with workplace ramifications, and thereby provide a framework for validating some employment exclusions, such studies are rare at the time of this drafting.}

The EEOC guidance then immediately moves forward in Section 6 with details on how to satisfy the second circumstance.

\textsuperscript{5} The guidance goes further to lay out in great detail how the employer could develop this policy according to the second circumstance:

- Develop a narrowly tailored written policy and procedure for screening applicants and employees for criminal conduct.
  - Identify essential job requirements and the actual circumstances under which the jobs are performed.
  - Determine the specific offenses that may demonstrate unfitness for performing such jobs.
  - Identify the criminal offenses based on all available evidence.
  - Determine the duration of exclusions for criminal conduct based on all available evidence.
  - Include an individualized assessment.
  - Record the justification for the policy and procedures.
  - Note and keep a record of consultations and research considered in crafting the policy and procedures.

This current report’s contribution fits within the EEOC’s first circumstance: using a criminal conduct screen that is validated per the standards in the Uniform Guidelines.

\textsuperscript{6} When there were only the three Green factors, employers could create matrices for different job categories. In each matrix for a given job, the rows could be categories of crime, organized by crime type (e.g., violent, property, drug) and seriousness (felony versus misdemeanor). The columns could be the time since the offense was committed. In each cell, the firm could make a decision about whether to hire an individual at a certain point. For example, for a given crime, the decision might be to not hire someone if the crime occurred within the past three years, but then hire after three years. Firms might create a middle category, in which a candidate would go through an additional layer of screening. If a candidate had more than one conviction, the firm would review each conviction and only extend a job offer if all elements in the matrix for each of these offenses indicated that the person should be hired.
For example, the slope of the hazard curves from Bureau of Justice Statistics (BJS) research on the recidivism rates of people exiting prison could be used to create nonlinear weights that reflect the rapid decline in risk for people who do not have another arrest for two or more years (Alper, Durose, and Markman, 2018). Employers could use their own expertise to assign weights to various crime types, using not only criminological literature on risk but their own assessments of brand exposure and legal liability. The end result would be a custom scale representing the employer’s own experience and the criminological literature assessment of relative risk. A cutoff point would then be set, allowing anyone who received a risk score below that cutoff to be hired and moving those with a score above the cutoff to a second-round process recommended by the EEOC. In this second round, candidates could provide additional information, such as evidence about performing the same type of work post-conviction with the same or a different employer with no known incidents of criminal conduct; the length and consistency of employment history before and after the offense or conduct; rehabilitation efforts (e.g., education or training); employment or character references and any other information regarding fitness for the particular position; and status on being bonded under a federal, state, or local bonding program. The first-round process could be automated or done via a worksheet, while the second-round process is determined by an individual or committee (for a description of this kind of review by government agencies charged with conducting the background check, see Denver and Ewald, 2018, and Kurlychek, Bushway, and Denver, 2019). We posit that firms that follow this approach have satisfied the second circumstance as outlined by the EEOC. Despite relying on valid research, however, these tools still do not allow objective comparisons of the risk associated with a criminal record to a risk of future offending. Instead, employers must rely on subjective judgment in the individualized assessment or appeals process to try to identify people who might in fact deserve another chance.

Developing Validated Screens Per the First Circumstance

But what about the first circumstance? Is it impossible to develop a validated screen or RAI for the employment setting? To the best of our knowledge, no employer has attempted to respond to the EEOC guidance by satisfying the first circumstance with a focus on recidivism risk. With increased specificity, such an approach would be beneficial to candidates with a criminal history who are seeking to reset their life course, to a society that benefits from reintegrating people who have criminal histories, and to employers who may be able to better distinguish candidates.

There are two primary issues that complicate the movement to statistical prediction models for recidivism in the hiring context: Statistical prediction models used in criminal justice have not traditionally been used in hiring, and such models used in the hiring context focus on recidivism risk instead of job performance.

Regarding the first issue, the hiring policies in most discussions of screening tools are pen-and-paper tests designed or used to measure an employee’s ability to perform a job’s particular tasks. These tests could be used to screen employees as part of what is sometimes called
personnel selection psychology (American Psychological Association, 2018). These tests could be then statistically linked to outcomes using the three kinds of validity checks described in the Uniform Guidelines on Employee Selection Procedures written in 1978. Those validity checks are criterion-related validity, content validity, and construct validity. Criterion-related validity is met when it is “demonstrated by empirical data showing that the selection procedure is predictive of or significantly correlated with important elements of work behavior” (29 C.F.R. Section 1607.16) In lieu of having data that allow for criterion-related validity, the employer can show that the content of the selection procedure captures important aspects of job performance (construct validity) or that the selection process measures the degree to which candidates have identifiable characteristics that have been determined theoretically or empirically to be important for successful job performance (content validity) (29 C.F.R. Section 1607.16).

Over the past decade, a new type of procedure, sometimes called an algorithmic tool, has been used extensively in hiring (Trindel, 2016). For example, employers empirically map elements of candidates’ résumés to successful job performance using machine-learning techniques and then screen for elements correlated with positive employment outcomes. These risk tools are directly analogous to the type of risk-prediction models used in criminal justice, with the difference that they are being used to predict job performance rather than recidivism risk. To enhance understanding of the implications of prediction models for equal employment opportunity law, the EEOC held a meeting in 2016 about the use of big data in the workplace (EEOC, 2016). With the change in presidential administrations, there appears to be some momentum for the EEOC to establish new regulations in this area (Newman, 2021). For the time being, however, the 1978 Uniform Guidelines are the relevant framework for evaluating new tools or instruments.

By definition, these new algorithms are built to predict employment performance, therefore meeting the definition of criterion-related validity (Barocas and Selbst, 2016; Dunleavy, 2016; Raghavan and Barocas, 2019). Kim (2017) is uncomfortable with what she refers to as a tautology in the context of the Uniform Guidelines—predictive tools meet the criterion-related validity test posed by the Uniform Guidelines by definition (see also Lundquist, 2016). This discomfort is caused in part by the fact that the Uniform Guidelines do not require a theoretical or causal link between the variables in the tool and the outcomes, only a statistical correlation (Kim, 2017; Trindel, 2016). It is not clear how such a requirement would affect the guidelines. It is clear, however, that the Uniform Guidelines did not anticipate prediction models of the kind currently in use, and the guidelines should be updated to reflect this new approach (Maurer, 2020).

---

7 The Uniform Guidelines identify construct validity as a distinct validation approach. Construct validity involves the theoretical linkage between the characteristics in the tool and the outcome of interest. In the guidelines, construct validity is an alternative to criterion-related validity and not an explicit part of criterion-related validity. Variables used in criterion-related validity tests do not need to be linked theoretically a priori to the outcome of interest.
The second issue with the use of statistical prediction models for recidivism in the hiring context involves the focus on recidivism risk prediction instead of job performance. The use of a criminal history screen, in the words of the Third Circuit:

> has nothing to do with the applicant’s ability to drive a paratransit bus; rather, it seeks to exclude applicants who, while able to drive a bus, pose too much of a risk of potential [479 F.3d 243] harm to the passengers to be trusted with the job. Thus, our standard of “minimum qualifications necessary for successful performance of the job in question” is appropriate in test-score cases, but awkward here because “successful performance of the job” in the usual sense is not at issue. See Lanning I, 181 F.3d at 482. SEPTA could argue that successful performance of the job includes not attacking a passenger and, therefore, that the standard is still appropriate. However, the standard is worded to address ability, not risk. Yet, the issue before us is the risk that the employee will harm a passenger, and the phrase “minimum qualification” simply does not fit, as it is hard to articulate the minimum qualification for posing a low risk of attacking someone. (El v. Southeastern Pennsylvania Transportation Authority, 2007)

The appeals court in *El v. SEPTA* made clear that it believes that it is proper for employers to focus on the risk of recidivism as a separate feature from job performance. Although our reading of the EEOC’s 2012 guidance suggests some resistance to this focus on recidivism as distinct from job performance, the models suggested under the second circumstance all focus on recidivism risk instead of job performance. As a result, the EEOC and the federal courts appear to have accepted recidivism risk as a legitimate outcome of concern to employers. The movement to statistical RAIs adapted for use by employers is straightforward once recidivism risk has been accepted as a valid outcome.

**Discrimination and Transparency**

There are two concerns in this new literature on algorithmic hiring that are relevant to developing recidivism risk models for employment. The first concern deals with the possibility of baking discrimination into the tools (Trindel, 2016). The second deals with transparency. Regarding the first, a broad concern around algorithmic hiring tools is that measures of job performance in discriminatory workplaces will be biased. Creating tools that predict job performance in discriminatory workplaces could bake in the bias that existed in those workplaces. This same problem has been raised in the context of risk tools for criminal justice (Eckhouse et al., 2019) and could be raised for criterion-related validation of standard pen-and-paper tests for hiring that are envisioned by the Uniform Guidelines. A criterion-related validity test correlates the results of the test with work performance as measured by that work environment. If that work environment has biased measures, the pen-and-paper test will also be biased.
The other concern involves transparency. It is common to refer to these predictive tools as black boxes because of the complex ways that the variables are used in the models to predict risk. In addition, the tools in criminal justice are often proprietary and therefore not available for review. The two problems are neither necessary nor insurmountable (Batterywala and Agarwal, 2020; Rudin and Radin, 2019). Our report starts from the premise that any recidivism risk tools for employment that are developed should be transparent and available for inspection, as proposed by Bushway, 2020. And, although the models are indeed complex, we present the implications or relationships among factors and outcomes in ways that allow inspection and assessment of bias. The remainder of this report is meant to push forward the discussion about key features of a risk-prediction model for criminal background checks.
3. The Reset Principle for Background Checks

In this chapter, we describe the reset principle that drives our basic approach to risk prediction for background checks. We define the reset principle as follows: any risk prediction model must be able to update risk estimates, i.e., be “reset,” at the time of the person’s criminal background check. This principle represents the foundational idea of this report. A background check often occurs many years after a person has interacted with the criminal justice system. Therefore, risk should be evaluated forward from the time of that interaction. This recognizes that a person who has been free in society for many years without another conviction is in a very different place in terms of risk to recidivate than a person who was only recently convicted. More important, a person who has not reoffended for many years is signaling important information: This person may have desisted and is now no longer at a higher risk for another conviction. Fundamentally, the fact that we are evaluating people at different points in time in relation to their interactions with the criminal justice system implies that models need to be able to consider the relative risk of a new conviction over some period both for people who have just been convicted and for people who were convicted many years ago.

This principle is easy to state and explain but surprisingly difficult to implement. Most of the available risk models in criminology that we are aware of are anchored on a specific conviction or arrest date and evaluated over a fixed period of time (e.g., predicting risk for pretrial detention or parole release). This makes sense for the criminal justice system, where the goal is often predicting future risk to make decisions about release into society (Slobogin, 2021). Criminal justice risk prediction is often most concerned with what will happen in the immediate aftermath following release or conviction. For example, in the case of parole, risk may be evaluated over a time horizon that represents the remainder of the supervision period—several years. In the case of pretrial detention, risk is evaluated over a time horizon that covers the resolution of the case—usually less than a year. Risk in the background-check environment is inherently different because there is additional information about an individual’s risk that is generated while living free in society.

This additional information implies that the time frame over which risk must be estimated for a background check needs to be much longer. Employers are often trying to evaluate the risk of a person who was last convicted many years ago. Recall that, in *El v. SEPTA*, SEPTA made the decision to fire someone with a 40-year-old conviction. Although this case was extreme, it makes sense that employers may be considering risk for at least ten years after the most recent conviction or release from prison. Even if the employer worries about risk only in the year after the background check, the employer needs to consider the years that have passed since the candidate last interacted with the criminal justice system.
The importance of considering the years that an individual has been free in society also can inform an employer’s choices among candidates. Employers are not simply making decisions at different points in time for everyone released from prison or convicted of crime. For example, an employer may specifically be interested in comparing the risk associated with someone who has managed to remain in the community for five years without a new conviction against a person who has been in the community for only one year. To fairly compare individuals, the employer can and should take the information revealed about this period of non-conviction into account, thereby resetting the risk calculation for each individual to occur at the time of the background check instead of from the time of the last interaction with the criminal justice system.

To further explore ideas of the reset principle in this report, we rely on methods from statistics known as survival analysis or, more generally, time-to-event analyses. To help readers understand simple ideas from survival analysis, we will provide a brief primer on concepts via figures in the remainder of the report.

**A Primer on Survival Analysis**

To understand how we define risk, here we provide an overview of survival functions (Klein and Moeschberger, 2003), which are mathematical tools that help illustrate the theoretical underpinning of the reset principle. Survival functions are typically estimated in analyses that aim to predict the time to another event—in our case, we aim to predict the amount of time until a person’s next conviction. An important aspect of estimators of survival functions is that they allow cases in which recidivism has not yet occurred to contribute information to the risk estimate. These types of observations are referred to as *censored* observations in time-to-event analyses. In our examples, such observations are important because we expect that many individuals will have had only one conviction in their criminal history. Because of this, we will not observe a time to their next conviction; the event will soon happen, will happen years after we stop following the individual, or may never happen. These censored observations will be common in our data sets and will provide critical information when predicting risk.

As noted, the survival function represents the probability that an individual “survives” longer than some predefined set of time. In the context of the criminal background checks, survival functions will represent the probability that an individual’s next conviction time occurs after some predefined time from their last conviction date or last release date. Figure 3.1 shows a single survival function.
Figure 3.1. Example of a Single Survival Function

NOTE: This represents a single hypothetical survival function for an individual. The y-axis represents the survival probability $S_i(t)$—i.e., the probability that the event time $T$ is greater than a chosen time $t$ or $Pr(T > t)$. The x-axis represents different time points $t$.

The x-axis represents specific times from a baseline event $t$ (e.g., date of last conviction or release from prison). The y-axis represents the survival probability, often denoted $S_i(t)$ (i.e., the probability that the next event time occurs after some time $t$ from a baseline event). Figure 3.2 provides more context about how to clearly read the figure for an individual’s one-year survival probability—i.e., how likely it is that conviction occurs after one year in the future from their original baseline event.

Figure 3.2. How to Interpret a Single Survival Function at a Given Point in Time

NOTE: At each point in time $t$ on the x-axis, the function can be read as “the probability the next conviction event occurs $t$ years after the baseline event.” That is, typical survival functions evaluate risk only as anchored from the baseline event. This figure represents the one-year survival probability.
Now consider the goal of comparing the risk level of individuals A and B using survival functions. Say that individual A’s one-year survival probability is lower than individual B’s. Curves that are higher on the y-axis represent lower risk (demonstrated in Figure 3.3). Therefore, it is possible to assert that individual B is less risky than individual A.

**Figure 3.3. Comparing Risk from Multiple Survival Functions**

![Graph showing survival functions for individuals A and B.](image)

**NOTE:** This represents hypothetical survival functions for two individuals. The y-axis represents the survival probability—i.e., the probability that the event time $T$ is greater than a chosen time $t$ or $Pr(T > t)$. The x-axis represents different time points $t$. If a survival function is “above” another at a given point, it can imply that the individual represented by that curve is less risky at that point in time.

Finally, a useful property of the survival function is that it can be updated through time with knowledge that an event has not occurred up to some point in the future. *It is this theoretical property that drives the need for the reset principle in criminal background checks.* Consider that we may have knowledge that an individual has spent time in the community without another conviction and therefore, up to that point in time, has desisted from crime to the best of our knowledge. This time without a conviction can be used to update the probabilities generated by the survival function with this new knowledge. Specifically, the probability that an individual had another conviction prior to the background check is now zero (and thus the survival probability is equal to one for these points). Figure 3.4 demonstrates the concepts for this update; Appendix B covers the math for accomplishing it.

19
NOTE: In the top graph, the original curves illustrate hypothetical times of background checks. Specifically, individual A has a background check at three years, and individual B has a background check at one year. When comparing both individuals at baseline, we might conclude that individual B is less risky. In the second graph, the survival functions are updated with knowledge of when the background check is actually occurring (for individual A, the update occurs three years from the last event, and for Individual B, the update occurs one year after the last event). This two-year difference makes it difficult to compare risk, so the last graph “realigns” the survival functions so that they can be fairly compared at the same new baseline date—i.e., when the background check occurs. The top graph demonstrates how individual B was lower risk than individual A at the original baseline. The last graph demonstrates how, with differing times for when their respective background checks occur, individual A is now slightly lower risk than individual B.
Demonstrating the Reset Principle in the North Carolina Data Set

The figures so far in this chapter outline a mechanism by which individuals may “flip” their risk designations as time progresses. Although this idea is compelling, it is only theoretically motivated. That is to say, the figures demonstrate only that, by using the math of survival analysis, it is possible to show that risk designations for two individuals can flip when applying the reset principle—not that it actually happens in real-world data sets. In this section of this chapter, we briefly demonstrate the reset principle in a simple survival analysis using the data from North Carolina previewed in Chapter 1.

Consider Figure 3.5, which represents survival functions estimated using the Kaplan-Meier (KM) estimator (Klein and Moeschberger, 2003) and done within groups of individuals determined by the number of conviction events that occurred in their past (i.e., number of prior convictions). KM estimators estimate the survival function, which specifically takes into account the fact that individuals may have censored observations. The curves demonstrate a strong association between information about criminal records, specifically number of prior conviction events and the likelihood that another event will occur in the future. The figure illustrates that people with more than ten prior convictions are higher risk than those with no prior convictions. It also shows a median survival time of just under three years for those with more than ten prior convictions, compared with a median survival time of more than 25 years for people with zero convictions.
As alluded to earlier, we can use properties of survival functions to quantify how risk changes for a hypothetical set of background check times. Specifically, for each group of prior number of convictions, we determine what the risk would be if we updated the survival probabilities at a hypothetical background check time. For each group of prior number of convictions, we reset risk estimates at zero, two, three, four, five, and six years into the future. Figure 3.6 shows the updated curves. We see that the survival probabilities are now closer for all of these groups and, in fact, some groups are lower risk than those with zero priors.
Figure 3.6. Demonstrating the Updating of Survival Functions in the NCDPS Data Set

![Graph showing survival functions for different numbers of prior convictions and update times.](image)

NOTE: Individuals are stratified by the number of prior convictions that they had (top legend) and updated at different times in the future (bottom legend).

This demonstrates that, if we consider risk, two individuals may now represent equivalent risk if we match them on their prior number of convictions and respective update times. This means the risk profiles of individuals with more than ten prior convictions may look similar to those of individuals with no prior convictions if those with more than ten convictions have been in the community for an extended period of time at the time of their background check.

To operationalize the idea of using the survival function to update risk at the time of the background check, we must estimate survival functions accurately using appropriate data. In Chapter 4, we lay out some of the important considerations that researchers will need to take into account when building a risk-prediction model from scratch. Some of these issues might also be relevant to building a Burgess model (or any other prediction model) that could be validated against actual outcome data (but we will not explore these ideas in this report). In Chapter 5, we demonstrate one particular model that we estimated, calibrated, and validated using the North Carolina data. Chapter 5 demonstrates a concrete example that illustrates some of the challenges in implementing these ideas and highlights some of the potential gains from this approach. The next two chapters are directed at researchers and policymakers who are interested in engaging with some of the challenges involved in building a risk-prediction model that adheres to the reset principle. Thus, readers focused less on methodology may wish to skip to Chapter 6, where we describe the workshop where this report was discussed, before the conclusion in Chapter 7, where we summarize the main findings and discuss the implications of the report.
4. Considerations for Building Recidivism Prediction Models That Follow the Reset Principle

As outlined in the previous chapter, satisfying the reset principle with survival models requires accurate estimations of individuals’ survival functions (the time between the next conviction and last major interaction with the criminal justice system). This estimation can be difficult to perform and often requires serious thought about many important statistical considerations, such as defining the population where the prediction model will be relevant, specifying the types of events that define how time is measured, specifying how to evaluate model “goodness of fit,” and distinguishing among candidate models. In this chapter, we discuss five important estimation considerations that researchers should focus on as they build risk-prediction models that incorporate the reset principle.¹ These considerations can be considered as good research practice, specifically for the estimation of recidivism risk in criminal background checks. This report contributes a discussion of these considerations in this specific policy context to the literature. The primary considerations are outlined in Figure 4.1, and many should be intuitive to those familiar with research design.

Note that the last estimation consideration, which focuses on the external validation of the risk-prediction model, directly addresses the concern discussed earlier about the near tautology for risk-prediction models (Lundquist, 2016). Lundquist expressed concern that a risk-prediction model may meet the Uniform Guidelines criteria for criterion-related validity by default because, unlike the hiring instruments envisioned by the Uniform Guidelines, risk-prediction tools are designed specifically to predict empirical risk. However, good statistical practice for creating prediction models requires a separate validation step on data that were not used to create the risk tool (Hastie, Tibshirani, and Friedman, 2009). In practice, this is usually a “test” sample drawn from the same data used to create the model, but cases are selected such that an individual’s observations can be only in the “training” or “test” data sets. Model performance can be asserted to be valid only if evaluated on separate test samples. Separating the training of the model from the validation of the model—both as distinct steps and with distinct data sets—eliminates the perceived tautology.

¹ For a similar effort for RAIs in the criminal justice system, see the guide by Slobogin, 2021.
In the remainder of this chapter, we will walk through a more detailed discussion of these considerations, with examples highlighting the importance of our choices where appropriate.

**Consideration 1**

Proper estimation requires that data sets accurately reflect the population for whom risk is being evaluated. This is a particularly important consideration in risk estimation because observational data sets that are often constructed to estimate risk contain information about individuals that contribute multiple observations to the data. This overrepresentation can complicate estimation.\(^2\) Similarly, not all events in these data sets may be useful for predicting risk—for example, “holdover” convictions, which are related to past offenses, may occur while an individual is incarcerated. Our goal is not to predict the time to these holdover convictions but rather convictions that result from new offenses. These considerations about data imply that it will be important to make assumptions on the data to perform estimation.\(^3\)

Estimating an individual’s risk to recidivate at the time of a background check requires knowing the types of individuals released around the time that the subject was released and

---

\(^2\) For an example in criminal justice risk estimation, see Rhodes et al., 2014.

\(^3\) The framing used by Rhodes et al., 2014, posited the choice of population as person-based versus event-based samples but is only part of the estimation problem. As a result, although we use Rhodes et al., 2014, as an exemplar for the idea that population matters, we start from first principles of identifying our population and working to an estimation frame.
observing the patterns of how individuals around them recidivate. Thus, an ideal (albeit unrealistic) setting to estimate the true risk of recidivating from the time of the background check would require a random sample of individuals (all with the same or very similar criminal histories) who are all reentering society on the same day and then observing the recidivism rates as they face the same temptations in society. Focusing on those reentering society—and not those who are incarcerated—is useful because their levels of risk are different than other groups; however, people on probation or parole can still commit new crimes. This ideal setting, demonstrated in Figure 4.2, would allow us to track a sample of similar individuals who are all entering society on the same day to observe rates of reoffending. Under this random sample, it would be possible to estimate the survival probabilities for each individual from their common baseline date.

**Figure 4.2. Hypothetical Random Sample of Individuals with Criminal Histories All Reentering Society on the Same Day**

![Graph showing individual re-entering society over time](image)

**NOTES:** The black x represents a “censored” observation in which the individual had not reoffended at the time that data collection was completed. Therefore, we do not know the outcome.

Having this information on the observed rates of recidivism would then allow us to assess how likely an individual is to recidivate in the future at specific points in time. Figure 4.3 shows the recidivism rate in the same sample of people that undergo a background check at four years and compares how many people have already recidivated prior to this point, and then assesses rates of failure from this point into the future.
As previously stated, although this ideal setting is clearly impossible to obtain (for both implementation reasons and obvious ethical concerns), a failure to attempt to construct a sample to estimate models as similar as possible to this ideal setting can lead to bias or misleading estimates of risk for individuals. Next, we discuss two important departure points from this assessment: (1) overrepresentation of individuals within the sample and (2) failure of data sets to include incarceration information.

Understanding the Impact of Individual Representation

In any given data set representing convictions, some people may have more than one conviction event. Figure 4.4, which is based on the North Carolina data, shows that, from 1995 to 2021, approximately 53 percent of individuals had convictions on one date (they may have had multiple sentences on one day, but additional convictions are not observed again for a subsequent set of events). Furthermore, approximately 72 percent of individuals only have one or two instances in the data.
Thus, if we naively estimate using all data that are based on conviction episodes, some people appear more than once and therefore will be overrepresented in the data. However, we require a sample that follows only a single history for each individual in our population when looking at our hypothetical sample.

To understand one potential issue with overrepresentation in the sample, consider the heuristic example in Table 4.1 (this is a simplified example that ignores such important concepts as censoring, which would be exhibited in the real world). This example represents two hypothetical individuals, each with a set of event times.

### Table 4.1. Importance of the Distinction Between People Versus Events

<table>
<thead>
<tr>
<th>Individual</th>
<th>Set of Event Times</th>
<th>Average of All Event Times</th>
<th>Average of Individual's Average Event Times</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>{5}</td>
<td>1/5(5 + 1 + 1 + 2 + 1) = 2</td>
<td>1/2(5 + 1/(1 + 1 + 2 + 1)) = 3.125</td>
</tr>
<tr>
<td>B</td>
<td>{1,1,2,1}</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTE: The two rightmost columns demonstrate the difference between the overall average time between events and average of individual average time between events.

If we were to focus on simply averaging the event times for these individuals, then we would assume that the average time between events would be approximately two years (see second column). Instead, if we focus on considering first the average time for each individual (to remove the effect of overrepresentation), and then averaging across individuals, we see that the time between events is approximately 3.125 years (see far right column). This difference is important for background checks and constructing samples for estimation. Removing the effects of overrepresentation in a sample allows us to understand how risky specific individuals are for any
given opportunity. Additionally, it allows us to assume that we are following a specific “group” of independent *individuals* through time, as we will only observe their events.

To further demonstrate this point, consider Figure 4.5, which shows a set of survival functions estimated using a simple KM estimator (Kaplan and Meier, 1958; Rich et al., 2010) on the NCDPS data.

**Figure 4.5. Demonstration of the Importance of Considering Individual Versus Event-Based Risk**

![Figure 4.5. Demonstration of the Importance of Considering Individual Versus Event-Based Risk](image)

NOTE: The blue curve naively assumes independence among all event times. The other curves take within-individual independence into account by (1) weighting by representation and (2) subsampling one observation from each individual. Using all events naively would indicate that 50 percent of individuals fail prior to five years, while the other methods suggest that the 25-year survival probability is greater than 60 percent. Ind = individual.

In the figure, the blue solid curve represents the survival probability when each *conviction event* is equally weighted (and thus does not represent the ideal population of interest because it overrepresents individuals with multiple convictions). The curve might suggest that the median survival time is approximately five years. The dashed and dotted lines tell a different story. These lines were created by giving each person equal weight to more accurately represent the ideal population by (1) sampling a single observation from each person and (2) weighing each observation proportional to its representation (i.e., if an individual has five events, then their weight for each event would be 1/5). We see that the effect is that the median survival time is greater than 25 years *within a population of individuals*. Furthermore, the probability that any individual’s next conviction time is greater than 25 years is approximately 60 percent. This

---

4 We are not the first to make this observation. See Rhodes et al., 2014.

5 Subsampling from each individual will be useful for training machine-learning models; therefore, we demonstrate how it affects assessments.
statistic suggests that many individuals actually may desist from crime after their first set of convictions (i.e., only 40 percent of convictions have a subsequent failure across all criminal histories). Predictive models for the background-check setting must be made in ways that take the potential issue of overrepresentation into account.

**Understanding the Impact of Incarceration Information**

The second important population consideration is that risk should be estimated only among individuals who are in society and free to commit a new offense. That is, the people in the ideal population cannot be incarcerated for portions of their history because they are typically not at risk while in prison. Additionally, people in prison or jail are generally not facing background checks by employers, volunteer organizations, or landlords. Unfortunately, this reasonable consideration creates a difficult data requirement. Most formal state criminal record repositories (i.e., rap sheets) do not collect information on incarceration spells (Alper, Durose, and Markman, 2018; Goggins and DeBacco, 2020). As a result, the data from those repositories should not be used to predict recidivism. Fortunately, the North Carolina data we identified provide information on prison spells (although they lack clear information on county jail periods).

This information allows us to again explore simple risk estimation using KM estimators that demonstrate survival functions. But our goal here is illustrate how individuals are not at risk for another conviction because they are in prison. To accomplish this, we perform a naive analysis that calculates risk erroneously while ignoring prison time (i.e., measuring the time between convictions but not time spent in prison), but these KM risk estimates are calculated within groups of actual observed prison time between the two conviction events. This will demonstrate how those with longer prison sentences can appear low risk over time, but this is because they are incarcerated. We show this in Figure 4.6, where the left panel estimates the raw survival probabilities as measured between conviction events and each curve in the right panel represents risk conditioning on years served in prison between conviction events.
Figure 4.6. Demonstration of the Necessity of Incarceration Time on Estimated Survival Functions, Part 1

NOTE: Risk estimates are provided by estimating the time between convictions and in bins/groupings based on actual incarceration time between those convictions: 0–1 years, 1–2 years, 2–5 years, 5–10 years, and more than ten years. The figure demonstrates that individuals in prison are not at risk for another conviction.

Figure 4.6 is illuminating in that, in the right panel, those with the longest amount of time served actually exhibit a long flat period of risk in the early years; this is masked in the left panel of Figure 4.6. The reason for this long, flat period is that these individuals were not really at risk for another conviction event for periods of time because they were incarcerated (in statistical terms, they are not in the risk set). This can potentially mislead risk assessments.

When we perform an analysis that considers the correct baseline event and starts from the most recent conviction or prison release and that assesses time to the next conviction (i.e., when the individual reenters society) from either of these baselines, as is presented in Figure 4.7, we see much more realistic risk patterns for individuals. It shows that those who have served long sentences still exhibit the lowest risk, but there are no longer flat spots in the estimated survival function. We caution readers not to take away a causal relationship between length of sentence and risk profile at a given point in time from this figure; the differences among these curves could be explained for many reasons (e.g., those who have served longer prison sentences are older at release, and age is related to likelihood of reoffending). The main takeaway from these figures is that the failure to account for incarceration spells will result in an underprediction of the actual risk presented by people who are undergoing for a background check.
Consideration 2

In its purest sense, this consideration is necessitated by the EEOC guidance (EEOC, 2012):

> [t]he fact of an arrest does not establish that criminal conduct has occurred. Arrests are not proof of criminal conduct. Many arrests do not result in criminal charges, or the charges are dismissed. Even if an individual is charged and subsequently prosecuted, he is presumed innocent unless proven guilty.

By contrast, a record of a conviction will usually serve as sufficient evidence that a person engaged in particular conduct, given the procedural safeguards associated with trials and guilty pleas.

The requirement that employers using criminal history information rely only on convictions, not arrests, affects many of the factors used to predict risk. For example, the crime-type factor must be affiliated with a conviction rather than an arrest. Likewise, the time since the last event must be based on conviction, not arrest. Most important, the factor for the total number of crimes must be based on counts of conviction rather than counts of arrest.

---

6 According to the EEOC, employers can use the underlying behavior for an arrest if they have independent confirmation of that underlying behavior. We will not consider that issue in this report because we are interested in the development of a recidivism risk tool that uses on information available in a formal criminal history record.
Although we recognize the influence that the EEOC guidance had on this second consideration, the focus on convictions is a good idea in other frameworks besides employment, even if not mandated by the legal context. For example, focusing on convictions partially addresses the concerns that recidivism risk models will bake in bias reflected in data from discriminatory environments. The criminological literature is quite clear that there is less evidence for racial bias in convictions than in arrests. It is now a routine finding that white people are more likely to be convicted after an arrest than people of color, a finding that suggests that arrests are racially biased relative to convictions (Kutateladze et al., 2014) and that the conviction process works to “unbake” some of the racial bias related to arrests (Beck and Blumstein, 2018; Kim and Kiesel, 2018). This does not guarantee that there is no racial bias in models that are based on convictions, but we can be confident that discrimination is less severe for models built on convictions.

The emphasis on convictions also has implications for the type of data that can be used to build recidivism risk models for employment. In the criminal justice context, RAIs are built using data from rap sheets. These repositories are anchored on the fingerprint cards filed with each arrest. Thus, rap sheets are fairly comprehensive measures of the arrests for individuals in a given state. However, they have a significant, well-established problem—the arrests are often missing disposition information about convictions (Goggins and DeBacco, 2020; U.S. Department of Justice, 2006). Rap sheets also miss convictions that do not start with arrests that result in fingerprint cards (Virginia State Crime Commission, 2018). As a result, the repositories are not comprehensive sources of conviction data and could introduce problems for estimating risk in background checks.

Consideration 3

In *El v. SEPTA*, a key issue was El’s risk of reoffending. As the judges pointed out in their ruling, the studies cited in support of the defendant’s exclusion policies only followed released prisoners for three years and thus was hardly convincing evidence for questions about El’s 40-year-old record. Although it might not be feasible to build models looking reliably at recidivism over 40 years with standard data, *El v. SEPTA* makes it clear that recidivism risk models need to contemplate the risk of conviction for many years after the last. Indeed, much of the criminological research generated in response to the *El v. SEPTA* ruling followed people for six or more years, with some extending for 20 years (Blumstein and Nakamura, 2009). Prior to *El v. SEPTA*, long-term recidivism models were rare because most recidivism tools were created for criminal justice purposes, which focus primarily on short-term recidivism.

---

7 Perhaps as a result, the BJS extended its follow-up period for its regular analysis of released prisoners to nine years (Alper, Durose, and Markman, 2018).
The choice of ten years is guided by several factors, including some understanding of available data and that fact that resetting risk at nine years for individuals and estimating an updated one-year survival probability requires at least ten years of data. An inspection of the North Carolina data demonstrates that this choice has at least face validity. Additionally, a closer look at Figure 4.7 shows that the survival curves appear to flatten after approximately ten years. Although most of the risk is “spent” in the first couple of years after the individual’s last criminal justice event, a full characterization of longer-term risk must cover at least ten years. This ten-year guidance is also supported by research showing that waiting for ten years before employing someone may not be necessary to prevent recidivism (Denver, 2017; Siwach, Bushway, and Kurlychek, 2017).

Consideration 4

This consideration is simple to say and more difficult to do. Essentially, this is where principles of good estimation enter the conversation. To estimate probabilities that reflect true rates of recidivating in the population, researchers must choose flexible models that can fit the data well and whose predictions are also well calibrated against true rates within the population. Calibration typically implies that, if the model asserts that 80 percent of individuals have a conviction after \( x \) years, then empirically we should see that approximately 80 percent do indeed have convictions after \( x \) years. To accomplish this, researchers should follow best practice on model selection and validation (as outlined in Goodfellow, Bengio, and Courville, 2016, and Hastie, Tibshirani, and Friedman, 2009) in which models are estimated using training and validation data sets. Then out-of-sample generalizability is assessed using an independent test data set (see Consideration 5). There are many models that could be used for estimation, but the ideas of model selection and assessment should be similar.

Consideration 5

This is, essentially, the criterion-related validity test referenced in the Uniform Guidelines. Having created a well-calibrated model in the preceding step, the researcher then validates the model by showing that it performs well on a second sample from the same population of data that were not used to build the model. Statisticians have developed sophisticated and rigorous techniques for validating a risk model using test samples that were not used in the construction of the tool. These statistical validation methods will continue to evolve, and it is important that validation tests in our context continue to evolve with the field of statistics.
5. A Risk-Prediction Model That Follows the Reset Principle

The empirical data and figures in the previous chapter point to some of the careful considerations that must be made when working with criminal justice data in the context of background checks. In this chapter, we detail our approach to modeling the entire survival function against the North Carolina data set and demonstrate the validity of our modeling approach to creating estimates that could be used in the background check context. It is important to note that we acknowledge that there are other approaches and models that could also be created to follow the reset principle. Our intent is not to provide a detailed survey of all possible approaches but to demonstrate the viability of building a valid model that could be used to predict risk.

We will document how we can provide risk estimates at the time when the background check occurs, a requirement of the reset principle. In doing so, we focus on nonparametric estimation and machine learning–based models and summarize findings from the example modes. Because this chapter aims to substantiate the technical viability of recidivism risk models for employment, it requires the use of concepts and terminology that may be unfamiliar to readers who do not have advanced knowledge of statistics. Although we have attempted to express technical concepts plainly here, this chapter is likely to be most useful to those who do have advanced knowledge of these areas of statistics. Additional details can be found in Appendix B.

An Approach to Modeling for Time-to-Event Analyses

In survival analyses, the goal for each individual \( i \) is to model and estimate the survival function \( S_i(t) \) that mathematically represents the probability that the next event for individual \( i \) occurs after some time \( t \) from baseline (if it occurs at all). In our example, baseline is the individual’s most recent conviction or prison release. We typically do not have enough observations on an individual to estimate their specific survival function \( S_i(t) \) directly. Thus, we rely on estimating survival functions in a population and then finding conditional survival functions given the individual’s set covariates \( X \) (e.g., information on past incarceration, previous number of convictions), which is expressed as \( S(t|X_i = x) \equiv S(t|x) \), and that is used to approximate \( S_i(t) \). In doing so, the conditional survival function assumes that individuals with similar covariates will have similar event times in the future and that the covariates are predictive of future conviction dates. These conditional survival functions will be important to estimate because knowing \( S(t|x) \) will allow us to further define updated conditional survival functions that will be used to reflect an individual’s risk profile when we incorporate the time of the background check (see Appendix B for a detailed derivation of how this updating occurs mathematically).
The estimation approach that we have focused on for measuring time to next conviction is that of nonparametric models, which adjust to the complexity of the data involved. Specifically, we are using a subclass of models referred to as *mixture-of-survival-expert models*. These models have seen a recent swell in use in the machine learning community (Bennis, Mouysset, and Serrurier, 2020; Erişoğlu, Erişoğlu, and Erol, 2011; Kuo and Peng, 2000; Nagpal et al., 2019; Raman et al., 2010). These models create estimates through a weighted average of many underlying functions. Figure 5.1 provides a heuristic example to illustrate how, by averaging two experts, we might be able to obtain flexible estimation.

**Figure 5.1. Heuristic Example to Illustrate Averaging Experts to Obtain Flexible Estimation**

The general form of the model we use is made up of \( k \) experts (sometimes referred to as *templates*) who are combined through a weighted averaging procedure to estimate an individual’s survival function. Specifically, the model is

\[
S(t|x) \approx S(t|x, \gamma, \theta) = \sum_k w_k(x|\gamma)S_k(t|\theta_k),
\]

where each \( w_k(x|\gamma) \) is referred to as a *mixing component* of the model with the property that \( \sum_k w_k(x|\gamma) = 1 \), and tells how each survival expert \( S_k(t|\theta_k) \) contributes to the prediction for an individual with covariates \( X \), and where \( (\gamma, \theta) \) are parameters that must be learned in the estimation phase (the statistical process wherein the value of some property of a model is calculated from observations of a sample). In our application, we have focused on estimating the weights \( w_k(\alpha|\gamma) \) using neural networks (see Goodfellow, Bengio, and Courville, 2016, for a methodological introduction); specifically, a simple sequential feedforward network. This provides a flexible method for choosing the experts given the covariate information. To choose the survival experts \( S_k(t|\theta_k) \), we focused on the Weibull family of distributions (Klein and Moeschberger, 2003), choosing a large grid of parameters \( \theta_k \) that represent a large diversity in expected failure times and that exhibit different shape parameters of the Weibull distributions.
(meaning some experts have shape parameters representing increasing hazard functions and others represent decreasing hazard functions).

The parameters of the model can be estimated using an optimization of a right-censored likelihood equation that we have implemented in TensorFlow\(^1\) (Abadi et al., 2016). The likelihood that is optimized is

\[
L(\gamma, \theta|x, y) = \prod_{i=1}^{N} f(t|x, \gamma, \theta)^{\delta_i} S(t|x, \gamma, \theta)^{1-\delta_i}
\]

\[
= \prod_{i=1}^{N} (\sum_k w_k(x|\gamma) f_k(t|\theta_k))^{\delta_i} (\sum_k w_k(x|\gamma) S_k(t|\theta_k))^{1-\delta_i},
\]

where \(f_k(t|\theta_k)\) is the density function of a Weibull distribution with parameters \(\theta_k\), and \(\delta_i\) is an indicator that the individual was observed to reoffend and is one if observed and zero otherwise.

Our proposed model comprises many parameters. To avoid overfitting of the model, we performed model selection for the parameters by partitioning the data set into training, validation, and test sets, where optimization is performed on the training data, model parameters are selected on the validation data to approximate out-of-sample generalizability, and finally the ultimate model performance was assessed on the test data set (see Goodfellow, Bengio, and Courville, 2016, chapter 5, for a methodological introduction to machine learning basics, and Hastie, Tibshirani, and Friedman, 2009, chapter 7, for a thorough overview on model assessment and selection).

Readers might note that there are alternative methods for estimating survival functions and comparing risk in this setting (see Klein and Moeschberger, 2003, for an extensive overview of survival analysis and estimation), such as parametric likelihood–based methods like accelerated failure time models, semi-parametric Cox proportional hazards (Cox-PH) models, and nonparametric estimators, such as the KM estimator. We focus on the mixture-model approach because many of these other methods may not directly provide calibrated probabilities and are thus less directly amenable to our updating approach. For further discussion about potential issues with other models, see Appendix B.

Data, Modeling, and Results

In this section, we present model results that were estimated using a subset of data from the NCDPS data set. Appendix C includes a description of how records were selected and a discussion of additional sampling to account for the overrepresentation in the NCDPS data set, where approximately 1.26 million conviction records represent approximately 524,000 individuals. Here, we focus on presenting general results from a subsample of \(N = 250,000\)

\(^1\) TensorFlow is an open-source software library for machine learning that focuses on training and inference of neural networks (Abadi et al., 2016)
people, related to events from 1995 to the present, that provides general insights on the performance of the model.

The model contains covariate information on individuals that includes

1. age-related information
2. information on the number of events that occur on the date of a record (i.e., number of convictions, how many of these convictions are felonies)
3. information on counts of offense types that an individual is convicted for on that day (e.g., larceny charges, sexual offenses)
4. minimum and maximum sentencing information for the set of convictions
5. partial penalty information (penalty information is specific to North Carolina and, to increase model generalizability, only a subset was included)
6. prison information that occurs between the date the record represents and the next conviction time
7. prior conviction counts
8. prior offense counts.

We do not include gender and race or ethnicity information in the model because these variables are not allowed in EEOC employment screens, even if such information is directly correlated with the outcome of interest (EEOC, 2012).

**Verification: Model Predictions Are Accurate and Fit the Data Well**

Overall, the model appears to fit the data well. Table 5.1 demonstrates the concordance index, or \( c \)-index, from the predicted survival estimates for three different points in time for each data set to illustrate performance of the estimated risk scores through time. The \( c \)-index is a measure of discriminatory ability of a risk score and provides an assessment of the number of “concordant pairs” when comparing pairs of predictions to observed conviction times (see Harrell, 2015, for an overview of diagnostic methods for survival analysis). Specifically, a pair of observations \( i \) and \( j \) will be concordant if \( T_i < T_j \) (the observed time between failures), and the estimated survival is \( S_i(t) < S_j(t) \), i.e., the higher-risk individual is observed to reconvict first. A \( c \)-index score of 0.5 implies that the model is no better than randomly guessing survival times; a \( c \)-index score of 1.0 implies that the model has perfect discrimination among individuals. Table 5.1 demonstrates that the model has a better-than-random discriminatory ability of approximately 0.7 (across all time points). This predictive ability generalizes to an independent test data set as well.
Table 5.1. Concordance Index Estimates of the Model Across Time

<table>
<thead>
<tr>
<th>Time Point Used for Survival Risk Estimates</th>
<th>Training Data</th>
<th>Validation Data</th>
<th>Test Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 year</td>
<td>0.707</td>
<td>0.703</td>
<td>0.706</td>
</tr>
<tr>
<td>5 years</td>
<td>0.709</td>
<td>0.706</td>
<td>0.709</td>
</tr>
<tr>
<td>10 years</td>
<td>0.709</td>
<td>0.705</td>
<td>0.708</td>
</tr>
</tbody>
</table>

The interpretation of a $c$-index of 0.7 depends on multiple factors. For instance, it can be difficult to fully understand the $c$-index because it is a single measure of the discriminatory ability of a model. The main point is to provide an assessment that the survival probabilities represent better-than-chance estimates of comparing individuals—that is, better than selecting an individual’s next conviction time randomly. In the next section, we explore a related concept of calibration that helps further contextualize the estimated $c$-index.

**The Model Is Well Calibrated**

Calibration attempts to assess how often individuals actually reoffend, conditioned on the estimated probability of reoffending, i.e., for a group of people with an average estimate of 40-percent one-year survival probability, it should be expected that 40 percent of them reoffend after one year (or possibly never). To assess the calibration of the estimates, we use the methods outlined in D’Agostino and Nam (2003) and Demler, Paynter, and Cook (2015) that assess calibration in censored data by leveraging the KM estimator and bins of estimated $x$-year survival probabilities. Figure 5.2 presents an example of a calibration curve for five-year survival probabilities for the training data, validation data, and independent test sets. It provides a clear visual display of calibration that also generalizes to an independent test data set. For example, we see that those with a predicted five-year survival probability of 0.5 from our model typically align with KM estimates among similar individuals where the survival probability is also 0.5. The results also present variation in predictions across wide time ranges. See Appendix D for additional calibration curves for one-year, ten-year, and 20-year survival probability estimates.
What the Model Might Be Telling Us: Insights About Important Factors in Risk

The results up to this point only provide information on the general “fit” of the model and do not inform insights that might be taken from the model. To further explore the model, we turn to a method called “permutation variable importance” (see Breiman, 2001, and Fisher, Rudin, and Dominici, 2019). The method randomly mixes the observations of specific variables and, if the performance of the model does not change, then the variables are deemed unimportant in the relationship. More specifically, the method evaluates the change in the loss function when groups of columns of covariates are permuted prior to making predictions. (The idea of permutation being that, for a given variable, such as “prior number of convictions,” two individuals will have their observations swapped, or permuted; this occurs for all observations in the data set for a given set of variables.) The intuition is that a set of variables will be unimportant to the model if evaluations of the loss function do not change under the permutation of observations from the selected columns. Logically, the method breaks the association between the selected covariate sets and the outcome while maintaining associations of the remaining columns. We performed this method for sets of variables and assessed the percentage change in the loss function (recall that we are using likelihood-based loss functions). We also provide an estimate for comparison if the outcomes (i.e., the next event times) themselves were permuted (i.e., we randomly permute the event times among the individuals and thus remove all associations between variables and the outcome in the data set).

In Table 5.2, the first row provides an upper bound of how much results can change if all of the event times were randomly assigned to sets of covariates, i.e., reassigning a time $Y_i$ from individual $i$ to an individual $j$ with covariates $X_j$. Specifically, it is creating an assumption that
there are no relationships between covariates and event times, and thus they can be swapped without degrading model performance. The performance of this model differs by approximately 20 percent, suggesting that the included variables are providing information about predicting event times and provides an upper bound on how much we can expect the loss to change. We can perform a similar assessment by permuting specific groups of columns as presented in the rest of the table. We see that current prison information changes performance the most (about 8 percent decrease in performance), followed by information related to current number of convictions (about 6 percent decrease in performance), noting that all results are conditioned on the fact that other variables are included in the model. The potentially surprising finding may be how little performance changes when offense information and prior incarceration information are permuted. We caution that these explanations are not to be interpreted causally, i.e., each variable may be important in another model if different covariates were chosen for prediction. These results are specifically related to the demonstration model presented in this chapter.

Table 5.2. Variable Importance from the Mixture-Model

<table>
<thead>
<tr>
<th>Rank</th>
<th>Groups of Covariates Permuted</th>
<th>Percentage Change in Loss Function (Negative Log Likelihood)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Permuted actual event times</td>
<td>20.39</td>
</tr>
<tr>
<td>2</td>
<td>Current prison information</td>
<td>7.62</td>
</tr>
<tr>
<td>3</td>
<td>Information on counts of events on date (convictions, felonies, misdemeanor)</td>
<td>5.87</td>
</tr>
<tr>
<td>4</td>
<td>Age information</td>
<td>3.71</td>
</tr>
<tr>
<td>5</td>
<td>Sentencing information</td>
<td>3.49</td>
</tr>
<tr>
<td>6</td>
<td>Information on counts of events prior to date (convictions, felonies, misdemeanor)</td>
<td>2.39</td>
</tr>
<tr>
<td>7</td>
<td>Information on prior offense types</td>
<td>2.31</td>
</tr>
<tr>
<td>8</td>
<td>Information on current offense types</td>
<td>0.95</td>
</tr>
<tr>
<td>9</td>
<td>Prior prison information: Prior prison time and number of prisons spells before this set of events (before this conviction and before the current prison spell)</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Limitations of the Model

Although the model is well calibrated, and the concordance results suggest an ability to differentiate risk at some level, there are a few limitations to this model and these limitations will be important for assessing the ultimate fairness of the model. The first limitation is that the model does not contain any information beyond that which concerns criminal justice, specifically, potentially important information that would demonstrate lower risk, ensuring fair assessments of risk. For example, differences in an individual’s socioeconomic conditions or post-release environment might be important for predicting an individual’s risk to recidivate.
Unfortunately, these data are not linked to this type of criminal history record information. Additionally, there may be moral or legal principles that obligate researchers to avoid certain information, such as socioeconomic conditions because they may be predictive because of known bias. More generally, we believe that there is room for continuing dialogue about the appropriate variables to include in a model. As noted in the EEOC’s 2016 meeting on big data, algorithmic models have significant flexibility about what variables can be added to a prediction model. The EEOC itself expanded the list of viable variables in its most recent guidance to employers on criminal background checks.

In the employment context, the additional factors not included in the first step of the formal background check are brought into play through an appeals process initiated by a person challenging an employment decision. It can take place during a personalized assessment process adjudicated by a decisionmaker at the firm who tries to take into account this new information. Although people who appeal may be fundamentally different than those who do not (Siwach, 2017), some evidence indicates that the experts on the appeal panels might not be able to usefully distinguish risk levels of those who do appeal (Denver, 2020). It is possible that such factors could be brought into the first-stage risk-prediction tool, although incorporating these factors makes both model-building and the background check more difficult to do because these data are more difficult to acquire. A second strategy might be to build a risk-prediction model for those who appeal. However, in a practical sense, this is a much harder population to define because it is hard to know a priori who is going to appeal conditional on the first decision made on the basis of criminal history records alone. As a result, any attempt to create risk prediction models may need to be very context specific.

Finally, the model we present in this chapter provides only point estimates and does not provide confidence intervals to further inform risk estimates from the model. To fully demonstrate that an individual’s risk assignment is fair, risk estimates should come with an estimate of the uncertainty, or confidence, in this assessment. This should be explored in future work.
Workshop Summary

On June 16, 2021, the RAND Corporation research team hosted the virtual workshop “Rethinking How Employers Understand Risk in Background Checks.” The goal was to bring policymakers, social scientists, organizations that specialize in helping people with records become employed, employers, legal professionals, and consumer reporting agencies (CRAs) together to discuss how our research (as documented in this report) could impact their fields. The workshop focused on three key questions related to this research:

- What criminal history information should employers consider during the hiring process?
- How can information generated during the time gap between a candidate’s last conviction or release from prison and when a candidate applies for a job be used to improve employers’ understanding of a candidate’s recidivism risk?
- What factors should employers consider during the background check?

The “Questions and Answers” section of this chapter summarizes the discussion on these issues.

Workshop Sessions

Jeremy Travis, executive vice president of criminal justice at Arnold Ventures, opened the workshop by contextualizing the research on the topic of background checks. Charlotte Burrows, chair of the EEOC, then delivered the keynote address, which is summarized in the next section of this chapter. Nina Hicks, the Detroit director of the Center for Employment Opportunities, provided insight about the discouraging effect of background checks on job searches for those who are involved in the justice system and are trying to move back into the labor market. She emphasized how a criminal conviction, particularly a felony conviction, essentially obscures any credentials from before the conviction and dominates any actions taken after the conviction that could demonstrate that a person has moved on from the activities that led to their conviction.

Following these speakers, we presented our research and answered questions from attendees using the question-and-answer feature of the web workshop. The workshop then held three panel discussions that focused on the research project’s social and policy implications, effects on employers and employees, and data implications. Finally, Jocelyn Fontaine, vice president of criminal justice research at Arnold Ventures, gave closing remarks. The workshop agenda, including the names and biographies of all speakers, can be found in the “Workshop Agenda” section of this chapter.
Keynote Address

EEOC chair Burrows delivered the keynote address. After graduating from Yale Law School, she clerked at the U.S. Court of Appeals for the Third Circuit, worked in private legal practice, and worked on Capitol Hill on a variety of legislative initiatives. She went on to serve as Associate Deputy Attorney General at the U.S. Department of Justice. Burrows was appointed to the EEOC in 2014 and named chair in 2021. Burrows’s full biography is provided via a link in the workshop agenda at the end of this chapter (Table 6.1, with minor edits from the original agenda).

The EEOC works to ensure that background checks are not used to deny equal employment opportunities to individuals with criminal records, which is particularly important given the number of people with criminal histories in the United States and the number of people currently in the U.S. prison system. Burrows emphasized that the issues discussed in the workshop concern civil rights and justice, and she highlighted the importance of creating screening policies that do not disproportionally impact individuals of certain races or ethnicities. She emphasized that insuring a lack of bias in screening policies is key to ending systemic discrimination and removing structural barriers to justice that can last generations. She also suggested that employers would benefit by gaining more access to this population of workers.

Key Workshop Questions

This section distills three key questions that arose in the workshop related to this research:

- What criminal history information should employers consider during the hiring process?
- How can information generated during the time gap between a candidate’s last conviction or release from prison and when a candidate applies for a job be used to improve employers’ understanding of a candidate’s recidivism risk?
- What factors should employers consider during background checks?

Question 1: Criminal History Information and the Hiring Process

EEOC guidance states that employers should use conviction data, not arrest data, in hiring decisions. However, all the recidivism research cited in the guidance uses criminal conviction information to predict the next arrest, not the next conviction. This reality is driven by a criminal justice system that views arrest as a relevant concern. As noted in this report, this focus on arrests has always been a bone of contention among advocates in this space, who have argued that if arrest information cannot be used to make decisions, then policymakers should not rely on research from criminal justice that uses arrest as the main outcome. At some level, this is a simple argument based on the legal claim, made by the EEOC, that arrests do not establish that criminal conduct has actually occurred. If this is true for prior records, it should also be true when trying to predict risk. This report also asserts that the criminological literature has shown that there is less evidence for racial bias in convictions than in arrest records. In keeping with the
goals of promoting civil rights and civil justice, reducing bias in any risk models used to inform hiring decisions is critical. The recidivism risk-prediction models that are based on the reset principle described in this report thus do not include arrest information. We received no negative feedback on this choice during the workshop.

Panelists pointed out that the structures currently in place for background checks, however, may include non-conviction information that negatively impacts job candidates with a criminal record. For instance, data repositories and CRAs (private companies that search for criminal history record information from public sources and provide them to clients, such as employers, for background checks) are not currently set up to exclude non-conviction information that may unfairly affect hiring decisions. As a result, employers often see this information even though they officially cannot use it. Roberta “Toni” Meyers Douglas, director of State Strategy and Reentry at the Legal Action Center, said that these data issues affect even employers who narrow their focus to the date of a candidate’s last conviction because employers often have access to other information, including arrest histories. This creates a problem of unconscious bias, where employers have to “unsee” arrest information that might cause them to question a candidate’s fitness for a position.

Multiple panelists and speakers said that, in order for risk-prediction models that adhere to the reset principle to succeed, innovation is needed in the data and policy environments. These discussions highlighted concerns about data quality that were not featured heavily in this report but have been highlighted elsewhere (e.g., Lageson, 2020). Meyers-Douglas went on to explain that there are a multitude of ways in which employers access sources of criminal record information. To help job applicants prepare for the interviewing and hiring process, she said that the Legal Action Center works with candidates to determine what those sources are and where there might be potential inaccuracies that they should be ready to fix if they are informed by the employer that their application has been rejected because of concerns about the criminal history record.

Sam Schaeffer, CEO of the Center for Employment Opportunities, emphasized the importance of companies carefully considering the information they need when making hiring decisions. He noted that if employers rely less on criminal history record information and more on information that describes the positive steps individuals take to “go straight,” more formerly incarcerated individuals will have immediate access to work and a paycheck, aiding their reintegration into their communities and potentially benefiting employers who are sometimes desperate for employees.

Toward the end of the workshop, we asked attendees to fill out an optional survey asking about their industry, thoughts about the report, and the workshop. When asked about their primary concerns or priorities around the use of criminal background checks, attendees overwhelmingly responded that they worry about fairness, accuracy and completeness of information, and uniformity of how records are maintained and disseminated across the country. These are valid concerns that affect employers’ hiring decisions. As stated, missing or
incomplete records can impact the accuracy of risk-prediction models that satisfy the reset principle. In sum, workshop participants noted the danger of employers considering information, such as arrest records, that is neither predictive nor legally relevant. The use of this information could potentially harm both applicants and employers.

**Question 2: Information During Gap Between Last Conviction or Release and Applying to a Job, Use to Improve Employers’ Understanding of Recidivism Risk**

The discussion about this question emphasized the idea that, after a certain period of time, the risk level of someone with a record can match that of someone without a criminal history. Jennifer Yeh, vice president and general counsel at Checkr, noted that current policy frameworks tend to lead employers to focus on the time of an applicant’s conviction or when the crime was committed, not what has happened since (a key focus of models that satisfy the reset principle). Current record-compilation efforts focus on finding out when the offense occurred, and there is not currently a policy motivation to look beyond that. The industry itself, she explained, does not have the experience or the direction to look for the information that the reset principle requires. Yeh also explained how difficult it is for CRAs to make use of incarceration data that are available from public-facing websites hosted by most states through their departments of corrections. The main challenge involves linking incarceration spells to conviction data. As a result, Yeh did not believe that CRAs would be able to provide credible data on incarceration spells for the foreseeable future.

David Roberts, executive director/CEO at SEARCH Group, Inc., elaborated on this idea. He pointed out that neither state repositories nor CRAs provide information about prison release or behavior on parole or probation. According to Roberts, both data providers and employers will struggle with implementing models or tools that satisfy the reset principle.

Adam Dean, a director at the New York State Division of Criminal Justice Services, pointed out that there is no system to track whether a person is on probation, in a halfway house, or back in the community. Dean expressed real doubts that this information would be included in repositories at any time in the immediate future. This creates a challenge in providing information that could describe how someone has spent their time while free. The reset principle argues that the time since release or conviction is the key thing that can predict risk—knowing what is happening during this stage (information that is not available in standard sources of criminal history record information) would dramatically improve the ability of researchers and employers to make intelligent decisions about risk. In this report, we argue that it is important to create new data infrastructure to support policies that respect the reset principle. The workshop discussion focused on this point by acknowledging that the current sources of criminal history record information do not meet the demands of the reset principle.
Question 3: Factors That Employers Consider in Background Checks

Employer decisions on whether to hire take into account both a candidate’s criminal history and such factors as their own risk tolerance for issues that include potential negligent hiring lawsuits and harm to brand reputation. Debbie Mukamal, executive director of the Stanford Criminal Justice Center at Stanford University, noted that many companies want to take on as little risk as possible for fear of negligent hiring lawsuits. From this perspective, avoiding negligent hiring lawsuits becomes the decision driver, not avoiding recidivism risk. If negligent hiring lawsuits are not driven by an accurate assessment of risk (Bushway and Kalra, 2021), then risk assessment itself is not the key to solving the problem. Instead, the goal needs to be structuring negligent hiring liability so that companies can tolerate more risk in their hiring practices.

Regarding factors related to a candidate’s personal history, Schaeffer supported the idea of investing in transitional work programs. He stressed that companies need to take into account not only the information in an applicant’s background check but participation in transitional work programs and other community services. Rod Fliegel, co-chair of the Privacy and Background Checks Practice Group at Littler Mendelson, explained how he talks his clients through background checks as a factor in hiring decisions. His starting point was that many companies could benefit from tapping this pool of workers if they could find ways to reliably identify the less-risky members in this pool. His clients often face worker shortages and are eager to increase the diversity of their workforce. He pointed out that although he recognized that many of his clients are legally constrained in their ability to hire individuals who have committed certain crimes (e.g., financial companies who cannot hire people with prior felonies), he encourages them to find ways to bring formerly incarcerated individuals into the workforce when it is legal to do so.

Fliegel further elaborated that he encourages his clients to look beyond background checks and, where possible, consider what applicants have been doing in the time since their release. It would be very helpful, he noted, for background checks to confirm if an individual has been employed or otherwise productive since release. This insight fits well with the reset principle that is outlined in this report.

This discussion highlighted how a focus on the time since release and conviction and the background check immediately shifts the focus of discussions about an applicant from criminal history to post-release behavior. This is precisely the point of the reset principle. The background check should be based on the risk that exists at the time of the background check, which requires information about criminal history and more current information. Unfortunately, more-current information is both more difficult to gather than criminal history records and more difficult to include in a risk-prediction model. Additional work is needed to think through plausible mechanisms whereby more time-proximate information can be added to risk models that satisfy the reset principle.
The Reset Principle and Background Checks

David Lopez, co-dean of Rutgers Law School, pointed out that the reset principle research is groundbreaking because it fundamentally challenges the underlying jurisprudence governing background checks and hiring, which is based on the three factors first established in the Green case. The court recognized the three main factors that should be used by employers when considering a candidate with a criminal record history: type of job, type of crime, and time since last conviction.

Lopez asserted that this research was important because it directly challenges the legitimacy of two of these factors—type of crime and type of job. He mentioned that challenges to these factors were not highlighted as much as could have been in our report. We find, like other researchers before us, that type of crime is simply not very predictive of future convictions. Lopez observed that, if employers are truly predicting risk of recidivism—even recidivism for certain serious crimes—then our research suggests, rather firmly, that the type of crime factor emphasized in the Green factors is not legitimate.

In addition, Lopez discussed how this report indicated that type of job factor does not help to predict risk of recidivism. He saw the importance of our argument that the type of job should only be relevant for choosing the risk threshold an employer applies to hiring for a certain position. This perspective asserts that a person who might be an acceptable risk in one environment might not be an acceptable risk in a different job. As a result, although the type of job is still relevant for setting the cutoff threshold, it is not a factor that estimates risk.

The final Green factor is time since last conviction/release from prison. Lopez recognized that this research clearly prioritizes this element and asserts, in part, that the current emphasis on time since last conviction is not strong enough. This assertion comes from the fact that the previous research cited by the EEOC centered on the conviction and release event instead of the background check. Once the decision is framed around the background check, time since last conviction or release from prison becomes even more predictive. At the same time, information about how time since conviction was spent (particularly since release from prison) becomes even more important.

Risk of Recidivism Versus Other Factors

The discussion about type of crime brought up an important issue that bears on factors employers use to make hiring decisions. This report focuses on risk of recidivism, but participants said that employers think that type of crime is the most important factor in background checks—even though it is not predictive of recidivism. According to these speakers, employers are connecting type of crime to extrinsic considerations, such as risk of negligent hiring liability and the effect of potential criminal acts by employees. The discussion strongly suggested that, under the current policy constraints, the report’s consideration of recidivism risk
should be only one element in a broader discussion about how employers should make decisions about those with criminal history records.
Table 6.1. Workshop Agenda

11:00 a.m. Welcome Remarks
Lee Remi, project manager, RAND Corporation
Shawn Bushway (https://www.rand.org/about/people/b/bushway_shawn.html), senior policy researcher, RAND Corporation

11:05 a.m. A Word from Arnold Ventures
Jeremy Travis (https://www.arnoldventures.org/people/jeremy-travis), executive vice president of Criminal Justice, Arnold Ventures

11:20 a.m. Keynote Address
This talk discussed the history of EEOC’s guidance in this area as well as some of the EEOC’s litigation victories in this space. The talk also highlighted some of the current challenges facing the EEOC around the current guidance and potential future actions by the EEOC.

11:45 a.m. Reshaping Hiring Practices: Creating High-Quality Talent Pipelines, Job Access, and Career Mobility
Nina Hicks (https://www.linkedin.com/in/nina-hicks-597a8918), Detroit director, Center for Employment Opportunities
Hicks addressed the challenge of background checks for people with criminal history records.

11:55 a.m. Short Break

Noon Plenary: The Reset Principle: Rethinking How We Understand Risk in Background Checks
Nidhi Kalra (https://www.rand.org/about/people/k/kalra_nidhi.html), senior information scientist, RAND Corporation
Brian Vegetabile (https://www.rand.org/pubs/authors/v/vegetabile_brian_g.html), statistician, RAND Corporation
Shawn Bushway (https://www.rand.org/about/people/b/bushway_shawn.html), senior policy researcher, RAND Corporation
This panel presented the RAND report that was to be discussed in the workshop.

12:55 p.m. Short Break

1:00 p.m. Panel 1: Social and Policy Implications
Moderator: Shawn Bushway, (https://www.rand.org/about/people/b/bushway_shawn.html), senior policy researcher, RAND Corporation
Panelists:
Debbie Mukamal, (https://law.stanford.edu/directory/debbie-mukamal), executive director of the Stanford Criminal Justice Center, Stanford University
Derek Cohen (https://www.texaspolicy.com/about/staff/derek-m-cohen), vice president of policy, Texas Public Policy Foundation
Sam Schaeffer (https://ceoworks.org/leadership/sam-schaeffer), CEO, Center for Employment Opportunities
This panel discussed the RAND report from the perspective of policy makers and advocates interested in getting more people with criminal history records into the labor market.
1:55 p.m.  Short Break

2:00 p.m.  Panel 2: Effects on Employers and Employees
Moderator: Esta Bigler (https://www.littler.com/people/rod-m-fliegel), director, Labor and Employment Law Program, Cornell School of Industrial and Labor Relations
Panelists:
Rod Fliegel (https://www.littler.com/people/rod-m-fliegel), co-chair, Privacy and Background Checks Practice Group, Littler
David Lopez (https://law.rutgers.edu/directory/view/dl903), co-dean, Rutgers Law School
Roberta “Toni” Meyers Douglas (https://www.lac.org/about/our-team/roberta-meyers), director of State Strategy and Reentry, Legal Action Center
Alex Alonso (https://www.shrm.org/about-shrm/Pages/AlexanderAlonso.aspx), chief knowledge officer, Society for Human Resource Management
This panel discussed the RAND report from the perspective of employers and employee advocates, with special attention to the legal issues raised by the report.

2:55 p.m.  Short Break

3:00 p.m.  Panel 3: Data implications for Courts, CRAs, and Repositories
Moderator: David Roberts (https://www.search.org/staff-bios/david-roberts), executive director/CEO at SEARCH Group, Inc.
Panelists:
Jennifer Yeh (https://www.linkedin.com/in/jennifer-yeh-3aab3a128), vice president and general counsel, Checkr
Richard Schauffler (https://www.linkedin.com/in/richard-schauffler-b47b703b), principal, Justice Solutions TRS, LLC
Sarah Lageson (https://rscj.newark.rutgers.edu/people/faculty/dr-sarah-e-lageson), assistant professor, Rutgers University
Adam Dean, director, Office of Criminal Justice Records, New York State Division of Criminal Justice Services
This panel discussed the data demands of the RAND report and reviewed the capabilities of the current data systems to meet these demands.

3:55 p.m.  Attendee Survey

4:05 p.m.  Closing Remarks
Jocelyn Fontaine (https://www.arnoldventures.org/people/jocelyn-fontaine), vice president of Criminal Justice Research, Arnold Ventures

4:15 p.m.  Adjourn
7. Providing Another Chance: The Reset Principle’s Implications for Policy

People desist from crime and, as a result, their risk of recidivism can change dramatically over time after a conviction or release from prison. Criminal background checks that take into consideration prior history need to reset to account for this changing risk because not everyone is at the same stage in their desistance process at the time of their background check. In Chapter 3, we described the reset principle that we argue should inform criminal background checks. This principle reanchors the risk prediction to occur at the time of the background check rather than at the time of the individual’s most recent involvement in the criminal justice system. We specifically focused on the employment context, although similar models could be used in other contexts, such as housing or volunteer screening. In Chapter 4, we discussed empirical considerations for researchers interested in building a risk-prediction model. In Chapter 5, we demonstrated the viability of creating a risk-prediction model that respects the reset principle using data from North Carolina. We show how these models could be accurate (i.e., calibrated) and validated on a separate test sample. The model-building in Chapter 5 provided a concrete demonstration of what we mean by model development and validation and how these ideas could be used to satisfy the reset principle. This exercise in model development should guide discussions around the development and use of these tools into the future and can guide policy decisions in this space. In Chapter 6, we reviewed the discussion during the workshop on the report. The reset principle and its implications resonated with the audience, who almost immediately began to grapple with what the reset principle means for current policy and the next steps that need to be taken to further develop this idea. The workshop participants also made clear that employers were concerned not only about recidivism risk but also about other avoidable costs. This is an important observation that should inform the use of recidivism risk estimation moving forward.

Primary Implications for Policymakers and Researchers

Here, we discuss some of the important implications and takeaways from this report for policymakers and researchers, who are interested in continuing to develop new and better risk-prediction models for background checks.

The Majority of People Who Get Convicted Are Not Reconvicted

Our report used a large data set from North Carolina with 20 years of follow-up information for recidivism and detailed information on prison incarceration. We show for a sample of people with convictions what Rhodes et al. (2016) showed for a sample of people released from
Most people do not reoffend again. The result is driven by focusing the analysis on the population of interest—people with a criminal history who are available for background checks and are not incarcerated. This result should move policymakers and practitioners away from the oft-cited finding, based on the well-known BJS recidivism analysis of a given prison release cohort, that (nearly) everyone fails (Alper, Durose, and Markman, 2018). In Chapter 3, we used North Carolina data to show that a person’s likelihood of reconviction declines rapidly over time, and even people with high risk of recidivism after a conviction may transition to levels of risk similar to those initially deemed the lowest risk. These results largely replicate work by other scholars who use arrest rather than conviction as a measure of recidivism (Bushway, Nieuwbeerta, and Blokland, 2011). The specific model developed in Chapter 5 provides a model that could be used to ensure that originally high-risk individuals can signal to employers through their behavior after a conviction that they are actually less risky than others.

**Models Can Be Used to Predict Risk of a New Conviction**

The EEOC worried that it would be impossible to develop a model for background checks that could be validated under the Uniform Guidelines. Some of the ambiguity centers on what exactly the employers should be predicting. The appeals panel in *El v. SEPTA* made it clear that it believed that recidivism risk itself is of concern for employers, independent of job performance. The EEOC (2012) showed a preference for a reliance on job performance but allowed models developed under the second circumstance of the Uniform Guidelines to use recidivism risk to guide model development. Of course, it is possible that employers might be concerned only about recidivism that occurs on the job. However, it is not hard to think of examples in which employers care about convictions that occur outside the job. For example, someone who works as a contractor for the federal government might lose their clearance or suitability rating after a conviction, regardless of whether the violation occurred on the job. In this report, we proceed as though the risk of a new conviction for a nontraffic offense was a valid concern for employers, even if it does not occur on the job.

One implication of this discussion is that the 1978 Uniform Guidelines might need to be updated (Maurer, 2020). There are two key issues: First, the guidelines did not consider recidivism risk as independent of job performance, but federal courts have declared that recidivism risk is a separate issue that deserves consideration. The research suggests that employers are concerned about convictions in general (Bushway and Kalra, 2021). To the extent to which this new classification is valid, it does not fit neatly into the box of job performance and may require additional clarification and conceptual development.

Second, the Uniform Guidelines did not envision the kinds of formal prediction or algorithmic models now being used in hiring. In this report, we consider prediction models for recidivism, but the same types of models are now being used to predict job performance. This is a major change from the environment in 2012, when these kinds of prediction models were unknown in the hiring context. That we are predicting convictions is a separate issue. However,
the good news is that model calibration and validation are well-developed methods in statistics. The EEOC can simply update the Uniform Guidelines to recognize this new class of models (Maurer, 2020).

**Risk-Prediction Models Do Not Define an Absolute Threshold for Acceptable Risk— Setting Acceptable Risk Thresholds Requires Policy Discussion and Decisions**

We have presented risk models that can compare the recidivism risk of different individuals. The next step is for an employer to define whether the level of risk is acceptable or low enough for a particular job. Recidivism risk models do not include characteristics of the job as a parameter because recidivism risk is not job-specific. Instead, characteristics of the job should define the threshold of acceptable recidivism risk. The risk cutoff—the level of acceptable risk—will also be a feature of the risk tool, and it needs to meet the same business necessity rules as other features of the employment screen. This cutoff is not determined by the tool but by the employer. Workshop participants highlighted the important innovation of the movement of job type from a factor that predicts risk to one that helps determine the level of acceptable risk.

The level of acceptable risk will depend on the nature of the job and the labor supply for an employer (DeWitt et al., 2017). Employers hiring individuals for well-supervised positions in which the employee has limited contact with clients might prefer a different cutoff level of risk than those whose employees are unsupervised and interact with clients (Berk, 2011; Bushway, 2013). Beyond determining the number of applicants who are eligible for a job, the choice of thresholds will also drive racial-equity assessments of the final employee pool. Without a decision about what that threshold would be, it is difficult to evaluate the effect of such models as those we developed on equity in hiring. The predicted risk from this kind of model is not uniformly distributed among subgroups, so different cutoff points may create different relative distributions of eligible people by race or ethnic groups. These different distributions then need to be evaluated according to different fairness and equity rules that reflect different and, at times, conflicting normative judgments about the precise meaning and definition of fairness and equity (Barocas and Hardt, 2017; Corbett-Davies et al., 2016; Dwork et al., 2012; Harris, Goss, and Gumbs, 2019; Mitchell et al., 2021; Pleiss et al., 2017).

Because our work does not set acceptable cutoffs for risk, to this point, it does not directly engage with questions of equity and bias that arise in much of the discussion of such statistical models. Those considerations must, however, be integrated into the development of tools that apply these methods. In addition, the promise of these statistical models is that they will be more fair and equitable compared with current methods (Bushway, 2020). An assessment of the type

---

1 Consider a cutoff very close to zero, where only people with no risk are considered for employment. Then consider another cutoff very far from zero, where only people with very high risk are eliminated from employment. These two rules will have very different racial profiles unless the distribution of risk by race is the same throughout the entire distribution of risk, a situation that is unlikely to occur.
of model we developed requires a comparison with the current state. We discuss this issue in the next section.

Data Quality Can Limit the Development of a Successful Recidivism Risk Model—Data Infrastructure Should Be of Primary Concern for the Equitable Development and Use of These Tools in the Future

Currently, most risk tools used in employment are created by consulting with experts about perceived risk. As a result, there is no infrastructure designed to collect well-curated data for the purpose of predicting the risk to recidivate for previous offenders. The data that are available were not created with this purpose in mind. Cleaning data to create models from secondary data sources for assessing risk to recidivate is a complex task. The data-cleaning step requires defining states, events, and the appropriate endpoints for time comparisons. In the case of the NCDPS data set used in this report, the data were originally collected for tracking offenders and inmates and for reporting on current trends in crime within North Carolina. The data are missing information on when individuals are in county jail, demonstrating how even very high-quality data can be missing important covariates for predicting the risk of another conviction. The process for creating a quality data set on which to build a model is complex. This may limit the utility of the model in the future, and future policy discussions should center around developing infrastructure that focuses on prospectively collecting data for addressing recidivism in the background-check setting.

There are not many existing data sets that could be used to create risk models for conviction events. It is our impression that the NCDPS data is unique in its ability to allow both tracking convictions through time and tying those convictions to incarceration spells. This and the aforementioned points suggest that efforts to create models for employment may require new initiatives to construct data sets with this intended goal in mind. Data sets should capture both prison and jail information and should be connected to important predictors on risk. We note that these types of data undertakings are difficult and, in many cases, can be considered by society as governmental overreach in data collection. There is a tension between requiring high-quality data sets to adequately predict risk and the construction of a surveillance infrastructure that violates personal privacy considerations. Again, this policy discussion is beyond the scope of this report, but we note that these models require high numbers of observations collected over extended periods of time to be useful for adequate prediction. This fact is at tension with their ultimate use.

---

2 For context, see the data documentation provided for the Criminal Justice Administrative Record System (CJARS), an effort to combine administrative criminal justice data with data from the U.S. Census Bureau. The researchers who built CJARS have had to address many of the same issues that we addressed in building a usable data set (Finlay and Mueller-Smith, 2021).
This tension is made all the more relevant by the litigious environment around background checks. Companies and other groups that conduct background checks have strong incentives to shield their data and decision rules to limit liability risk. But the best environment for developing valid and useful tools may require more transparency and shared data. One possible solution might be to move to a model where the government takes on more of the responsibility for background checks and, therefore, ultimately builds the risk-prediction models based on government data. Groups that follow the government recommendations could be offered indemnification from negligence lawsuits (Bushway and Kalra, 2021). Further discussion about the best way to build valid models for background checks should include discussions about the best institutional frameworks for these checks, in addition to discussions about the best models and data sets.

*Prediction-Based Models Are Associational, Not Causal; They Reflect Historical Trends on Recidivism to Predict Future Outcomes Grounded in a Particular Place and Time*

Factors used in these models can exhibit complex relationships, and different factors can matter depending on data availability. One reason we undertook the complex modeling approach in our research is that is not easy to *a priori* to know the relationships among variables and thus either specify empirical Burgess models or design regression models by hand that will be the most predictive of the risk to recidivate. Similarly, the variables in the data may interact in ways that are not obvious to researchers. That means that the ability to assess these interactions will depend on the variables that are at researchers’ disposal when creating models. Similarly, the predictive power of variables in a model will be conditioned on the type of model chosen and other variables in the data set. Therefore, it is difficult to both understand the model and create comparisons among models that are fair.

That said, in our model, information on time served in the most recent prison spell was among the set of most important predictors of recidivism risk. Most criminal record checks used by employers do not have information available about time served in prison or jail. Technically, the EEOC does not list prison time as a relevant factor, but we believe that the EEOC was assuming that time since event would, in fact, take into consideration the time spent in prison. We base this claim on the fact that the title of the relevant section in the 2012 guidance is “The Time That Has Passed Since the Offense, Conduct and/or Completion of the Sentence” (EEOC, 2012). There is no discussion about what the relevant frame of reference is or whether this information was available, but in the end, the lack of information about prison sentences in most background checks is a major hurdle for further progress in this space. The considerations discussed in Chapter 4 and the model estimated in Chapter 5 suggest that the lack of information on time served in prison will cause substantial errors in the risk-prediction exercise. Data-collection efforts designed to support the development of recidivism risk models for employment must include information on time served in prison and, to a lesser extent, jail.
Our approach also found that offense type was not particularly important in the model for predicting recidivism risk (given the other factors present, such as sentencing information). In some ways, this is not particularly surprising. Crime type has not been found to be a particularly strong predictor of recidivism in criminal justice models (Bushway and Kalra, 2021). Future research should further interrogate the role of crime type.

Use of a Risk-Prediction Model in Employment or Other Contexts Could Alter Its Performance

The risk models that we developed do not currently influence the observed rates of recidivism. However, if the model we developed based on North Carolina data were to be developed into a tool used to inform hiring in North Carolina, the predictive ability of the model could change (Bushway and Smith, 2007; Kleiman, Ostrom, and Cheesman, 2007). This is because employment affects recidivism, and this model could ultimately affect who gets hired. As a result, it would be neither surprising nor problematic that the risk-prediction model would need to be updated if it were used to make decisions systematically. This is not a bug but rather a feature of signaling systems that use information to systematically make inferences about an unobservable characteristic, such as recidivism risk or desistance (Bushway and Apel, 2012).

Consider the application of a risk model in the context of European background checks. Western European governments limit the power to conduct background checks to the same government agency that collects information about criminal history (Jacobs, 2015). Employers with authorization can ask the agency to verify that the person can work in a given job. The agency provides a yes or no decision without having to provide the actual criminal history. Currently, these decisions appear to be driven by simple waiting periods that are conditioned on the nature of the offense and the job in question (Jacobs, 2015). Suppose instead that the agency did a risk-prediction exercise and only people who met a certain level of risk passed, with the level of risk depending, in part, on the job. If these characteristics are static, then there will be nothing that the applicant can do except wait. However, if dynamic characteristics are included, then applicants interested in employment could take steps to reduce their perceived risk. In other words, the incentives to acquire those signals will change, thereby changing their value for discerning risk.

There could be a nefarious element to this. For example, in the hiring context, firms are currently looking at which elements of a résumé are correlated with job performance. Savvy applicants could be taught to add these elements to their résumés strategically. In the long term, researchers and policymakers should expect to update any risk-prediction models both to guard against gaming and take into account the dynamic relationship between observable behavior and risk that is itself affected by decisionmakers using the risk-prediction tool. This further emphasizes the need for data systems and infrastructure that can facilitate model and tool updating.
Exploring and Stressing Risk Prediction Models for Bias Will Be Crucial

Risk recidivism models that are organized on the risk principle are useful because they provide a mechanism through which an individual who is declared high risk at their last release or conviction can demonstrate how their risk level has changed over time. The models also begin a conversation on ways that this change in risk can be demonstrated mathematically. Like any decision rule, however, it could create bias. Tools developed based on models of this type need to be evaluated for fairness throughout each iteration of their life cycle. As discussed, questions of fairness can help inform which variables are used (e.g., arrest versus conviction), but ultimately the fairness of any given tool relies on many different factors. It is largely for this reason that the model outlined in Chapter 5 is a starting point for discussion, not a tool for immediate use. Future work should also further center these tools and concepts within concepts of algorithmic fairness and algorithmic equity (Barocas and Hardt, 2017; Dwork et al., 2012; Mitchell et al., 2020; Pleiss et al., 2017).

In each case, these tools need to be compared with standard practice. We are aware of few cases where standard background checks are evaluated for bias (for an exception, see Kurlychek, Bushway, and Denver, 2019). Moreover, there is evidence that current rules that rely heavily on crime type could, in fact, create racial bias without adding value in terms of risk avoidance. The promise of these tools is that they can improve on current practice, not that they will achieve perfection (Siwach, Bushway, and Kurlychek, 2017). It is easier to debug an algorithm than human decisionmaking, in part because algorithms are explicit and ultimately transparent (Mullainathan, 2019). The need for new tools is urgent given the number of people who have convictions in the United States, and we look forward to an active discussion about the fairness and equity of models and tools anchored in the reset principle.

Risk of Recidivism Is Not the Only Factor That Employers Consider When Setting Standards for Hiring While Using a Criminal Background Check

Employers consider other things besides the risk of a new crime, including the reputational risk associated with hiring someone who has damaged the community, statutory or licensing restrictions, and negligent hiring concerns. Although the latter two are ostensibly tied to recidivism risk, they are often not informed by recidivism research and therefore may have features that clash with recommendations of an approach based on the reset principle. The most obvious conflict comes from reliance on type of crime, which features heavily in current practice but is not particularly informative for risk prediction. This does not mean that employers should not consider crime type. It just means that they should explicitly link these considerations to some other concern besides recidivism risk. In other words, an employer might have a multistep decision process in which they take into consideration recidivism risk in one step and then take into account legal concerns about brand value, regulatory compliance, and liability concerns. If these latter concerns change, the decision rules can be changed without conflating these other
legitimate business concerns with the legitimate business concern of recidivism. Distinguishing between the reasons for the decision is only a first step. Future research is needed to understand how these competing concerns can be effectively integrated into a coherent decision framework, particularly if a concern, such as negligent hiring, uses a biased perception about actual recidivism risk (Bushway and Kalra, 2021).

Final Thoughts

This report represents both a culmination and an initiation. Following the Third Circuit ruling in *El v. SEPTA*, researchers and policymakers began to grapple seriously with the limitations of criminal justice research on recidivism for policies governing background checks. This ruling noted, pointedly, that the recidivism research used by the defense followed people for only three years after they were released from prison. Although this research was appropriate for use by criminal justice decisionmakers who were concerned about behavior following release from prison (its original purpose), it was not appropriate for employers who were trying to make decisions ten or more years after someone had been convicted or released from prison.

The short-term effect of the *El v. SEPTA* case was a rapid increase in research on recidivism that was more relevant for the background-check process. However, this research still used the sampling frames, modeling architecture, and variable measures initially created for the criminal justice context rather than the background check context. This report’s development of the reset principle makes a clear statement that there needs to be a new start for research on recidivism to help employers and others make informed decisions about whom to hire and to reimagine, from the beginning, the relevant sampling frames, modeling architecture, and variable measures that are appropriate for the background-check question. This is, in a sense, the endgame of the movement started by *El v. SEPTA*’s skepticism about the relevance of criminal justice research on recidivism for the background check context.

However, we would be beyond naive if we claimed to have solved all the problems or created a clear solution for employers and others who would like to make better background check decisions. We are well aware that much more work needs to be done. At that same time, we are confident that this report provides a useful roadmap for future research and policy development in this space. We look forward to seeing both research and policy around background checks evolve using samples, models, and data that are created specifically to inform the decisions facing the thousands of employers who are conducting millions of background checks every year.
Appendix A. Additional Data Considerations

North Carolina Data Structure

Throughout this report, we illustrated the importance of the principle that risk must be reset at the time of the background check using a data set on criminal sentencing events from North Carolina. In this appendix, we outline some issues that we encountered that could be useful for researchers constructing data sets.

Recall that the data set is from the NCDPS Offender Public Information Search downloads webpage\(^1\) and represents public information related to convictions, sentencing events, and prison incarceration spells dating back to 1972.\(^2\) The portal provides access to the data files, and the files underlying the database can be downloaded for use by researchers. Of interest to our analysis are the files listed at the NCPDS download portal pertaining to the “Offender Profile (OFNT3AA1),” “Sentence Component (OFNT3CE1),” and “Sentence Computations (IMT4BB1).” For our analysis, we restricted the data to cover events occurring after 1995, events for individuals who are between 18 and 70 years of age, and people who were born in North Carolina. Finally, the analysis is focused on records through April 8, 2021, the date on which we downloaded the data.

\textit{Structuring the Data for Analysis}

The data were designed and structured for tracking the criminal history of people convicted in North Carolina, not for predicting risk. In this subsection, we will illustrate three ideas related to the complex longitudinal structure of criminal history data:

1. How convictions “cluster” on particular days and how this affects analysis.
2. How the longitudinal structure of the data necessitates complex decisions in defining different discrete “events” (conviction, release, etc.) and continuous periods of different “states” (e.g., in prison versus not in prison) that must be parsed to adequately define the passage of time.

---

\(^1\) From NCDPS, undated-a:

The files . . . are provided to facilitate users interested in data analysis of NC DPS Division of Adult Correction offender related data. Users may download these database files which can then be imported into their own database software. The files contain all public information on all NC DPS Division of Adult Correction offenders convicted since 1972.

The bottom of the referenced page states:

The North Carolina Department of Public Safety does not offer any technical support or systems analysis assistance to users wishing to work with these files.

See the help page for more information.

\(^2\) See NCDPS, undated-b, for the offender public information search.
3. How variables in these databases can change over time and affect decisions that must be made in modeling, specifically the changing nature of “offense codes.”

Complication: Conviction Clustering

To demonstrate how conviction and sentencing events cluster on specific dates and such events interplay with incarceration periods, Figures A.1, A.2, A.3, and A.4 provide examples of patterns of criminal history and a subset of information from the data set. Such clusters can arise when a person receives multiple sentences on a single day. Figures A.1 and A.3 display smaller subsets of covariate information about a person’s conviction dates (age, number of events related to that date, and number of felonies). Figures A.2 and A.4 demonstrate additional covariate information specific to the types of offenses that the individual was convicted of on each date.

Figures A.1 and A.2, which depict Example 1, demonstrate an individual who had four total convictions on three specific dates: Their last set of convictions occurred in 2016, and their last release from prison was in 2017. The individual had no felonies but ultimately served a prison sentence after receiving a driving while intoxicated (DWI) conviction. This series of events is relatively simple, but it demonstrates that individuals can have multiple convictions and that groups of these convictions can occur on the same date. In the background-check setting, it is necessary to consider the time between clusters of conviction events that occur together on the same date because employers should be interested in the time to the next overall set of instances. That said, we do not completely remove this time and account for the total number of convictions to use as a covariate through time to achieve better precision for predicting an individual’s time to their next conviction.

---

3 Although these figures demonstrate many events for each individual, we note that many individuals actually have only one sentencing record in the data set (a fact that will be explored further later in this chapter).
4 We leveraged information available from NCDPS, Office of Research and Planning, undated, to group categories of offense codes using the sentence components in way that was consistent with the NCDPS system.
5 Analytically, this will also be important, or there will be instances in which the time between events is zero, which should trigger a reconsideration of how to break ties between events in individual’s history. Thus, the focus should be time between clusters of sentencing events that occur on the same conviction date.
Figure A.1. Example 1 of an Individual's Sentencing History, Including a Subset of Covariate Information

NOTE: Blue dots represent clusters (or sets) of convictions that occur on the same date (e.g., four sentence records all may reference the same conviction date). These can be thought of as “events” in a person’s history. Red diamonds represent start and end points for incarceration periods further illuminated by the light red box to indicate that the individual is incarcerated (in prison) and thus generally not at risk, although new convictions may arise. These can be thought of as “states” in a person’s history. The y-axis information provides covariate information (age, number of events, and how many events were felonies) related to the date that the event occurred. In this example, the individual has two events in 2016: one DWI and another non-DWI–related traffic offense. This is shown in Figure A.2.

Figure A.2. Additional Information on Example 1’s Criminal History Demonstrating Offense Codes That Occur Through Time

NOTE: Purple squares represent when events occurred and the counts of that offense type on that date. Red bands again represent incarceration spells. The y-axis reflects additional covariate information specific to the types of offenses that the person was convicted for at each date.
Figures A.3 and A.4, representing Example 2, demonstrate a more complex case. This individual had 12 convictions occurring on ten dates. We see that conviction dates do not necessarily line up with incarceration spells (consider the cases around 2010 for this individual). This illustrates that nontrivial decisions must be made to understand how to track events through time and how to incorporate incarceration periods into an analysis that tries to predict the amount of time that is expected to pass until the next event. We believe that these types of considerations will be in most secondary data sets that are used for risk prediction. An additional consideration arises related to time spent in county jail rather than state prison: The NCDPS data sets do not contain direct information on county jail incarceration. This problem will appear in many data sets and may cause issues for accurately quantifying risk.

**Figure A.3. Example 2 of an Individual’s Sentencing History, Including a Subset of Covariate Information**

*NOTE: Blue dots and red diamonds are “events” (convictions and prison entry and release, respectively); red bands running vertically through the figure represent multiple incarceration spells (or states).*
Figure A.4. Additional Information on Example 2's Criminal History Demonstrating Offense Codes That Occur Through Time

NOTE: Purple squares represent when events occurred and the counts of that offense type on that date. Red bands represent incarceration spells.

Complication: Events Versus States

The complex nature of individual criminal trajectories demonstrated in the previous figures show the importance of defining the time span between which baseline event and which final event in the context of recidivism risk models for criminal background checks. Figure A.5 captures interdependencies between what can be called “events” and “states” in the NCDPS data set. Each row in the figure represents how an individual can transition among distinct states and events in the NCDPS data sets. For example, the first row illustrates a simple example in which time is measured between convictions and no events or states occur between them. The second row demonstrates a conviction, after which the individual immediately enters prison. Here, time is measured from the release date to the next conviction. The third row adds a complication to the second row, where the individual does not immediately begin their prison sentence, but for the intent of a criminal background check, we only want to measure time from their release—the most relevant time for the background check is the time since the last major interaction with the criminal justice system. The final row demonstrates how convictions that occur while incarcerated are handled and, similar to the third row, we only measure time from release.
Generally, individual events (convictions, release dates, new convictions) can occur discretely through time, while during that time, individuals can also be represented in different complex states (e.g., in prison versus not in prison, in jail versus not in jail, on parole versus not on parole, in the community without supervision versus under supervision) in many data sets. Similarly, events are not mutually exclusive to particular states, as shown in Figure A.5. For example, new convictions can occur while an individual is incarcerated (as demonstrated in the fourth row of Figure A.5). Additionally, a new event does not always immediately cause a transition to a new state. For example, in the criminal justice system, sentences often may be suspended and a conviction will not immediately send an individual to prison (see third row of Figure A.5). How we choose to define events, transitions among states, and timing between all of these can complicate modeling choices and data representation, making it difficult for models to satisfy the reset principle.

**Figure A.5.** Possible Transitions Among Events and States That Are Considered in These Analyses and How We Track Time Between Events

NOTES: The first row demonstrates a case in which an individual might have a conviction and be released into the community on probation (i.e., a noncustodial sentence) and ultimately has another conviction; we account for the time between these events. The next rows demonstrate cases in which there is an incarceration spell. In these cases, an individual has a conviction and ultimately enters prison; in each of these instances, we account for the time from prison release to their next conviction. For each of these cases, there are corresponding outcomes (not shown), where the time for the next event is not observed (i.e., they may never fail again). When this occurs, there individuals are considered “censored.”

For the background-check setting, the timing between events should capture the time when the individual is in the community and therefore most able to commit another offense. Similarly, this is the period over which individuals are most likely to be looking for employment or other opportunities. Therefore, it is important to define the time between events as the time between some initiating conviction event or incarceration release and the time to another conviction.
Convictions occurring during incarceration periods should be ignored because they often represent holdover sentences from earlier crimes or are convictions that are conditioned on the incarceration sentence (i.e., there would not have been a conviction but for the incarceration sentence). In the NCDPS data set, we only have conviction dates and periods during which individuals are in prison and thus have focused our analyses on time between the initiating conviction or prison release and the next conviction date.

Complication: Changing Offense Codes

The final issue related to the construction of data sets is the changing nature of some variables in a person’s criminal history over time—specifically, offense codes. Consider Figure A.6, which demonstrates how references to DWI and driving under the influence (DUI) charges change over time in the NCDPS data set. From approximately 1972 to the mid-1980s, the NCDPS data set tracked impaired driving with codes related to “Driving Under the Influence (DUI).” Then, around 1985, the offense codes began referring to these charges as “Driving While Impaired (DWI).” Around 1995, the DWI charges were split into different levels (I through V) to capture the severity of the offense. Such alterations in scope and distinctions of offenses are expected to continue sporadically through time as laws and norms change in society. These changes can cause issues for modeling and predicting outcomes, especially over the time periods that are important to measure risk.

When creating complex models, if specific variables are restricted to occurring at certain points in time, then the generalizability of their use in models through time can be limited. For example, if a simple factor for “DUI” was included in the model, it would be only relevant from approximately 1985 to 1995, limiting its use in predicting events in the future. Therefore, it may be important to create aggregate groupings on specific criminal domains (as was done in Figures A.2 and A.4 for each crime time) to help models achieve good predictive properties through time. Often, these decisions can be done only in concert with domain experts, which can mean a time-consuming process of data cleaning to maximize predictive utility.
Data Structure Takeaways

All the examples in our discussion of data structure were chosen because they represent unique features that can make modeling complex and, if not accounted for, predictions inaccurate. Similarly, the complex modeling decisions that result from these features may limit generalizability of the model to new populations and domains. For those interested in developing prediction models, these issues illustrate some, but not all, of the complications that available data may present. Given the importance of data quality to a model’s predictive quality, it is essential to account for such nuances.
Appendix B. More on Survival Analysis

Description of Survival Analysis and Relationship to the Reset Principle

To define risk and illustrate the empirical underpinnings of the reset principle, we use a mathematical object called survival functions. Survival functions are useful in analyses in which the goal is to predict the time to some event. In statistics, estimators of survival functions are particularly important in that they allow us to handle “censored” observations in analyses (Klein and Moeschberger, 2003)—i.e., cases in which an outcome (another conviction) has not yet been observed. In our examples, the concept of censoring will be important in cases where an individual may have had only one conviction in their criminal history and we have not observed a time for their next conviction.

The survival function represents the probability that an individual “survives” longer than some predefined set of time. More explicitly, for an individual $i$, let $T_i$ be a random variable for that individual’s next event time, let the survival function be denoted as $S_i(t)$, and represent the probability that the next event time occurs after some time as $t$, i.e., $T_i > t$, from a baseline event. This is represented mathematically as $S_i(t) \equiv \text{Pr}(T_i > t)$.

In the case of the criminal background-check context, survival functions will represent the probability that an individual’s next conviction time occurs after some predefined set of time from their last conviction date or last release date. These survival functions are useful for quantifying risk in that, by comparing two individual’s survival functions, we might say that Individual A’s one-year survival probability is greater than Individual B’s one-year survival probability (i.e., $S_A(1) > S_B(1)$). Therefore, Individual A is less risky than Individual B.

A useful property of the survival function that relates to the reset principle is that it can be updated through time simply with the knowledge that an event has not occurred at some point in the future. This time without a conviction can be used to update the probabilities generated by the survival function with this new knowledge. This can be represented mathematically, where if we define $t_b$ to be the time of the background check and $\tau \equiv t - t_b$ as the additional time past the background check for which we are interested in evaluating risk (i.e., asking “what is the $\tau$-year survival probability given that this person has not had another event up to time $t_b$ when their background check occurs”), then the updated survival probability can be written as:

$$S(\tau|t_b) \equiv \text{Pr}(T > t_b + \tau|T > t_b) = \frac{p_r(T > t_b + \tau)}{p_r(T > t_b)} \equiv \frac{S(t_b + \tau)}{S(t_b)}.$$

The function $S(\tau|t_b)$ is defined to be one if $t < t_b$ (or $\tau < 0$), i.e., asserting that the probability of surviving longer than $t$ given that we know a person actually survived to $t_b > t$ is one. This updating is the most important aspect of the reset principle and has the potential to help demonstrate how high-risk individuals transition to low-risk individuals over time. Essentially, this means that we can make comparisons on updated survival probabilities using the actual time
of the background check for each individual. This can be done by making comparisons on
updated one-year survival probabilities and querying “Is Individual A (who has a background
check at one year) less risky than Individual B (who has their background check at six years)
over the next one-year” by mathematically asking is \( S_A(1|t_{b,A}) > S_B(1|t_{b,B}) \)?

An additional advantage of the updated survival function is that it is possible to define
thresholds of acceptable risk and then estimate when an individual might cross that threshold.
That is, by knowing \( S(t) \), we can construct additional queries, such as asking if we set \( \tau = 1 \),
can we find all times \( t_b \), such that \( S(\tau|t_b) > \mathcal{R}_s \)? In other words, can we find the time when a
background check could occur, such that an individual’s additional one-year survival probability
is above some threshold of acceptable risk \( \mathcal{R}_s \)? This last query may be similar to what was
desired in \( Ei \ v. SEPTA \), an absolute standard of risk (\textit{Douglas Ei} \ v. \textit{Southwestern Pennsylvania Transportation Authority}, 2005), and could prompt a future policy discussion about how to set an
acceptable threshold \( \mathcal{R}_s \).

**Survival Functions Estimated by Conditioning on Covariate Information**

In survival estimation, the goal is to model and estimate, for each individual \( i \), the survival
function \( S_i(t) \) (as described earlier) that mathematically represents the probability that the next
event for individual \( i \) occurs after some time \( t \) from baseline (if it occurs at all). To estimate
\( S_i(t) \), we rely on estimating \textit{conditional survival functions} given the individual’s set covariates \( X \)
(e.g., information on past incarceration, previous number of convictions), which is expressed as
\( S(t|X_i = x) \equiv S(t|x) \). In approximating \( S_i(t) \) with \( S(t|x) \), the conditional survival function
assumes that individuals with similar covariates will have similar event times in the future and
that the covariates are predictive of future conviction dates.

The conditional survival function is important to the reset principle in practice in that it will
allow us to define updated conditional survival functions that will be used to reflect an
individual’s risk profile when we incorporate the time of the background check \( t_b \). The updated
conditional survival function that can answer this question can be written as

\[
S(\tau|x, t_b) \equiv \Pr(T > t_b + \tau|T > t_b, X = x) = \frac{\Pr(T > t_b + \tau|X = x)}{\Pr(T > t_b|X = x)} = \frac{s(t_b + \tau|x)}{s(t_b|x)}.
\]

In this equation, \( S(\tau|x, t_b) \) is defined to have a value of one if \( t < t_b \), because we know that
the individual has not had another event up to time \( t_b \). Given a survival function \( S(t|x) \), we can
immediately find \( S(\tau|x, t_b) \) for any \( t_b \) and \( \tau \). Thus, we should focus our prediction on estimating
\( S(t|x) \) as accurately as possible. This is the reason that Chapter 5 focuses on accurate estimation
directly.

**A Mixture of Experts Model for Survival Estimation**

As described in Chapter 5, our focus was on nonparametric models often referred to as a
mixture-of-survival-expert models (Bennis, Mouysset, and Serrurier, 2020; Erişoğlu, Erişoğlu,
and Erol, 2011; Kuo and Peng, 2000; Nagpal et al., 2019; Raman et al., 2010). The general form of the model comprises $k$ experts (sometimes these experts may be referred to as templates) that are combined through a weighted averaging procedure, i.e.,

$$S(t|x) \approx S(t|x, y, \theta) = \sum_{k} w_k(x|y) S_k(t|\theta_k),$$

where each $w_k(x|y)$ are the mixing components of the model and the restriction that $\sum_{k} w_k(x|y) = 1$. The parameters $(y, \theta)$ must be learned from data. In our model, we impose a restriction that the covariates will only select the weights and that the experts themselves are not a function of the observed covariates (only $w_k(x|y)$ is a function of $x$). We note that there are other mixture-model–based approaches that are similar to this form of the model that allow both components to be functions of $x$, but we do not consider them here.

To estimate the weights $w_k(x|y)$, we used neural networks (see Goodfellow, Bengio, and Courville, 2016, for a methodological introduction), specifically a simple sequential feedforward network. This provides a flexible method for choosing the experts given the covariate information. The survival experts $S_k(t|\theta_k)$ were chosen to be from the Weibull family of distributions (Klein and Moeschberger, 2003) and were selected by constructing a large grid of parameters $\theta_k$ that represent a range of expected failure times (a grid of approximately 100 defined means) and diversity in shape parameters. Some experts have shape parameters that represent increasing hazard functions, and others represent decreasing hazard functions, for a grid of 50 shape parameters—25 as increasing and 25 as decreasing.

Parameters of the model were estimated using an optimization of a right-censored likelihood equation that we have implemented in TensorFlow (Abadi et al., 2016). The likelihood that was optimized is

$$L(y, \theta|x, y) = \prod_{i=1}^{N} f(t|x, y, \theta)^{\delta_i} S(t|x, y, \theta)^{1-\delta_i}$$

$$= \prod_{i=1}^{N} \left( \sum_{k} w_k(x|y) f_k(t|\theta_k) \right)^{\delta_i} \left( \sum_{k} w_k(x|y) S_k(t|\theta_k) \right)^{1-\delta_i},$$

where $f_k(t|\theta_k)$ is the density function of a Weibull distribution with parameters $\theta_k$, and $\delta_i$ is an indicator that the individual was observed to reoffend—it is one if observed and zero otherwise.

To train the model, data were split into three partitions: (1) training data, (2) validation data, and (3) test data. Candidate models were trained using gradient-descent methods on the training data. Parameters that minimize the likelihood equation on the validation data were selected. Performance of the model was ultimately assessed on the independent test set.

Note that there are alternative methods for estimating survival functions and comparing risk in this setting (see Klein and Moeschberger, 2003, for an extensive overview of survival analysis and estimation), such as parametric likelihood-based methods. These include accelerated failure time models, semi-parametric Cox-PH models, and non-parametric estimators, such as the KM
estimator. We focus on the mixture-model approach because many of these other methods can be restricted in one way or another when estimating the survival function directly. For example, parametric models and Cox-PH models often require correctly specifying the functional form of the model. The KM-estimator can be difficult to estimate when conditioning on many parameters. Thus, we have focused on methods that remove the need to directly specify a functional form and that can condition on many variables.
Appendix C. Description of Data Used for Modeling

In this appendix, we present a detailed description of the data that were used to estimate the model in Chapter 5. The results were estimated on a subset of data from the NCDPS data set. The data are restricted to those events occurring after 1995 until April 8, 2021 (when the data were downloaded), events for individuals who are between 18 and 70 years of age, and people born in North Carolina. In modeling risk for individuals, we remove records of those who have been given a life sentence or if they are currently serving a prison sentence on April 8, 2021.

We restricted analyses to those born in North Carolina to guard against potential biases that might arise if individuals leave the state (in essence, leaving the risk set). We are making an assumption that those born in North Carolina are more likely to stay in the state over time. This focus on individuals who are born in North Carolina allows us to make an implicit assumption that individuals with long survival times actually have desisted from crime and have not moved to another state.

These restrictions that we place on the NCDPS data set create a subset of approximately 1.26 million conviction records for approximately 524,000 individuals. To remove the correlation among observations (i.e., overrepresentation of some individuals many times in the data set), we randomly sampled one observation from each individual before performing estimation when modeling. Finally, we focused on presenting general results from only \( N = 250,000 \) individuals, which provides general insights on the performance of the model. These data were split into 200,000 observations for model training (160,000 for the training data set, 40,000 for the validation data set) and 50,000 observations for the test data set.

We needed to create a set of covariates that were potentially predictive of recidivism and used in estimation. Table C.1 outlines the factors that were parsed from the NCDPS data set and constructed into a clean analytic file. The factors are grouped together into the following categories:

1. age-related information
2. information on the number of events that occur on the date of a record (i.e., number of convictions, how many of these convictions are felonies)
3. information on counts of offense types that an individual is convicted for on that day (e.g., larceny charges, sexual offenses)
4. minimum and maximum sentencing information for the set of convictions
5. partial penalty information (penalty information is specific to North Carolina and, to increase model generalizability, only a subset was included)
6. prison information that occurs between the date the record represents and the next conviction time
7. prior conviction counts
8. prior offense counts.
Table C.1. Detailed Description of Variables Included in Model

<table>
<thead>
<tr>
<th>Covariate Group</th>
<th>Covariate Label</th>
<th>Covariate Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age information</td>
<td>Age at first conviction</td>
<td>Age of first conviction in the NCDPS data set</td>
</tr>
<tr>
<td></td>
<td>Age at conviction</td>
<td>Age at conviction date</td>
</tr>
<tr>
<td>Current event information</td>
<td>Convictions on date</td>
<td>Number of convictions that align on this event date</td>
</tr>
<tr>
<td></td>
<td>Number of felonies</td>
<td>Of the number of convictions, how many were felonies?</td>
</tr>
<tr>
<td></td>
<td>Number of misdemeanors</td>
<td>Of the number of convictions, how many were misdemeanors?</td>
</tr>
<tr>
<td>Current offense information</td>
<td>Many categories</td>
<td>For each offense type (see Table C.2), how many offense types were counted on this date?</td>
</tr>
<tr>
<td>Current sentencing information</td>
<td>Total sum of minimum sentence information</td>
<td>Total sum of all minimum sentences on this date</td>
</tr>
<tr>
<td></td>
<td>Average of minimum sentence information</td>
<td>Average of all minimum sentences on this date</td>
</tr>
<tr>
<td></td>
<td>Total sum of maximum sentence information</td>
<td>Total sum of all maximum sentences on this date</td>
</tr>
<tr>
<td></td>
<td>Average of maximum sentence information</td>
<td>Average of all maximum sentences on this date</td>
</tr>
<tr>
<td></td>
<td>Any sentences labeled consecutive?</td>
<td>Count of convictions labeled “consecutive” on this date</td>
</tr>
<tr>
<td></td>
<td>Any sentences labeled concurrent?</td>
<td>Count of convictions labeled “concurrent” on this date</td>
</tr>
<tr>
<td>Partial penalty information</td>
<td>All sentences of type “community?”</td>
<td>A community sentence is the lowest level of supervision in North Carolina. Are all convictions on this date related to a sentence at this lowest level?</td>
</tr>
<tr>
<td></td>
<td>Count of DWI penalties</td>
<td>Was the penalty a DWI-specific penalty?</td>
</tr>
<tr>
<td>Current prison information</td>
<td>From prison?</td>
<td>Is the time between the last event and the next event calculated from a set of conviction dates, or a release from prison date?</td>
</tr>
<tr>
<td></td>
<td>Total incarceration time</td>
<td>Total amount of time incarcerated between this set of convictions and the next set of convictions</td>
</tr>
<tr>
<td>Prior event information</td>
<td>Total number of convictions prior to date</td>
<td>Total number of convictions prior to this date (i.e., all events, not just conviction dates)</td>
</tr>
<tr>
<td></td>
<td>Total number of felonies prior to date</td>
<td>Total number of felonies prior to this date</td>
</tr>
<tr>
<td></td>
<td>Total number of misdemeanors prior to date</td>
<td>Total number of misdemeanors prior to this date</td>
</tr>
<tr>
<td>Prior prison information</td>
<td>Total incarceration time prior to this date</td>
<td>Total incarceration time prior to this date</td>
</tr>
<tr>
<td>Prior offense information</td>
<td>Number of prison spells prior to this date</td>
<td>Total number of contiguous periods incarcerated prior to this date</td>
</tr>
<tr>
<td>--------------------------</td>
<td>------------------------------------------</td>
<td>------------------------------------------------------------------</td>
</tr>
</tbody>
</table>

- Many categories

For each offense type, how many offense types were counted prior to this date?

### Table C.2. List of Offense Types

<table>
<thead>
<tr>
<th>Offense Types from NCDPS Automated Query System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abandonment and non-support</td>
</tr>
<tr>
<td>Assault</td>
</tr>
<tr>
<td>Auto theft</td>
</tr>
<tr>
<td>Breaking entering</td>
</tr>
<tr>
<td>Burglary</td>
</tr>
<tr>
<td>Burnings</td>
</tr>
<tr>
<td>Driving while impaired</td>
</tr>
<tr>
<td>Drugs (trafficking), Drugs (non-trafficking)</td>
</tr>
<tr>
<td>Forgery</td>
</tr>
<tr>
<td>Fraud</td>
</tr>
<tr>
<td>Kidnapping and abduction</td>
</tr>
<tr>
<td>Larceny</td>
</tr>
<tr>
<td>Manslaughter</td>
</tr>
<tr>
<td>Murder (first degree)</td>
</tr>
<tr>
<td>Murder (second degree)</td>
</tr>
<tr>
<td>Not observed</td>
</tr>
<tr>
<td>Not reported/undefined</td>
</tr>
<tr>
<td>Other alcohol offense</td>
</tr>
<tr>
<td>Other offense against person</td>
</tr>
<tr>
<td>Other property</td>
</tr>
<tr>
<td>Other public order</td>
</tr>
<tr>
<td>Other sexual offense</td>
</tr>
<tr>
<td>Other traffic violations</td>
</tr>
<tr>
<td>Post-release offense</td>
</tr>
<tr>
<td>Robbery</td>
</tr>
<tr>
<td>Sexual assault</td>
</tr>
<tr>
<td>Violation of sexual offense conditions</td>
</tr>
<tr>
<td>Weapons offenses</td>
</tr>
<tr>
<td>Worthless checks</td>
</tr>
</tbody>
</table>
Appendix D. Additional Model Calibration Curves

As described in the main body of the report, *calibration* assesses how well probability estimates agree with reality. That is, if a group of individuals are predicted to have an 80-percent five-year survival probability, then we should expect that approximately 80 percent survive longer than five years and that 20 percent should fail prior to that point. Figures D.1, D.2, and D.3 present additional comparisons of calibration curves for the model described in Chapter 5 for one-year, ten-year, and 20-year survival probabilities for the training data, validation data, and independent test sets. The results provide a clear visual display of calibration that also generalizes to an independent test data set. Additionally, the results present diversity in predictions across wide time ranges (e.g., a large degree of diversity when 20-year predictions are considered).

Figure D.1. One-Year Survival Model Calibrations
Figure D.2. Ten-Year Survival Model Calibrations

Figure D.3. 20-Year Survival Model Calibrations
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BJS</td>
<td>Bureau of Justice Statistics</td>
</tr>
<tr>
<td>Cox-PH</td>
<td>Cox proportional hazards</td>
</tr>
<tr>
<td>CRA</td>
<td>consumer reporting agency</td>
</tr>
<tr>
<td>DUI</td>
<td>driving under the influence</td>
</tr>
<tr>
<td>DWI</td>
<td>driving while impaired</td>
</tr>
<tr>
<td>EEOC</td>
<td>Equal Employment Opportunity Commission</td>
</tr>
<tr>
<td>KM</td>
<td>Kaplan-Meier</td>
</tr>
<tr>
<td>NCDPS</td>
<td>North Carolina Department of Public Safety</td>
</tr>
<tr>
<td>RAI</td>
<td>risk-assessment instrument</td>
</tr>
<tr>
<td>SEPTA</td>
<td>Southeastern Pennsylvania Transportation Authority</td>
</tr>
</tbody>
</table>
References


http://www.albanylawreview.org/Articles/Vol74_3/74.3.1087%20BUSHWAY.EE.CDL.pdf


https://doi.org/10.1177/0093854896023001009


https://core.ac.uk/download/pdf/151523478.pdf

Corbett-Davies, Sam, Emma Pierson, Avi Feller, and Sharad Goel, “A Computer Program Used for Bail and Sentencing Decisions Was Labeled Biased Against Blacks. It’s Actually Not


https://www.ojp.gov/pdffiles1/ojjdp/209158.pdf

https://doi.org/10.1177/0011128705281748

https://supreme.justia.com/cases/federal/us/401/424/

Harrell, Frank E., Jr., *Regression Modeling Strategies with Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*, 2nd ed., Cham, Switzerland: Springer International Publishing Switzerland, 2015. As of November 30, 2021:
https://link.springer.com/content/pdf/10.1007%2F978-3-319-19425-7.pdf


https://doi.org/10.1177/0963721418763931

https://www.vera.org/publications/for-the-record-unjust-burden


As of November 30, 2021:
http://www.jstor.org/stable/2281868

https://doi.org/10.1111/puar.12842

https://scholarship.law.wm.edu/cgi/viewcontent.cgi?article=3680&context=wmlr

https://doi.org/10.1177/0011128706294442


https://doi.org/10.1108/EDI-09-2018-0176


https://doi.org/10.1111/lsi.12066

https://doi.org/10.1093/sf/sso092


https://doi.org/10.1177/0093854884011003007

https://doi.org/10.1146/annurev-statistics-042720-125902


National Reentry Resource Center, “Collateral Consequences Inventory,” National Inventory of Collateral Consequences of Conviction webpage, undated. As of December 1, 2021: https://niccc.nationalreentryresourcecenter.org/consequences

NCDPS—See North Carolina Department of Public Safety.


86


Rudin, Cynthia, and Joanna Radin, “Why Are We Using Black Box Models in AI When We Don’t Need To? A Lesson from an Explainable AI Competition,” Harvard Data Science Review, No. 1.2, Fall 2019. As of December 1, 2021: https://hdsr.mitpress.mit.edu/pub/f9kuryi8/release/6


https://doi.org/10.1111/1745-9125.12256


State of Nevada, “Health Facilities - Background Check FAQ’S,” Department of Health and Human Services, Nevada Division of Public and Behavioral Health webpage, July 13, 2015. As of December 1, 2021:
http://dpbh.nv.gov/Reg/HealthFacilities/dta/FAQs/Health_Facilities_-_Background_Check_FAQ_s/


89