

Center for Evidence-Based  
Corrections

Department of Criminology, Law  
& Society

School of Social Ecology

# Validation of the PSA in Los Angeles County

**James Hess and Susan Turner**

September 2020 (Updated January 2021)

# Contents

List of Tables .....	iii
List of Figures.....	v
Glossary.....	vi
Executive Summary .....	vii
1. Introduction .....	1
Goal: PSA Validation .....	1
2. Risk Assessment .....	2
PSA Tool .....	3
3. Methods.....	4
Study Design.....	4
Data .....	4
Bookings.....	5
History.....	5
Warrants .....	6
Demographics .....	6
Tables for Classification of Offenses .....	6
Violent Offenses.....	6
Exclusions.....	6
Offense Classification.....	7
Data Limitations .....	7
Data Processing for PSA Scale Construction .....	8
Measures of Predictive Ability .....	9
Recidivism by Risk Score .....	9
Odds Ratio.....	9
Understanding the AUC .....	10
4. Results.....	12
Sample Description .....	12
Individuals .....	12
Bookings.....	13
Outcomes by Release Type.....	20

NCA Outcome Offense Category by Release Type.....	21
All Charges and Most Serious Charge .....	21
Validation Options .....	25
Suggested Choices among Validation Options .....	26
Descriptive Table All Validation Options: Demographics, Data Completeness and Quality, PSA Factors and Outcomes .....	27
Option A: Core Complete Data .....	31
Option A: Parameters .....	31
Option A: Outcome by Risk Score .....	32
Option A: Odds Ratios and AUCs .....	35
Option B: Core Data, Plus NULL and Missing Charge Level .....	37
Option B: Parameters.....	37
Option B: Outcome by Risk Score .....	38
Option B: Odds Ratios and AUCs .....	41
Option C: Imputation of Sentence Date.....	42
Option C: Parameters.....	42
Option C: Outcome by Release Type .....	43
Option C: Outcome by Risk Score .....	44
Option C: Odds Ratios and AUCs.....	47
Option D: Recovery of Charges.....	48
Option D: Parameters.....	48
5. Discussion of Options A, B, and C. ....	54
Exclusions.....	57
6. Conclusions.....	59
Appendix A: LJAF Documentation of Risk Factor and Outcome Definitions and Measurement.....	61
Appendix B: Violent Offenses.....	64
Appendix C: Exclusions.....	67
Appendix D: Offense Classification.....	72
Appendix F: Option D: Supplementary Analysis.....	76
Appendix G: Calibration.....	88

## List of Tables

Table S.1: Outcome by Release Type .....	viii
Table S2: Comparative PSA Performance, Receiver Operating Characteristic Area Under the Curve .....	xi
Table 2.1: PSA Risk Factors .....	3
Table 3.1: LA Court-Provided Data Sets .....	4
Table 3.2: Booking Year by Study Group.....	5
Table 4.1: Individual Demographics.....	13
Table 4.2.a: Booking Demographics (All Bookings).....	14
Table 4.2.b.1: Bookings with History: Data Completeness .....	16
Table 4.2.b.2: Bookings with History: Record .....	17
Table 4.2.b.3: Bookings with History: Outcomes.....	19
Table 4.3.a: Outcomes by Release Type (Bookings with Sentence Date).....	20
Table 4.3.b: Outcomes by Release Type (Include Bookings with Imputed Sentence Date) .....	21
Table 4.4.a: New Arrests: All Charges by Offense Type and Release Type.....	22
Table 4.4.b: New Arrests Most Serious Charge by Offense Type and Release Type .....	24
Table 4.5.1: Validation Options: Demographics .....	27
Table 4.5.2: Validation Options: Data Completeness .....	28
Table 4.5.3: Validation Options: Record .....	29
Table 4.5.4: Options: Outcomes .....	30
Table 4.A.1: Option A: Outcomes by Release Type.....	31
Table 4.A.2: Option A Failure to Appear by Risk Score .....	32
Table 4.A.3: Option A New Criminal Activity by Risk Score .....	33
Table 4.A.4.a: Option A New Violent Criminal Activity by Risk Score.....	34
Table 4.A.4.b: Option A New Violent Criminal Activity by Violence Flag.....	35
Table 4.A.5: Option A PSA Predictive Ability, Failure to Appear.....	35
Table 4.A.6: Option A PSA Predictive Ability, New Criminal Activity.....	36
Table 4.A.7: Option A PSA Predictive Ability, New Violent Criminal Activity .....	36
Table 4.B.1: Option B Outcomes by Release Type.....	37
Table 4.B.2: Option B Failure to Appear by Risk Score .....	38
Table 4.B.3: Option B New Criminal Activity by Risk Score.....	39

Table 4.B.4.a: Option B New Violent Criminal Activity by Risk Score .....	40
Table 4.B.4.b: Option B New Violent Criminal Activity by Violence Flag .....	40
Table 4.B.5: Option B PSA Predictive Ability, Failure to Appear .....	41
Table 4.B.6: Option B PSA Predictive Ability, New Criminal Activity .....	41
Table 4.B.7: Option B PSA Predictive Ability, New Violent Criminal Activity.....	42
Table 4.C.1: Option C Outcomes by Release Type .....	43
Table 4.C.2: Option C Failure to Appear by Risk Score .....	44
Table 4.C.3: Option C New Criminal Activity by Risk Score.....	45
Table 4.C.4: Option C New Violent Criminal Activity by Risk Score .....	46
Table 4.C.4.b: Option C New Violent Criminal Activity by Violence Flag .....	46
Table 4.C.5: Option C PSA Predictive Ability, Failure to Appear .....	47
Table 4.C.6: Option C PSA Predictive Ability, New Criminal Activity .....	47
Table 4.C.7: Option C PSA Predictive Ability, New Violent Criminal Activity.....	48
Table 4.D.1: Option D: Outcomes by Release Type .....	49
Table 4.D.2: Option D Failure to Appear by Risk Score.....	50
Table 4.D.3: Option D New Criminal Activity by Risk Score .....	51
Table 4.D.4: Option D New Violent Criminal Activity by Risk Score .....	52
Table 4.D.4.b: Option D New Violent Criminal Activity by Violence Flag .....	52
Table 4.D.5: Option D PSA Predictive Ability, Failure to Appear .....	53
Table 4.D.6: Option D PSA Predictive Ability, New Criminal Activity.....	53
Table 4.D.7: Option D PSA Predictive Ability, New Violent Criminal Activity .....	53
Table 5.1: Sample Sizes Before and After Restoring Excluded Cases .....	58
Table 5.2: Risk Scales Before and After Restoring Excluded Cases.....	58
Table 6.1: Comparative PSA Performance, Receiver Operating Characteristic Area Under the Curve.....	60
<i>Table A.1 LJAF Documentation of Risk Factor and Outcome Definitions and Measurement .....</i>	<i>61</i>
<i>Table A.2: Rescaling Total Points for PSA Final Risk Scores .....</i>	<i>63</i>
<i>Table F.1.1 Supplementary Analysis: Demographics .....</i>	<i>77</i>
<i>Table F.1.2 Supplementary Analysis: Data Completeness .....</i>	<i>78</i>
<i>Table F.1.3 Supplementary Analysis: Record .....</i>	<i>79</i>
<i>Table F.1.4 Supplementary Analysis: Outcomes .....</i>	<i>80</i>
<i>Table F.2 Supplementary Analysis: Outcomes by Release Type.....</i>	<i>81</i>
<i>Table F.3 Supplementary Analysis: Failure to Appear by Risk Score .....</i>	<i>82</i>

<i>Table F.4 Supplementary Analysis: New Criminal Activity by Risk Score</i> .....	83
<i>Table F.5 Supplementary Analysis: New Violent Criminal Activity by Risk Score</i> .....	84
<i>Table F.6 Supplementary Analysis: PSA Predictive Ability, Failure to Appear</i> .....	85
<i>Table F.7 Supplementary Analysis: PSA Predictive Ability, New Criminal Activity</i> .....	86
<i>Table F.8 Supplementary Analysis: PSA Predictive Ability, New Violent Criminal Activity</i> .....	87
<i>Table G.1: Risk Scales Before and After Rescaling Raw Scores</i> .....	89

## List of Figures

Figure S.1: Failure to Appear by Risk Score.....	x
Figure S.2: New Criminal Activity by Risk Score .....	x
Figure S.3: New Violent Criminal Activity by Risk Score .....	xi
Figure 3.1. Receiver Operating Characteristic Curves and Area Under the Curve .....	11
Figure 5.1: Failure to Appear by Risk Score .....	55
Figure 5.2: New Criminal Activity by Risk Score.....	56
Figure 5.3: New Violent Criminal Activity by Risk Score .....	57
<i>Figure G.1: Failure to Appear by Risk Score Rescaled</i> .....	89
<i>Figure G.2: New Criminal Activity by Risk Score Rescaled</i> .....	90
<i>Figure G.3: New Violent Criminal Activity by Risk Score Rescaled</i> .....	90

## Glossary

AUC	Area Under the Curve
FTA	Failure to Appear
LJAF	Laura and John Arnold Foundation
LAC	Los Angeles Courts
NCA	New Criminal Activity
NVCA	New Violent Criminal Activity
PSA	Public Safety Assessment

# Executive Summary

## Background

Jurisdictions across the country have joined a movement to rethink how individuals are handled at the pretrial stage of case processing. Although alternatives to cash bail systems have been around since the 1960s,<sup>1</sup> renewed interest has focused on the use of risk assessment algorithms to help determine which pretrial individuals might be released safely into the community. These types of tools hold promise as a means to move away from “debtor prisons” for individuals who do not have the financial resources to pay for their release. However, the field is still in the relatively early stage of testing these tools for predictive ability, potential racial bias in administration, as well as whether their use actually reduces incarceration.<sup>2</sup>

California has recently entered the pretrial risk assessment arena. Senate Bill 10 was passed in 2018 to change from a cash-based pretrial system to a risk-based release and detention system; although it is on hold until November 2020 when California voters determine its fate.<sup>3</sup> However, legislation passed as part of the 2019 Budget Act created a pilot program to test the use of various risk assessment tools in a number of counties across California. This report presents findings from the Los Angeles pilot effort under the Act to validate the Public Safety Assessment (PSA).

The PSA is a risk assessment instrument developed by the Laura and John Arnold Foundation to inform pretrial judicial decisions on whether to release or detain a defendant. The tool predicts three outcomes after pretrial release: Failure to Appear (FTA); New Criminal Activity (NCA, arrest on any misdemeanor or felony charge); and New Violent Criminal Activity (NVCA, arrest on a violent misdemeanor or felony charge). The tool’s nine risk factors include prior convictions, incarceration, and failures to appear, violent offenses, pending cases at the time of arrest and age. Risk factor counts are weighted by an integer multiplier and summed to create a risk score. Several sets of adjacent scores are collapsed together into one score to produce a final 6-point risk scale for each of the outcomes.

---

<sup>1</sup> Goldkamp, J. S., & Gottfredson, M. R. (1979). Bail decision-making and pretrial detention. *Law and Human Behavior*, 3(4), 227-249.

<sup>2</sup> Desmarais, S. L., Zottola, S. A., Duhart Clarke, S. E., & Lowder, E. M. (2020). Predictive Validity of Pretrial Risk Assessments: A Systematic Review of the Literature. *Criminal Justice and Behavior*, 0093854820932959; Bechtel, K., Holsinger, A. M., Lowenkamp, C. T., & Warren, M. J. (2017). A meta-analytic review of pretrial research: Risk assessment, bond type, and interventions. *American Journal of Criminal Justice*, 42(2), 443-467; Viljoen, J. L., Jonnson, M. R., Cochrane, D. M., Vargen, L. M., & Vincent, G. M. (2019). Impact of risk assessment instruments on rates of pretrial detention, post conviction placements, and release: A systematic review and meta-analysis. *Law and Human Behavior*, 43(5), 397-420. <http://dx.doi.org/10.1037/lhb0000344>

<sup>3</sup> <http://www.courts.ca.gov/pretrial.htm>



## Validation Sample

The sample for validation was based on bookings for arraignment and trial in the Los Angeles Courts arising from arrests made during the calendar years 2015 through 2018. LA Courts (LAC) elected to include releases through Bail, Bond, Cited and Released, and Own Recognizance, providing a validation sample of 724,524 cases.

The LAC provided four data sets, covering bookings, criminal history, bench warrants, and subject demographics. Ideally, all bookings would be tied to specific cases for the validation; however, data could not be matched this way for a large percentage of the cases. Therefore, the current analysis utilizes four approaches to deal with various “unknowns” present in the data. Option A represents Core Complete data and is limited to bookings where all charges and convictions are available, the sentence date is available, and the disposition date is in 2018 or earlier. Option B includes Core Complete Data and adds cases with incomplete charge data, as well as disposition dates later than December 2018. This provides a more robust assessment with more data. Option C adds cases where we imputed a sentence date for cases that did not have a link between booking number and case history record. Option D explored recovering ill-formed charges and imputing level of charge from offense information.

**Table S.1: Outcome by Release Type**

<b>Outcome by Release Type (%)</b>				
<b>(Include Bookings with Imputed Sentence Date)</b>				
<b>Release Type</b>	<b>FTA</b>	<b>NCA</b>	<b>NVCA</b>	<b>Number</b>
				<b>% of Total</b>
<b>Bail</b>	512	487	117	1463
	35.00	33.29	8.00	0.74
<b>Bond</b>	17865	19492	4819	48917
	36.52	39.85	9.85	24.81
<b>Cite</b>	71810	58530	10901	104144
	68.95	56.20	10.47	52.81
<b>OR</b>	20706	15913	3481	42670
	48.53	37.29	8.16	21.64
<b>Total</b>	110893	94422	19318	197194
	56.24	47.88	9.80	100.00

Table S.1 presents the distribution for the three major outcomes, by release type for the validation sample.<sup>4</sup> The vast majority of cases were cite and release, followed by OR, bond and bail. Cite and release also had the highest percent of recidivism for the three outcomes – FTA, NCA and NVCA. NCVA rates were relatively low, generally under 10 percent.

## Results

Figures F1, F2 and F3 show the percent of adverse outcomes at each level of the PSA risk score for each of the three outcomes for the different options that were tested.<sup>5</sup> For example, Figure 1 presents the percent of cases with an FTA at each of the six PSA scale scores for Options A, B, and C. Ideally, we would want to see the lowest levels of recidivism/FTAs for the lowest PSA scale values with an increasing rate of recidivism as the PSA scale scores increase. Results showed that across all three outcomes, discrimination is strongest in the middle range of the PSA, declining sharply at the extremes, with the rate of recidivism sometimes falling slightly as PSA scores go from 5 to 6 points on the scale. An examination of the percentage scoring at the top end of the scale reveal the highest risk group was near to or less than one percent of the sample, much lower than the 7-8 percent in the Kentucky validation. The drop in recidivism at the high end of the scale does not appear to be the result of exclusion of offenders booked with violent charges under PC 1270.1 and PC 1319.5. However, it is clear that the population assessed in this validation are dissimilar to previous validations. Very serious and the very low-level offenders are both missing from this sample relative to other jurisdictions, making discrimination between outcome rates more difficult. The decline in outcome rates at the high end of the scales has little impact on overall predictive ability as so few cases are involved. In practice, this minor weakness is overcome if the distinction between a risk score of 5 and 6 is not used to change release decisions.

To summarize results for validation purposes we use the AUC, which is a measure of the overall ability of the PSA to predict outcomes across all levels of the PSA instrument (Table S2). We highlight Option B because it is likely the best match to the operational context, with scoring based on records which include missing charges and a definite follow-up window from release to final sentencing rather than an imputed median time from release to case disposition.<sup>6</sup> We find that the PSA as validated in this study compares very favorably with that found in other jurisdictions.

---

<sup>4</sup> This table is based on Option C, discussed in the full report. Option C includes the highest number of cases.

<sup>5</sup> Option D is not included for presentation clarity, as results were almost exactly the same as for Option C.

<sup>6</sup> Its weakness is the addition of convictions within the follow-up window to compensate for the limitation of bookings coverage, which a comparison of NCA and NCVA rates with and without convictions suggests adds 4 percentage points, overestimating the recidivism rate.

Figure S.1: Failure to Appear by Risk Score

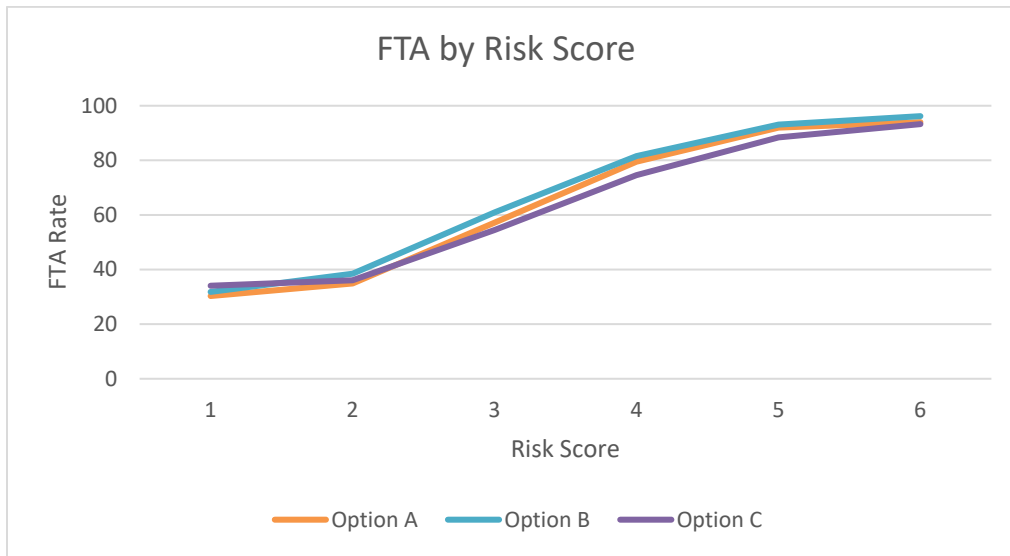
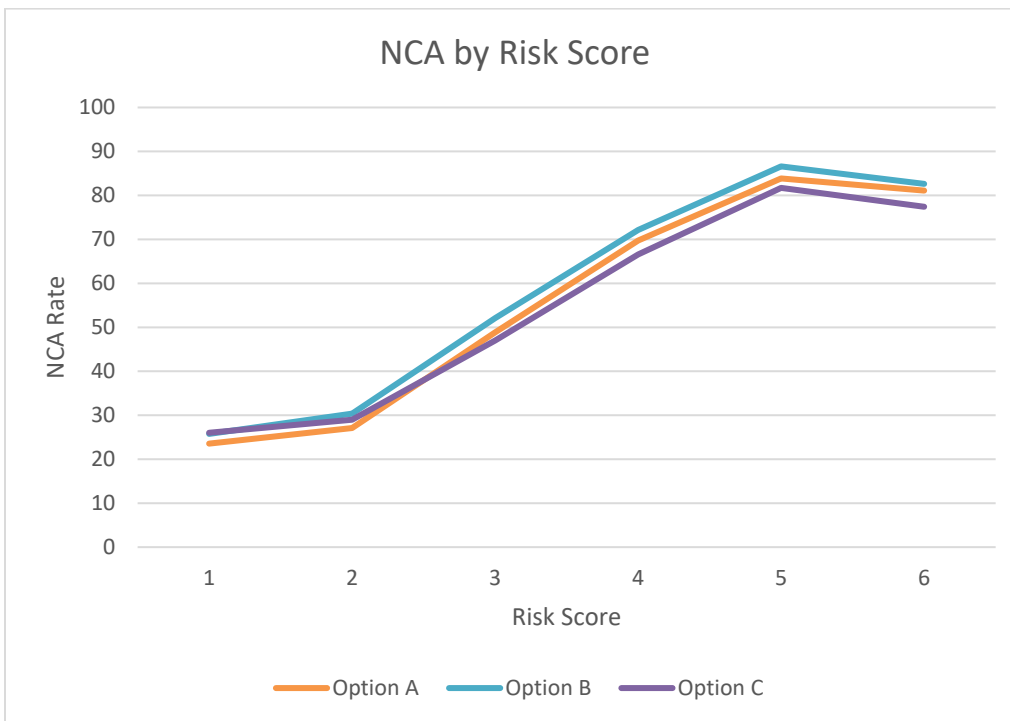
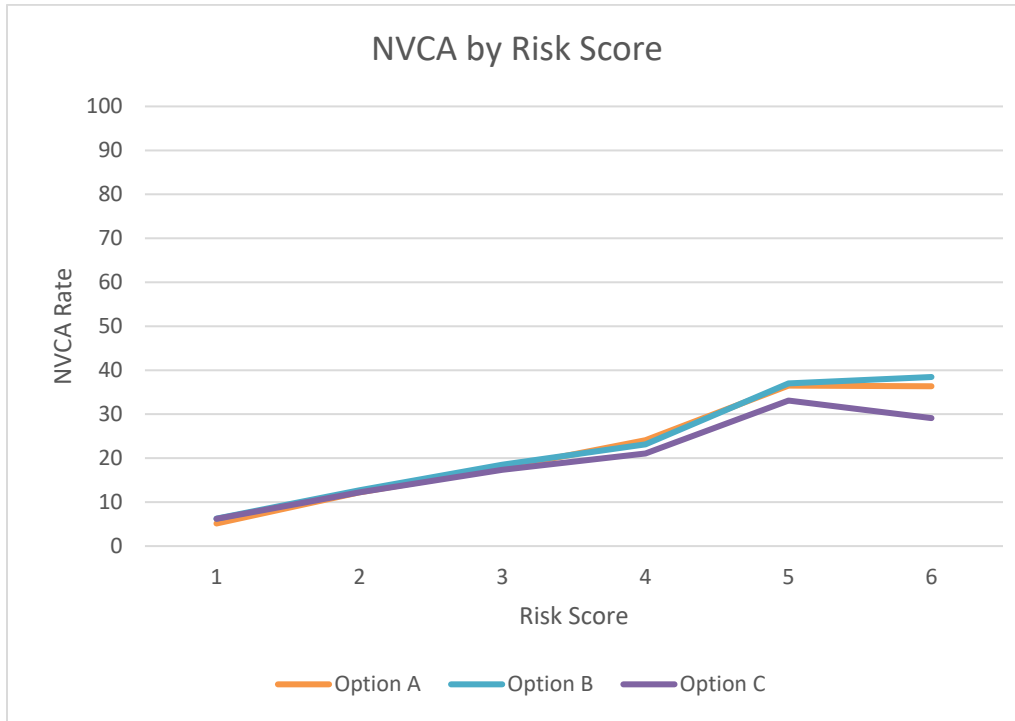


Figure S.2: New Criminal Activity by Risk Score



**Figure S.3: New Violent Criminal Activity by Risk Score**



**Table S2: Comparative PSA Performance, Receiver Operating Characteristic Area Under the Curve**

Jurisdiction	FTA	NCA	NVCA
Kentucky	0.646	0.650	0.664
Los Angeles (Option A)	0.733	0.721	0.670
Los Angeles (Option B)	0.746	0.722	0.654
Los Angeles (Option C)	0.721	0.701	0.650
Los Angeles (Option D)	0.721	0.700	0.651

## Conclusion

Comparing AUCs across the analysis options, we examined points to another important feature of the PSA: Its performance is generally robust, holding up despite variations in data quality and completeness. Its weakest performance is in the assessment of risk for new violent criminal activity, which shows the

strongest performance in Kentucky.<sup>7</sup> For FTA and NCA, the PSA achieves strong performance in this evaluation. Overall, we find that the PSA is a valid risk assessment tool for use with the Los Angeles Courts population, exhibiting robust and moderate to strong predictive performance. Local decisions as to cut-point values for high risk may need to be considered (combining highest risk groups) as well as which data recovery method (E.g., Option A, B, C or D) best suits the local county effort.

---

<sup>7</sup> In the Kentucky validation study, 18% of the sample was classified as at risk by the NVCA flag based on a score of 5 or 6 compared to less than 2% with comparable scores in our samples. The Kentucky NVCA baseline rate was only 1.4% compared to 10-11% in this study; the NVCA rate in their flagged group was 3% compared to our rate of 28-30% here. The differences in these figures suggest a deep qualitative difference in the two populations.

## 1. Introduction

Legislation passed as part of the 2019 Budget Act created a pilot program to test the use of various risk assessment tools in counties across California. The Los Angeles pilot includes use of two risk assessment tools. First, prior to arraignment, in-custody defendants are subject to a change in bail and possible release from custody under PC 1269a, also known as “bail deviation.” This process uses the Public Safety Assessment (PSA). Second, at arraignment, the Los Angeles County pilot uses the Criminal Court Assessment Tool for those defendants who request a bail hearing. The present study focuses on the PSA.

The PSA is a risk assessment instrument developed by the Laura and John Arnold Foundation to inform pretrial judicial decisions on whether to release or detain a defendant. The PSA is used by magistrate judges who review the bookings of such defendants 24 hours a day. The Superior Court of California for the County of Los Angeles (LA Courts) contracted with the Center for Evidence-Based Corrections (CEBC) of the University of California, Irvine to validate the predictive ability of the PSA for the court’s correctional population.

The use of the PSA at Bail Deviation influences our sampling strategy. As we discuss in more detail below, PC 1270.1 and PC 1319.5 preclude the release of persons arrested for certain charges prior to arraignment. Such persons are not eligible for release at Bail Deviation and thus are not to be assessed by the PSA nor reviewed by the magistrate judge. They are excluded from our sample as discussed below. Severe jail overcrowding means that low-level offenses also are less likely be represented in the sample. For instance, turnstile-jumpers are rarely arrested in Los Angeles County, compared with other jurisdictions in which zero-tolerance policing results in custodial detention for many such violators. The use of the PSA pre-arraignment, and the California statutory exclusion of violent offenses from pre-arraignment release, means that the sample does not include the full spectrum of criminally-involved individuals. The Los Angeles County population in our sample is very different than the population of released defendants in the jurisdictions on which previous validations were based. In theory, this makes the job of discriminating high, middle, and low risk more challenging for any assessment tool.

This report begins with a short background on risk assessment, a description of the PSA, validation methods, a description of the sample, and validation of the three major outcomes which the PSA predicts: Failure to Appear (FTA), New Criminal Activity (NCA, arrest on any misdemeanor or felony charge), and New Violent Criminal Activity (NVCA, arrest on a violent misdemeanor or felony charge).<sup>1</sup>

### Goal: PSA Validation

The PSA is an actuarial risk assessment instrument. Actuarial risk assessments make predictions about future behavior based on statistical regularities found in past behavior in a population used for instrument development. It is often found that the predictive ability of an actuarial assessment weakens when applied to another population. There are several reasons for this. In criminal justice

---

<sup>1</sup> This analysis does not test whether the PSA is similarly predictive for subgroup populations (e.g., groups defined by gender, race/ethnicity, offense type).

applications, each jurisdiction is different, with a particular population and social and economic environment. Criminal justice agencies have different policies and practices. A conviction or arrest is a result not only of offender opportunity and proclivity but also the processes and policies that lead to reports, arrests, prosecutions, and sentencing. These also change over time, notably in California with sweeping code changes such as AB109 and Proposition 47. It can't be assumed that results in Los Angeles County will match those in Kentucky, or that results in 2020 will match results in 2025.

Validation is the process of demonstrating that expected results are achieved when employing the risk instrument within the population and context where it will be used. Thus, although the PSA was developed with data from over 300 jurisdictions with a published validation study based on use by Kentucky's statewide pretrial services agency,<sup>2</sup> PSA guidelines strongly recommend validation of the PSA in each jurisdiction adopting it, with revalidation every few years.<sup>3</sup>

## 2. Risk Assessment

Risk assessment as a tool for criminal justice applications has a long history. It stands in contrast with offense-based, "let the punishment fit the crime" methods by assigning sanctions or case handling procedures on the basis of a prediction of the likelihood to re-offend and the severity of the (likely) new offense rather than on the severity of the incident offense and/or offense history. This contrast is often described as "risk versus stakes." Despite its long history, applications of risk assessment have become widespread only in the past two decades.<sup>4</sup> Today risk assessment has widespread use across the justice system - in pre-trial processes, sentencing, incarceration and supervision in the community.

Risk assessment methods were developed largely piecemeal; efforts to clarify the history have applied post-hoc categorization into four generations.<sup>5</sup> The first generation is typified by informal procedures such as clinical assessment based on expert judgment. The second generation is characterized by the introduction of formalized actuarial risk assessment. Starting with objective data points, actuarial methods apply statistical analysis to determine which are associated with the outcome. These data points are then used as predictive risk factors, weighted by the strength of their relationship to the outcome and combined to generate a risk score. The PSA is a second generation instrument.

The third generation shifts from instruments depending only on "static" risk items such as criminal history and demographic factors to "dynamic" factors that can change looking forward, e.g., education, employment, housing, and relationships. Many of these dynamic items are considered indicators of

---

<sup>2</sup> DeMichele, M., Baumgartner, P., Wenger, M., Barrick, K., & Comfort, M. (2020). Public safety assessment: Predictive utility and differential prediction by race in Kentucky. *Criminology & Public Policy*. Also available online at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3168452](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3168452)

<sup>3</sup> Advancing Pretrial Policy and Research, nd. Guide to Outcomes and Oversight.

<https://advancingpretrial.org/guide/guide-to-outcomes-and-oversight/> (Site registration required.)

<sup>4</sup> Harris, G. T., & Rice, M. E. (2007). Characterizing the value of actuarial violence risk assessments. *Criminal Justice and Behavior*, 34(12), 1638-1658.

<sup>5</sup> Andrews, D. A., Bonta, J., & Wormith, J. S. (2006). The recent past and near future of risk and/or need assessment. *Crime & Delinquency*, 52(1), 7-27.

criminogenic needs that can be reduced by rehabilitative interventions.<sup>6</sup> Fully-developed third generation instruments attempt to be more comprehensive, targeting risk, needs, and responsivity, with responsivity aimed at identifying which interventions would be most effective for which individuals, based on receptivity to the interventional approach. Fourth-generation instruments integrate case-management, guiding supervision and service programming through case closure. The Los Angeles Courts (LAC) has also adopted a third-generation risk and needs instrument, the Criminal Court Assessment Tool, which will be validated in the next phase of this project.

## PSA Tool

The PSA aims to predict three outcomes after pretrial release: Failure to Appear (FTA); New Criminal Activity (NCA, arrest on any misdemeanor or felony charge); and New Violent Criminal Activity (NVCA, arrest on a violent misdemeanor or felony charge). As the goal is to measure only the additional burden on courts and public safety arising from release prior to sentencing, the follow-up window for recidivism or failure to appear is the period between booking and disposition of all charges in the court case filing. Nine risk factors were selected for their ability to predict these outcomes. Risk factors include prior convictions, incarceration, and failures to appear, violent offenses, pending cases at the time of arrest and age (see Table 2.1). Risk factor counts are weighted by an integer multiplier and summed to create a risk score. Several sets of adjacent scores are collapsed together into one score to produce a final 6-point risk scale for each of the outcomes. (See Appendix A for details.)

**Table 2.1: PSA Risk Factors**

Risk Factor	Pretrial Outcome		
	FTA	NCA	NVCA
1. Age at current arrest		✓	
2. Current violent offense			✓
2a. Current violent offense and 20 years old or younger			✓
3. Pending charge at the time of the offense	✓	✓	✓
4. Prior misdemeanor conviction		✓	
5. Prior felony conviction		✓	
5a. Prior conviction	✓		✓
6. Prior violent conviction		✓	✓
7. Prior failure to appear pretrial in past 2 years	✓	✓	
8. Prior failure to appear pretrial older than 2 years	✓		
9. Prior sentence to incarceration		✓	
FTA = Failure to Appear; NCA = New Criminal Activity; NVCA = New Violent Criminal Activity			

<sup>6</sup> Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal justice and Behavior*, 17(1), 19-52; Bonta, J. (1996). Risk-needs assessment and treatment. In A. T. Harland (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply* (pp. 18-32). Thousand Oaks, CA: Sage



## 3. Methods

### Study Design

The unit of analysis for validation of a pretrial release risk assessment is a case (complaint or indictment) in the Trial Court Information System (TCIS).<sup>7</sup> A person may have multiple records representing multiple cases; current charges and prior history are specific to a case booking and the date of the arrest. In consultation between LA Courts and CEBC personnel, the sample for validation was based on bookings for arraignment and trial in the court arising from arrests made during the calendar years 2015 through 2018. This extended period was chosen to maximize the number of cases and achieve more precise estimates of predictive ability. However, due to the changes in charging and incarceration following the 2014 passage of Proposition 47, we judged that the cases brought to the court before 2015 could be misleading in assessing the current population. The cutoff of December 31, 2018 allows for a year of follow-up to capture new FTAs and arrests, thereby reducing the problem of right-censorship of the outcome events of interest.<sup>8</sup> The recruitment sample consists of 1,054,278 cases filed against 476,480 individuals.<sup>9</sup> To draw a validation from this sample it is necessary to select individuals who were released pretrial and therefor at risk for a new arrest or a failure to appear. It is also desirable to select releases representative of those likely to be considered for pretrial release when the tool is used in practice. The LAC elected to include releases through Bail, Bond, Cited and Released, and Own Recognizance. This provides an initial validation sample of 724,524 cases.

### Data

The LAC provided UCI with four data sets, covering bookings, criminal history, bench warrants, and subject demographics. (Table 3.1) We were also provided spreadsheets listing offenses classified as violent according to the criteria of the Judicial Council of California, and offenses disqualifying a defendant from consideration for pre-trial release from custody (Appendix Tables B.1, C.1, D.1).

**Table 3.1: LA Court-Provided Data Sets**

Data Set	Coverage	Records	Cases	Subjects
Bookings	Jan 2015- Dec 2018	1,765,119	1,054,278	476,480
History	June 1962 - Feb 2020	4,941,991	1,629,842	190,129
Warrants	Oct 1980 - Jan 2020	1,707,248	952,478	86,329
Demographic	NA	476,480	NA	476,480

<sup>7</sup> TCIS is the Court's case management system.

<sup>8</sup> After data were received, it was possible to calculate the time from booking to sentencing. The mean is 208 days with a standard deviation of 198 days. The median shows that 50% of cases are sentenced within 142 days; 85% are disposed of within one year.

<sup>9</sup> Individuals had between one and 85 bookings in the four-year sample window.

## Bookings

The Booking dataset identifies membership in the study cohort and provides information on current charges (risk factors) and arrests and violent arrests after release (outcomes). It includes bookings based on arrests from January 2015 through December 2018. It includes arrest date, booking date, release date, release reason, charges, and booking number. Release reason is used (in combination with exclusions, see below) to identify people with Bail, Bond, Cite and Release, and Own Recognizance release who are included in the validation sample for this study. Booking number is the key field necessary to link bookings to offender criminal history. Bookings data includes almost 1.8 million records for over one million bookings of 476,480 individuals.

Table 3.2 shows the percentage of bookings in the validation sample (PSAPop) versus non-eligible bookings (NonPSAPop) by year of arrest.<sup>10</sup> The “All” column show the yearly totals as a number and as a percent of the 4-year total. There is no suggestion of a trend over the years.

**Table 3.2: Booking Year by Study Group**

Arrest Year	Study Group		
	NonPSAPop	PSAPop	All
<b>2015</b>	262233	181702	443935
	59.07	40.93	25.15
<b>2016</b>	265950	182406	448356
	59.32	40.68	25.40
<b>2017</b>	259575	179780	439355
	59.08	40.92	24.89
<b>2018</b>	252745	180636	433381
	58.32	41.68	24.55
<b>Total</b>	1040503	724524	1765027
	58.95	41.05	100.00

## History

The History dataset provides prior misdemeanor, felony, and violent convictions and previous incarceration sentences. It also provides case filing date, charges, and sentence and probation information. The disposition and the disposition date of each charge are used to identify the follow-up

---

<sup>10</sup> In order to promote clarity and simplify interpretation, we have left unnecessary frequencies or counts out of tables. For groups, we use column or row percents as appropriate to compare distributions across groups. For outcomes, we show the percent within the group (a release type, a risk score) experiencing the event, which is a row percent when both Yes and No columns are used. However, the rightmost or Total column uses column percents, showing the distribution of the group by risk score or release type.

window, which is used for capturing failures of early release outcomes and for identifying cases pending final disposition at time of an arrest. According to LA Court project staff, there is no nexus between bookings and cases filed. The LA Court project staff made great efforts to link history and cases, and assigned a booking number to a case when possible. Approximately 235,000 cases appeared in both bookings and history. This is 22% of bookings. In some analyses, we use the measured follow-up time where a booking was linked to a history case to impute follow-up time for bookings lacking this link. The History dataset also includes CASE\_ID, used to link warrants to specific cases. History data includes nearly 5 million records from 1.6 million cases covering 190,129 individuals.

## Warrants

The Warrants dataset consists of bench warrants issued for failure to appear. It includes the warrant issue date and warrant quashed date, allowing us filter out of warrants quashed the same day as issued per PSA scoring rules. The CASE\_ID is used to link warrants to cases and determine whether post release warrants are counted as Failure to Appear because they are issued before disposition of the final outstanding charge or whether they fall outside of the follow-up window. Warrants data includes 1.7 million records from 952,000 cases issued for 86,329 individuals.

## Demographics

The Subjects dataset includes date of birth, used along with date of arrest to determine age at arrest. Along with gender and descent, age is also used to describe the population in the study. OffenderId is the link to all other datasets. Subjects includes the 476,480 individuals found in the Bookings data.

## Tables for Classification of Offenses

### Violent Offenses

Two risk factors and one outcome in the PSA require identification of violent offenses. The PSA was developed with a specific classification of offenses as violent; variations from this classification could affect predictive performance. Implementation support includes a Guide to the PSA Violent Offense List. The Judicial Council developed a consensus list of 226 violent offenses specific to the legal context and statute code of the state of California. (Appendix Table B.1) These were identified by specification of the statute code section, paragraph, and subsection as appropriate. Comparing these designations to how offenses were recorded in the records supplied to the CEBC, it was determined that offenses covered by the consensus list might be specified differently, with more or less specificity as to what part of the statute defined the offense. Comparing the consensus list to the provisions in the code and in consultation with LA Courts project staff, we added an additional 120 statute code specifications to the list. This allows the list of violent offense to be merged with the lists of prior or current offenses in the data to flag those to count for the PSA violent risk factors.

### Exclusions

Certain offenders are ineligible for release prior to a hearing in open court, meaning they cannot be released at Bail Deviation and must await their arraignment. The LA Courts provided the CEBC with listings of 136 offenses excluded under statutes PC 1270.1 and 85 statutes excluded under PC 1319.5. As there is some overlap between these two lists, the CEBC combined them into one list of 187 statute code offenses. While most offenses listed in the exclusions are sufficient on their own to block release,

the offenses listed under PC 1319.5 are exclusionary only if there have been 3 or more FTAs within the 3 years preceding the instant arrest; thus exclusions are classified by type. The exclusions list (Appendix Table C.1) is merged with the charges in bookings and used to divert PSA results for excluded cases into their own output file, where they may be analyzed separately if desired in the future.

### Offense Classification

Offenses are recorded in the Bookings and History datasets by California statute code section and statute. Several hundred are found in the data. As not all of these are common and broadly recognized, we used a system developed for California risk assessment instruments that maps these statutes to 20 broad offense categories. Categories include felonies and misdemeanors, supervision violations, and a miscellaneous category that covers charges for offenses such as public disorder, fish and game violations, and infractions (Appendix Table D.1). This classification is used to report the new offenses charged in the follow-up window by charged offenders released into the community.

### Data Limitations

Due to the nature and structure of the data, measuring the outcomes presents challenges. The PSA counts events occurring in a window between the booking date and the date at which all charges in a case have been sentenced. That is, the purpose of the PSA is not to predict general recidivism or FTAs that may occur during probation, but only the additional burdens that might arise from not holding charged offenders in custody between booking and sentencing. The first challenge is determining when all counts in a booking were finally sentenced. LAC colleagues attempted to match bookings to cases where possible, allowing determination of the window for counting new criminal activity. Not all bookings were matched. Only 22.3% of bookings have a link to a case with a sentence date. Within the validation sample, however, 66.4% of cases have a sentence date. In these cases where a sentence date is not available, we estimated the window from release to sentence using the median number of days between these two events for the specific release type.<sup>11</sup>

The additional challenge for new criminal activity is found in collecting all records of arrests within the follow-up window. PSA instructions define arrests as measure of new criminal activity, favoring use of bookings. Booking records, however, were selected by date to fall within the recruitment window of the study. Therefore, we have no arrest data from January 1, 2019 onward, leading to right-censorship of outcomes if we use bookings with a sentencing date outside of the recruitment window. One solution, adopted in one of the analysis options below, is to drop bookings with a sentencing date outside of the recruitment window from the analysis.

Another approach is to use charges brought in cases from the history records as the measure of new criminal activity. This raises another challenge. The criminal history does not include date of arrest, forcing the use of the case filing date to estimate whether the offense may have occurred in the window between a booking and the date when the case based on the booking was finally sentenced. As a case might not be filed for a year or more after a booking, many NCAs/NVCAs can be missed. However, the

---

<sup>11</sup> The median number of days from release to sentencing of the final charge filed in a case is: Bail, 208; Bond, 206; Cite, 217; Own Recognizance, 156.

Courts might judge that a charge brought to trial is a more relevant outcome than an arrest charge. This would favor the use of criminal history records, at the cost of missing new arrests made inside the follow-up window but filed after the sentence date of the charges in the incident booking. It is also necessary to filter out case filings based on the incident booking and (given that some cases lack the booking number necessary to link them to a booking) any charges matching the incident booking.

To compensate for the shortcomings of both datasets, we can also combine history cases and bookings arrests. Some charges may occur in both, but this is not a problem. If the PSA outcomes counted the number of new arrests, this could lead to double-counting. However, the PSA uses instead a dichotomous Yes/No measure. Thus, a charge is found in bookings records or case history records or both produces the same outcome. The problem is not double counting, but whether a charge may have changed between booking and filing. For example, reviewing records identified cases in which weapons offenses charged on booking were dropped in the case filing.

Thus, there is not one determinative measure of new criminal activity outcomes, but three possible measures, based on bookings records, history records, or both. We will adopt the precautionary approach, which is to use a new offense found in either the bookings or criminal history records, in one of our analysis below.

Two other limitations in the available data will affect the selection of the sample of bookings to use for analysis. One is based on the fact that not all charges are properly formatted in the records, due to data entry errors and non-standard formats used by agencies outside of Los Angeles County. In these records, the value of "NULL" is entered in the CCHG\_CJIS\_CHARGE\_CODE\_TYPE field. We recorded how many charges in the pre-booking history had a bad entry as a quality control measure; 3.4% of all history records have a "NULL" Code Type. The second restriction is based on the need for Level of Charge to identify prior felonies and misdemeanors. Since some offenses can be charged at either the felony or misdemeanor level (so-called "wobblers"), a separate indicator of charge level is needed; 4.2% of all history records are missing this information. Since the average individual has multiple records in the history data, these enter into many criminal history records. Of 183,737 individuals in the dataset ready for analysis, 31.9% have a "NULL" Code Type in one or more records; 28.1% have one or more records missing Level of Charge. UCI has some experience with recovering charges from improperly formatted records. We address these problems in the analysis options presented below.

## Data Processing for PSA Scale Construction

There are two required stages in processing the data to create PSA risk items, risk scores, and outcome measures. In the first stage, steps 1 to 3 below, each of the three main data files is processed separately to prepare it for the next stage and operations usually involve one or two fields at a time. In the second stage, the three files are combined and event records compared across source files to create risk items and outcomes.

- 1) History is processed to flag violent offenses, classify offense type, and determine the sentencing date of the last count adjudicated.

- 2) Warrants are processed to drop all warrants quashed or recalled the same day as issued, and flag warrants issued during the follow-up window.
- 3) Bookings are processed to
  - a. flag violent offenses,
  - b. classify offense type,
  - c. flag disqualifying offenses
  - d. attach the sentencing date when the booking can be linked to a case in the history file
  - e. estimate a sentencing date when the booking cannot be linked to a case in the history file, based on the median time to the sentencing date by release type
  - f. merge with demographic data and use birthdate to calculate age at arrest
  - g. aggregate information across the counts of the booking to flag current violent offenses and current violent offenses at 20 or younger
  - h. Select Bail, Bond, Cites, and Own Recognizance releases for study inclusion.
- 4) Merge bookings selected for study inclusion with history files to create PSA risk predictors (pending cases, prior offenses and incarceration), and PSA outcomes based on post-booking offenses.
- 5) Merge selected bookings with offenses in the bookings files to create PSA outcomes.
- 6) Merge selected bookings with warrants to count prior FTAs, post-booking FTAs, and all FTAs as required for one of the disqualifying offense categories.
- 7) Merge the data across the source files and collect the risk factors and outcomes for each booking.
- 8) Evaluate the disqualification criteria and filter such bookings out of the analysis data set.

## Measures of Predictive Ability

We use three assessments of the predictive utility of the PSA pretrial risk assessment tool: comparison of the rate of each outcome to the risk score; the Odds Ratio, or the average increase in odds of failure by score; and the Area Under the Curve (AUC), a statistic that provide an overall measure of predictive ability and allows comparison of different assessments across different jurisdictions.

### Recidivism by Risk Score

These tables show how the rate of FTA and recidivism changes with the PSA risk score for each of the three outcome types: Failure to Appear (FTA); New Arrest on any charge (NCA); and New Arrest on a violent charge (NVCA). If the risk assessment tool is working as intended, the rate of the outcome should rise with each increase in the risk score. It is desirable to be able to differentiate between low, medium, and high rates, and achieve relatively stable increases with each risk score increment and a somewhat equal distribution of the population across the scores. Some suggest that a good target is outcome rates ranging from one-half to twice the base rate.

### Odds Ratio

Other measures provide a more global assessment of the relationship between predictors and an outcome. Odds Ratio Estimates gives us the average increase in the odds of failure with each one

point increase in the risk score. An Odds Ratio of 1.5 tells us that the average increase is 50%.<sup>12</sup> The Confidence Limits tell us that we can be 95% confident that the true value of the odds ratio in the population lies within the range between the lower and upper limit.

### Understanding the AUC

Most risk assessment instruments produce a risk score, where a lower score indicates low risk and the risk rises with the score. In use, policy makers decide what level of risk should correspond with what kind of treatment, and select cut-point scores to divide the population into appropriate groups. Only after the cut-points are chosen and individuals are assigned to groups can we make a prediction and measure the accuracy of the prediction. For example, we judge that people with a score under five are unlikely to recidivate and can potentially be released from custody. Then we can measure how many in the low risk group succeed in avoiding a future arrest, and how many fail. The proportion of correct predictions (of both success and failure) is the accuracy of a prediction. In this application of the PSA, the Superior Courts do not set a cut-point score that determines release, instead using risk scores as an additional factor to guide the decision of judicial officers. In this case, accuracy and related statistics are not directly applicable.

Many measures of predictive accuracy are affected by the base rate<sup>13</sup> of the outcome in question and the choice of a cut-point and therefore can't be used to compare instruments across different outcomes or jurisdictions. The Area Under the Curve (AUC) of the Receiver Operating Characteristic Curve<sup>14</sup> is a measure of the overall ability of the risk score to distinguish one outcome from another regardless of base rate or cut-point. Therefore, it's very useful for comparing the predictive ability of risk instruments across different outcomes and jurisdictions.

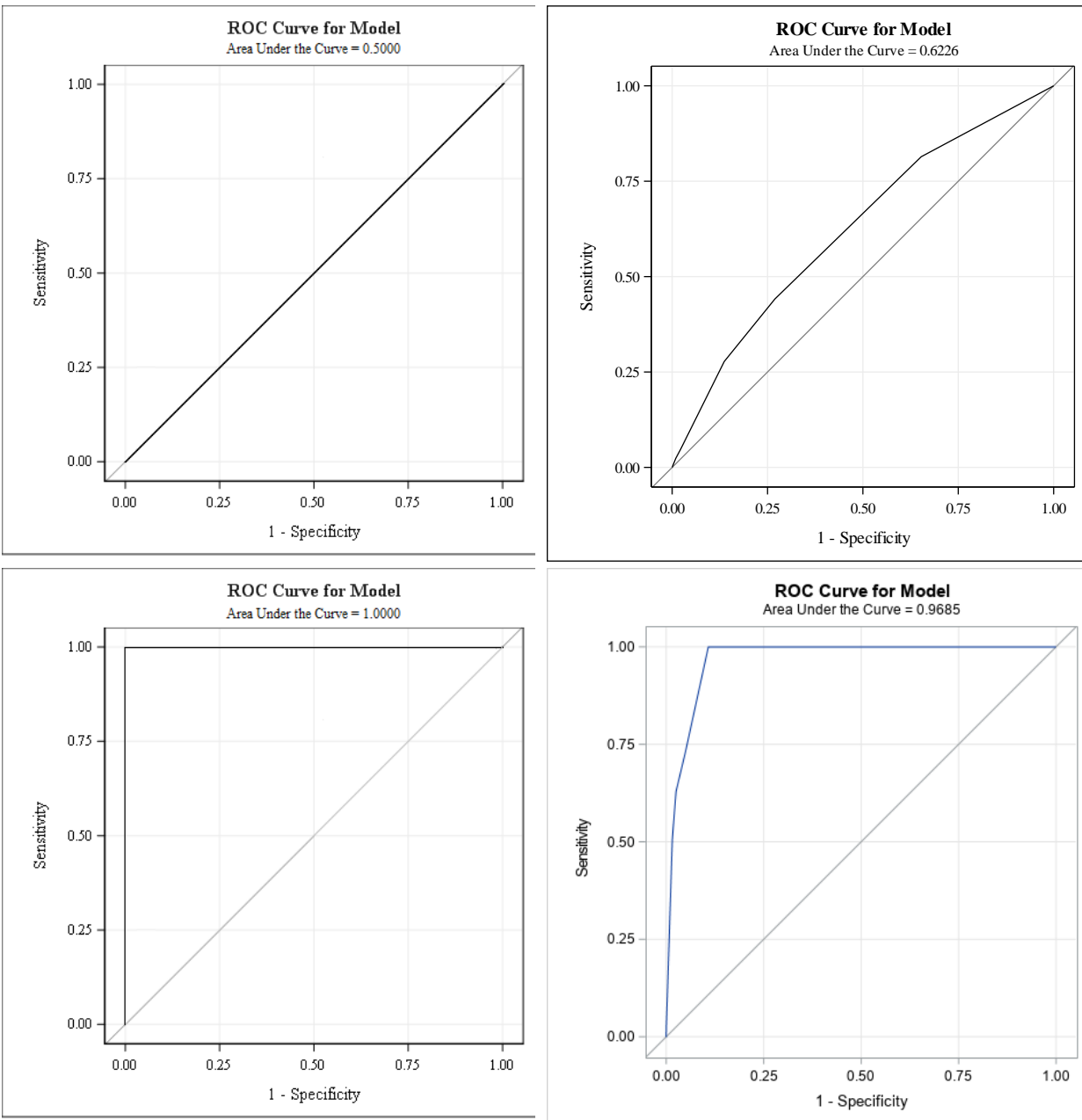
---

<sup>12</sup> "Odds" are different from risk. Risk is measured as the probability of failure. The odds are the probability of failing divided by the probability of not failing. The odds ratio is the odds of failure at score X+1 divided by the odds of failure at score X.

<sup>13</sup> That is, the overall rate of occurrence in the population sample.

<sup>14</sup> The Receiver Operating Characteristic Curve is so named as it was developed in WWII to evaluate the ability of radio receivers to discriminate signal from noise and interference.

Figure 3.1: Receiver Operating Characteristic Curves and Area Under the Curve



Examples of the ROC Curve are displayed in Figure 3.1. The curves show accuracy across all possible cut-points in the risk score. If we start from a very high cutoff score, we will not predict that anyone will recidivate. This is the 0,0 point on the graph. If we lower the cut-point to 0, we predict that everyone will recidivate - this is the 1,1 point. Intuitively, as we lower the cut-point, we want the proportion of true positive predictions (sensitivity, the vertical axis) to increase faster than the number of false positive predictions (1-specificity, the horizontal axis). If true and false positives increase at the same rate, the curve follows the diagonal and the results are pure chance. If we manage to predict all true positives



before adding any false positives, the curve climbs vertically from 0,0 to 0,1 before moving horizontally to the right, and the entire area of the graph is under the curve. An AUC of 1.0 indicates perfect prediction; an AUC of .5 indicates predictions no better than chance. Although there are no hard and fast rules for what are considered strong versus weak tools, AUCs over .71 are considered strong; tools under .56 are weak.<sup>15</sup>

## 4. Results

### Sample Description

#### Individuals

While the unit of analysis and hence the population is bookings, the demographic characteristics of the individuals showing up in bookings is of interest. First we need to classify bookings into two categories, those eligible for release at Bail Deviation and thus given a PSA assessment, and those not eligible. An important detail is that some bookings which lead to a release, however, occur on a booking that is ineligible for release at Bail Deviation and thus are excluded from this study. For convenience, we label these two groups “PSA Eligible” and “PSA Ineligible”.

Individuals can have more than one booking. Since some bookings of an individual can lead to a release leading to inclusion in the validation sample while others may not, the individual could be classified with either the PSA Eligible or the PSA Ineligible population for purposes of sample description. As we are interested in maximizing the numbers in the validation sample, we group them with the PSA Eligible Population if any booking qualifies. We also break out the individuals with a release into the validation sample who also have records in the history dataset, as the history of convictions is necessary to use the PSA tool. It is possible that for many of the 60.1% of individuals without history records, the booking was their first arrest, a possibility further suggested by the fact that the age distribution of the subsample with history skews older than the overall population. Lacking confirmation at the time of this report, we proceeded with a validation sample only including individuals with history records in the data available as a prudential step. This avoids the possibility of falsely attributing no prior convictions or pending cases, leading to an incorrect risk score.

Comparing the group with an eligible release to the ineligible population, we see modest differences (Table 4.1). The eligibles are somewhat younger (46% under 20 versus 38%), slightly more female (24% versus 21%), more likely to have a descent other than Black (80% versus 73%). It is also informative to compare those in the validation sample with the total recruitment sample. Fewer than half as many in the validation sample are under 30 (19%) as in the total population (43%). While the eligible population is a close match to the total on sex, females were 6 percentage points lower in the group with history.<sup>16</sup>

---

<sup>15</sup> Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior*, 29(5), 615-620; Skeem, J. L., & Lowenkamp, C. T. (2016). Risk, race, and recidivism: Predictive bias and disparate impact. *Criminology*, 54(4), 680-712.

<sup>16</sup> This could occur if females are less likely to be arrested and tried and build a case history.

The disparity on descent is slightly smaller than the 7 percentage point difference in the eligible versus non-eligible contrast; Black descent is 26% of the validation sample versus 22% of the total sample.

**Table 4.1: Individual Demographics**

Values	Release Group (%)			Study Sample Total n=476,480
	PSA Ineligible Population n= 154,910	PSA Eligible Population All n=321,570	PSA Eligible Population With History n=116,915	
<b>Age</b>				
<b>Under 18</b>	0.01	0.01	0.01	0.01
<b>18-19</b>	4.52	5.89	0.09	5.44
<b>20-29</b>	33.74	39.78	18.83	37.81
<b>30-39</b>	27.29	26.42	34.71	26.70
<b>40-49</b>	17.58	15.25	24.42	16.01
<b>50-59</b>	12.47	9.35	16.54	10.37
<b>60-69</b>	3.71	2.77	4.71	3.08
<b>70+</b>	0.68	0.54	0.69	0.58
<b>Sex</b>				
<b>Female</b>	21.26	23.93	17.56	23.06
<b>Male</b>	78.74	76.07	82.44	76.94
<b>Other<sup>1</sup></b>	0.00	0.00	0.00	0.00
<b>Descent</b>				
<b>Asian</b>	0.96	1.46	0.97	1.30
<b>Black</b>	27.11	19.78	25.67	22.16
<b>Hispanic</b>	50.16	53.02	51.88	52.09
<b>Other</b>	4.24	5.77	3.08	5.27
<b>White</b>	17.52	19.97	18.40	19.17

Table Note 1: “Other” includes 11 individuals who requested another designation.

### Bookings

The population of bookings (Table 4.2.a) mirrors the patterns of the population of individuals, weighted by the number of bookings an individual or demographic group has. The average individual has 2.2 bookings; 66% have 1, 90% have 3 or fewer, 99% have 9 or fewer, and the remaining 1% have up to 85

bookings in the 4-year span of the recruitment window. The overall distribution for age and descent are about the same, with differences of only a few percentage points. The proportion of female bookings overall is down from 23% to 20%, suggesting 15% fewer bookings per female. The number of bookings in 4 years is patterned as one might expect; those eligible for release have fewer bookings than those not eligible; among those eligible, those who can be connected to a case history have more frequent bookings than those who can't. "Exclusions" cover those with an excluded offense among the booking charges. Of the bookings supplied for this study, only 7.8% are actually excluded. Some are excluded for offenses covered by PC 1270.1. Exclusions under PC 1319.5 are only effective if the individual also has 3 or more FTAs in the prior 3 years; this is common especially among the PSA eligible with history, but this exclusion type is rare (1% of all bookings). This criterion uses all FTAs including ones issued after sentencing; as it is counted differently from FTAs used as PSA risk factors, we tabulate this field separately here. The first table of descriptive statistics for bookings also includes the two PSA risk factors based on the current/incident offense charges, violent offense and violent offense under 20. We see that both are less common among those considered for release than those not eligible, and a couple of percentage points rarer still among those with history.

**Table 4.2.a: Booking Demographics (All Bookings)**

Values	Release Group (%)			Total n=1,054,630
	PSA Ineligible Population n=570,067	PSA Eligible Population With History		
		All n=484,563	n=197,225	
<b>Age</b>				
<b>Under 18</b>	0.00	0.00	0.01	0.00
<b>18-19</b>	3.99	5.25	0.08	4.57
<b>20-29</b>	38.25	40.60	20.23	39.33
<b>30-39</b>	28.56	27.18	36.03	27.93
<b>40-49</b>	16.13	15.08	23.67	15.65
<b>50-59</b>	10.25	9.01	15.40	9.68
<b>60-69</b>	2.48	2.45	4.05	2.47
<b>70+</b>	0.33	0.42	0.52	0.37
<b>Sex</b>				
<b>Female</b>	18.28	22.33	16.73	20.14
<b>Male</b>	81.71	77.67	83.27	79.85
<b>Other</b>	0.01	0.00	0.00	0.00

**Table 4.2.a: Booking Demographics (All Bookings) continued**

Values	Release Group (%)			Total n=1,054,630
	PSA Ineligible Population n=570,067	PSA Eligible Population		
		All n=484,563	With History n=197,225	
<b>Descent</b>				
Asian	0.77	1.24	0.90	0.99
Black	26.11	20.04	24.17	23.32
Hispanic	51.61	53.21	52.27	52.35
Other	3.12	4.91	2.81	3.94
White	18.40	20.60	19.84	19.41
<b>Bookings</b>				
1	35.45	56.65	46.33	45.19
2-5	46.67	32.15	38.23	40.00
6-20	17.15	10.40	14.23	14.05
21-50	0.70	0.77	1.17	0.73
51-85	0.03	0.03	0.04	0.03
<b>Exclusions</b>				
No	90.52	94.10	94.94	92.17
Yes	9.48	5.90	5.06	7.83
<b>FTAs Before Current Offense</b>				
0-2	62.69	75.77	40.46	68.70
3+	37.31	24.23	59.54	31.30
<b>Violent Current Offense</b>				
No	77.21	85.94	88.29	81.22
Yes	22.79	14.06	11.71	18.78
<b>Violent Current Offense at &lt;=20</b>				
No	97.95	98.71	99.97	98.30
Yes	2.05	1.29	0.03	1.70

The description fields for the unit of analysis above depend only on the data found in the bookings records. The subsequent fields depend on data from the history records, so Table 4.2.b includes only the bookings of individuals with case records in the history dataset. The first two fields (4.2.b.1) cover the

quality and completeness of the data in the history records. Overall 35% of bookings have unrecovered charges in the history records; 6% have 5 or more. Here, “unrecovered” means that we were not able to match them to the California Department of Justice’s list of offense-defining statute codes. Where these still have level of offense, we are able to use them for prior felony and misdemeanor convictions, but not for violent offenses. Level of offense is missing for 30% of the overall number of bookings with history. Many of the unrecovered charges arise from violations of municipal codes and are unlikely to rise to the level of a misdemeanor or violent offense. Others are ill-formed charges, perhaps data entry errors, which on inspection designate sometimes violent or otherwise serious offenses. We make an effort to recover these, as noted above.

**Table 4.2.b.1: Bookings with History: Data Completeness**

Values	Release Group (%)		Total
	PSA Ineligible Population	PSA Eligible Population	
	n= 302,724	n=197,204	n=499,928
<b>Unrecovered Charges</b>			
<b>0</b>	62.76	68.54	65.04
<b>1-2</b>	24.51	21.18	23.19
<b>3-4</b>	6.70	5.33	6.16
<b>5-110</b>	6.04	4.95	5.61
<b>Missing Level of Charge</b>			
<b>0</b>	68.62	72.11	69.99
<b>1-2</b>	17.27	16.10	16.81
<b>3-4</b>	6.68	5.88	6.37
<b>5-100</b>	7.43	5.91	6.83

The next 8 fields show the counts of the PSA risk factors (4.2.b.2). Here we note mainly the high rate of priors. Incarceration is defined for the purpose of scoring the PSA as a term of 14 days or more. Flash incarceration for a weekend does not count. When a sentence included multiple terms of incarceration, we assumed they were consecutive (rather than concurrent) with the guidance of LAC personnel; the difference of assuming concurrent terms was only one percentage point.

**Table 4.2.b.2: Bookings with History: Record**

Values	Release Group (%)		Total
	PSA Ineligible Population n= 302,724	PSA Eligible Population n=197,204	
<b>Pending Case at Arrest</b>			
No	97.52	96.49	97.11
Yes	2.48	3.51	2.89
<b>FTA Within Past 2 Years</b>			
0	61.42	67.98	64.01
1	17.55	13.57	15.98
2+	21.03	18.45	20.01
<b>FTA Older than 2 Years</b>			
No	33.25	39.64	35.77
Yes	66.75	60.36	64.23
<b>Prior Misdemeanor Conviction</b>			
No	8.32	12.32	9.90
Yes	91.68	87.68	90.10
<b>Prior Felony Conviction</b>			
No	30.30	45.95	36.47
Yes	69.70	54.05	63.53
<b>Prior Conviction For Misdemeanor or Felony</b>			
No	3.42	8.36	5.37
Yes	96.58	91.64	94.63
<b>Prior Violent Conviction</b>			
0	51.37	63.83	56.29
1	13.71	10.90	12.60
2	6.59	4.57	5.80
3+	28.32	20.70	25.32
<b>Prior Incarceration (if Consecutive Sentences)</b>			
No	15.47	28.92	20.77
Yes	84.53	71.08	79.23

The subsequent 12 fields show the results of four different measurement options for the three outcome measures of FTA, NCA, and NVCA (4.2.b.3). The first set is the most conservative. It uses only bookings

with a link to a case in the history records, allowing determination of a date for the final sentencing of all charges in the case. The second set, “Sentence Date Imputed”, imputes the median follow-up window based on release type for bookings without a sentence date.<sup>17</sup> The third set uses charges from court convictions to supplement arrests, given that arrests data is not available for the entire follow-up window for bookings later in the recruitment window. The convictions are filtered to make sure they don’t come from a case filed on the basis of the incident booking or replicate a charge in the incident booking. We find that that adding convictions only increase the rate only two to five percentage points for any new arrest and about one percentage point for violent arrests, suggesting that this is a reasonable step unlikely to make a major difference in the analysis.

It is also of note that the Ineligible Population shows a rate of new criminal charges of about 5% and about 1% for violent criminal charges. Since these individuals are presumably in jail or under other supervision, it may indicate criminal activity occurring under these conditions, or possibly charges from and ongoing investigation now filed. Presumably, these represent costs to law enforcement more than public safety. It is a policy question as to whether to balance these offenses under supervision against the burden of offenses on release.

---

<sup>17</sup> The median was chosen rather than the mean because the PSA is based on counting charges, not the time to failure. Therefore, a median assures us that while some charges will be lost when they occur after the imputed sentencing date, an equal number will have been added before the imputed date. If this imputation fails to balance false positives and false negatives, we would expect to see a substantial effect on the AUC.

**Table 4.2.b.3: Bookings with History: Outcomes**

Values	Release Group (%)		
	Ineligible Population n= 302,724	PSA Eligible Population n=197,204	Total n=499,928
<b>New FTA</b>			
No	62.30	43.76	54.99
Yes	37.70	56.24	45.01
<b>New Arrest (Sentence Date Available)</b>			
No	95.38	68.03	84.59
Yes	4.62	31.97	15.41
<b>New Violent Arrest (Sentence Date Available)</b>			
No	99.09	93.82	97.01
Yes	0.91	6.18	2.99
<b>New Arrest (Sentence Date Imputed)</b>			
No	95.38	52.12	78.32
Yes	4.62	47.88	21.68
<b>New Violent Arrest (Sentence Date Imputed)</b>			
No	99.09	90.20	95.59
Yes	0.91	9.80	4.41
<b>New Arrest or Conviction (Sentence Date Available)</b>			
No	94.96	65.62	83.39
Yes	5.04	34.38	16.61
<b>New Violent Arrest or Conviction (Sentence Date Available)</b>			
No	98.85	93.09	96.58
Yes	1.15	6.91	3.42
<b>New Arrest or Conviction (Sentence Date Imputed)</b>			
No	94.96	46.73	75.94
Yes	5.04	53.27	24.06
<b>New Violent Arrest or Conviction (Sentence Date Imputed)</b>			
No	98.85	88.74	94.86
Yes	1.15	11.26	5.14



## Outcomes by Release Type

Tables of outcomes by release type allow comparison and judgement of whether some subgroup of the released population presents a greater risk than another. The first table compares outcomes where sentence date is available; the second table uses the larger number of bookings available when sentence date imputation is applied. Both tables are based on arrests from the bookings records. The tables both show a similar pattern, with lower rates of risk for bail and own recognizance releases and the highest rates among citation-only releases. Bonded releases are relatively low risk for failure to appear, in the middle for a new arrest, and higher for a new violent arrest. Comparing the two tables, the risk appears slightly lower (7%) when using imputed sentence dates, suggesting interpreting results of analysis with a bit of caution when these bookings are included. The differences in outcomes arise mainly from the Cite release type group, which may suggest further investigation.

**Table 4.3.a: Outcomes by Release Type (Bookings with Sentence Date)**

<b>Outcome by Release Type (%)</b> (Bookings with Sentence Date)				
<b>Release Type</b>	<b>FTA</b>	<b>NCA</b>	<b>NVCA</b>	<b>Number % of Total</b>
<b>Bail</b>	313	316	75	815
	38.40	38.77	9.20	0.68
<b>Bond</b>	11146	12494	3010	27361
	40.74	45.66	11.00	22.88
<b>Cite</b>	48775	41221	7262	67726
	72.02	60.86	10.72	56.62
<b>OR</b>	11287	9018	1850	23708
	47.61	38.04	7.80	19.82
<b>Total</b>	71521	63049	12197	119610
	59.80	52.71	10.20	100.00

**Table 4.3.b: Outcomes by Release Type (Include Bookings with Imputed Sentence Date)**

<b>Outcome by Release Type (%)</b>				
<b>(Include Bookings with Imputed Sentence Date)</b>				
<b>Release Type</b>	<b>FTA</b>	<b>NCA</b>	<b>NVCA</b>	<b>Number % of Total</b>
<b>Bail</b>	512	487	117	1463
	35.00	33.29	8.00	0.74
<b>Bond</b>	17865	19492	4819	48917
	36.52	39.85	9.85	24.81
<b>Cite</b>	71810	58530	10901	104144
	68.95	56.20	10.47	52.81
<b>OR</b>	20706	15913	3481	42670
	48.53	37.29	8.16	21.64
<b>Total</b>	110893	94422	19318	197194
	56.24	47.88	9.80	100.00

### **NCA Outcome Offense Category by Release Type**

#### **All Charges and Most Serious Charge**

Tables 4.4.a and 4.4.b show the crimes charged to individuals released before trial during the follow-up window from release to final sentencing based on new arrests in the bookings records. Table 4.4a uses all charges brought; Table 4.4b selects only the most serious offense charged. Missing/Invalid covers charges not found in our list of California crime-defining statute codes described above, including municipal code violations and data entry errors. Looking at all charges, the most common category is Misdemeanor Drug (31%), followed by Misdemeanor Property (16%), Felony Drug (7%), Miscellaneous and Felony Property (6%). Discounting possible felonies among Miscellaneous and Missing charges, felonies account for 21% of the total and misdemeanors for 62%. Drugs comprise 40% of the total, followed by property crimes (exclusive of violent property) at 22% and violent crimes (including violent property) at 16%. The subset of violent felonies comprises 8% of the total, with one quarter of those (2 percentage points) coming from Felony Weapon charges.

Table 4.4.a: New Arrests: All Charges by Offense Type and Release Type

Offense Type by Release Type (%)					
Offense Type	Release Type				
	Bail	Bond	Cite	OR	Total
Missing/Invalid	245	7712	56076	10659	74692
	8.73	8.26	10.48	9.09	9.98
Miscellaneous	269	8789	31068	6091	46217
	9.59	9.41	5.81	5.20	6.17
Felony Homicide	7	418	382	158	965
	0.25	0.45	0.07	0.13	0.13
Felony Sex	15	522	470	170	1177
	0.53	0.56	0.09	0.15	0.16
Felony Violent Property	117	4795	13367	4137	22416
	4.17	5.13	2.50	3.53	2.99
Felony Assault not Domestic	105	3498	9381	3008	15992
	3.74	3.75	1.75	2.57	2.14
Felony Domestic Assault/Violent	81	2144	3118	1052	6395
	2.89	2.30	0.58	0.90	0.85
Felony Weapon	116	6182	6362	1595	14255
	4.14	6.62	1.19	1.36	1.90
Felony Property	291	12280	24050	7455	44076
	10.37	13.15	4.49	6.36	5.89
Felony Drug	216	10600	32925	6904	50645
	7.70	11.35	6.15	5.89	6.77
Felony Escape	28	957	983	304	2272
	1.00	1.02	0.18	0.26	0.30
Misdemeanor Assault not Domestic	87	1758	15015	3606	20466
	3.10	1.88	2.81	3.08	2.73
Misdemeanor Domestic Assault/Violent	99	3429	11571	3639	18738
	3.53	3.67	2.16	3.10	2.50

**Table 4.4.a: New Arrests: All Charges by Offense Type and Release Type, continued**

<b>Offense Type by Release Type (%)</b>					
<b>Offense Type</b>	<b>Release Type</b>				
	<b>Bail</b>	<b>Bond</b>	<b>Cite</b>	<b>OR</b>	<b>Total</b>
<b>Misdemeanor Sex</b>	25 0.89	914 0.98	6031 1.13	3380 2.88	10350 1.38
<b>Misdemeanor Other Domestic Violent</b>	6 0.21	348 0.37	1043 0.19	376 0.32	1773 0.24
<b>Misdemeanor Weapon</b>	29 1.03	995 1.07	4320 0.81	1035 0.88	6379 0.85
<b>Misdemeanor Property</b>	343 12.23	9359 10.02	90004 16.82	23583 20.12	123289 16.47
<b>Misdemeanor Drug</b>	503 17.93	12179 13.04	190326 35.57	27628 23.57	230636 30.81
<b>Misdemeanor Escapes</b>	98 3.49	2295 2.46	25357 4.74	8146 6.95	35896 4.80
<b>Misdemeanor Alcohol</b>	111 3.96	3590 3.84	10439 1.95	3696 3.15	17836 2.38
<b>Supervision Violations</b>	14 0.50	629 0.67	2800 0.52	578 0.49	4021 0.54
<b>Total</b>	2805 0.37	93393 12.48	535088 71.49	117200 15.66	748486 100.00

Considering only the most serious charges in a booking, the most common category is again Misdemeanor Drug (17%) followed by Misdemeanor Property (14%), but now Felony Property (12%) moves to 3<sup>rd</sup> most common, and Felony Violent Property (10%) moves up to 4<sup>th</sup>. Felony Drug (8%) slips to 5<sup>th</sup>. Felonies and misdemeanors each now account for 47% of the total. Violent crimes now predominate with 37% of the total. Drugs drop by a quarter to 29% of the total, followed closely by property crimes at 26%, with violent felonies up to 25% of the total, with one fifth of those coming (5 percentage points) from Felony Weapon charges.

Table 4.4.b: New Arrests Most Serious Charge by Offense Type and Release Type

Offense Type by Release Type (%)					
Offense Type	Release Type				Total
	Bail	Bond	Cite	OR	
Missing/Invalid	32	1079	2721	1035	4867
	3.34	3.21	2.52	2.99	2.75
Miscellaneous	48	1116	2864	823	4851
	5.01	3.32	2.66	2.38	2.74
Felony Homicide	7	400	364	155	926
	0.73	1.19	0.34	0.45	0.52
Felony Sex	13	464	383	151	1011
	1.36	1.38	0.36	0.44	0.57
Felony Violent Property	93	3689	10886	3471	18139
	9.70	10.98	10.10	10.02	10.25
Felony Assault not Domestic	79	2459	6304	2191	11033
	8.24	7.32	5.85	6.33	6.23
Felony Domestic Assault/Violent	61	1738	2464	895	5158
	6.36	5.17	2.29	2.58	2.91
Felony Weapon	64	3735	3860	1091	8750
	6.67	11.12	3.58	3.15	4.94
Felony Property	147	5669	12033	4183	22032
	15.33	16.87	11.16	12.08	12.45
Felony Drug	88	4364	7774	2660	14886
	9.18	12.99	7.21	7.68	8.41
Felony Escape	7	384	349	143	883
	0.73	1.14	0.32	0.41	0.50
Misdemeanor Assault not Domestic	27	538	4794	1368	6727
	2.82	1.60	4.45	3.95	3.80
Misdemeanor Domestic Assault/Violent	43	1524	4606	1723	7896
	4.48	4.54	4.27	4.98	4.46
Misdemeanor Sex	12	330	2008	1301	3651
	1.25	0.98	1.86	3.76	2.06

**Table 4.4.b: New Arrests Most Serious Charge by Offense Type and Release Type, continued**

Offense Type by Release Type (%)					
Offense Type	Release Type				
	Bail	Bond	Cite	OR	Total
<b>Misdemeanor Other Domestic Violent</b>	1 0.10	83 0.25	253 0.23	101 0.29	438 0.25
<b>Misdemeanor Weapon</b>	8 0.83	330 0.98	1349 1.25	392 1.13	2079 1.17
<b>Misdemeanor Property</b>	78 8.13	1799 5.35	17576 16.30	5397 15.59	24850 14.04
<b>Misdemeanor Drug</b>	98 10.22	2360 7.02	22820 21.17	5269 15.22	30547 17.26
<b>Misdemeanor Escapes</b>	14 1.46	200 0.60	1094 1.01	458 1.32	1766 1.00
<b>Misdemeanor Alcohol</b>	36 3.75	1253 3.73	3138 2.91	1750 5.05	6177 3.49
<b>Supervision Violations</b>	3 0.31	84 0.25	158 0.15	69 0.20	314 0.18
<b>Total</b>	959 0.54	33598 18.98	107798 60.91	34626 19.56	176981 100.00

### Validation Options

We have noted several limitations in the data available for analyses. We can partially redress this challenge by presenting several analyses using different subsets of the data or different ways of obtaining outcomes. Constraining the analysis to bookings with the cleanest and most complete data gives the best estimate of the performance of the PSA tool in itself. Changing the constraints can show how well the PSA copes with missing data and how well it works in a real-world application. For example, how sensitive is it to having complete information on the level of a charge? A tool that gives good reliable results across different scenarios is said to be more robust.

To summarize, the factors we consider in choosing analysis options are:

- The availability of a link between history cases and bookings to allow determination of a follow-up window. Our choices are whether to limit the analysis to bookings with a determinable disposition date or to increase the number of bookings available for the analysis by imputing a disposition date.

- Missing charges and level of charge will affect the identification of felonies and misdemeanors and violent offenses. We will present analysis in which these are not part of the record in contrast to others where they are present.
- Arrests as a measure of new criminal activity may be right-censored in later bookings. We will present analysis both restricting bookings to those with a disposition date within the range of the bookings data, and analysis where more recent bookings are allowed with charges from arrests supplemented by charges from case filings.
- It may be possible to recover some of the records missing a useable charge or level of charge. This may improve reliability and performance, support analysis of a larger set of bookings, and offer the possibility of incorporation into LAC data processing resulting in more complete and useful data.

### Suggested Choices among Validation Options

Combining all factor possibilities would produce hundreds of options for analysis. We suggested several of them with an eye to answering likely questions about the ability of the PSA to support pretrial release decisions the courts. Four analyses were agreed on, presented below.

#### *Option A: Core Complete Data*

Option A is intended to test the intrinsic predictive ability of PSA by using the subset of bookings with the cleanest and most complete data. It is limited to bookings where Sentence Date is available and the disposition date is in 2018 or earlier, letting us use arrests alone for measuring new criminal activity per design intent. Bookings with NULL Code Types and missing Level of Charge are dropped.

#### *Option B: Core Data, Plus NULL and Missing Charge Level*

Option B recognizes that the PSA will be applied in less than ideal conditions where missing data is inevitable and provide an estimate of how much predictive ability is affected. Bookings with NULL Code Types and missing Level of Charge are included. Like Option A, it is limited to bookings where Sentence Date available, but disposition dates later than December of 2018 are allowed to include more cases for a more robust assessment. As a consequence, new criminal activity is picked up both from arrest charges in bookings records and filed charges from case history records.

#### *Option C: Imputation of Sentence Date*

Option C allows the use of the largest sample for validation by using imputation of a Sentence Date where the lack of a link via Booking Number to a court case history record prevents direct ascertainment. Like Option B, bookings with NULL Code Types and missing Level of Charge are included. Disposition dates later than December of 2018 are allowed and new criminal activity is picked up both from arrest charges in bookings records and filed charges from case history records.

#### *Option D: Imputation for Bad or Missing Data*

Option D is intended first to test the possibility of recovering ill-formed charges and imputing Level of Charge from offense information. The second intent is to assess whether this effort improves PSA performance. Most risk factors only count the first occurrence of an offense; adding a second instance

doesn't change the risk score. It is probable that most charges in a record for anyone with more than a few arrests will be redundant as far as the PSA prediction is concerned; recovering one more charge won't make a difference. Prior violent convictions is the one exception; the PSA counts up to three. A missing Level of Charge has a wider impact as the PSA only counts misdemeanors and felonies; unless a charge missing this information is flagged as a violent offense, it will not be counted among the priors. How far charge and level of charge recovery can succeed and whether it will change scores and predictive ability are empirical questions. To answer them, Option D will replicate Option C with imputation for bad and missing value, with additional supplemental analysis of the effect of imputation on the scores and predictions of those whose prior records have changed.

### Descriptive Table All Validation Options: Demographics, Data Completeness and Quality, PSA Factors and Outcomes

Table 4.5 describes the validation sample under the analysis options A, B, C, and D, with demographics, data quality and completeness, and PSA risk factors and outcomes measures. The demographics of all option groups (4.5.1) is similar, but collectively differ from the whole PSA eligible group in being older and proportionally more male by about five percentage points. The one selection criterion operating across all options is that they could be linked to criminal history files available to LA Courts staff. It seems reasonable to speculate that these characteristics are associated with a higher probability of having a criminal history in official records.

**Table 4.5.1: Validation Options: Demographics**

Values	Option A n=56,465	Option B n=112,440	Option C n=183,749	Option D n=183,663	PSA Eligible n=484,563	All n=1,054,630
<b>Age</b>						
<b>Under 18</b>	0.01	0.01	0.01	0.01	0.00	0.00
<b>18-19</b>	0.11	0.08	0.08	0.08	5.25	4.57
<b>20-29</b>	26.86	20.37	19.97	19.96	40.60	39.33
<b>30-39</b>	37.86	36.20	35.95	35.95	27.18	27.93
<b>40-49</b>	21.27	23.60	23.82	23.81	15.08	15.65
<b>50-59</b>	10.68	15.39	15.59	15.59	9.01	9.68
<b>60-69</b>	2.79	3.90	4.09	4.09	2.45	2.47
<b>70+</b>	0.42	0.46	0.49	0.49	0.42	0.37



**Table 4.5.1: Validation Options: Demographics continued**

Values	Option A n=56,465	Option B n=112,440	Option C n=183,749	Option D n=183,663	PSA Eligible n=484,563	All n=1,054,630
<b>Sex</b>						
Female	17.53	16.19	16.81	16.81	22.33	20.14
Male	82.47	83.80	83.19	83.18	77.67	79.85
Other	0.01	0.00	0.00	0.00	0.00	0.00
<b>Descent</b>						
Asian	1.03	0.99	0.92	0.92	0.90	0.99
Black	18.51	22.17	23.67	23.67	20.04	23.32
Hispanic	59.79	53.58	52.58	52.58	53.21	52.35
Other	3.21	2.86	2.79	2.79	4.91	3.94
White	17.45	20.41	20.04	20.04	20.60	19.41
<b>Bookings</b>						
1	67.73	60.32	58.69	58.69	56.65	45.19
2-5	28.21	32.93	34.84	34.84	32.15	40.00
6-20	4.00	6.45	6.18	6.18	10.40	14.05
21-50	0.06	0.29	0.26	0.26	0.77	0.73
51-85	0.00	0.02	0.02	0.02	0.03	0.03

**Table 4.5.2: Validation Options: Data Completeness**

Values	Option A n=56,465	Option B n=112,440	Option C n=183,749	Option D n=183,663	PSA Eligible n=484,563	All n=1,054,630
<b>Unrecovered Charges</b>						
0	100.00	68.97	68.09	90.31	68.54	65.04
1-2	0.00	20.97	21.41	8.22	21.18	23.19
3-4	0.00	5.18	5.41	1.06	5.33	6.16
5-110	0.00	4.88	5.09	0.41	4.95	5.61
<b>Missing Level Of Charge</b>						
0	100.00	72.62	71.94	72.86	72.11	69.99
1-2	0.00	16.01	16.16	16.03	16.10	16.81
3-4	0.00	5.71	5.92	5.66	5.88	6.37

The risk factors (4.5.3) and outcome measures (4.5.4) are also similar across Options B, C, and D groups. Option A consistently has lower percentages measuring positive for the risk factors and outcomes. It may be that they have more complete records because they have shorter records and fewer chances for a poorly recorded charge to be entered.

Table 4.5.3: Validation Options: Record

Values	Option A n=56,465	Option B n=112,440	Option C n=183,749	Option D n=183,663	PSA Eligible n=484,563	All n=1,054,630
<b>Violent Current Offense</b>						
No	93.37	94.08	92.83	92.69	85.94	81.22
Yes	6.63	5.92	7.17	7.31	14.06	18.78
<b>Violent Current Offense at &lt;=20</b>						
No	99.97	99.98	99.98	99.98	98.71	98.30
Yes	0.03	0.02	0.02	0.02	1.29	1.70
<b>Pending Case At Arrest</b>						
No	97.24	96.64	96.41	96.41	96.49	97.11
Yes	2.76	3.36	3.59	3.59	3.51	2.89
<b>FTA Within Past 2 Years</b>						
0	75.60	69.42	66.80	66.80	67.98	64.01
1	10.94	12.36	13.93	13.93	13.57	15.98
2+	13.46	18.22	19.27	19.27	18.45	20.01
<b>FTA Older Than 2 Years</b>						
No	49.47	39.67	38.27	38.27	39.64	35.77
Yes	50.53	60.33	61.73	61.73	60.36	64.23
<b>Prior Misdemeanor Conviction</b>						
No	18.53	12.28	12.01	12.01	12.32	9.90
Yes	81.47	87.72	87.99	87.99	87.68	90.10
<b>Prior Felony Conviction</b>						
No	57.58	45.86	45.63	45.58	45.95	36.47
Yes	42.42	54.14	54.37	54.42	54.05	63.53
<b>Prior Conviction for Misdemeanor Or Felony</b>						
No	13.78	8.67	8.17	8.16	8.36	5.37
Yes	86.22	91.33	91.83	91.84	91.64	94.63
<b>Prior Violent Conviction</b>						
0	71.90	64.24	63.89	63.76	63.83	56.29
1	8.79	10.85	10.89	10.84	10.90	12.60
2	3.27	4.37	4.54	4.58	4.57	5.80
3+	16.04	20.54	20.67	20.83	20.70	25.32
<b>Prior Incarceration (If Consecutive Sentences)</b>						
No	38.17	28.71	28.41	28.41	28.92	20.77
Yes	61.83	71.29	71.59	71.59	71.08	79.23

We may also note two counter-intuitive aspects of Option D. Although the inclusion criteria are the same as for Option C, there are a few dozen fewer cases in Option D. In these cases, one of the

recovered charges matched one of the exclusion criteria and they were dropped from the analysis group. Recovered charges are also behind the minimal decline of New Criminal Activity for Option D. The measure of new criminal activity is a new misdemeanor or felony; formally, this only applies to offenses punishable by incarceration but in this application all misdemeanors and felonies are counted. In adding convictions from the court history records to new arrests found in bookings, it became necessary to avoid counting cases filed on the basis of the incident booking. This was operationalized as a case filing listing the same booking number, or a case filing on the same charge within a year of post-booking release. Poorly formatted charges from history records could bypass this second filter and show up as new criminal activity; in some 75 cases the recovered charge now matched the booking charge, excluding it from counting as new activity.

**Table 4.5.4: Options: Outcomes**

Values	Option A n=56,465	Option B n=112,440	Option C n=183,749	Option D n=183,663	PSA Eligible n=484,563	All n=1,054,630
<b>New FTA</b>						
No	45.74	38.40	41.81	41.81	43.76	54.99
Yes	54.26	61.60	58.19	58.19	56.24	45.01
<b>New Arrest (Sentence Date Available)</b>						
No	50.78	na	na	na	68.03	84.59
Yes	49.22	na	na	na	31.97	15.41
<b>New Violent Arrest (Sentence Date Available)</b>						
No	91.00	na	na	na	93.82	97.01
Yes	9.00	na	na	na	6.18	2.99
<b>New Arrest (Sentence Date Imputed)</b>						
No	na	na	na	na	52.12	78.32
Yes	na	na	na	na	47.88	21.68
<b>New Violent Arrest (Sentence Date Imputed)</b>						
No	na	na	na	na	90.20	95.59
Yes	na	na	na	na	9.80	4.41
<b>New Arrest or Conviction (Sentence Date Available)</b>						
No	na	43.02	na	na	65.62	83.39
Yes	na	56.98	na	na	34.38	16.61
<b>New Violent Arrest or Conviction (Sentence Date Available)</b>						
No	na	89.32	na	na	93.09	96.58
Yes	na	10.68	na	na	6.91	3.42
<b>New Arrest or Conviction (Sentence Date Imputed)</b>						
No	na	na	46.78	46.84	46.73	75.94
Yes	na	na	53.22	53.16	53.27	24.06
<b>New Violent Arrest or Conviction (Sentence Date Imputed)</b>						
No	na	na	89.57	89.49	88.74	94.86
Yes	na	na	10.43	10.51	11.26	5.14

## Option A: Core Complete Data

### Option A: Parameters

To recapitulate, Option A is intended to test the intrinsic predictive ability of PSA by using the subset of bookings with the cleanest and most complete data. It is limited to bookings where Sentence Date available and the disposition date is in 2018 or earlier, letting us use arrests alone for measuring new criminal activity per design intent. Bookings with NULL Code Types and missing Level of Charge are dropped.

**Table 4.A.1: Option A: Outcomes by Release Type**

<b>Outcome by Release Type (%)</b>				
<b>(Bookings with Sentence Date)</b>				
<b>Release Type</b>	<b>FTA</b>	<b>NCA</b>	<b>NVCA</b>	<b>Total</b>
<b>Bail</b>	133	132	31	373
	35.66	35.39	8.31	0.66
<b>Bond</b>	5006	5419	1039	12171
	41.13	44.52	8.54	21.55
<b>Cite</b>	20631	18308	3250	31593
	65.30	57.95	10.29	55.95
<b>OR</b>	4870	3934	762	12328
	39.50	31.91	6.18	21.83
<b>Total</b>	<b>30640</b>	<b>27793</b>	<b>5082</b>	<b>56465</b>
	<b>54.26</b>	<b>49.22</b>	<b>9.00</b>	<b>100.00</b>

The pattern in Outcomes by Release Type (Table 4.A.1) is similar to the results for the PSA eligible population overall, with lower rates of risk for bail and own recognizance releases and the highest rates among citation only releases. Bonded releases are in the upper middle for all outcomes. For FTA, the range is 30 percentage points from 36% to 65% or from 66% to 120% of the overall rate for the group. The pattern is similar for NCA, with a range of 26 percentage points from 32% to 58% or from 65% to 118% of the overall rate. With a lower base rate of offending, for NVCA the range is only 4 percentage points from 6% to 10%, which still works out as from 69% to 114% of the group rate. Again the Cite group is the main driver of the differences, with the exception of NCA where there is still a 40% difference between OR and Bond releases.

## Option A: Outcome by Risk Score

**Table 4.A.2: Option A Failure to Appear by Risk Score**

FTA Risk Score	Failure to Appear after Release		
	No	Yes	Total
<b>1</b>	4438 69.69	1930 30.31	6368 11.28
<b>2</b>	11790 65.09	6323 34.91	18113 32.08
<b>3</b>	7665 42.85	10222 57.15	17887 31.68
<b>4</b>	1329 20.48	5161 79.52	6490 11.49
<b>5</b>	587 8.00	6754 92.00	7341 13.00
<b>6</b>	16 6.02	250 93.98	266 0.47
<b>Total</b>	<b>25825</b> <b>45.74</b>	<b>30640</b> <b>54.26</b>	<b>56465</b> <b>100.00</b>

The table of FTA by Risk Score shows good discrimination with a range of 64 percentage points from 30% to 94% or from 56% to 173% of the base rate, an average increase of 13 percentage points per increase in risk score. The discrimination is strongest in the middle range, declining sharply at the extremes, and scoring only 0.5% of the validation sample in the highest risk group.

**Table 4.A.3: Option A New Criminal Activity by Risk Score**

NCA Risk Score	New Arrest after Release		
	No	Yes	Total
<b>1</b>	5146 76.45	1585 23.55	6731 11.92
<b>2</b>	9354 72.89	3479 27.11	12833 22.73
<b>3</b>	8948 51.16	8541 48.84	17489 30.97
<b>4</b>	4455 30.27	10263 69.73	14718 26.07
<b>5</b>	691 16.14	3590 83.86	4281 7.58
<b>6</b>	78 18.89	335 81.11	413 0.73
<b>Total</b>	<b>28672</b> <b>50.78</b>	<b>27793</b> <b>49.22</b>	<b>56465</b> <b>100.00</b>

The table of NCA by Risk Score shows mostly good discrimination with a range of 58 percentage points from 24% to 81% or from 48% to 165% of the base rate, an average increase of 12 percentage points per increase in risk score. The discrimination is again strongest in the middle range, declining sharply at the extremes, with the rate of failure actually falling by 3 percentage points from 5 to 6 points on the scale, again scoring only 0.7% of the validation sample in the highest risk group. Rescaling the final risk score by changing the way the raw scores are collapsed to achieve a range of 1 to 6 may be able to increase the size and the failure rate at the top risk score for a more linear relation of NCA rate to risk.

**Table 4.A.4.a: Option A New Violent Criminal Activity by Risk Score**

NVCA Risk Score	New Violent Arrest after Release		
	No	Yes	Total
<b>1</b>	35165 94.86	1904 5.14	37069 65.65
<b>2</b>	6885 87.76	960 12.24	7845 13.89
<b>3</b>	8329 82.30	1791 17.70	10120 17.92
<b>4</b>	583 75.91	185 24.09	768 1.36
<b>5</b>	414 63.50	238 36.50	652 1.15
<b>6</b>	7 63.64	4 36.36	11 0.02
<b>Total</b>	<b>51383</b> <b>91.00</b>	<b>5082</b> <b>9.00</b>	<b>56465</b> <b>100.00</b>

With low base rates as in the case of NVCA, it is difficult to improve on predicting the most frequent outcome; in this case, with only 9% overall failing the accuracy of predicting no new violent offense would be 91%.

Initially the table of NVCA by Risk Score shows good discrimination with a range of 31 percentage points from 5% to 36% or from 57% to 404% of the base rate, with an average increase of 6 percentage points per increase in risk score. The discrimination is more even from a score of 1 through the middle range, declining very sharply at the high end; as with NCA there is an actual (if small) decline in the rate of NVCAs at the highest risk score. There is also a weakness in the distribution of scores in the validation sample; 66% have a score of 1; less than 3% score 4 or above and adding scores of 3 only increases the proportion of the sample to 20%. This is a shortcoming of using only the range of failure rates across risk scores to evaluate predictive ability unless one also looks at the distribution; the effects will show up in the AUC.

The NCA and NVCA outcomes by risk score in Option A raise two issues we will note also in the other options. We see very few cases at the top end of the risk scores, and small increases or even declines in the offense rate as we move from a risk score of 5 to 6. First, recalling that exclusions target violent offenders, are the low numbers with high scores due to the exclusion criteria applied to the PSA Eligible sample? Second, given that the PSA was developed using a wide variety of jurisdictions that may not

match Los Angeles County very well, is the PSA well-calibrated for Los Angeles?<sup>18</sup> We will take these issues up in the discussion section rather than with each observation.

**Table 4.A.4.b: Option A New Violent Criminal Activity by Violence Flag**

Violence Flag	New Violent Arrest after Release		
	No	Yes	Total
<b>No</b>	50379 91.54	4655 8.46	55034 97.47
<b>Yes</b>	1004 70.16	427 29.84	1431 2.53
<b>Total</b>	<b>51383</b> <b>91.00</b>	<b>5082</b> <b>9.00</b>	<b>56465</b> <b>100.00</b>

Table 4.A.4.b shows the comparative recidivism rates obtained when selecting only the top 3% for a prediction of failure. PSA guidelines call for collapsing the NVCA risk score by applying a cut-point to the scale between 3 and 4, dichotomizing the results into a Yes/No Violence Flag. NVCA's are more than three times as frequent in the Yes row versus the No row. However, the accuracy (now calculable given a cut-point for making yes/no predictions is 90% - down from the 91% accuracy one would obtain by predicting "No" for all cases.

### Option A: Odds Ratios and AUCs

**Table 4.A.5: Option A PSA Predictive Ability, Failure to Appear**

Sample	No FTA	One or More FTAs	Total
Observations Used	25,825	30,640	56,465
Statistic	Point Estimate	95% Confidence Interval	
		Lower Bound	Upper Bound
Odds Ratio	2.345	2.302	2.388
Area Under the Curve	0.733	0.729	0.737

<sup>18</sup> Calibration is concerned with whether the assessment score corresponds to the probability of the outcome. The AUC is a measure of discrimination – the ability to distinguish those with the outcome from those without. It is not a measure of calibration.



**Table 4.A.6: Option A PSA Predictive Ability, New Criminal Activity**

Sample	No NCA	One or More NCAs	Total
Observations Used	28,672	27,793	56,465
Statistic	Point Estimate	95% Confidence Interval	
		Lower Bound	Upper Bound
Odds Ratio	2.159	2.122	2.197
Area Under the Curve	0.721	0.717	0.725

**Table 4.A.7: Option A PSA Predictive Ability, New Violent Criminal Activity**

Sample	No NVCA	One or More NCVAs	Total
Observations Used	51,383	5,082	56,465
Statistic	Point Estimate	95% Confidence Interval	
		Lower Bound	Upper Bound
Odds Ratio	1.887	1.837	1.937
Area Under the Curve	0.670	0.663	0.678
Violence Flag			
Odds Ratio	4.604	4.095	5.176
Area Under the Curve	0.532	0.528	0.536

Tables 4.A.5, 4.A.6, and 4.A.7 show the odds ratios and AUCs for the three risk score predictions of outcome. With similar base rates the odds ratios for FTA and NCA are comparable (2.35, 2.16 respectively) and both show a substantial increase in the odds of failure versus success as risk scores increase. Given the low base rate of NVCA, the odds ratio (1.89) is also comparable. The AUCs take the number of individuals at each score into account. Here the FTA risk score has the strongest predictive ability with a strong AUC of 0.733. The NCA risk score is also a strong predictor with an AUC of 0.721. The NVCA risk score is only a moderate predictor with an AUC of 0.670.

As one would anticipate from the great difference of failure percentages for the Violence Flag, it has a high odds ratio. Given that only less than 3% score a 4 or above, this is perhaps not very useful. Looking back to Table 4.A.4.b, using the flag will only achieve 29% true positive predictions of a NVCA, at a cost

of 71% positive predictions where the outcome is negative.<sup>19</sup> The flag will also only capture about 8% of those who go on to a new violent arrest. This explains the very low AUC of 0.532. Lowering the cut-point to capture more of those new violent offense arrests will necessarily increase the rate of false positives. Given the low percentage flagged for a likely violent arrest, the cost of false positives may seem acceptable. Given the low rate of serious violent offenses, the cost of missing most of them may also be acceptable. Judicial discretion should be informed by the relation of NCVA outcomes to both risk scores and the Violent Flag.

## Option B: Core Data, Plus NULL and Missing Charge Level

### Option B: Parameters

Option B is intended as an evaluation of the predictive ability of the PSA when applied in less than ideal conditions where missing data is inevitable. Bookings with NULL Code Types and missing Level of Charge are included. Like Option A, it is limited to bookings where Sentence Date available, but disposition dates later than December of 2018 are allowed to include more cases for a more robust assessment. As a consequence, new criminal activity is picked up both from arrest charges in bookings records and filed charges from case history records.

**Table 4.B.1: Option B Outcomes by Release Type**

<b>Outcome by Release Type (%)</b>				
<b>(Bookings &amp; History with Sentence Date)</b>				
<b>Release Type</b>	<b>FTA</b>	<b>NCA</b>	<b>NVCA</b>	<b>N</b>
<b>Bail</b>	289	302	63	699
	41.34	43.20	9.01	0.62
<b>Bond</b>	9713	10820	2077	22093
	43.96	48.97	9.40	19.65
<b>Cite</b>	48399	43700	7985	67027
	72.21	65.20	11.91	59.61
<b>OR</b>	10862	9243	1878	22621
	48.02	40.86	8.30	20.12
<b>Total</b>	<b>69263</b>	<b>64065</b>	<b>12003</b>	<b>112440</b>
	<b>61.60</b>	<b>56.98</b>	<b>10.68</b>	<b>100.00</b>

<sup>19</sup> The rate at which positive predictions turn out to be wrong is called the False Discovery Rate. The denominator is positive predictions. This is distinct from the False Positive Rate, which is the number of false positives divided by the number of false positives plus true negatives, i.e., the denominator is negative outcomes. The False Positive Rate for the NCVA Flag is 2%.

The most noticeable change from Option A is the higher base rate of the outcome, up 7 to 8 percentage points for FTA and NCA, 2 percentage points for NVCA. About half of this increase comes from the addition of court case records to arrests from bookings; the remainder is due to the higher rates in the bookings added to the sample. The pattern in Outcomes by Release Type (Table 4.B.1) remains similar to the results for Option A and the PSA eligible population overall, with lower rates of risk for Bail and Own Recognizance releases and the highest rates among Citation only releases. Bond releases are in the upper middle for most outcomes, with the exception of FTA, where Own Recognizance releases are now higher than Bond. For FTA, the range is 31 percentage points from 41% to 72% or from 67% to 117% of the overall rate for the group. The pattern is similar for NCA, with a range of 24 percentage points from 41% to 65% or from 72% to 114% of the overall rate. With the lower base rate of offending, for NVCA the range is only 3 percentage points from 8% to 12%, which works out as from 78% to 112% of the group rate. Again, the Cite group is the main driver of the differences, for NCA the difference between Bond and OR releases is down to 20%.

### Option B: Outcome by Risk Score

**Table 4.B.2: Option B Failure to Appear by Risk Score**

FTA Risk Score	Failure to Appear after Release		
	No	Yes	Total
<b>1</b>	5313 68.18	2480 31.82	7793 6.93
<b>2</b>	18403 61.54	11503 38.46	29906 26.60
<b>3</b>	15324 39.08	23884 60.92	39208 34.87
<b>4</b>	2735 18.47	12074 81.53	14809 13.17
<b>5</b>	1370 6.89	18515 93.11	19885 17.68
<b>6</b>	32 3.81	807 96.19	839 0.75
<b>Total</b>	<b>43177</b> <b>38.40</b>	<b>69263</b> <b>61.60</b>	<b>112440</b> <b>100.00</b>

The table of FTA by Risk Score again shows good discrimination with a range of 64 percentage points from 32% to 96% or from 52% to 156% of the base rate, an average increase of 13 percentage points per increase in risk score. The discrimination is strongest in the middle range, declining sharply at the extremes, and scores only 0.8% of the validation sample in the highest risk group. The rate of FTAs and the risk scores have both increased. With higher risk scores, the distribution across risk scores has shifted up. In Option A, 43% had a score of 1 or 2; now it is down to 34%. The inclusion of bookings with unrecovered charges or missing Level of Charge has not led to lower scores. Rather, the rising scores indicates that bookings added to this validation sample have more extensive criminal history, and that the appearance of unrecovered charges and levels of charge may be a consequence.

**Table 4.B.3: Option B New Criminal Activity by Risk Score**

NCA Risk Score	New Arrest after Release		
	No	Yes	Total
<b>1</b>	6237 74.24	2164 25.76	8401 7.47
<b>2</b>	13913 69.65	6062 30.35	19975 17.77
<b>3</b>	16230 47.92	17636 52.08	33866 30.12
<b>4</b>	10060 27.84	26069 72.16	36129 32.13
<b>5</b>	1718 13.40	11103 86.60	12821 11.40
<b>6</b>	217 17.39	1031 82.61	1248 1.11
<b>Total</b>	<b>48375</b> <b>43.02</b>	<b>64065</b> <b>56.98</b>	<b>112440</b> <b>100.00</b>

The table of NCA by Risk Score is similar to Option A, again with increased outcome rates and a shift up in the distribution of booking cases across risk scores. The proportion with a score of 1 or 2 has declined from 35% to 25%. Again, the risk score shows mostly good discrimination with a range of 57 percentage points from 26% to 83% or from 45% to 145% of the base rate, an average increase of 11 percentage points per increase in risk score. The discrimination is again strongest in the middle range, declining sharply at the extremes, with the rate of failure now falling by 4 percentage points as the risk score

increases from 5 to 6, with only 1.1% of the validation sample scoring in the highest risk group. This again raises questions about exclusions and calibration.

**Table 4.B.4.a: Option B New Violent Criminal Activity by Risk Score**

NVCA Risk Score	New Violent Arrest after Release		
	No	Yes	Total
<b>1</b>	61815 93.75	4121 6.25	65936 58.64
<b>2</b>	16287 87.32	2365 12.68	18652 16.59
<b>3</b>	20004 81.49	4543 18.51	24547 21.83
<b>4</b>	1384 76.85	417 23.15	1801 1.60
<b>5</b>	931 62.99	547 37.01	1478 1.31
<b>6</b>	16 61.54	10 38.46	26 0.02
<b>Total</b>	<b>100437</b> <b>89.32</b>	<b>12003</b> <b>10.68</b>	<b>112440</b> <b>100.00</b>

**Table 4.B.4.b: Option B New Violent Criminal Activity by Violence Flag**

Violence Flag	New Violent Arrest after Release		
	No	Yes	Total
<b>No</b>	98106 89.89	11029 10.11	109135 97.06
<b>Yes</b>	2331 70.53	974 29.47	3305 2.94
<b>Total</b>	<b>100437</b> <b>89.32</b>	<b>12003</b> <b>10.68</b>	<b>112440</b> <b>100.00</b>

The table of NVCA by Risk Score again shows higher outcome rates and a high proportion of the sample in the low scores range, now with less of a shift to higher scores than for FTAs and NCAs. The proportion with a score of 1 or 2 has declined from 80% to 75%. The increase in outcome rates over Option A is only a couple of percentage points, but proportional to the increases found for FTA and NCA (ranging from 14% to 18% higher). The range of NVCA across scores is now down to 32 percentage points difference from a low of 6% to a high of 38%, or from 59% to 360% of the base rate, with an average increase of 6 percentage points per increase in risk score. The pattern is similar to Option A. Increases in recidivism are more regular between scores of 1 through 4, but jump to 14 percentage points from 4 to 5 and decline to only 1 percentage point between scores of 5 and 6. The modest decrease in the proportion with low scores does not repair the poor distribution among high scores; still less than 3% score 4 or above and adding scores of 3 only increases the proportion of the sample to 25%. The performance of the Violence Flag is essentially the same as in Option A.

### Option B: Odds Ratios and AUCs

**Table 4.B.5: Option B PSA Predictive Ability, Failure to Appear**

Sample	No FTA	One or More FTAs	Total
Observations Used	43,177	69,263	112,440
Statistic	Point Estimate	95% Confidence Interval	
		Lower Bound	Upper Bound
Odds Ratio	2.464	2.431	2.498
Area Under the Curve	0.746	0.744	0.749

**Table 4.B.6: Option B PSA Predictive Ability, New Criminal Activity**

Sample	No NCA	One or More NCAs	Total
Observations Used	48,375	64,065	112,440
Statistic	Point Estimate	95% Confidence Interval	
		Lower Bound	Upper Bound
Odds Ratio	2.20	2.176	2.232
Area Under the Curve	0.723	0.720	0.726

**Table 4.B.7: Option B PSA Predictive Ability, New Violent Criminal Activity**

Sample	No NVCA	One or More NVCA's	Total
Observations Used	100,437	12,003	112,440
Statistic	Point Estimate	95% Confidence Interval	
		Lower Bound	Upper Bound
Odds Ratio	1.763	1.732	1.795
Area Under the Curve	0.655	0.650	0.660
Violence Flag			
Odds Ratio	3.718	3.442	4.017
Area Under the Curve	0.529	0.527	0.532

Tables 4.B.5, 4.B.6, and 4.B.7 show the odds ratios and AUCs for the three risk score predictions of outcome. With similar base rates the odds ratios for FTA and NCA are comparable (2.46, 2.20 respectively) and both show a substantial increase in the odds of failure versus success as risk scores increase. With the low base rate of NVCA, the odds ratio (1.76) is also comparable. The FTA risk score again has the strongest predictive ability with an AUC of 0.746. The NCA risk score is also a strong predictor with an AUC of 0.723. The NVCA risk score is only a moderate predictor with an AUC of 0.655. The AUCs split around the AUCs from Option A; FTA is higher, NCA the same, and NVCA lower. This strengthens the argument that the inclusion of cases with missing information on convictions has not harmed PSA performance.

### Option C: Imputation of Sentence Date

#### Option C: Parameters

Option C allows the use of largest sample for a more precise estimate and will show if there is evidence of bias affecting sentence date availability or time to disposition and support evaluation of robustness of PSA in the face of missing and imputed data. Imputation of a Sentence Date increases the number of useable bookings by 63%. Like Option B, bookings with NULL Code Types and missing Level of Charge are included. Disposition dates later than December of 2018 are allowed and new criminal activity is picked up both from arrest charges in bookings records and filed charges from case history records.

## Option C: Outcome by Release Type

**Table 4.C.1: Option C Outcomes by Release Type**

<b>Outcome by Release Type (%)</b> (Bookings & History with Imputed Date)				
<b>Release Type</b>	<b>FTA</b>	<b>NCA</b>	<b>NVCA</b>	<b>N</b>
<b>Bail</b>	472	477	101	1257
	37.55	37.95	8.04	0.68
<b>Bond</b>	15459	17784	3657	39123
	39.51	45.46	9.35	21.29
<b>Cite</b>	71085	62956	11868	102650
	69.25	61.33	11.56	55.86
<b>OR</b>	19914	16579	3535	40719
	48.91	40.72	8.68	22.16
<b>Total</b>	<b>106930</b>	<b>97796</b>	<b>19161</b>	<b>183749</b>
	<b>58.19</b>	<b>53.22</b>	<b>10.43</b>	<b>100.00</b>

Option C has base rates roughly halfway in between Options A and B, up 4 percentage points from Option A for FTA and NCA, 1 percentage point for NVCA. Most of this increase can be attributed to the addition of court case records to arrests from bookings. The proportion of Cite releases in the sample is about the same as in Option A, partially accounting for the decline in outcome rate from Option B. The pattern in Outcomes by Release Type (Table 4.C.1) remains similar to previous results, with lower rates of risk for Bail and Own Recognizance releases, again with the exception of FTA, where Own Recognizance releases are remain higher than Bond. For FTA, the range is 32 percentage points from 38% to 69% or from 65% to 119% of the overall rate for the group. For NCA, the range is 23 percentage points from 38% to 61% or from 71% to 115% of the overall rate. For NVCA the range is 4 percentage points from 8% to 12%, which works out as from 77% to 111% of the group rate.



## Option C: Outcome by Risk Score

**Table 4.C.2: Option C Failure to Appear by Risk Score**

FTA Risk Score	Failure to Appear after Release		
	No	Yes	Total
<b>1</b>	7125 65.92	3684 34.08	10809 5.88
<b>2</b>	29648 63.92	16737 36.08	46385 25.24
<b>3</b>	29127 45.52	34854 54.48	63981 34.82
<b>4</b>	6853 25.45	20073 74.55	26926 14.65
<b>5</b>	3968 11.61	30224 88.39	34192 18.61
<b>6</b>	98 6.73	1358 93.27	1456 0.79
<b>Total</b>	76819 41.81	106930 58.19	183749 100.00

The table of FTA by Risk Score again shows good discrimination with a slight decline to a range of 59 percentage points from 34% to 93% or from 59% to 160% of the base rate, and an average increase of 12 percentage points per increase in risk score. The discrimination remains strongest in the middle range, declining most sharply at the lower end, and scoring only 0.8% of the validation sample in the highest risk group. The distribution across risk scores has not shifted up much from Option B. In Option B, 34% had a score of 1 or 2; now it is down to 31%. The addition of 63 percent more bookings which could not be linked to a court case has not led to lower scores. It has changed the outcome base rate to one halfway in between Options A and B despite the highest risk scores, suggesting possibly weaker performance.

The use of the median for imputation of the follow-up window means that across the sample, if the cases added are similar to those with full information, the number of post-release FTAs undercounts will be balanced by the over-counts. Therefore, the new bookings added by way of imputation have lower FTA rates than the bookings added in Option B, or the median follow-up time from the cases with history linkage are too short for the imputed cases.

**Table 4.C.3: Option C New Criminal Activity by Risk Score**

NCA Risk Score	New Arrest after Release		
	No	Yes	Total
<b>1</b>	8922 73.98	3138 26.02	12060 6.56
<b>2</b>	23051 71.01	9410 28.99	32461 17.67
<b>3</b>	29506 53.01	26155 46.99	55661 30.29
<b>4</b>	20085 33.44	39979 66.56	60064 32.69
<b>5</b>	3916 18.29	17491 81.71	21407 11.65
<b>6</b>	473 22.57	1623 77.43	2096 1.14
<b>Total</b>	<b>85953</b> <b>46.78</b>	<b>97796</b> <b>53.22</b>	<b>183749</b> <b>100.00</b>

Comparing the table of NCA by Risk Score to previous options shows a similar pattern with a slight weakening of discrimination. The outcome rate is again halfway between Options A and B, with similar implications. The distribution across scores is very close to Option B. The rate of NCAS doesn't increase quite as fast as before. Risk score discrimination is still good but the range, which was 57-58 percentage points is down to 52 (from 26% to 77%), stretching from 49% to 145% of the base rate. The average increase per risk score step is 10 percentage points. As usual, the discrimination is strongest in the middle range, declining sharply at the extremes, with the rate of failure again falling by 4 percentage points from 5 to 6 points on the scale and 1.1% of the validation sample scoring in the highest risk group. This continues to raise questions about the calibration of the final risk score.

**Table 4.C.4: Option C New Violent Criminal Activity by Risk Score**

NVCA Risk Score	New Violent Arrest after Release		
	No	Yes	Total
<b>1</b>	99301 93.80	6566 6.20	105867 57.62
<b>2</b>	26871 87.76	3747 12.24	30618 16.66
<b>3</b>	33485 82.63	7041 17.37	40526 22.06
<b>4</b>	2757 78.93	736 21.07	3493 1.90
<b>5</b>	2118 66.90	1048 33.10	3166 1.72
<b>6</b>	56 70.89	23 29.11	79 0.04
<b>Total</b>	<b>164588</b> <b>89.57</b>	<b>19161</b> <b>10.43</b>	<b>183749</b> <b>100.00</b>

**Table 4.C.4.b: Option C New Violent Criminal Activity by Violence Flag**

Violence Flag	New Violent Arrest after Release		
	No	Yes	Total
<b>No</b>	159657 90.20	17354 9.80	177011 96.33
<b>Yes</b>	4931 73.18	1807 26.82	6738 3.67
<b>Total</b>	<b>164588</b> <b>89.57</b>	<b>19161</b> <b>10.43</b>	<b>183749</b> <b>100.00</b>

The table of NVCA by Risk Score shows an outcome rate between Options A and B. It shows the same high proportion of the sample scored in the low range, again closely matching Option B although not as high a proportion as in Option A. The range of NVCA across scores is now down to 23 percentage points different from a low of 6% to a high of 29%, or from 59% to 279% of the base rate, with an average

increase in the rate of NCVAs of 5 percentage points per increase in risk score. The pattern is similar to Options A and B with relatively even discrimination of failure rates from a score of 1 through the middle range, and it again swings negative at the high end as between scores of 5 and 6 the rate of NVCA declines 4 percentage points. Less than 4% score 4 or above and adding scores of 3 only increases the proportion of the sample to 26%. Taking score distributions across all options and considering the declines in the NVCA outcome rate from a score of 5 to 6, rescaling the final score might be explored for NVCA as well as NCA. The Violence flag performance is substantially the same as before.

### Option C: Odds Ratios and AUCs

**Table 4.C.5: Option C PSA Predictive Ability, Failure to Appear**

Sample	No FTA	One or More FTAs	Total
Observations Used	76,819	106,930	183,749
Statistic	Point Estimate	95% Confidence Interval	
		Lower Bound	Upper Bound
Odds Ratio	2.161	2.140	2.182
Area Under the Curve	0.721	0.719	0.723

**Table 4.C.6: Option C PSA Predictive Ability, New Criminal Activity**

Sample	No NCA	One or More NCAs	Total
Observations Used	85,953	97,796	183,749
Statistic	Point Estimate	95% Confidence Interval	
		Lower Bound	Upper Bound
Odds Ratio	2.024	2.004	2.043
Area Under the Curve	0.701	0.698	0.703

**Table 4.C.7: Option C PSA Predictive Ability, New Violent Criminal Activity**

Sample	No NVCA	One or More NVCA's	Total
Observations Used	164,588	19,161	183,749
Statistic	Point Estimate	95% Confidence Interval	
		Lower Bound	Upper Bound
Odds Ratio	1.687	1.664	1.709
Area Under the Curve	0.650	0.646	0.653
Violence Flag			
Odds Ratio	3.373	3.189	3.568
Area Under the Curve	0.532	0.530	0.534

Tables 4.C.5, 4.C.6, and 4.C.7 show the odds ratios and AUCs for the three risk score predictions of outcome. With comparable base rates the odds ratios for FTA and NCA again are similar but slightly lower (2.16, 2.02 respectively). The odds ratio for NVCA (1.69) is lower than Option A. The FTA risk score again has the best predictive ability with a lower but still strong AUC of 0.721, followed by the NCA risk score with an AUC of 0.701. The NVCA risk score AUC is now 0.650. The odds ratio for Violence Flag has again declined, now to 3.37, with an AUC of 0.532.

Option C uses all the data available by accepting records with missing values and estimating an average follow-up window in the absence of actual dates. AUCs across all outcomes are lower than with Options A or B. Mean risk scores are as high as B and slightly higher than A, so missing information on risk factors is not likely the cause. However, even if imputation of a median follow-up window balances out undercounts and over counts across the sample, it is bound to lead to errors in individual cases. Some predicted to fail, for example, may have already been sentenced to incarceration and off the street before their imputed window ends. It is not surprising that there would be a weaker association between risk scores and outcomes and a lower AUC. However, performance is still comparable to results when used with higher quality data.

## **Option D: Recovery of Charges**

### **Option D: Parameters**

Like Option C, Option D uses the largest sample through imputation of missing sentence dates to expand the number of useable cases. Disposition dates later than December of 2018 are allowed and new criminal activity is picked up both from arrest charges in bookings records and filed charges from case history records. The goal is a more complete ascertainment of risk factors and outcomes. Here, bookings with NULL Code Types and missing Level of Charge are included, but an effort is made to use available

information in CHARGE\_CODE to impute CHARGE\_CODE\_TYPE and CHARGE\_LEVEL\_CODE. CHARGE\_CODE entries are processed to regularize poorly formatted charges so that they can be matched to standard tables classifying the offense and flagging violent offenses.<sup>20</sup>

Recovery efforts are able to significantly reduce the prevalence of unrecovered charges; the achievement for level of charge is minor. Of 190,129 individuals with history records, 32.8% have a “NULL” Code Type in one or more History records; 30.9% have one or more records missing Level of Charge. After recovery, only 20.4% have a “NULL” Code Type in one or more history records, but for CHARGE\_LEVEL\_CODE there are still 30.3% with one or more records missing; adequate information is lacking or not yet incorporated into our recovery algorithm. Of the 476,480 individuals with bookings records, 16.9% have a “NULL” in one or more records; 59 have one or more records missing Level of Charge. For bookings after recovery, only 5.6% have a “NULL” Code Type in one or more records; 46 still have one or more records missing CHARGE\_LEVEL\_CODE. The ability to recover charges is limited by the presence many charges based on the local codes and regulations of cities and counties; it is time-prohibitive to attempt to map all these to a useful list of offenses.

**Table 4.D.1: Option D: Outcomes by Release Type**

<b>Outcome by Release Type (%)</b>				
<b>(Bookings &amp; History with Imputed Date)</b>				
<b>Release Type</b>	<b>FTA</b>	<b>NCA</b>	<b>NVCA</b>	<b>N</b>
<b>Bail</b>	472	476	101	1255
	37.61	37.93	8.05	0.68
<b>Bond</b>	15437	17725	3671	39075
	39.51	45.36	9.39	21.28
<b>Cite</b>	71062	62884	11974	102623
	69.25	61.28	11.67	55.88
<b>OR</b>	19908	16552	3556	40710
	48.90	40.66	8.73	22.17
<b>Total</b>	<b>106879</b>	<b>97637</b>	<b>19302</b>	<b>183663</b>
	<b>58.19</b>	<b>53.16</b>	<b>10.51</b>	<b>100.00</b>

Option D base rates are less than two-tenths of a percent different from those of Option C; the additional exclusions from recovered charges don’t change the outcomes in any significant sense. The differences in FTAs post release are due solely to the change in the sample due to additional exclusions. Comparing NCA and NVCA across options D and C show that we are now measuring a few additional

<sup>20</sup> The data processing steps are outlined in the appendix.

violent arrests and convictions in the follow-up window (NVCA +141) but due to exclusions have fewer overall (NCA: -159).

**Table 4.D.2: Option D Failure to Appear by Risk Score**

FTA Risk Score	Failure to Appear after Release		
	No	Yes	Total
<b>1</b>	7104 65.88	3679 34.12	10783 5.87
<b>2</b>	29652 63.93	16732 36.07	46384 25.25
<b>3</b>	29115 45.52	34840 54.48	63955 34.82
<b>4</b>	6850 25.45	20066 74.55	26916 14.66
<b>5</b>	3965 11.60	30206 88.40	34171 18.61
<b>6</b>	98 6.74	1356 93.26	1454 0.79
<b>Total</b>	<b>76784</b> <b>41.81</b>	<b>106879</b> <b>58.19</b>	<b>183663</b> <b>100.00</b>

The tables for Outcome by Risk Score again show a close match to Option C with differences in the distribution of risk scores and outcome rates at each risk score within a few hundredths of a percentage point. Only in NVCA is there any sign of possibly systematic change to very slightly higher risk scores and outcomes.

**Table 4.D.3: Option D New Criminal Activity by Risk Score**

NCA Risk Score	New Arrest after Release		
	No	Yes	Total
<b>1</b>	8912 74.03	3126 25.97	12038 6.55
<b>2</b>	23038 71.06	9381 28.94	32419 17.65
<b>3</b>	29478 53.08	26052 46.92	55530 30.23
<b>4</b>	20176 33.57	39923 66.43	60099 32.72
<b>5</b>	3947 18.38	17528 81.62	21475 11.69
<b>6</b>	475 22.60	1627 77.40	2102 1.14
<b>Total</b>	<b>86026</b> <b>46.84</b>	<b>97637</b> <b>53.16</b>	<b>183663</b> <b>100.00</b>



Table 4.D.4: Option D New Violent Criminal Activity by Risk Score

NVCA Risk Score	New Violent Arrest after Release		
	No	Yes	Total
<b>1</b>	98924 93.79	6548 6.21	105472 57.43
<b>2</b>	26755 87.64	3774 12.36	30529 16.62
<b>3</b>	33652 82.55	7115 17.45	40767 22.20
<b>4</b>	2791 78.69	756 21.31	3547 1.93
<b>5</b>	2183 66.78	1086 33.22	3269 1.78
<b>6</b>	56 70.89	23 29.11	79 0.04
<b>Total</b>	<b>164361</b> <b>89.49</b>	<b>19302</b> <b>10.51</b>	<b>183663</b> <b>100.00</b>

Table 4.D.4.b: Option D New Violent Criminal Activity by Violence Flag

Violence Flag	New Violent Arrest after Release		
	No	Yes	Total
<b>No</b>	159331 90.14	17437 9.86	176768 96.25
<b>Yes</b>	5030 72.95	1865 27.05	6895 3.75
<b>Total</b>	<b>164361</b> <b>89.49</b>	<b>19302</b> <b>10.51</b>	<b>183663</b> <b>100.00</b>

**Table 4.D.5: Option D PSA Predictive Ability, Failure to Appear**

Sample	No FTA	One or More FTAs	Total
Observations Used	76,784	106,879	183,663
Statistic	Point Estimate	95% Confidence Interval	
		Lower Bound	Upper Bound
Odds Ratio	2.161	2.140	2.182
Area Under the Curve	0.721	0.719	0.723

**Table 4.D.6: Option D PSA Predictive Ability, New Criminal Activity**

Sample	No NCA	One or More NCAs	Total
Observations Used	86,026	97,637	183,663
Statistic	Point Estimate	95% Confidence Interval	
		Lower Bound	Upper Bound
Odds Ratio	2.022	2.002	2.041
Area Under the Curve	0.700	0.698	0.703

**Table 4.D.7: Option D PSA Predictive Ability, New Violent Criminal Activity**

Sample	No NVCA	One or More NVCA's	Total
Observations Used	163,361	19,302	183,663
Statistic	Point Estimate	95% Confidence Interval	
		Lower Bound	Upper Bound
Odds Ratio	1.688	1.666	1.711
Area Under the Curve	0.651	0.647	0.655
Violence Flag			
Odds Ratio	3.391	3.208	3.584
Area Under the Curve	0.533	0.531	0.535

Tables 4.D.5, 4.D.6, and 4.D.7 show the odds ratios and AUCs for the three risk score predictions of outcome. As anticipated in comparing outcomes by risk scores, any differences are minimal changes in the fourth decimal place. In the aggregate, recovery of offense charges and imputation of missing level of charge do not change the predictive ability of the PSA.

## 5. Discussion of Options

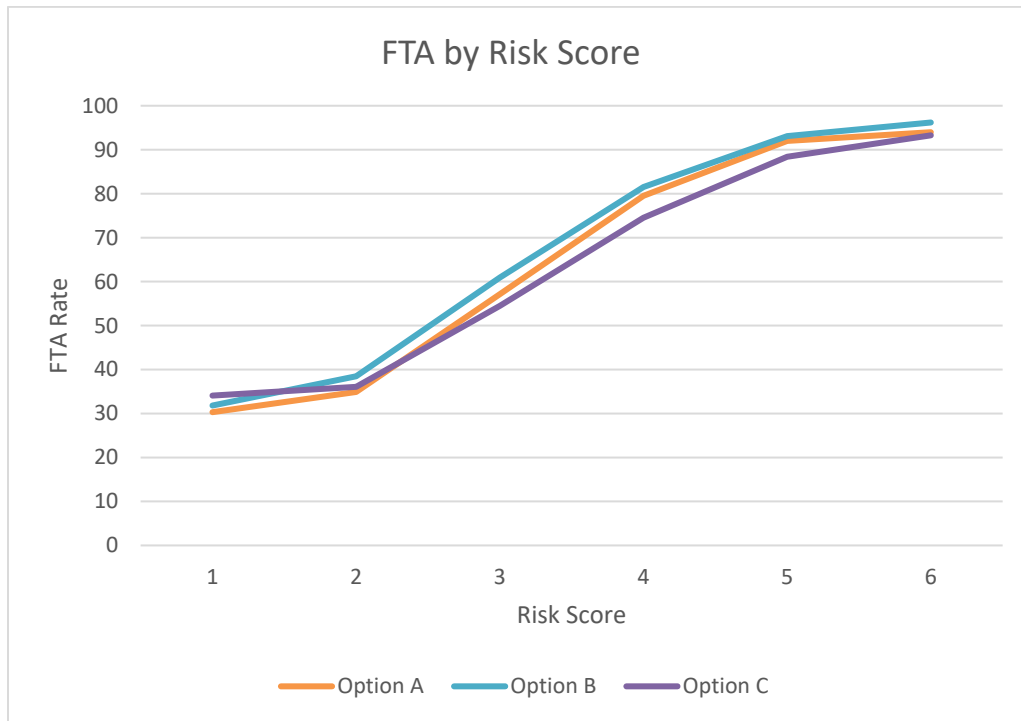
It was anticipated that due to the way the PSA counts only one occurrence of an event in calculating risk scores and the extensive history of many booked individuals that Option D would make limited changes to most risk scores. The results bear this out. As Option D results closely track those from Option C, it is most informative to limit the main discussion to Options A, B, and C. We present the results of charges recovery in a supplemental analysis of the effect of imputation on the scores and predictions of those whose prior records have changed in Appendix E.

Summarizing results across options, we first find that the inclusion of bookings with missing data on charges does not have a deleterious effect on scoring the PSA. If so, we would expect scores to fall; instead scores rise. It seems that missing data is offset by more extensive criminal records and that perhaps the longer record exposes them to more chances of getting incomplete data or unresolved charges in a record. Table 4.5 provides evidence. Compared to Option A, Options B and C are 25% more likely to have multiple bookings in the recruitment window. Bookings missing information on convictions are most likely only missing one or two charges or levels of charge, so it takes few additional records to replace the missing charges. Option B has 8% more prior misdemeanors than A and 28% more felonies, with increases for other priors falling between these two. The bookings in Option A account for one-half of the total in Option B; if we were to look solely at the additional bookings added in B, we would find even higher increases in priors.

Outcome failure rates also rise in Options B and C in comparison to A. Controlled comparison on Option A cases with and without convictions added to arrests in the follow-up window show these add about 2 percentage points to NCA rates; with Option B it is about 4 percentage points. The increase in NCA rates in Option B is twice this, 8 percentage points, which suggests that NCAs are genuinely about 4 percentage points higher in the Option B sample and is also consistent with higher risk scores. This relationship of average risk scores to outcomes moderates in Option C as risk scores are as high as in Option B but outcome rates fall in between Options A and B.

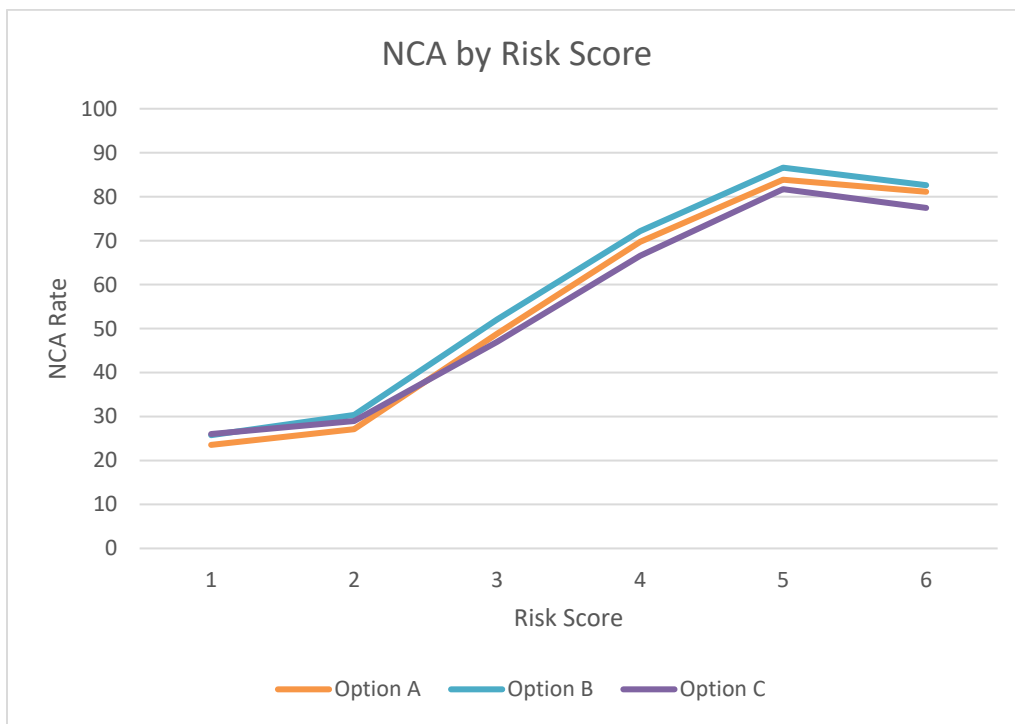
Figures 5.1, 5.2, and 5.3 plot outcome rates by risk score for Options A, B, and C on line graph to facilitate assessment of whether PSA performance differs across options. Option D tracks Option C so closely that it overlays the C line, so it is left out for clarity. A difference in the slope or the height of the lines indicates a difference in the relationship of outcome to risk score.

Figure 5.1: Failure to Appear by Risk Score



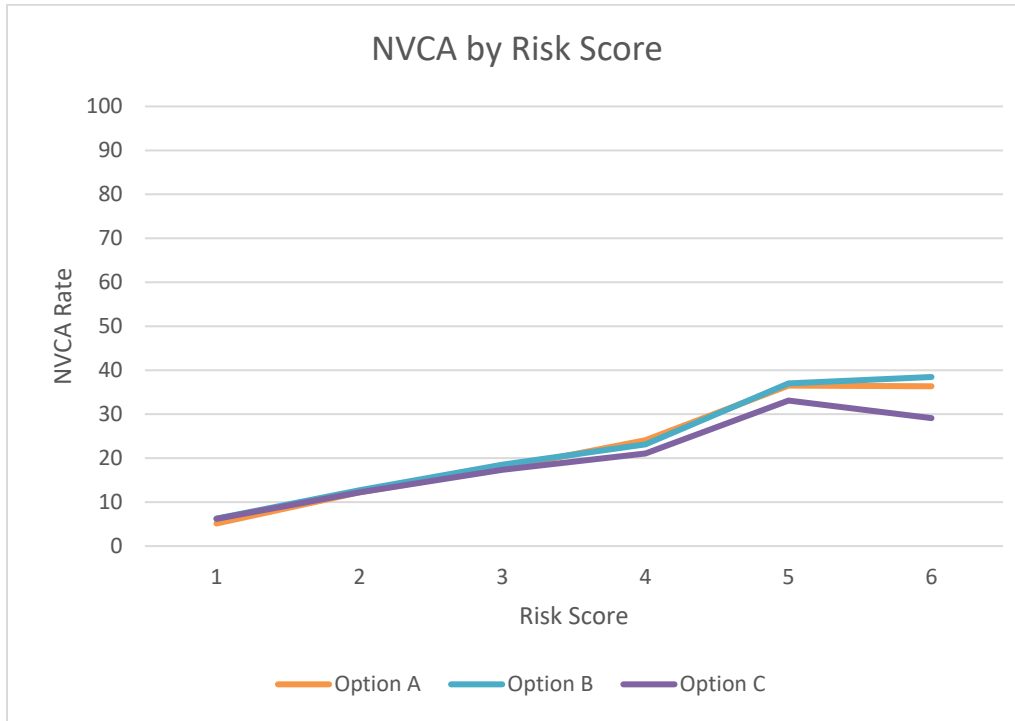
Qualitatively we see little difference in slopes or height between Options A and B; PSA performance is unaffected. Option C actually has a slightly more linear relationship of FTA to risk score, along with very slightly milder slope and lower height. These are not big enough to affect our evaluation of the PSA's predictive ability.

Figure 5.2: New Criminal Activity by Risk Score



For NCA, we see small differences in the height of the lines of a few percentage points and a very slightly lower slope for Option C, but we still judge the relationship of risk score to outcome to be essentially the same across options. We also see the decline in NCA rates from a score of 5 to 6 noted in the tables above. This suggests that the LA county courts population is sufficiently different from the development population that the risk factor scoring is not well calibrated for LA. The AUC is a measure of discrimination, not calibration, and doesn't measure the linearity of a relationship; given that only about 1% of cases are affected this imposes only a small penalty on the AUC statistic.

Figure 5.3: New Violent Criminal Activity by Risk Score



Evaluating the graphs of NVCA by risk score, the options again show substantially the same relationship, with the divergence from linearity at the high extreme as noted above. While the effect is pronounced for Option C and notable in Option B, the signs of it are already noticeable in Option A as the curve almost flatlines from a score of 5 to 6. This suggests that like NCA, NVCA risk factor scoring is not well calibrated for LA. The slope of the line for Option C is lower, showing that the PSA is less able to differentiate NVCA for this group than the others. Review of the risk factors and outcome measures for Option C does not suggest an explanation for this difference. It may be that imputation holds up fairly well for the more frequent FTA and NCA events but is more sensitive to random variation with rarer events in the small groups scoring 5 or 6 on the NVCA risk score.

### Exclusions

Tables of outcomes by risk scores consistently show a very low proportion of the sample with a score of 6, usually near to or less than one percent. In contrast, DeMichele *et al.* in their Kentucky validation report 7% to 8%.<sup>21</sup> For NVCA, they only report numbers after dichotomizing score into the Violent Flag Yes/No classifications; with scores of 5 and 6 flagged,<sup>22</sup> the proportion with the flag is 18%; in our samples this is less than 2%. Is the exclusion of offenders booked with violent charges under PC 1270.1

<sup>21</sup> Ibid.

<sup>22</sup> Current guidance calls for flagging scores of 4 and above; instructions may have changed. Even with the broader coverage, in Los Angeles the flag only captures about 4% of the sample.

and PC 1319.5 eliminating a disproportionate number of those with more extensive records, leading to the lower proportion with higher scores?

To answer this question, we re-ran our tables while retaining the usually excluded bookings (Tables 5.1, 5.2). Table 5.1 shows the number in our standard samples with violent cases excluded (N), the number with exclusions restored (N+E), and the percentage of cases thus added. This runs 6% to 7% of the standard sample; these cases would have to be extremely concentrated among the high risk scores to come close to the Kentucky scores distributions.

**Table 5.1: Sample Sizes Before and After Restoring Excluded Cases**

Excluded Sample Size			
Sample	N	N+E	Δ%
A	56465	60332	7%
B	112440	119522	6%
C	183749	197225	7%

Table 5.2 shows the changes in the key statistics for judging whether the exclusions are responsible for the risk by score distributions. We look first at the percent of the risk scores in the top half of the range (ideally the midpoint of the scores would split the distribution in half, 50-50). For the NCA assessments across Options A to C, the percentage in the top half actually falls by about one percent. Looking at the baseline percentage failing with a new arrest or conviction, restoring the exclusions leads this rate to decline. Finally, we look at whether exclusion restoration leads to a rise in recidivism rates from a risk score of 5 to 6. (Signs of the percentage change are set so a positive change shows a shift to higher outcome rates, including a shift from a negative change before restoration to a less negative change after.) For Option A, the NCA rate actually moves in the wrong direction. For the other two options, although restoring exclusions ameliorates the problem, the NCA rate still falls as the risk score increases. Exclusions are not the source of the undesirable properties for the NCA rates by risk score distributions.

**Table 5.2: Risk Scales Before and After Restoring Excluded Cases**

Option	Restore Exclusion								
	% High Scores			% Fail			Rate Change		
	Pre	Post	Δ%	Pre	Post	Δ%	Pre	Post	Δ%
<b>NCA</b>									
A	34.4	34.1	-1%	49.2	48.9	-1%	-2.8	-3.0	-9%
B	44.6	44.2	-1%	57.0	56.5	-1%	-4.0	-3.6	11%
C	45.5	45.0	-1%	53.2	52.5	-1%	-4.3	-4.2	1%
<b>NVCA</b>									
A	2.5	3.8	51%	9.0	9.8	9%	-0.1	8.3	6043%
B	2.9	4.4	48%	10.7	11.4	7%	1.5	7.9	446%
C	3.7	5.5	49%	10.4	11.1	6%	-4.0	-3.6	9%

For the NVCA assessments across Options A to C, the percentage of the scores in the top half of the range improves by about 50%, but starting from a very low proportion this is still far from the Kentucky report. Looking at the baseline percentage failing with a new arrest or conviction, restoring the exclusions leads this rate to rise by 6 to 9 percent. The rate of NVCA in the Option A exclusions, for example, is 22% - clearly the exclusions change the risk by scores distribution. Finally, we look at whether exclusion restoration leads to a rise in recidivism rates from a risk score of 5 to 6. For Options A and B, the NCA rate now shows a substantial rise to about 8 percentage points. For Option C, although restoring exclusions ameliorates the problem, the NCA rate still falls as the risk score increases. We conclude that exclusions contribute the undesirable properties for the NVCA rates by risk score distributions, but are not the sole source.

## 6. Conclusions

In the introduction we noted a couple of factors impinging on the performance and validation of a risk assessment instrument in the context of Los Angeles County. Generally, the population in our sample is very different than the population of released defendants in the jurisdictions on which previous validations were based, in two ways:

- Severe jail overcrowding means that released individuals in our validation sample are less likely than elsewhere to be accused of low-level crimes. For instance, turnstile-jumpers are rarely arrested in Los Angeles County, compared with other jurisdictions in which zero-tolerance policing results in custodial detention for many such violators.
- The use of the PSA pre-arraignment, and the California statutory exclusion of violent offenses from pre-arraignment release, means that the sample does not include the full spectrum of criminally involved individuals.

As a result, the very serious and the very low-level offenders are both missing from this sample relative to other jurisdictions, making prediction more difficult. In this context, the predictive ability of the PSA as validated in this study compares very favorably with that found in other jurisdictions (Table 6.1). We highlight Option B because it is likely the best match to the operational context, with scoring based on records which include missing charges and a definite follow-up window from release to final sentencing rather than an imputed median. Its weakness is the addition of convictions within the follow-up window to compensate for the limitation of bookings coverage, which a comparison of NCA and NCVA rates with and without convictions suggests adds 4 percentage points, overestimating the recidivism rate.



**Table 6.1: Comparative PSA Performance, Receiver Operating Characteristic Area Under the Curve**

Jurisdiction	FTA	NCA	NVCA
Kentucky	0.646	0.650	0.664
Los Angeles (Option A)	0.733	0.721	0.670
Los Angeles (Option B)	0.746	0.722	0.654
Los Angeles (Option C)	0.721	0.701	0.650
Los Angeles (Option D)	0.721	0.700	0.651

Comparing AUCs across options points to another important feature of the PSA: Its performance is generally robust, holding up despite variations in data quality and completeness. Its weakest performance is in the assessment of risk for new violent criminal activity, which shows the strongest performance in Kentucky.<sup>23</sup> For FTA and NCA, the PSA achieves strong performance in this evaluation. We find that the PSA is a valid risk assessment tool for use with the Los Angeles Courts population, exhibiting extremely robust and moderate to strong predictive performance. Local decisions as to cut-point values for high risk may need to be considered (combining highest risk groups) as well as which data recovery method (E.g., Option A, B, C or D) best suits the local county effort.

---

<sup>23</sup> In the Kentucky validation study, 18% of the sample was classified as at risk by the NVCA flag based on a score of 5 or 6 compared to less than 2% with comparable scores in our samples. The Kentucky NVCA baseline rate was only 1.4% compared to 10-11% in this study; the NVCA rate in their flagged group was 3% compared to our rate of 28-30% here. The differences in these figures suggest a deep qualitative difference in the two populations.

## Appendix A

### LJAF Documentation of Risk Factor and Outcome Definitions and Measurement

*Table A.1*

Risk Factor	Weights (scores)
<b><u>Predicting Failure to Appear (0-7 points)</u></b>	
The person has a pending charge at time of the instant case	0: No 1: Yes
The person has any type of conviction prior to arrest on the instant case	0: No 1: Yes
The person failed to appear pretrial in the past 2 years	0: None 2: Once 4: Twice or more
The person failed to appear pretrial more than 2 years prior to the instant case	0: None 1: Yes
<b><u>Predicting a New Criminal Offense (0-13 points)</u></b>	
Age at arrest on the instant case	0: 23 years or older 2: 22 years or younger
The person has a pending charge at time of the instant case	0: No 3: Yes
The person has a misdemeanor conviction prior to arrest on the instant case	0: No 1: Yes
The person has a felony conviction prior to arrest on the instant case	0: No 1: Yes
The person has a conviction on a violent offense prior to arrest on the instant case	0: No 1: One or two 2: Three or more
The person failed to appear pretrial in the past 2 years	0: None 1: Once 2: Twice or more

The person was previously sentenced to incarceration 0: No  
2: Yes

---

**Predicting a New Violent Criminal Offense (0-7 points)**

Current charged offense is for violent offense 0: No  
2: Yes

Current charged offense is for violent offense AND person is 20 years old or younger 0: No  
1: Yes

The person has a pending charge at time of the instant case 0: No  
1: Yes

The person has any type of conviction prior to arrest on the instant case 0: No  
1: Yes

The person has a conviction on a violent offense prior to arrest on the instant case 0: None  
1: Once or twice  
2: Three or more

From: "Data Requirements for Validation of PSA Risk Assessment Tool". Bryan Borys, Director, Research and Data Management, PREP Pilot Project Manager

*Table A.2: Rescaling Total Points for PSA Final Risk Scores*

Total Points			Rescaled Risk Score
FTA	NCA	NVCA	
0	0	0 or 1	1
1	1 or 2	2	2
2	3 or 4	3	3
3 or 4	5 or 6	4	4
5 or 6	7 or 8	5	5
7	9 to 13	6 or 7	6

**Documentation of risk factor and outcome definitions and measurement:**

Advancing Pretrial Policy and Research, nd. Risk Factors Public Safety Assessment Implementation Guide 11: Guide to PSA Risk Factors. At: <https://advancingpretrial.org/implementation/guides/> (Registration required)

## Appendix B Violent Offenses

CODE_TYPE	STATUTE_CODE	PC	207(C)	PC	241
PC	11413(A)	PC	207(D)	PC	241(A)
PC	11413(B)	PC	208(B)	PC	241(B)
PC	11418(B)	PC	209(A)	PC	241(C)
PC	11418(C)	PC	209(B)(1)	PC	241.1
PC	11418(D)	PC	209.5(A)	PC	241.2
PC	136.1(C)(1)	PC	210.5	PC	241.2(A)
PC	140(A)	PC	211	PC	241.2(A)(1)
PC	148(B)	PC	212	PC	241.3
PC	148(C)	PC	212.5	PC	241.3(A)
PC	148(D)	PC	212.5(A)	PC	241.4
PC	148.10(A)	PC	212.5(B)	PC	241.5
PC	149	PC	212.5(C)	PC	241.6
PC	151	PC	214	PC	241.7
PC	151(A)(1)	PC	215	PC	241.8(A)
PC	18540(A)	PC	215(A)	PC	242
PC	186.26(C)	PC	217.1(A)	PC	243
PC	187(A)	PC	217.1(B)	PC	243(A)
PC	18740	PC	218	PC	243(B)
PC	18745	PC	218.1	PC	243(C)
PC	18750	PC	219	PC	243(C)(1)
PC	18755	PC	219.1	PC	243(C)(2)
PC	18755(B)	PC	219.2	PC	243(D)
PC	191.5(A)	PC	220	PC	243(E)
PC	192(A)	PC	220(A)	PC	243(E)(1)
PC	192(B)	PC	220(A)(1)	PC	243.1
PC	192(C)(1)	PC	220(A)(2)	PC	243.2(A)(1)
PC	192(C)(3)	PC	220(B)	PC	243.25
PC	192.5(A)	PC	222	PC	243.3
PC	192.5(B)	PC	236	PC	243.35
PC	192.5(C)	PC	236.1	PC	243.35(A)
PC	203	PC	236.1(A)	PC	243.4
PC	205	PC	236.1(B)	PC	243.4(A)
PC	206	PC	236.1(C)	PC	243.4(B)
PC	207(A)	PC	236.1(C)(1)	PC	243.4(C)
PC	207(A)/208(B)	PC	236.1(C)(2)	PC	243.4(D)
PC	207(A)/208(D)	PC	237(A)	PC	243.4(D)(1)
PC	207(B)	PC	240	PC	243.4(E)(1)

PC	243.5(A)(1)	PC	262(A)(3)	PC	286(C)(1)
PC	243.6	PC	262(A)(4)	PC	286(C)(2)
PC	243.65(A)	PC	262(A)(5)	PC	286(C)(2)(A)
PC	243.7	PC	262(A)(6)	PC	286(C)(2)(B)
PC	243.8(A)	PC	264.1	PC	286(C)(2)(C)
PC	243.9(A)	PC	264.1(A)	PC	286(C)(3)
PC	244	PC	264.1(B)(2)	PC	286(D)
PC	244.5(B)	PC	266A	PC	286(D)(1)
PC	244.5(C)	PC	266B	PC	286(D)(2)
PC	245(A)(1)	PC	266C	PC	286(D)(3)
PC	245(A)(2)	PC	266H(B)	PC	286(F)
PC	245(A)(3)	PC	266H(B)(1)	PC	286(G)
PC	245(A)(4)	PC	266H(B)(2)	PC	286(H)
PC	245(B)	PC	266I(B)	PC	286(I)
PC	245(C)	PC	266I(B)(1)	PC	286(J)
PC	245(D)(1)	PC	266I(B)(2)	PC	286(K)
PC	245(D)(2)	PC	266J	PC	287(B)
PC	245(D)(3)	PC	267	PC	287(B)(1)
PC	245.2	PC	269(A)	PC	287(B)(2)
PC	245.3	PC	269(A)(1)	PC	287(C)
PC	245.5(A)	PC	269(A)(2)	PC	287(C)(1)
PC	245.5(B)	PC	269(A)(3)	PC	287(C)(2)(A)
PC	245.5(C)	PC	269(A)(4)	PC	287(C)(2)(C)
PC	245.6	PC	269(A)(5)	PC	287(D)
PC	245.6(A)	PC	273.4	PC	287(F)
PC	246	PC	273.5(A)	PC	287(G)
PC	246.3(A)	PC	273.5(E)	PC	287(H)
PC	246.3(B)	PC	273.5(E)(1)	PC	287(I)
PC	261(A)	PC	273.6(B)	PC	287(J)
PC	261(A)(1)	PC	273.6(D)	PC	287(K)
PC	261(A)(2)	PC	273A(A)	PC	288(A)
PC	261(A)(3)	PC	273A(A)(1)	PC	288(B)
PC	261(A)(4)	PC	273A(B)	PC	288(B)(1)
PC	261(A)(5)	PC	273AB(A)	PC	288(B)(2)
PC	261(A)(6)	PC	273AB(B)	PC	288(C)
PC	261(A)(7)	PC	273D(A)	PC	288(C)(1)
PC	261.5(A)	PC	278	PC	288(C)(2)
PC	261.5(B)	PC	278(A)	PC	288.2(A)
PC	261.5(C)	PC	285	PC	288.2(A)(1)
PC	261.5(D)	PC	286(B)	PC	288.2(A)(2)
PC	26100(C)	PC	286(B)(1)	PC	288.3
PC	262(A)(1)	PC	286(B)(2)	PC	288.3(A)
PC	262(A)(2)	PC	286(C)	PC	288.3(B)

PC	288.4	PC	289.6(A)(3)	PC	647.6(C)
PC	288.4(A)(1)	PC	311.4(A)	PC	647.6(C)(1)
PC	288.4(B)	PC	311.4(B)	PC	647.6(C)(2)
PC	288.5(A)	PC	311.4(C)	PC	664/187(A)
PC	288.7(A)	PC	347(A)	PC	664/211
PC	288.7(B)	PC	347(A)(1)	PC	667.61(D)(2)
PC	288A(B)	PC	368(B)	PC	667.61(D)(3)
PC	288A(B)(1)	PC	368(B)(1)	PC	667.61(E)(1)
PC	288A(B)(2)	PC	368(C)	PC	667.61(E)(2)
PC	288A(C)	PC	368(F)	PC	667.61(E)(4)
PC	288A(C)(1)	PC	404(A)	PC	667.61(E)(5)
PC	288A(C)(2)	PC	417(A)	PC	667.8
PC	288A(C)(2)(A)	PC	417(A)(1)	PC	667.85
PC	288A(C)(2)(B)	PC	417(A)(2)	PC	674
PC	288A(C)(2)(C)	PC	417(A)(2)(A)	PC	69
PC	288A(C)(3)	PC	417(A)(2)(B)	PC	836.6(C)
PC	288A(D)	PC	417(B)	VC	2800.3(A)
PC	288A(D)(1)	PC	417(C)	VC	2800.3(B)
PC	288A(D)(2)	PC	417.3		
PC	288A(D)(3)	PC	417.8		
PC	288A(F)	PC	422.6(A)		
PC	288A(G)	PC	4500		
PC	288A(H)	PC	4501		
PC	288A(I)	PC	4501(A)		
PC	288A(J)	PC	4501(B)		
PC	288A(K)	PC	4501.1(A)		
PC	289	PC	4501.5		
PC	289(A)	PC	4503		
PC	289(A)(1)	PC	451(A)		
PC	289(A)(1)(A)	PC	451(B)		
PC	289(A)(1)(B)	PC	451.1		
PC	289(A)(1)(C)	PC	451.5(A)		
PC	289(A)(2)	PC	452(A)		
PC	289(B)	PC	452(B)		
PC	289(C)	PC	4530(A)		
PC	289(D)	PC	4532(A)(2)		
PC	289(D)(4)	PC	4532(B)(2)		
PC	289(E)	PC	455		
PC	289(F)	PC	455(A)		
PC	289(G)	PC	646.9(A)		
PC	289(H)	PC	646.9(B)		
PC	289(I)	PC	647.6(A)(1)		
PC	289(J)	PC	647.6(B)		

## Appendix C Exclusions

### PC 1270.1

<b>Penal Code</b>	186.22(e)(5)	192.5(a)	245(b)	9(a)(2)
37	186.22(e)(6)	192.5(b)	245(c)	269(a)(3)
37(a)	186.22(e)(7)	192.5(c)	245(d)	269(a)(4)
37(b)	186.22(e)(8)	203	245(d)(1)	269(a)(5)
128	186.22(e)(9)	205	245(d)(2)	269(b)
136.1	186.22(e)(10)	206	245(d)(3)	269(c)
136.1(a)	186.22(e)(11)	207	245.2	273ab
136.1(a)(1)	186.22(e)(12)	207(a)	245.3	273ab(a)
136.1(a)(2)	186.22(e)(13)	207(b)	245.5	273ab(b)
136.1(a)(3)	186.22(e)(14)	207(c)	245.5(a)	273.4
136.1(b)	186.22(e)(15)	207(d)	245.5(b)	273.5
136.1(b)(1)	186.22(e)(16)	207(e)	245.5(c)	273.5(a)
136.1(b)(2)	186.22(e)(17)	207(f)	245.5(d)	273.5(b)
136.1(b)(3)	186.22(e)(18)	207(f)(1)	245.6(d)	273.5(b)(1)
136.1(c)	186.22(e)(19)	207(f)(2)	246	273.5(b)(2)
136.1(c)(1)	186.22(e)(20)	209	246.3(a)	273.5(b)(3)
136.1(c)(2)	186.22(e)(21)	209(a)	247	273.5(b)(4)
136.1(c)(3)	186.22(e)(22)	209(b)	247(a)	273.5(c)
136.1(c)(4)	186.22(e)(23)	209(b)(1)	247(b)	273.5(d)
136.1(d)	186.22(e)(24)	209(b)(2)	261	273.5(e)
136.1(e)	186.22(e)(25)	209(c)	261(a)	273.5(f)
136.1(f)	186.22(e)(26)	209(d)	261(a)(1)	273.5(f)(1)
140	186.22(e)(27)	209.5	261(a)(2)	273.5(f)(2)
140(a)	186.22(e)(28)	209.5(a)	261(a)(3)	273.5(g)
140(b)	186.22(e)(29)	209.5(b)	261(a)(4)	273.5(h)
148.10(a)	186.22(e)(30)	209.5(c)	261(a)(5)	273.5(h)(1)
182/serious felony	186.22(e)(31)	211	261(a)(6)	273.5(h)(2)
182(a)	186.22(e)(32)	212.5	261(a)(7)	273.5(h)(3)
182(a)1	186.22(e)(33)	212.5(a)	261(b)	273.5(i)
182(a)2	186.22(e)(f)	212.5(b)	261(c)	273.5(i)(1)
182(a)3	186.22(e)(g)	212.5(c)	262	273.5(i)(2)
182(a)4	186.22(e)(h)	214	262(a)	273.5(j)
182(a)5	186.22(e)(i)	215	262(a)(1)	273.5(k)
182(a)6	186.22(e)(j)	215(a)	262(a)(2)	273.6
182(b)	186.22(e)(k)	215(b)	262(a)(3)	273.6(a)
186.22(a)	187	215(c)	262(a)(4)	273.6(b)
186.22	187(a)	217.1(b)	262(a)(5)	273.6(c)
186.22(a)	187(b)	218	262(b)	273.6(c)(1)
186.22(b)	191.5	219	262(c)	273.6(c)(2)
186.22(b)(1)	191.5(a)	220	262(d)	273.6(c)(3)
186.22(b)(2)	191.5(b)	220(a)	262(d)(1)	273.6(c)(4)
186.22(b)(3)	191.5(c)	220(a)(1)	262(d)(2)	273.6(d)
186.22(b)(4)	191.5(c)(1)	220(a)(2)	264.1	273.6(e)
186.22(b)(5)	191.5(c)(2)	220(b)	264.1(a)	273.6(f)
186.22(c)	191.5(d)	236.1(c)(2)	264.1(b)	273.6(g)
186.22(d)	191.5(e)	236.4(b)	264.1(b)(1)	273.6(g)(1)
186.22(e)	191.5(f)	243(d)	264.1(b)(2)	273.6(g)(2)
186.22(e)(1)	191.5(g)	243(e)(1)	264.1(b)(3)	273.6(h)
186.22(e)(2)	192(a)	244	269	273.6(h)(1)
186.22(e)(3)	192(b)	245(a)(1)	269(a)	273.6(h)(2)
186.22(e)(4)	192(c)(1)	245(a)(2)	269(a)(1)	273.6(i)
	192(c)(3)	245(a)(3)	26	286(c)



286(c)(1)	451.1(a)(2)	667.61(d)(4)	12022.53(a)(8)	<b>Military And Vets</b>
286(c)(2)	451.1(a)(3)	667.61(d)(5)	12022.53(a)(9)	<b>Code</b>
286(c)(3)	451.1(a)(4)	667.61(d)(6)	12022.53(a)(10)	1670
286(d)	451.1(a)(5)	667.61(d)(7)	12022.53(a)(11)	1671
286(d)(1)	451.1(b)	667.61(e)	12022.53(a)(12)	1672
286(d)(2)	451.5	667.61(e)(1)	12022.53(a)(13)	1672(a)
286(d)(3)	451.5(a)	667.61(e)(2)	12022.53(a)(14)	1672(b)
286(d)(4)	451.5(a)(1)	667.61(e)(3)	12022.53(a)(15)	1674
287(c)	451.5(a)(2)	667.61(e)(4)	12022.53(a)(16)	<b>Welfare And Inst</b>
287(c)(1)	451.5(a)(3)	667.61(e)(5)	12022.53(a)(17)	<b>Code</b>
287(c)(2)	451.5(b)	667.61(e)(6)	12022.53(a)(18)	1768.8(b)
287(c)(3)	451.5(c)	667.61(e)(7)	12022.53(b)	<b>Election Code</b>
287(d)	451.5(d)	667.61(f)	12022.53(c)	18540
287(d)(1)	452(a)	667.61(g)	12022.53(d)	18540(a)
287(d)(2)	452(b)	667.61(h)	12022.53(e)	18540(b)
287(d)(3)	452(c)	667.61(i)	12022.53(e)(1)	
287(d)(4)	453	667.61(j)	12022.53(e)(2)	
288	453(a)	667.61(j)(1)	12022.53(f)	
288(a)	453(b)	667.61(j)(2)	12022.53(g)	
288(b)	453(b)(1)	667.61(k)	12022.53(h)	
288(b)(1)	453(b)(2)	667.61(l)	12022.53(i)	
288(b)(2)	453(b)(3)	667.61(m)	12022.53(j)	
288.5	453(c)	667.61(n)	12022.53(k)	
288.5(a)	455	667.61(n)(1)	12022.53(l)	
288.5(b)	455(a)	667.61(n)(2)	12022.55	
288.5(c)	455(b)	667.61(n)(3)	12022.7	
288.7	487(d)(2)	667.61(n)(4)	12022.7(a)	
288.7(a)	550(g)	667.61(n)(5)	12022.7(b)	
288.7(b)	600(d)	667.61(n)(6)	12022.7(c)	
289(a)	626.9(d)	667.61(o)	12022.7(d)	
289(a)(1)	646.9	667.75	12022.7(e)	
289(a)(1)(A)	646.9(a)	667.8	12022.7(f)	
289(a)(1)(B)	646.9(b)	667.8(a)	12022.7(g)	
289(a)(1)(C)	646.9(c)	667.8(b)	12022.7(h)	
289(a)(1)(D)	646.9(c)(1)	667.8(c)	12022.8	
289(a)(2)	646.9(c)(2)	667.8(c)(1)	12022.9	
289(j)	646.9(d)	667.8(c)(2)	12022.95	
347(a)(2)	646.9(e)	667.8(c)(3)	18740	
368(b)(2)	646.9(f)	667.85	18745	
368(b)(2)(A)	646.9(g)	4500	18750	
368(b)(2)(B)	646.9(h)	4501(a)	18755	
368(b)(3)	646.9(i)	4501.1(a)	18755(a)	
368(b)(3)(A)	646.9(j)	4503	18755(b)	
368(b)(3)(B)	646.9(k)	11413(a)	26100(c)	
404.6(c)	646.9(k)(1)	11418(b)	26100(d)	
417(b)	646.9(k)(2)	11418(b)(1)	<b>Vehicle Code</b>	
417(c)	646.9(l)	11418(b)(2)	2800.3	
417.3	646.9(m)	11418(b)(3)	2800.3(a)	
417.6	664	11418(b)(4)	2800.3(b)	
417.6(a)	667.61	11418(c)	2800.3(c)	
417.6(b)	667.61(a)	12022(b)	2800.3(d)	
417.6(c)	667.61(b)	12022(b)(1)	23104(b)	
417.8	667.61(c)	12022(b)(2)	23153	
422	667.61(c)(1)	12022.3(a)	23153(a)	
422(a)	667.61(c)(2)	12022.5(a)	23153(b)	
422(b)	667.61(c)(3)	12022.5(b)	23153(c)	
422(c)	667.61(c)(4)	12022.5(d)	23153(d)	
451	667.61(c)(5)	12022.53	23153(e)	
451(a)	667.61(c)(6)	12022.53(a)	23153(f)	
451(b)	667.61(c)(7)	12022.53(a)(1)	23153(g)	
451(c)	667.61(c)(8)	12022.53(a)(2)	<b>Health And</b>	
451(d)	667.61(c)(9)	12022.53(a)(3)	<b>Safety Code</b>	
451(e)	667.61(d)	12022.53(a)(4)	11353	
451.1	667.61(d)(1)	12022.53(a)(5)	11380	
451.1(a)	667.61(d)(2)	12022.53(a)(6)	25189.7	
451.1(a)(1)	667.61(d)(3)	12022.53(a)(7)		

## PC 1270.1 Attempts

<b>Penal Code</b>	664/186.22(e)(25)	664/212.5(c)	664/264.1(b)(1)	664/286(d)(1)
664/37	664/186.22(e)(26)	664/214	664/264.1(b)(2)	664/286(d)(2)
664/37(a)	664/186.22(e)(27)	664/215	664/264.1(b)(3)	664/286(d)(3)
664/37(b)	664/186.22(e)(28)	664/215(a)	664/269	664/286(d)(4)
664/128	664/186.22(e)(29)	664/215(b)	664/269(a)	664/287(c)
664/136.1	664/186.22(e)(30)	664/215(c)	664/269(a)(1)	664/287(c)(1)
664/136.1(a)	664/186.22(e)(31)	664/217.1(b)	664/269(a)(2)	664/287(c)(2)
664/136.1(a)(1)	664/186.22(e)(32)	664/218	664/269(a)(3)	664/287(c)(3)
664/136.1(a)(2)	664/186.22(e)(33)	664/219	664/269(a)(4)	664/287(d)
664/136.1(a)(3)	664/186.22(e)(f)	664/236.1(c)(2)	664/269(a)(5)	664/287(d)(1)
664/136.1(b)	664/186.22(e)(g)	664/236.4(b)	664/269(b)	664/287(d)(2)
664/136.1(b)(1)	664/186.22(e)(h)	664/243(d)	664/269(c)	664/287(d)(3)
664/136.1(b)(2)	664/186.22(e)(i)	664/243(e)(1)	664/273ab	664/287(d)(4)
664/136.1(b)(3)	664/186.22(e)(j)	664/244	664/273ab(a)	664/288
664/136.1(c)	664/186.22(e)(k)	664/245(a)(1)	664/273ab(b)	664/288(a)
664/136.1(c)(1)	664/187	664/245(a)(2)	664/273.4	664/288(b)
664/136.1(c)(2)	664/187(a)	664/245(a)(3)	664/273.5	664/288(b)(1)
664/136.1(c)(3)	664/187(b)	664/245(b)	664/273.5(a)	664/288(b)(2)
664/136.1(c)(4)	664/191.5	664/245(c)	664/273.5(b)	664/288.5
664/136.1(d)	664/191.5(a)	664/245(d)	664/273.5(b)(1)	664/288.5(a)
664/136.1(e)	664/191.5(b)	664/245(d)(1)	664/273.5(b)(2)	664/288.5(b)
664/136.1(f)	664/191.5(c)	664/245(d)(2)	664/273.5(b)(3)	664/288.5(c)
664/140	664/191.5(c)(1)	664/245(d)(3)	664/273.5(b)(4)	664/288.7
664/140(a)	664/191.5(c)(2)	664/245.2	664/273.5(c)	664/288.7(a)
664/140(b)	664/191.5(d)	664/245.3	664/273.5(d)	664/288.7(b)
664/148.10(a)	664/191.5(e)	664/245.5	664/273.5(e)	664/289(a)
664/186.22	664/191.5(f)	664/245.5(a)	664/273.5(f)	664/289(a)(1)
664/186.22(a)	664/191.5(g)	664/245.5(b)	664/273.5(f)(1)	664/289(a)(1)(A)
664/186.22(b)	664/192(a)	664/245.5(c)	664/273.5(f)(2)	664/289(a)(1)(B)
664/186.22(b)(1)	664/192(b)	664/245.5(d)	664/273.5(g)	664/289(a)(1)(C)
664/186.22(b)(2)	664/192(c)(1)	664/245.6(d)	664/273.5(h)	664/289(a)(1)(D)
664/186.22(b)(3)	664/192(c)(3)	664/246	664/273.5(h)(1)	664/289(a)(2)
664/186.22(b)(4)	664/192.5(a)	664/246.3(a)	664/273.5(h)(2)	664/289(j)
664/186.22(b)(5)	664/192.5(b)	664/247	664/273.5(h)(3)	664/347(a)(2)
664/186.22(c)	664/192.5(c)	664/247(a)	664/273.5(i)	664/368(b)(2)
664/186.22(d)	664/203	664/247(b)	664/273.5(i)(1)	664/368(b)(2)(A)
664/186.22(e)	664/205	664/261	664/273.5(i)(2)	664/368(b)(2)(B)
664/186.22(e)(1)	664/206	664/261(a)	664/273.5(j)	664/368(b)(3)
664/186.22(e)(2)	664/207	664/261(a)(1)	664/273.5(k)	664/368(b)(3)(A)
664/186.22(e)(3)	664/207(a)	664/261(a)(2)	664/273.6	664/368(b)(3)(B)
664/186.22(e)(4)	664/207(b)	664/261(a)(3)	664/273.6(a)	664/404.6(c)
664/186.22(e)(5)	664/207(c)	664/261(a)(4)	664/273.6(b)	664/417
664/186.22(e)(6)	664/207(d)	664/261(a)(5)	664/273.6(c)	664/417(b)
664/186.22(e)(7)	664/207(e)	664/261(a)(6)	664/273.6(c)(1)	664/417(c)
664/186.22(e)(8)	664/207(f)	664/261(a)(7)	664/273.6(c)(2)	664/417.3
664/186.22(e)(9)	664/207(f)(1)	664/261(b)	664/273.6(c)(3)	664/417.6
664/186.22(e)(10)	664/207(f)(2)	664/261(c)	664/273.6(c)(4)	664/417.6(a)
664/186.22(e)(11)	664/209	664/262	664/273.6(d)	664/417.6(b)
664/186.22(e)(12)	664/209(a)	664/262(a)	664/273.6(e)	664/417.6(c)
664/186.22(e)(13)	664/209(b)	664/262(a)(1)	664/273.6(f)	664/417.8
664/186.22(e)(14)	664/209(b)(1)	664/262(a)(2)	664/273.6(g)	664/422
664/186.22(e)(15)	664/209(b)(2)	664/262(a)(3)	664/273.6(g)(1)	664/422(a)
	664/209(c)	664/262(a)(4)	664/273.6(g)(2)	664/422(b)
	664/209(d)	664/262(a)(5)	664/273.6(h)	664/422(c)
	664/209.5	664/262(b)	664/273.6(h)(1)	664/451
	664/209.5(a)	664/262(c)	664/273.6(h)(2)	664/451(a)
	664/209.5(b)	664/262(d)	664/273.6(i)	664/451(b)
	664/209.5(c)	664/262(d)(1)	664/286(c)	664/451(c)
	664/211	664/262(d)(2)	664/286(c)(1)	664/451(d)
	664/212.5	664/264.1	664/286(c)(2)	664/451(e)
	664/212.5(a)	664/264.1(a)	664/286(c)(3)	664/451.1
	664/212.5(b)	664/264.1(b)	664/286(d)	664/451.1(a)

664/451.1(a)(1)	664/667.61(d)(7)	664/12022.53(a)(7)	664/23153(c)
664/451.1(a)(2)	664/667.61(e)	)	664/23153(d)
664/451.1(a)(3)	664/667.61(e)(1)	664/12022.53(a)(8)	664/23153(e)
664/451.1(a)(4)	664/667.61(e)(2)	)	664/23153(f)
664/451.1(a)(5)	664/667.61(e)(3)	664/12022.53(a)(9)	664/23153(g)
664/451.1(b)	664/667.61(e)(4)	)	<b>Health And Safety</b>
664/451.5	664/667.61(e)(5)	664/12022.53(a)(1	<b>Code</b>
664/451.5(a)	664/667.61(e)(6)	0)	664/11353
664/451.5(a)(1)	664/667.61(e)(7)	664/12022.53(a)(1	664/11380
664/451.5(a)(2)	664/667.61(f)	1)	<b>Military And</b>
664/451.5(a)(3)	664/667.61(g)	664/12022.53(a)(1	<b>Veterans Code</b>
664/451.5(b)	664/667.61(h)	2)	664/1670
664/451.5(c)	664/667.61(i)	664/12022.53(a)(1	664/1671
664/451.5(d)	664/667.61(j)	3)	664/1672
664/452(a)	664/667.61(j)(1)	664/12022.53(a)(1	664/1672(a)
664/452(b)	664/667.61(j)(2)	4)	664/1672(b)
664/452(c)	664/667.61(k)	664/12022.53(a)(1	664/1674
664/453	664/667.61(l)	5)	<b>Welfare And Inst</b>
664/453(a)	664/667.61(m)	664/12022.53(a)(1	<b>Code</b>
664/453(b)	664/667.61(n)	6)	664/1768.8(b)
664/453(b)(1)	664/667.61(n)(1)	664/12022.53(a)(1	<b>Elections Code</b>
664/453(b)(2)	664/667.61(n)(2)	7)	664/18540
664/453(b)(3)	664/667.61(n)(3)	664/12022.53(a)(1	664/18540(a)
664/453(c)	664/667.61(n)(4)	8)	664/18540(b)
664/455	664/667.61(n)(5)	664/12022.53(b)	
664/455(a)	664/667.61(n)(6)	664/12022.53(c)	
664/455(b)	664/667.61(o)	664/12022.53(d)	
664/487(d)(2)	664/667.75	664/12022.53(e)	
664/550(g)	664/667.8	664/12022.53(e)(1	
664/600(d)	664/667.8(a)	)	
664/626.9(d)	664/667.8(b)	664/12022.53(e)(2	
664/646.9	664/667.8(c)	)	
664/646.9(a)	664/667.8(c)(1)	664/12022.53(f)	
664/646.9(b)	664/667.8(c)(2)	664/12022.53(g)	
664/646.9(c)	664/667.8(c)(3)	664/12022.53(h)	
664/646.9(c)(1)	664/667.85	664/12022.53(i)	
664/646.9(c)(2)	664/4500	664/12022.53(j)	
664/646.9(d)	664/4501(a)	664/12022.53(k)	
664/646.9(e)	664/4501.1(a)	664/12022.53(l)	
664/646.9(f)	664/4503	664/12022.55	
664/646.9(g)	664/11413(a)	664/12022.7	
664/646.9(h)	664/11418(b)	664/12022.7(a)	
664/646.9(i)	664/11418(b)(1)	664/12022.7(b)	
664/646.9(j)	664/11418(b)(2)	664/12022.7(c)	
664/646.9(k)	664/11418(b)(3)	664/12022.7(d)	
664/646.9(k)(1)	664/11418(b)(4)	664/12022.7(e)	
664/646.9(k)(2)	664/11418(c)	664/12022.7(f)	
664/646.9(l)	664/12022(b)	664/12022.7(g)	
664/646.9(m)	664/12022(b)(1)	664/12022.7(h)	
664/667.61	664/12022(b)(2)	664/12022.8	
664/667.61(a)	664/12022.3(a)	664/12022.9	
664/667.61(b)	664/12022.5(a)	664/18740	
664/667.61(c)	664/12022.5(b)	664/18745	
664/667.61(c)(1)	664/12022.5(d)	664/18750	
664/667.61(c)(2)	664/12022.53	664/18755	
664/667.61(c)(3)	664/12022.53(a)	664/18755(a)	
664/667.61(c)(4)	664/12022.53(a)(1	664/18755(b)	
664/667.61(c)(5)	)	664/26100(c)	
664/667.61(c)(6)	664/12022.53(a)(2	664/26100(d)	
664/667.61(c)(7)	)	<b>Vehicle Code</b>	
664/667.61(c)(8)	664/12022.53(a)(3	664/2800.3	
664/667.61(c)(9)	)	664/2800.3(a)	
664/667.61(d)	664/12022.53(a)(4	664/2800.3(b)	
664/667.61(d)(1)	)	664/2800.3(c)	
664/667.61(d)(2)	664/12022.53(a)(5	664/2800.3(d)	
664/667.61(d)(3)	)	664/23104(b)	
664/667.61(d)(4)	664/12022.53(a)(6	664/23153	
664/667.61(d)(5)	)	664/23153(a)	
664/667.61(d)(6)		664/23153(b)	

## PC 1319.5

Penal Code		
166	243(G)	244.5(C)
166(A)	243.1	244.5(D)
166(B)	243.2	245
166(C)	243.2(A)	245(A)
166(D)	243.2(B)	245(B)
166(E)	243.2(C)	245(C)
186.26	243.25	245(D)
186.28	243.3	245(E)
186.33	243.35	245(F)
240	243.35(A)	245.6
241	243.35(B)	245.6(A)
241(A)	243.5	245.6(B)
241(B)	243.5(A)	245.6(C)
241(C)	243.5(B)	245.6(D)
241(D)	243.6	245.6(E)
241.1	243.65	245.6(F)
241.2	243.65(A)	246.3
241.2(A)	243.65(B)	246.3(A)
241.2(B)	243.7	246.3(B)
241.2(C)	243.8	246.3(C)
241.3	243.8(A)	247
241.3(A)	243.8(B)	247(A)
241.3(B)	243.83	247(B)
241.3(C)	243.83(A)	247.5
241.4	243.83(B)	248
241.5	243.83(C)	459
241.5(A)	243.83(D)	12021.5
241.5(B)	243.83(E)	12022(A)
241.6	243.83(F)	12022(B)
241.7	243.9	12022(C)
241.8	243.9(A)	12022.2(A)
241.8(A)	243.9(B)	12022.3(B)
241.8(B)	243.9(C)	12022.5
242	243.9(D)	653M
243	243.9(E)	653M(A)
243(A)	243.10	653M(B)
243(B)	243.10(A)	653M(C)
243(C)	243.10(B)	653M(D)
243(D)	243.15	653M(E)
243(E)	244.5	653M(F)
243(F)	244.5(A)	653M(G)
	244.5(B)	

## Appendix D

### Offense Classification

The offense classification is based on the California Department of Justice Master Offense Table, which is used for reporting criminal justice statistics. It was adapted for use for the California Static Risk Assessment in 2008 and has been updated several times since then to add new statutory definitions of criminal acts. It maps over 4,000 California code sections to 19 general categories of felony and misdemeanor offenses and supervision violations. The full mapping table is too long to include; it may be requested from the U.C. Irvine Center for Evidence-Based Corrections. The offense categories are:

Felony Homicide

Felony Sex

Felony Violent Property

Felony Assault - not domestic

Felony Domestic Assault/Violation

Felony Weapon

Felony Property

Felony Drug

Felony Escape

Misdemeanor Assault - not domestic

Misdemeanor Domestic Assault/Violation

Misdemeanor Sex

Misdemeanor Other Domestic Violation

Misdemeanor Weapon

Misdemeanor Property

Misdemeanor Drug

Misdemeanor Escape

Misdemeanor Alcohol

Supervision Violations

## Appendix E

### Charges Recovery by Imputation

Option D is intended first to test the possibility of recovering ill-formed charges and imputing Level of Charge from offense information. We focus on the main source of problematic charges, records in the History and Bookings file where the value of “NULL” is entered in the CJIS\_CHARGE\_CODE\_TYPE fields. To recap, 3.4% of all history records have a “NULL” Code Type and 4.2% of all history records are missing CHARGE\_LEVEL\_CODE. As the average individual has multiple records in the history data, these enter into many criminal history records. Of 190,129 individuals with history records, 32.8% have a “NULL” Code Type in one or more History records; 30.9% have one or more records missing Level of Charge. In Bookings records, 8.2% have a “NULL” Code Type but less than less than 0.01% are missing CHARGE\_LEVEL\_CODE. Of the 476,480 individuals with bookings records, 16.9% have a “NULL” in one or more records; 59 have one or more records missing Level of Charge. For history, after recovery 20.4% have a “NULL” Code Type in one or more history records; 30.3% have one or more records missing CHARGE\_LEVEL\_CODE. For bookings after recovery, 5.6% have a “NULL” Code Type in one or more records; 46 have one or more records missing CHARGE\_LEVEL\_CODE.

The data processing to recover poorly formatted charges by imputation on the basis of available evidence was initially developed by UCI for the California Static Risk Assessment under contract to the California Department of Corrections. Successful charge recovery helps alleviate the possibility that someone might receive lower risk level than warranted. The process is somewhat complex with multiple subroutines designed to address different kinds of formatting problems and using tables of authorities to fill in missing information when possible. Perl script functions and calls are used to parse text strings and identify charge code strings and their variants. Here we present a brief sketch of the process.

- 1) As History and Bookings records have different field names and slightly different formatting problems, they are imported and processed with different codes at the first stage. Nevertheless, the steps perform similar functions.
  - a) Extract records with a “NULL” CJIS\_CHARGE\_CODE\_TYPE.
  - b) Standardize two-character strings in the CHARGE\_CODE field used to identify the government origin of the statutes used to record offenses. Extract these strings and use them to replace the “NULL” in CJIS\_CHARGE\_CODE\_TYPE with these statute code designations or their type. These are:
    - i) Major divisions of California statute codes, e.g. “PC” identifies Penal Codes. Statute code designations are used to replace “NULL”.
    - ii) The two-character string “US” in the CHARGE\_CODE field used to identify federal charges and is used to replace “NULL”.
    - iii) Identify two-character strings in the CHARGE\_CODE field used to identify municipalities or counties when the offense charge is based on local ordinances and regulations, e.g., “LB” for

Long Beach. No further effort is made to identify these offenses as it would be prohibitive to investigate hundreds of municipal codes. Most of these records have CHARGE\_LEVEL\_CODE filled in already, and these offenses will not include those listed as exclusion criteria or other violent offenses needing to be flagged. Hence, the “NULL” in CJIS\_CHARGE\_CODE\_TYPE with these municipal code designations is replaced with “MUNI”.

- c) Flag records where the initial character of the string in CHARGE\_CODE is alpha rather than numeric, as these are likely to be text descriptions of the offense rather than statute code references, and export a table with these descriptions and the frequency of their occurrence to facilitate manual mapping of high impact offenses to violent or serious statute codes.
- 2) Records from both History and Bookings are now sufficiently standardized to allow further processing of both sources by the same code. First, records separated into 3 datafiles for format-specific processing:
  - a) Records based on California and Federal statute-based charges;
  - b) Records based on municipal or county code-based charges;
  - c) Records with text descriptions of charges
- 3) California and federal charges are parsed to identify deviations from the standard formatting such as misplacement of parenthesis, substitution of other punctuation for periods, etc.
- 4) When multiple charges are found in one record, they are split out into one record each before further processing.
  - a) Charges with an alpha character following the numeric statute code number are matched to a table of legitimate charges with this variant format; ones that are not verified are adjusted to turn the alpha character into a subsection designation between parentheses.
  - b) Charges still missing a valid in CJIS\_CHARGE\_CODE\_TYPE are matched to a list of statute code designations which exist in only one of the major code sections; this allows determination of the origin so the code type can be filled in.
  - c) CJIS\_CHARGE\_CODE\_TYPE and CHARGE\_CODE are checked against the California Department of Justice Master Offense Table. Matching codes are assumed to be successfully recovered. This is an iterative process:
    - i) Charges are first matched at the full level of specificity in which they are records, i.e., at code type, section, subsection, and sub-subsection;
    - ii) If a match is not achieved at this level of specificity, the match is next attempted at the subsection, then the section level.
    - iii) A search is made for a match to any statute, not limited to the DOJ list, with a unique section designation that allows a positive match.
  - d) Records with PC 17 are used to ascertain charge level where multiple levels are possible. Where multiple levels are possible and PC 17 is not found, the precautionary principle is used to fill in CHARGE\_LEVEL\_CODE with the highest level.
- 5) Import the table of manually mapped text offenses and merge it with the text datafile to incorporate the mapping.
- 6) Reassemble a History or Booking file of the recovered charges and levels of charge by aggregating the individual files of California and Federal charges, text charge descriptions, and municipal and

county charges. Check for a pre-existing level of charge to ensure that an imputed value is not used to override a recorded on.

- 7) Merge the files of recovered charges with the original files using the unique ROW\_NUM field to maintain alignment with the appropriate events and dates. Use the recovered level of charge to overwrite the CHARGE\_LEVEL\_CODE where it has a missing or invalid value.



## Appendix F

### Option D: Supplementary Analysis

Because the risk factor counting rules limit the impact of adding felony and misdemeanor convictions, we found little impact on predictive performance from charge recovery. This leaves open a second question: What is the effect of charge recovery when it does change risk factors? Does it improve predictive ability? To answer this, we present this supplementary analysis of the subset of 703 booking cases in Options C and D where recovery changed risk scores.

Tables show demographics (F.1.1), data completeness (F.1.2) risk factors (F.1.3), and outcomes (F.1.4) for this subset, replicating the presentation of Table 4.5. The values before and after the recovery operation are labeled Option C and Option D respectively. Age, sex, descent, and the number of bookings for the individuals are not affected by charge recovery; hence, they are the same in the both columns.

*Table F.1.1 Supplementary Analysis: Demographics*

Values	Option C n=703	Option D n=703
<b>Age</b>		
<b>Under 18</b>	0.00	0.00
<b>18-19</b>	0.00	0.00
<b>20-29</b>	11.38	11.38
<b>30-39</b>	31.86	31.86
<b>40-49</b>	24.18	24.18
<b>50-59</b>	24.47	24.47
<b>60-69</b>	5.41	5.41
<b>70+</b>	2.70	2.70
<b>Sex</b>		
<b>Female</b>	12.23	12.23
<b>Male</b>	87.77	87.77
<b>Descent</b>		
<b>Asian</b>	0.85	0.85
<b>Black</b>	33.57	33.57
<b>Hispanic</b>	40.68	40.68
<b>Other</b>	2.13	2.13
<b>White</b>	22.76	22.76
<b>Bookings</b>		
<b>1</b>	65.29	65.29
<b>2-5</b>	29.16	29.16
<b>6-20</b>	5.55	5.55
<b>21-50</b>	0.00	0.00
<b>51-85</b>	0.00	0.00

The group in question is not a random subset of the bookings in Options C and D. They are older, with 57% over the age of 40, compared to only 44% in the larger group. There are about 28% fewer women (down from 17% to 12%). The proportion of Whites increases about 2 percentage points to 22% and Blacks by about 10 percentage points to 34%, the proportion of Hispanics declining an equivalent amount to 41%. This group is counterintuitively less likely to have two or more bookings, which would provide more opportunity for accumulating charges with municipal regulation violations, incomplete data, or data entry problems. They do start with many more unrecovered charges than others in Option C by over 40 percentage points (76% to 32%). Similarly, they are more likely to have at least one missing

level of charge (34% versus 28%). After recovery efforts, they match the whole of the Option D sample on unrecovered charges (~10%) but still fall about 4 percentage points behind on missing level of charge (31% versus 27%).

*Table F.1.2 Supplementary Analysis: Data Completeness*

Values	Option C n=703	Option D n=703
<b>Unrecovered Charges</b>		
<b>0</b>	24.47	90.61
<b>1-2</b>	26.60	7.25
<b>3-4</b>	15.65	1.14
<b>5-110</b>	33.29	1.00
<b>Missing Level Of Charge</b>		
<b>0</b>	66.00	68.85
<b>1-2</b>	18.92	17.50
<b>3-4</b>	6.54	6.12
<b>5-100</b>	8.53	7.54

Moving on to risk factors, we now identify 41% with a violent offense in the booking charges, 8 times as many as before. None of these were under 20 years old at arrest. Pending cases at arrest are not affected by recovery so the figures in both columns match, and are essentially the same as for the larger group. Prior incarcerations are also unaffected; but the analysis subgroup group has 85% with this risk factor compared to 72% of the larger group. Before charges recovery, they also start with more past FTAs in the past two years (39% versus 33%) but slightly fewer older than two years (59% versus 62%). Prior misdemeanor convictions are slightly higher than the larger group (92% versus 88%) and the match is essentially unchanged by charge recovery. This group also had slightly more felony convictions than the larger group before recovery (58% versus 54%); charge recovery finds a prior felony for a few more bookings, bringing the figure up to 65%. Looking at prior convictions for either a felony or misdemeanor shows a similar but less dramatic pattern; the supplemental analysis group starts out with a few more priors than the larger group (95% versus 92%); recovery brings it up to 97%. The effect on prior violent offense convictions is bigger. Before recovery, 43% the selected group had a prior compared to 36% of the larger group. After recover, 80% of the selected group has a prior violent conviction as a risk factor.

*Table F.1.3 Supplementary Analysis: Record*

Values	Option C n=703	Option D n=703
<b>Violent Current Offense</b>		
No	94.59	58.75
Yes	5.41	41.25
<b>Violent Current Offense @&lt;=20</b>		
No	100.00	100.00
Yes	0.00	0.00
<b>Pending Case At Arrest</b>		
No	96.73	96.73
Yes	3.27	3.27
<b>FTA Within Past 2 Years</b>		
0	71.12	71.12
1	11.52	11.52
2+	17.35	17.35
<b>FTA Older Than 2 Years</b>		
No	40.54	40.54
Yes	59.46	59.46
<b>Prior Misdemeanor Conviction</b>		
No	7.54	6.97
Yes	92.46	93.03
<b>Prior Felony Conviction</b>		
No	41.82	35.28
Yes	58.18	64.72
<b>Prior Conviction for Misdemeanor Or Felony</b>		
No	5.26	2.84
Yes	94.74	97.16
<b>Prior Violent Conviction</b>		
0	57.33	20.06
1	21.62	16.50
2	10.53	10.95
3+	10.53	52.49

*Table F.1.3 Supplementary Analysis: continued*

Values	Option C n=703	Option D n=703
<b>Prior Incarceration (If Consecutive Sentences)</b>		
No	21.76	21.76
Yes	78.24	78.24

Looking at PSA outcome measures, the supplementary analysis subgroup and larger groups are similar on the percent with new FTAs (55% versus 58%), which is unaffected by charges recovery. The supplementary group starts with as similar rate of new arrests and convictions (51% versus 53%) and recovery doesn't add to the proportion with a new arrest. On new violent arrests and convictions, the supplementary group starts with a nearly 50% higher proportion (15% versus 10%); recovery only adds one percentage point, about a 6% increase.

*Table F.1.4 Supplementary Analysis: Outcomes*

Values	Option C n=703	Option D n=703
<b>New FTA Incarceration (If Consecutive Sentences)</b>		
No	44.67	44.67
Yes	55.33	55.33
<b>New Arrest or Conviction (Sentence Date Imputed)</b>		
No	48.65	48.65
Yes	51.35	51.35
<b>New Violent Arrest or Conviction (Sentence Date)</b>		
No	85.49	84.64
Yes	14.51	15.36

Given the close match of outcome measures before and after charges recovery, there is not much to see in the outcomes of Table F.2. There is a modest shift of the supplementary group distribution across the release types. Bail and Own Recognizance contribute about the same proportion to the distribution as in the larger group, but the proportions coming in with Cite and Release has declined from 56% to 50% with a corresponding increase of the Bonded Out group from 21% to 26% of the total.

*Table F.2 Supplementary Analysis: Outcomes by Release Type*

Release Type	Outcome by Release Type					Total
	Both	Option C		Option D		
	FTA	NCA	NVCA	NCA	NVCA	
<b>Bail</b>	4	4	4	4	4	11
	36.36	36.36	36.36	36.36	36.36	1.56
<b>Bond</b>	63	73	27	73	28	180
	35.00	40.56	15.00	40.56	15.56	25.60
<b>Cite</b>	229	207	60	207	64	351
	65.24	58.97	17.09	58.97	18.23	49.93
<b>OR</b>	93	77	11	77	12	161
	57.76	47.83	6.83	47.83	7.45	22.90
<b>Total</b>	<b>389</b>	<b>361</b>	<b>102</b>	<b>361</b>	<b>108</b>	<b>703</b>
	<b>55.33</b>	<b>51.35</b>	<b>14.51</b>	<b>51.35</b>	<b>15.36</b>	<b>100.00</b>

Table F.3 shows that the supplementary subgroup starts with somewhat lower FTA risk scores than the larger group with 71% scoring one to three points versus 66%. Charge recovery, adding a few convictions to the risk score calculation, shifts two percentage points (16 cases) up from a score of one point to two points and moves another case from a score of three to four. This doesn't improve the performance; the range of rates of new FTAs is across risk scores is reduced slightly from 59 to 55. As a percentage of the baseline rate, the range before recovery is from 39% to 145%; after recovery it is from 45% to 145%.

*Table F.3 Supplementary Analysis: Failure to Appear by Risk Score*

FTA Risk Score	Failure to Appear after Release					
	Option C			Option D		
	No	Yes	Total	No	Yes	Total
<b>1</b>	22	6	28	9	3	12
	78.57	21.43	3.98	75.00	25.00	1.71
<b>2</b>	129	75	204	142	78	220
	63.24	36.76	29.02	64.55	35.45	31.29
<b>3</b>	124	140	264	123	140	263
	46.97	53.03	37.55	46.77	53.23	37.41
<b>4</b>	19	64	83	20	64	84
	22.89	77.11	11.81	23.81	76.19	11.95
<b>5</b>	19	100	119	19	100	119
	15.97	84.03	16.93	15.97	84.03	16.93
<b>6</b>	1	4	5	1	4	5
	20.00	80.00	0.71	20.00	80.00	0.71
<b>Total</b>	<b>314</b>	<b>389</b>	<b>703</b>	<b>314</b>	<b>389</b>	<b>703</b>
	<b>44.67</b>	<b>55.33</b>	<b>100.00</b>	<b>44.67</b>	<b>55.33</b>	<b>100.00</b>

The NCA risk scores (Table F.4) of the supplementary subgroup begin with a fairly close split among the top and bottom half of the range with 53% in the bottom half compared to 55% for all of Option C bookings. However, they're more concentrated in the middle third of the range by 75% to 63%. Charges recovery adds enough convictions on misdemeanors and felonies to push 68% into the top half of the range. Before recovery, the range between the lowest and highest NCA rates was 33 percentage points, down from 51 for Option C as a whole. After recovery, the range is up to 48 percentage points and the variation around the baseline rate has increased from a range of 65% to 130% before recovery to 52% to 146% after. The average increase in rates per step in risk scores is up from 7 percentage points to 10. The inversion of expected increases in the rates between scores one and two before recovery goes away after, and the inversion from a score of 5 to 6 is reduced from a severe -14% to -5%, close to the -4% found for the entire set of Option C cases.

*Table F.4 Supplementary Analysis: New Criminal Activity by Risk Score*

NCA Risk Score	New Arrest or Conviction after Release					
	Option C			Option D		
	No	Yes	Total	No	Yes	Total
<b>1</b>	20 66.67	10 33.33	30 4.27	11 73.33	4 26.67	15 2.13
<b>2</b>	68 71.58	27 28.42	95 13.51	50 71.43	20 28.57	70 9.96
<b>3</b>	149 59.60	101 40.40	250 35.56	86 63.70	49 36.30	135 19.20
<b>4</b>	94 34.06	182 65.94	276 39.26	165 48.82	173 51.18	338 48.08
<b>5</b>	9 19.57	37 80.43	46 6.54	27 20.30	106 79.70	133 18.92
<b>6</b>	2 33.33	4 66.67	6 0.85	3 25.00	9 75.00	12 1.71
<b>Total</b>	<b>342</b> <b>48.65</b>	<b>361</b> <b>51.35</b>	<b>703</b> <b>100.00</b>	<b>342</b> <b>48.65</b>	<b>361</b> <b>51.35</b>	<b>703</b> <b>100.00</b>

The first thing noticed in reviewing risk score versus outcome rates for NVCA (Table F.5) is that the subgroup starts with no one scoring a 5 or 6. This is understandable given that we are working with a small 0.4% subset of the larger group and that less than 2% start with a NVCA risk score above 4 in Option C. The 97% with risk scores between one and three is very close to the larger group's 96%. After charges recovery, only 75% have risk scores between one and three and 25% have risk scores of 4 or more. Charges recovery has added a few NCVA outcomes. The range between NVCA rates at the lowest and top scores has increased 3 times from 9 percentage points to 30, also an improvement over the 23 percentage point range of Option C. The variation around the baseline rate has increased from a range of 84% to 145% before recovery, achieving 0% to 192% after. The average increase in rates per step in risk scores is up from 2 percentage points to 6.



*Table F.5 Supplementary Analysis: New Violent Criminal Activity by Risk Score*

NVCA Risk Score	New Violent Arrest or Conviction after Release					
	Option C			Option D		
	No	Yes	Total	No	Yes	Total
<b>1</b>	326 87.87	45 12.13	371 52.77	16 100.00	0 0.00	16 2.28
<b>2</b>	186 86.92	28 13.08	214 30.44	132 91.67	12 8.33	144 20.48
<b>3</b>	74 74.75	25 25.25	99 14.08	315 86.54	49 13.46	364 51.78
<b>4</b>	15 78.95	4 21.05	19 2.70	58 78.38	16 21.62	74 10.53
<b>5</b>	0 0.00	0 0.00	0 0.00	74 70.48	31 29.52	105 14.94
<b>6</b>	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00
<b>Total</b>	<b>601</b> <b>85.49</b>	<b>102</b> <b>14.51</b>	<b>703</b> <b>100.00</b>	<b>595</b> <b>84.64</b>	<b>108</b> <b>15.36</b>	<b>703</b> <b>100.00</b>

Tables F.6, F.7, and F.8 show the odds ratios and AUCs for the three risk score predictions of outcome. Given the similarity of the FTA outcome by risk score tables before and after charges recovery, the close match on odds ratios and AUCs is expected. For NCA, the outcome rate is also the same and the changes in the distribution of risk scores and rates by risk score again aren't sufficient to make much difference on odds ratios and AUCs. Note that with the smaller number of cases in the analysis, the confidence intervals are much wider than the differences in odds ratios or AUCs – despite small differences, they are effectively the same.<sup>24</sup> For NVCA, the differences are substantial, and for the AUCs, statistically significant. With an incomplete record of prior violent convictions, the predictive ability of the PSA's NVCA risk score is weak. With charges recovery, it achieves moderate predictive ability.

<sup>24</sup> The AUC is very slightly lower for the Option D despite a better range as the distribution leads to more “tied” pairs, where a randomly selected pair of recidivist and non-recidivist have the same scores, which dilutes the ability to discriminate as measured by the AUC.

*Table F.6 Supplementary Analysis: PSA Predictive Ability, Failure to Appear*

<b>Option C</b>			
<b>Sample</b>	<b>No FTA</b>	<b>One or More FTAs</b>	<b>Total</b>
Observations Used	314	389	703
<b>Statistic</b>	<b>Point Estimate</b>	<b>95% Confidence Interval</b>	
		Lower Bound	Upper Bound
Odds Ratio	2.126	1.809	2.498
Area Under the Curve	0.706	0.670	0.742

<b>Option D</b>			
<b>Sample</b>	<b>No FTA</b>	<b>One or More FTAs</b>	<b>Total</b>
Observations Used	314	389	703
<b>Statistic</b>	<b>Point Estimate</b>	<b>95% Confidence Interval</b>	
		Lower Bound	Upper Bound
Odds Ratio	2.145	1.818	2.530
Area Under the Curve	0.702	0.666	0.739

*Table F.7 Supplementary Analysis: PSA Predictive Ability, New Criminal Activity*

<b>Option C</b>			
<b>Sample</b>	<b>No NCA</b>	<b>One or More NCAs</b>	<b>Total</b>
Observations Used	342	361	703
<b>Statistic</b>	<b>Point Estimate</b>	<b>95% Confidence Interval</b>	
		Lower Bound	Upper Bound
Odds Ratio	2.000	1.674	2.388
Area Under the Curve	0.676	0.639	0.713

<b>Option D</b>			
<b>Sample</b>	<b>No NCA</b>	<b>One or More NCAs</b>	<b>Total</b>
Observations Used	342	361	703
<b>Statistic</b>	<b>Point Estimate</b>	<b>95% Confidence Interval</b>	
		Lower Bound	Upper Bound
Odds Ratio	1.982	1.665	2.359
Area Under the Curve	0.673	0.637	0.710

*Table F.8 Supplementary Analysis: PSA Predictive Ability, New Violent Criminal Activity*

Option C			
Sample	No NVCA	One or More NVCA's	Total
Observations Used	601	102	703
Statistic	Point Estimate	95% Confidence Interval	
		Lower Bound	Upper Bound
Odds Ratio	1.417	1.115	1.801
Area Under the Curve	0.574	0.515	0.623

Option D			
Sample	No NVCA	One or More NVCA's	Total
Observations Used	595	108	703
Statistic	Point Estimate	95% Confidence Interval	
		Lower Bound	Upper Bound
Odds Ratio	1.701	1.388	2.085
Area Under the Curve	0.640	0.587	0.693

Those affected enough by charge recovery are a bit different from the non-affected population – more males, more Blacks and Whites, older, with fewer bookings. Their records show prior incarcerations and FTAs within the past two years, and more unrecovered charges with the charge type recorded as “NULL”. Charges recovery in particular increases the number with prior violent charges and other felonies, increases the percentage with a non-excluded violent current offense by 8 times. It has little effect on FTA risk scores and outcomes. For NCA and NVCA, it increases the average risk score and the range of recidivism rates between low and high risk groups, rather dramatically for NVCA. AUCs for FTA and NVA are effectively the same. However, it changes the NVCA risk score from a weak to a moderate predictor. In summary, charges recovery only changes risk predictions for a very small group, principally by adding violent offenses and other felonies to their records, and improving the prediction of risk for future violent criminal activity. Outside of this group, the main effect is to increase the number of cases excluded by a few dozen. The main utility of charges recovery might be as insurance against release of someone later found to have a serious offense against public safety among the unrecovered charges, assuming it is not found otherwise in the vetting process.

## Appendix G

### Calibration

Calibration is frequently ignored in assessing risk instruments. Compared with some domains where it quantifies how well a measurement instrument corresponds to a gold standard, it is somewhat more amorphous in criminology where there is no gold standard. It could be argued that desirable properties are that the middle of the scale predicts the baseline rate of the outcome and each step has the same effect, corresponding to the same change in the rate of the outcome or dividing the population into equally-sized groups. It is rare that our constructed scales can achieve such fine discrimination throughout their range. Creating a scale involves a certain amount of craft on top of the science with consideration to clarity for the end users and how it is used in practice.

The PSA initially produces point scores with a range of 0 to 7 (FTA and NVCA) or 0 to 13 (NVA). The APPR provides instructions on how to collapse these scores into the 6 categories of each of the final scales and the two categories of the Violence Flag.<sup>25</sup> In reviewing the outcome by risk scores results, we have noted that these final scores produce skewed score distributions, with many ending up with low or middle scores and extremely few high scores. We have also noted that often there is little increase in the rate of the outcome between score of 5 and 6 and how for some scales, the failure rate actually falls as the risk score increases. We investigated whether these issues could be the result of excluding bookings based on violent offenses from the test samples, and found that it couldn't account for the noted weaknesses of risk scale calibration. Therefore, we experimented with other ways of collapsing the raw point totals into 6 point risk scales. The results are in Table G.1.

Across the different samples in the Options, there is a common pattern showing that it is possible to increase the number of cases in the higher end of the scales and assure that a higher risk score corresponds to a higher failure rate. This is not without some cost. For FTAs, in order to push more individuals into scores of 5 and 6, the proportion of cases scoring 4 has been severely reduced to about four percent of the total. For FTA and NVCA, in order to achieve better discrimination at the high end of the risk scale, we would sacrifice discrimination at the low end of the scale, with differences of only a few points between scores of one and two. Perhaps in use, there is little need to distinguish between cases with risk score of 5 and 6 and judges will pay more attention to differences in low risk groups, in which case rescaling would be counter-productive. Whether these tradeoffs are desirable or not is a policy decision.

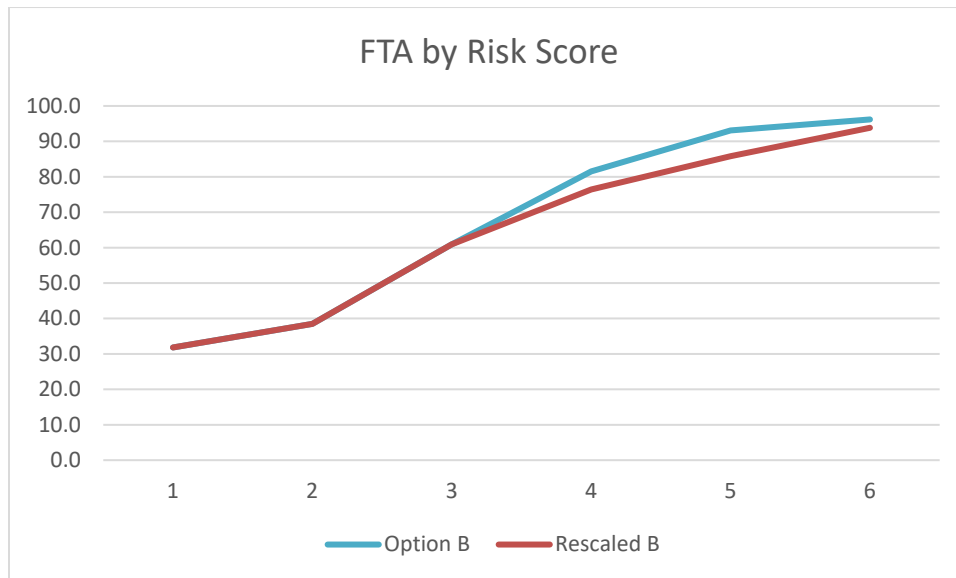
---

<sup>25</sup> Advancing Pretrial Policy and Research, nd. 14E\_PSA\_PointsandScales\_2020April.pdf. <https://advancingpretrial.org/implementation/guides/> (Site registration required.)

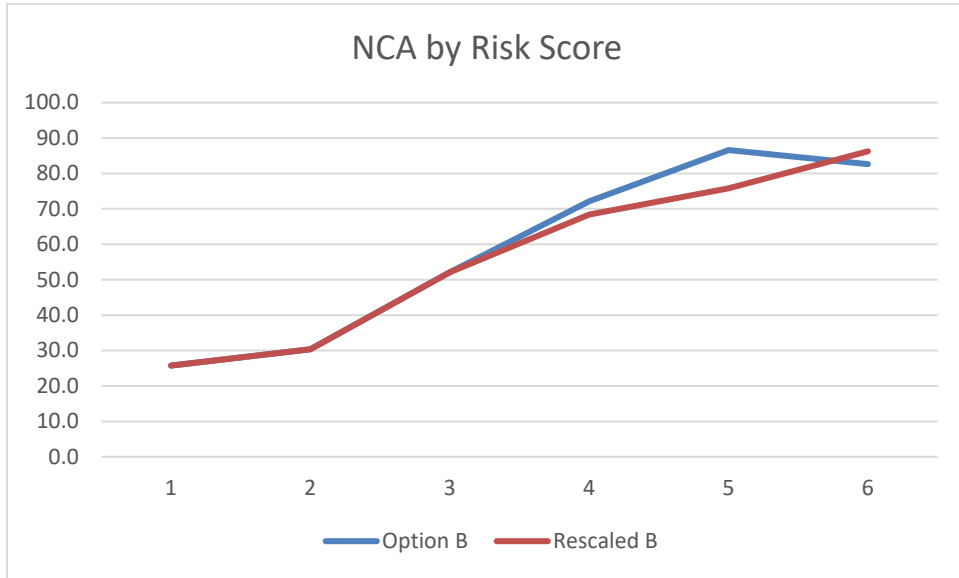
*Table G.1: Risk Scales Before and After Rescaling Raw Scores*

Outcomes by Standard and Rescaled Risk Scores												
Score	Option A				Option B				Option C			
	Distribution		Failure Rate		Distribution		Failure Rate		Distribution		Failure Rate	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
<b>FTA</b>												
1	11.3	11.3	30.3	30.3	6.9	6.9	31.8	31.8	5.9	5.9	34.1	34.1
2	32.1	32.1	34.9	34.9	26.6	26.6	38.5	38.5	25.2	25.2	36.1	36.1
3	31.7	31.7	57.2	57.2	34.9	34.9	60.9	60.9	34.8	34.8	54.5	54.5
4	11.5	4.3	79.5	75.5	13.2	4.3	81.5	76.4	14.7	4.8	74.6	69.5
5	13.0	10.5	92.0	84.1	17.7	12.3	93.1	85.8	18.6	13.6	88.4	79.1
6	0.5	10.2	94.0	93.1	0.8	15.0	96.2	93.8	0.8	15.6	93.3	89.6
<b>NCA</b>												
1	11.9	11.9	23.6	23.6	7.5	7.5	25.8	25.8	6.6	6.6	26.0	26.0
2	22.7	22.7	27.1	27.1	17.8	17.8	30.4	30.4	17.7	17.7	29.0	29.0
3	31.0	31.0	48.8	48.8	30.1	30.1	52.1	52.1	30.3	30.3	47.0	47.0
4	26.1	13.5	69.7	65.8	32.1	15.8	72.2	68.4	32.7	16.2	66.6	62.8
5	7.6	12.5	83.9	74.0	11.4	16.4	86.6	75.8	11.7	16.5	81.7	70.3
6	0.7	8.3	81.1	83.6	1.1	12.5	82.6	86.3	1.1	12.8	77.4	81.3
<b>NVCA</b>												
1	65.7	12.1	5.1	4.3	58.6	7.6	6.3	5.3	57.6	7.0	6.2	5.5
2	13.9	53.5	12.2	5.3	16.6	51.1	12.7	6.4	16.7	50.6	12.2	6.3
3	17.9	13.9	17.7	12.2	21.8	16.6	18.5	12.7	22.1	16.7	17.4	12.2
4	1.4	17.9	24.1	17.7	1.6	21.8	23.2	18.5	1.9	22.1	21.1	17.4
5	1.2	1.4	36.5	24.1	1.3	1.6	37.0	23.2	1.7	1.9	33.1	21.1
6	0.0	1.2	36.4	36.5	0.0	1.3	38.5	37.0	0.0	1.8	29.1	33.0

*Figure G.1. Failure to Appear by Risk Score Rescaled*



**Figure G.2. New Criminal Activity by Risk Score Rescaled**



**Figure G.3. New Violent Criminal Activity by Risk Score Rescaled**

