Document Title: National Institute of Justice's Forecasting Recidivism Challenge: Team "DEAP" (Final Report)

Author(s): David B. Wilson, Evan M. Lowder, Peter Phalen, Ashley Rodriguez

Document Number: 305037

Date Received: July 2022

Award Number: NIJ Recidivism Forecasting Challenge Winning Paper

National Institute of Justice's Forecasting Recidivism Challenge: Team "DEAP" (Final Report)

David B. Wilson[1], Evan M. Lowder[1], Peter Phalen[2], and Ashley Rodriguez[1]

[1]George Mason University

[2]University of Maryland School of Medicine

Author Note

Submitted to NIJ on August 31, 2021. Contact email: elowder@gmu.edu

National Institute of Justice’s Forecasting Recidivism Challenge: Team “DEAP” (Final Report)

Introduction

The science and practice of predicting recidivism risk is a growing enterprise. There are now decades of research identifying risk and protective factors for criminal recidivism (Dowden & Brown, 2002; Gendreau et al., 1996; Scott & Brown, 2018; Yukhnenko et al., 2020). These risk and protective factors have been used to develop risk assessment tools to predict recidivism at various stages of correctional system involvement and aid correctional practices in release and supervision decision-making. There are now at least 19 different risk assessment systems that have been developed, implemented, and validated across correctional settings, including an additional 47 assessments developed for local use (Desmarais & Singh, 2013). In addition, half of local jurisdictions have implemented risk assessments at the pretrial stage (Lattimore et al., 2020), 18 states have adopted risk assessment tools for statewide sentencing decisions (Casey et al., 2017), and the majority of states have adopted risk assessments tools more broadly for some form of statewide use (Wachter, 2015).

Criminal justice practitioners and policymakers are invested in accurately predicting whether an individual involved in the justice system will recidivate to the system. The criminal-legal system is expensive, incurring an estimated $182 billion annual costs to society (Wagner & Rabuy, 2017). Further, incarceration harms families, health outcomes, and socioeconomic stability upon release (Schnittker & John, 2007; Turanovic et al., 2012; Western et al., 2001). Finally, recidivism contributes to high incarceration rates in the United States, surpassing those of other developed nations (World Prison Brief, 2018). Risk assessments can help reduce recidivism by triaging finite resources to the highest risk individuals, informing interventions to address individuals’ criminogenic needs, and delivering interventions in a way that is responsive to an individual’s background (Andrews, Bonta, et al., 1990; Andrews, Zinger, et al., 1990; Andrews & Bonta, 2010).

At the same time, there has been a growing chorus of critics arguing that the risk prediction enterprise may exacerbate structural inequalities (Holder, 2014; Starr, 2014). These critics argue that by measuring factors that reflect systemic disadvantage for racial minorities, particularly Black individuals, these tools will produce less accurate assessments of recidivism risk for these individuals. Disproportionately higher risk scores then increase the likelihood that criminal justice systems will impose more restrictive conditions on these individuals, furthering the harmful effects of system involvement (Harcourt, 2015; Minow et al., 2019; Pretrial Justice Institute, 2020). These concerns have prompted scholarly attention on issues of algorithmic fairness in risk assessment. Many complementary and competing definitions of fairness have been proposed (Berk et al., 2017; Kleinberg et al., 2016; Skeem & Lowenkamp, 2016). However, there is a unanimous consensus that when rates of misconduct differ across groups, achieving both accuracy and fairness is a zero-sum endeavor (Corbett-Davies et al., 2017; Kleinberg et al., 2016). There is no single definition of fairness and balancing one will come at the expense of another. Thus, a focus of the risk assessment enterprise has become the goal of achieving a balance between accuracy and fairness that is acceptable to local stakeholders (Berk et al., 2017).

In consideration of these issues, the National Institute of Justice (NIJ) developed the *Forecasting Recidivism Challenge* to accelerate technical and substantive knowledge on predicting recidivism risk. As part of this competition, NIJ released data on Georgia parolees in three stages and challenged researchers to predict recidivism during the 1-year, 2-year, and 3-year periods following release from custody. The competition was judged on overall accuracy (Brier Score) and a combined measure of accuracy and fairness as operationalized by the difference in false positive rates between White and Black parolees. This report documents our efforts as a "Small Team" participating in this challenge. We note that we are not machine learning scholars; rather, we are a group of applied statisticians and researchers with experience in predictive analysis and risk assessment. As a result, our approach and results reflect in part our learning process during this competition.

Study Purpose

The purpose of this study was to predict recidivism in a sample of parolees from Georgia following release from prison custody. The goal of our modeling strategy was to maximize predictive accuracy while minimizing differences in false-positive rates between Black and White individuals, consistent with contest guidelines for the National Institute of Justice’s *Forecasting Recidivism Challenge*. We generated predicted probabilities for three discrete time periods (Year 1, Year 2, and Year 3 following release) for parolees without misconduct in the preceding year. Below we discuss our model building process and results in further detail.

Method

Data Sources

The primary data for our models, of course, was the person-level data for Georgia State parolees provided by NIJ to contestants for this competition. We added to this PUMA-level census data, as discussed in the “PUMA Variables” section below. We also created several person-level composite variables, also discussed below.

PUMA Variables

We reviewed nearly 500 Public Use Microdata Areas (PUMA) data tables on the Census website (US Census Bureau, 2021). PUMA’s are non-overlapping, statistical geographic areas that divide a state into areas with at least 100,000 persons in each. We identified 64 tables that represented constructs related to socioeconomic status that we believed might be relevant to the prediction of recidivism. These constructs and associated tables are listed in Appendix B, Table B1.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

We merged these data tables into a single file and created variables for analysis. The 64 tables included 509 separate variables, the majority of which reflected counts within specific sub-categories. We reviewed the 509 variables and identified a smaller sample of unique categories that could be combined with sufficient counts to produce meaningful rates for prediction modeling. As necessary, we merged categories to reduce the number of variables calculated and increase the rates. Where average values were provided, we used these variables directly without manipulation. Count values were transformed into rates using the total unit count provided within each table and PUMA code. The final set of variables, definitions, and variable types are provided in Appendix B, Table B2. 111 PUMA variables were available to our model generating process.

We initially explored other geographic-level variables for inclusion in our modeling. However, the provided dataset was limited in that the only geographic identifier was the PUMA code. Other geographic data sources that may have been useful for modeling include crime data collected at the county level (Georgia Data, 2021; Georgia's Uniform Crime Reporting Program, 2013). However, counties were not neatly nested within PUMA identifiers, limiting the utility of these additional data sources.

Data Manipulations

A handful of variables in the dataset were ordinal. The top value for these variables reflected a range of possible values, such as "3 or more" or "10 or more". These were all converted into numeric scales, with the top value being the next in the series (i.e., the lowest value in the range). Thus, "3 or more" was coded as 3, and "10 or more" was coded as 10. The variables manipulated in this way were:

Dependents,
Prior_Arrest_Episodes_Felony,
Prior_Arrest_Episodes_Misd,
Prior_Arrest_Episodes_Violent,
Prior_Arrest_Episodes_Property,
Prior_Arrest_Episodes_Drug,
Prior_Arrest_Episodes_PPViolation,
Prior_Conviction_Episodes_Felony,
Prior_Conviction_Episodes_Misd,
Prior_Conviction_Episodes_Prop, and
Prior_Conviction_Episodes_Drug.

We could have converted several additional variables to numerical ordinal scales, but we left them as nominal variables and dummy coded them according. These variables were Age_at_Release, Education_Level, and Prison_Years. Our leaving these as nominal factors and dummy coding them for modeling allowed for complex nonlinear relationships to emerge, if present.

Composite Variables

We created several composite variables at the level of the parolee. The literature on risk-prediction and recidivism informed our construction of these composites. The goal was to assist the machine learning algorithm in locating potentially useful interactions among variables. We briefly discuss each of these composites below.

Female_Dependents: The effect of having dependents on recidivism may depend on whether the parolee is male or female. Some research suggests women may be less likely to recidivate in the community if they have children (Barrick et al., 2014; Benda, 2005; Harm & Phillips, 2001), including in parole contexts (Huebner et al., 2010). Thus, we created a composite to indicate if a parolee had dependents and was female.

MHSA_Priors: We hypothesized that the prior criminal history for those parolees conditionally released to mental health or substance abuse programming might relate to recidivism differentially relative to those not so conditionally released. This decision was based on theoretical and empirical scholarship suggesting that the presence of both criminogenic needs (i.e., behavioral health problems) and criminogenic risk (i.e., prior criminal justice involvement) may amplify to affect general recidivism outcomes (Andrews & Bonta, 2010b; Walters, 2015). Thus, we created an interaction term for `Condition_MH_SA` and the sum of `Prior_Arrest_Episodes_Felony` and `Prior_Arrest_Episodes_Misd`. This variable was 0 for all parolees with no conditional release to mental health or substance abuse programming and the sum of prior arrests for those with such a conditional release.

`CogED_Priors`: As with mental health and substance abuse, we hypothesized that prior criminal history for those parolees conditionally released to cognitive skills or educational programming might be related to recidivism differentially relative to those not so conditionally released. Thus, we created an interaction term for `Condition_Cog_Ed` and the sum of `Prior_Arrest_Episodes_Felony` and `Prior_Arrest_Episodes_Misd`. This variable was 0 for all parolees with no conditional release to cognitive skills or educational programming and the sum of prior arrests for those with such a conditional release.

`Male_PriorViolent`: Based on evidence that prior history of violence may be a risk factor for men in particular (Benda, 2005; Collins, 2010), we created a composite variable to reflect whether a parolee was both male and had a prior arrest episode for a violent offense.

`MHSA_Drug`: Previous research suggests individuals with a prior history of drug use or drug-related arrests are less likely to complete drug treatment successfully (Hohman et al., 2000) and may be at higher risk for recidivism. We created a composite of the interaction of a conditional release to mental health or substance abuse programming and the number of prior drug arrests. This variable was 0 for all parolees with no conditional release to mental health or substance abuse programming and the number of prior drug arrests for those with such a conditional release.

`MHSA_PrisonYears`: Because length-of-detention has been shown to affect behavioral health symptoms in individuals with a history of incarceration (Porter & DeMarco, 2019), we created a composite to capture any interaction between a conditional release to a mental health or substance abuse program and whether prison years were less than or greater than one year.

`DrugsEmployedMoves`: We created an indicator variable that was true if a parolee had more than 10% (a proportion of .1) of their drug tests positive for THC and the percent days they were employed was less than 50% (a proportion of .5), and they moved more than once. This variable was only available for the Year 2 and Year 3 predictions.

`DelinqGangs`: Hypothesizing that parolees who had a history of gang affiliation and were delinquent on parole would have higher rates of recidivism, we created an indicator variable that was true if a parolee had more than 2 delinquency reports and was gang-affiliated. This variable was only available for the Year 2 and Year 3 predictions.

`DrugsGangs`: We created an indicator variable that was true if a parolee was gang affiliated and they have more than 10% of their drug tests positive for THC. This variable was only available for the Year 2 and Year 3 predictions.

`MHSADrug`: We created a composite interaction term for whether a parolee was conditionally released to a mental health or substance abuse program and the number of positive drug tests for THC. This variable was 0 for those not conditionally released to a mental health or substance abuse program and was the proportion of positive THC tests otherwise. This variable was only available for the Year 2 and Year 3 predictions.

`Violations`: Hypothesizing that individuals with a greater number of parole violations would be more likely to recidivate, we created a composite of dummy variables reflecting different types of parole violations. This variable could take on the value of 0 (no violations) through 4 (true for all four violation types). The relevant variables were: `Violations_ElectronicMonitoring`, `Violations_FailToReport`, `Violations_Instruction`, and `Violations_MoveWithoutPermission`. This variable was only available for the Year 2 and Year 3 predictions.

Missingness

The following variables had missing data: Supervision_Risk_Score_First (2%), Gang_Affilited (12%), Supervision_Level_First (6.7%), and Prison_Offense (13%). To address missingness for the Supervision_Risk_Score_First variable, we imputed the median value of 6 for missing cases and created an additional variable indicating whether this variable was missing or not (a 0/1 dummy variable), called Supervision_Risk_Score_First_isNA. For the remaining variables, we created an additional category called “Unknown”. This unknown category was incorporated into the dummy coding for each of these variables.

Normalization and Dummy Coding

For both the xgboosted decision tree models and the neural networks models, all nominal variables (factor variables in R) were dummy coded into a set of numeric 0/1 indicator variables. For neural network models, variables must be normalized so that they share a common scale. For these models, we followed advice from a post-mortem written by a “Kaggle” winner named Michael Jahrer and used a Gaussian rank-transformation plus mean-centering to normalize all variables introduced into the model, which appeared to perform better than z-scoring. This step involved our writing a custom function because we could not find a package-based implementation.

Prior Probabilities

The models for Years 2 and 3 included our prior Year model probabilities. Models including these prior probabilities outperformed models without these probabilities.

Modeling and workflow

Our initial modeling method was to use gradient boosted decision trees estimated using the XGBoost algorithm as implemented in the “xgboost” R package (https://CRAN.R-project.org/package=xgboost). As a team, we were new to using this machine learning method and as such our approach evolved as we progressed from Year 1 to Year 2 to Year 3 predictions. We will discuss each prediction model separately.

Note that we never calculated statistical significance at any point in our workflow. Null hypothesis testing was not relevant to the purpose of the contest. Models were always provided with all available variables described in the "Variables" section above. Statistical algorithms assign less weight to variables with less predictive power; therefore, some variables had little or no influence on predictions despite being included in the models. In addition, some variables were not selected by the xgboost algorithm to produce the final predicted probabilities. We provide information below on the variables that contributed most to the models.

Year 1 Prediction Model

The model inputs for the Year 1 predictions included all available person-level variables, the PUMA-level census variables, and a set of fixed-effect dummy-codes for each PUMA area. Several parameters control the estimation of a gradient boosted trees model. Two of these parameters are related to the structure of the model: the maximum depth of any tree and the number of trees. Of the other possible parameters that can be set, we tuned the eta and lambda parameters. The eta parameter affects the step size in the iterative modeling process and helps prevent overfitting. The lambda parameter sets the L2 regularization term. Larger values produce a more conservative model. In a regression context, L2 regularization moves regression coefficients towards zero.

To determine the optimal parameters for our xgboosted decision trees, we used k-fold validation as implemented by the "mle" R package. This approach divides the training data into k equally sized random subsets. Next, $k-1$ subsets are used to train the model with a given set of model parameters. Each of these models is then validated on the remaining subset. Finally, the performance of each model is averaged to determine the overall performance of a given set of parameters.

We performed 20,000 k-fold cross-validation with $k = 4$. These 20,000 iterations of the k-fold cross-validation assessed the performance of a random search of the following parameter space: (a) number of trees: 2 to 100; (b) depth of trees: 1 to 10; (c) eta: .01 to .50; and (d) lambda: 10^{-1} to $10^{.5}$ or .1 to 3.16. With a random seed of 1, the optimal parameters were: number of trees = 48, depth of trees = 3, eta = .147, lambda = 2.51.

As we increased our knowledge on how to tune xgboosted decision trees, we learned that our random search of the parameter space was ineffective at exploring the full range of possible parameter combinations given the infinitely large number of possible values for eta and lambda (we were searching a continuous range for each). Therefore, in our Year 2 and Year 3 models, we did a grid search of a discrete list of parameter values, increasing the likelihood of identifying an optimal set of parameters. Had we have more time or greater computer power, we could have refined our tuning of these models even further.

We tweaked the predicted Year 1 probabilities to help equate the false positive between White and Black parolees. We did this by subtracting .005 from the probabilities of Black parolees, effectively increasing the threshold for Black parolees to .505 while maintaining the threshold at .50 for White parolees. While advocated by some, this approach is likely to be unconstitutional (see, for example, Huq, 2019; Mayson, 2018). However, our goal was to maximize the *Contest Score* that was part of this competition. Unfortunately, this model failed to rank in the top four in any category of the competition.

Year 2 Prediction Model

We built our Year 2 prediction much the same as with Year 1. We used xgboosted gradient trees and included all variables included in the Year 1 predictions. As mentioned previously, we also included the Year 1 predicted probability for each parolee. The Year 2 data also had several additional new variables available that reflected parolee behavior while on parole. All of these were used, including the composite variables, as discussed above.

As we increased our knowledge on how to tune xgboosted decision trees, we learned that our random search of the parameter space was ineffective at exploring the full range of possible parameter combinations given the infinitely large number of possible values for eta and lambda (we were searching a continuous range of values for each). In our Year 2 and Year 3 xgboost models, we instead performed a grid search of a discrete list of parameter values, increasing the likelihood of identifying an optimal set of parameters. We also changed our tuning of the xgboosted parameters by using 6-fold instead of 4-fold cross-validation when assessing the model's performance under different parameters. The values used for the grid search were: (a) number of trees: 2 to 60; (b) depth of trees: 2 to 4; (c) eta: .1, .15, .2, .25, .3, .35, .4, .45, .5; and (d) lambda: 1.0, 1.1, 1.2, 1.4, 1.6, 1.8, 2.0, 2.3, 2.6, and 2.9.

The probabilities generated from this model had roughly equivalent false positives for Black and White parolees. As such, no tweaking of these probabilities by race was performed. In the competition, this model ranked fourth place for female parolees and fourth place for overall accuracy.

Year 3 Prediction Model

For our Year 3 predictions, we introduced three primary innovations that we believe further improved our performance, contributing to first- and second-place finishes in two Year 3 categories: male parolees and average accuracy. These three procedures are described here:

1. *Model-stacking.* Rather than attempting to select a single "best" model, as we had done for our Year 1 and Year 2 submissions, we used model stacking. Model stacking means taking an average (or weighted average or nonlinear combination) of the predictions of more than one model to produce a single prediction informed by

multiple models. Model stacking can reduce overfitting and often outperforms the individual models that compose it. For our final submission, we used an unweighted average of predictions from two xgboost models (one with our Year 1 and 2 model probabilities included as predictors and one without) and two neural networks (one with Year 1 and 2 probabilities and one without).

2. *Neural networks.* We fit two neural networks to the data, one that included our Year 1 and Year 2 model probabilities as input variables and one that did not. Parameters for the neural network were tuned using five-fold cross-validation, which ultimately selected only a single layer: a simple model structure equivalent to logistic regression under some specifications. An alternative grid search approach selected a more complex 5-layer network, but this model did not outperform the single-layer network when evaluated using our simulation-based approach described below.

3. *Simulation-based assessment of our workflow.* We began assessing our complete workflow using simulations. Specifically, this involved removing a 20% holdout of the training dataset and then running our entire workflow (viz., data preprocessing, parameter tuning of xgboost and neural network models, cross-validation-based selection of competing models, model-stacking, posterior predictive checks, and assessment of competing models using the contest score equation) using the remaining 80% of the training dataset. This allowed us to assess the performance of our workflow using a 20% holdout that did not factor into the workflow in any way and for which we knew the ground truth outcome. This step simulated the NIJ *Forecasting Recidivism Challenge* within which we were competing, allowing for a more realistic assessment of our overall approach to the contest.

We made decisions about our final submission using the simulation-based assessment of our workflow as described in (3). We ultimately selected a model stack consisting of the unweighted average of predictions from two xgboost models (one with

Year 1 and 2 probabilities and one without) and two neural networks (one with Year 1 and 2 probabilities and one without). This unweighted stack outperformed all individual models in the stack as well as the stacks of each neural network and each xgboost model taken separately, which served to finalize our decision to stack all four models in this way.

Results

Feature (Variable) Importance

The top 30 variables in terms of importance for each xgboost model are listed in Appendix A. These are sorted from the most important to the least important within each table.

Table A1 shows that the most important variables to the Year 1 model are gang affiliation, prior arrest record for a property offense, the supervision risk score, and 5 or more probation or parole violations.

For the Year 2 model, by far the most important variable is the predictive probability from the Year 1 model (see Table A2. The following three variables were jobs per year, days employed, and average days on parole between drug tests.

Recall that for Year 3, we stacked four different models, two xgboosted models and two neural net models. We show the contribution of variables to the xgboosted model that excluded prior-year probabilities in Table A3, whereas Table A4 shows the contribution of variables for the xgboosted model that included prior probabilities. The most important variables for the former were prior arrest episodes for probation and parole violations, percent days employed, average days on parole between drug tests, and jobs per year. For the latter, the predicted probabilities for Year 1 and Year 2 were the two most important variables, followed by jobs per year and average days on parole between drug tests. Unfortunately, neural-net models do not generate such straightforward output that allows you to assess which variables contribute most to the model. However, manual graphical

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

posterior predictive checks can serve a similar purpose, and we used this technique to interrogate these models during the fitting and selection process.

Model Predictions

Table 1 shows the model performance statistics (accuracy, sensitivity, specificity, false-positive rate, and Brier Score) for both the training and testing (i.e., holdout) data across each of the years by gender. While these models have acceptable accuracy and very good specificity, they have low sensitivity. The low sensitivity is in part due to the low base rate of recidivism across the dataset. The male and female *Contest Score* for each dataset (training and testing) and year are shown in Table 2.

Table 3 shows the various confusion tables (2 by 2 contingency tables between observed and predicted recidivism, using a .5 threshold for the latter). Note that with a .5 threshold, no parolees were predicted to recidivate during Year 3 for either the training or testing datasets. This absence of positive predicted cases reflects that the predicted probabilities were all below the .5 threshold. Also, note the large number of false negatives across these tables, driving the low sensitivity values.

Table 4 shows the mean, minimum, and maximum predicted recidivism probabilities for individuals in the dataset. The average predicted probability of recidivism ranged from a high of .30 for Year 1 for both testing and training to a low of .20 and .19 for Year 3 for testing and training, respectively. The maximum values for Years 1 and 2 were between .72 to .83, whereas Year 3 had much lower maximum probabilities. These were .46 for testing and .49 for training. Note that for Year 3, no individual is predicted to recidivate with a threshold of .50, and therefore there were no false positives and no possible racial bias penalty as defined by the *Contest Score*.

Examination of Predicted Probabilities Compared to the Supervision Risk Score

An interesting question is how our decision-tree-based risk predictions performed in comparison with Georgia’s risk assessment score. Existing risk assessment tools do not typically predict a binary outcome based on a .5 threshold; instead, they produce an ordinal scale of risk from low to high. The number of levels and the cut-points for these levels vary across risk assessment tools.

Georgia’s Supervision Risk Score has 10 levels of ostensibly increasing risk. However, as shown in Figure 1, this Risk Score does not monotonically increase with actual rates of recidivism. Recidivism decreases initially from .24 for the lowest Risk Score of 1, to .19 for Scores 2 and 3. At the highest Risk Score, the proportion recidivating is only .40. Thus, actual recidivism risk does not change substantially across the 10 levels of the Supervision Risk Score.

In order to produce a comparable graph for our Year 1 model probabilities, we divided the model-predicted Year 1 probabilities into 10 equally sized groups or deciles. Figure 2 shows the proportion of the training sample that recidivated by the end of Year 1 across these 10 deciles. There is a clear monotonic relationship between these deciles and the actual proportion of parolees recidivating. Furthermore, these deciles outperform Georgia’s Supervision Risk Score. For example, the first decile has a recidivism proportion of .08, whereas the top decile has a recidivism proportion of .54. From a policy perspective, it is interesting to note that our model only predicts an approximately 50/50 chance of recidivism for parolees in the top decile.

Additionally, the correlation between the model-predicted probabilities and the Supervision Risk Score is modest at best. On the training set across the three years, the correlations between Risk Score and our model probabilities were 0.44, 0.35, and 0.39, respectively. On the testing set across the three years, the correlations were 0.45, 0.36, and

0.39, respectively. These correlations suggest there is room for improvement in Georgia's risk tool, assuming they are still using it.

Future Considerations

We want to begin by crediting NIJ for creating this competition which allowed for a much more rigorous treatment of recidivism risk assessment than is typically achieved in the literature. The use of a holdout sample dramatically reduces over-fitting and reinforces accuracy instead of overconfidence, which is a contrast to the incentive structure that exists in academic publishing and even the broader prison-industrial complex. Our initial consideration is that this field and others like it would benefit from the more frequent use of holdout samples to more accurately evaluate the performance of any risk assessment tool proposed for use. Our learning process during this competition informs several suggestions we have for the future development, research, and implementation of risk assessments in correctional settings.

First, we want to note that this contest illustrated some of the difficulties inherent in predicting recidivism. Humans are complex, and the world they live in is even more so. Therefore, it should come as no surprise that none of the participating teams achieved highly accurate predictions about outcomes for individuals in these datasets. The structure of this contest, which used a blinded test dataset and evaluated contestants based on accuracy, ensured that statistical models developed by participating teams would be appropriately humble about their ability to predict future recidivism, which is refreshing and better reflects the true uncertainty inherent in the future lives of individuals involved in the criminal justice system. These risk assessment tools attempt to predict future behavior, and that behavior is likely to be influenced by numerous contextual factors and future life events that are unknown and uncertain. Thus, these models are not likely to ever be highly accurate.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Unfortunately, in academic and practical settings, risk assessment tools are not typically presented with such humility. In many ways, the *Supervision Risk Score* used by the State of Georgia is the worst of all possible worlds, both because it is a poor predictive variable and because the ordinal 1–10 encoding expresses no information about the actual risk entailed by each score. These two issues compound themselves, as the categorical score hides the fact that people with high Supervision Risk Scores are not actually much higher risk in either absolute or relative terms. Among people scoring the lowest possible Supervision Risk Score of 1: 25% were rearrested in Year 1. Among those scoring the highest possible Supervision Risk Score of 10: 41% were rearrested in Year 1. The use of categorical Risk Scores obfuscates the relatively small difference in actual risk associated with such an extreme difference in Risk Scores and hides the fact that people on both extremes of the scale will most likely not recidivate. Figure 2 illustrates the fact that our own model outperforms Supervision Risk Score when used as an ordinal predictor of risk, but even here the process of discretization is undesirable because it hides the actual model output, which takes the form of true probabilities. When communicating risk to decision-makers, we recommend the use of a probability range (e.g., “20 to 30% chance of rearrest in 3 years”) rather than an ordinal category, in order to avoid reinforcing capricious supervision decisions.

A related point that is often lost in risk assessment scholarship is that the body of work on contemporary actuarial risk assessment grew primarily out of the Risk-Need-Responsivity (RNR) model (Andrews, Bonta, et al., 1990; Andrews & Bonta, 2010a). The RNR model recognized that criminal justice agencies have limited resources and are constantly faced with decisions about prioritizing the distribution of those resources. Thus, this model’s foundational “risk principle” argues that agencies should assess for criminogenic risk and devote more resources to their highest-risk population. In this way, this model does not assume a binary level of recidivism risk that would denote a binary decision (i.e., to release or not); instead, it argues that individuals can be placed into risk bins based on their probability of misconduct to direct more resources to those in the higher probability bins. However, particularly concerning definitions of fairness, the use of single thresholds may have limited utility for assessing the potential differential impact of risk assessments in practice.

A specific goal of this competition was to identify geographic-level predictors of recidivism outcomes. Over 100 PUMA variables (see Appendix B2) were used in machine learning models and fewer than 10 emerged as important features. Broadly, these variables could be divided into three categories: dependents in the household (e.g., PUMA 14; PUMA 164); sources of familial income and employment (e.g., PUMA 6; PUMA 95); and housing affordability and stability (e.g., PUMA 19; PUMA 49; PUMA 221). Employment and housing instability are common dynamic risk factors assessed on risk assessment tools (Desmarais & Singh, 2013). Our findings suggest community economic stability is important to individuals' community reintegration while on supervised release. However, we note that these variables were not the strongest predictors of recidivism in our models. Further, as a geographic measure, the PUMA represented larger and often more diverse geographic regions than other geographic-level units commonly used in research, including county, zip code, or census tract. It is possible that smaller geographic-level measures may have more predictive utility.

Including geographic variables into a risk prediction for an individual involved in the criminal justice system is also potentially problematic, depending on how the risk score is used. As an ethical principle, no one should be punished for where they happen to live. As such, incorporating information about the geographic area that an individual will return to on release into a release decision, even if it improves the predictive validity of the risk assessment, strikes us as unjust. In contrast, the inclusion of geographic area information into a risk and *needs* tool to identify the services and supports an individual may need to be successful on probation or parolee (assuming these services and supports are viewed positively by the individual) could be potentially beneficial, ethical, and responsive to individuals' needs. Ensuring the latter and avoiding the former, however, might be difficult in practice.

The contest also elicits reflections on race and discrimination. For example, we noticed a technical issue in the fact that the dataset provided to contestants was approximately evenly divided between White and Black parolees (actual numbers were 43% White and 57% Black), whereas the State of Georgia is only about 33% Black. By providing contestants with a racially balanced dataset, the NIJ artificially minimized the likelihood of algorithmic racial bias because statistical models with strong racial bias in model accuracy will, therefore, ipso facto, tend to perform poorly on the dataset as a whole. For example, if the dataset had been minority Black, then a statistical model could have more easily traded higher false-positive rates (FPR) for Black people in exchange for lower FPR among White people to achieve a lower overall FPR, but this was not the case. Ironically, the racial bias of the justice system in Georgia, which substantially enriched the representation of Black people among the population of parolees, thereby reduced the potential for algorithmic racial bias for this contest. The NIJ could have considered presenting a dataset with racial proportions that approximated the demographics of the State of Georgia or perhaps could have dropped the "Accounting for Racial Bias" category and focused on accuracy instead.

The use of race as a potential predictor in all of the statistical models developed for this competition diverges from most contemporary approaches to risk assessment. However, we note that most contemporary risk assessments include measures of criminal history, which have been argued to serve as "proxies" for race, given discrimination in the criminal justice system (Harcourt, 2015). Race only ranked in the top 30 variables in terms of importance in our Year 2 model and had a low importance weight even then. Furthermore, the limited number of racial groups included in the Georgia data (Black and White) is at odds with reality. There are other racial and ethnic identities within the

United States, and many individuals have multiple such identities. This complexity dramatically complicates the use of race/ethnicity in a risk algorithm.

Legal scholars have argued that the focus on “equality” and not “equity” in criminal-legal processing is unlikely to solve the systemic discrimination that plagues the system—and our society (Huq, 2019). There is a growing conversation in the academic community that disentangling systemic disparities from risk assessments is a near-impossible endeavor (Mayson, 2018; Vincent & Viljoen, 2020). However, advancing an equity framework would mean that criminal-legal systems focus on achieving equitable outcomes rather than an equal process (Huq, 2019). In a risk assessment context, this may be an argument for including race and its proxies in machine learning approaches to adjust for the reality that systemic racism means that risk factors will not be as salient predictors across racial groups. However, whether this would be constitutionally acceptable or acceptable to relevant stakeholders remains uncertain. Additionally, there would need to be a more robust empirical basis for the inclusion of race in risk assessment models, whereas, for this contest, neither race nor its interactions were strong predictors of outcome.

Regardless, as scientists, our objectives are to work within existing constraints to develop fairer and more effective tools. In developing risk assessment tools, large and heterogeneous validation samples are essential. These samples should be representative of the population for which the tool is designed but also diverse enough to allow for the investigation of predictive validity across demographic sub-groups. Further, although cross-validation and pre-registered predictions using holdout samples are essential components of actuarial test development (Schumacher et al., 1997; Wollert, 2002), it is not common in academic risk assessment development. The use of cross-validation and holdout samples allows researchers to assess how the model likely will perform in other settings and to anticipate and attempt to solve problems of racial bias.

An important consideration for any risk assessment tool is its acceptability to the various stakeholders, including those on whom to tool is being used (Berk et al., 2017). As such, the algorithm must not be so complex as to prevent a non-statistician from understanding how a risk score is produced. This principle of acceptability to stakeholders argues against using neural networks because the underlying model is relatively opaque. The highly complex decision trees that we generated for this competition are also problematic in this regard—in our Year 1 model, the decision tree produced by xgboost comprised more than forty parallel three-step decisions. However, if pruned to a small number of trees with limited depth, decision trees can be presented in a manner that is easily understood by the general public. We did not test how well a simplified decision-tree would perform relative to our complex models, but we expect, based on the rapid drop-off in importance gain values beyond the first 5 or so variables, that simplified models are likely to perform nearly as well in practice, just not well enough to have won this competition. Additional work should explore how well simplified models perform and, more importantly, the acceptability of decision-tree-based algorithms and presentation methods that are understandable to the general public.

Finally, we provide one note of caution for scholars working to advance discussions and findings on the accuracy and fairness of risk assessment tools. This competition relied on a dichotomized risk classification (i.e., false-positive rates) to calculate contest scores. This practice is common in many discussions of fairness in risk assessment tools (for a review, see Skeem & Lowenkamp, 2020). However, the dichotomization of risk assessment information limits the generalizability of findings to practice. One example of an alternative would have been to use differences in Brier scores between White and Black parolees rather than differences in false-positive rates. Interestingly, our Year 3 model made no predictions of >50% risk for any individual and therefore had no false positives. Therefore, the model was guaranteed to receive no racial bias penalty under the contest definition, whereas a continuous measure such as differences in Brier scores could have allowed for continued sensitivity to racial bias in the context of low base rates.

Conclusion

Overall, we have several concluding thoughts after participating in this competition. Primarily, we acknowledge that the types of models required to be competitive in this competition may be difficult to apply in practice. Researchers looking to use machine learning models to develop risk assessment tools face the challenge of simplifying models for use in practice without overly compromising predictive accuracy, but this contest’s scoring criteria placed no special value on simplicity. Relatedly, we recommend caution in using some of the variables included in our models for practical risk assessment development. In the early stages of the competition, our team had several lengthy discussions about the ethics of using variables like race and geographic region in our modeling strategy. We ultimately included these variables but recognized that they may not be appropriate for risk assessments tools in practice, and that the small incremental gains in predictive accuracy which motivated their inclusion in the contest models may easily be outweighed by other practical and ethical considerations. Another common discussion point throughout this competition was the importance of translating risk assessment information into practice. As we discussed earlier, there is room to improve the communication of uncertainty in risk estimates. In many cases, even “high” risk individuals are more likely *not* to offend than offend, supporting the use of risk assessment information to facilitate the provision of services to address criminogenic needs rather than restrictive placements. Finally, this competition used data on parolees, a fairly entrenched population in the criminal-legal system. Replication of this competition with a broader justice-involved population (e.g., pretrial defendants, probationers) could have a greater impact on correctional practice.

References

Andrews, D. A., & Bonta, J. (2010a). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy, and Law, 16*(1), 39–55. https://doi.org/10.1037/a0018362

Andrews, D. A., & Bonta, J. (2010b). *The psychology of criminal conduct.* Routledge.

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior, 17*(1), 19–52. https://doi.org/10.1177/0093854890017001004

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology, 28*(3), 369–404. https://doi.org/10.1111/j.1745-9125.1990.tb01330.x

Barrick, K., Lattimore, P. K., & Visher, C. A. (2014). Reentering women: The impact of social ties on long-term recidivism. *The Prison Journal, 94*(3), 279–304. https://doi.org/10.1177/0032885514537596

Benda, B. B. (2005). Gender differences in life-course theory of recidivism: A survival analysis. *International Journal of Offender Therapy and Comparative Criminology, 49*(3), 325–342. https://doi.org/10.1177/0306624X04271194

Berk, R., Heidari, H., Jabbari, S., Kearns, M., & Roth, A. (2017). *Fairness in criminal justice risk assessments: The state of the art.* ArXiv. https://arxiv.org/abs/1703.09207v2

Casey, P., Elek, J., & Warren, R. (2017). *Use of risk and needs assessment information in state sentencing proceedings. Center for Sentencing Initiatives, National Center for State Courts.* https://www.ncsc.org/___data/assets/pdf_file/0025/25297/ebs-rna-brief-sep-2017.pdf

Collins, R. E. (2010). The effect of gender on violent and nonviolent recidivism: A meta-analysis. *Journal of Criminal Justice, 38*(4), 675–684. https://doi.org/10.1016/j.jcrimjus.2010.04.041

Corbett-Davies, S., Pierson, E., Feller, A., Goel, S., & Huq, A. (2017). Algorithmic decision making and the cost of fairness. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* 797–806. https://doi.org/10.1145/3097983.309809

Desmarais, S. L., & Singh, J. P. (2013). *Risk assessment instruments validated and implemented in correctional settings in the United States.* Council of State Governments Justice Center.

Dowden, C., & Brown, S. L. (2002). The role of substance abuse factors in predicting recidivism: A Meta-analysis. *Psychology, Crime & Law, 8*(3), 243–264. https://doi.org/10.1080/10683160208401818

Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology, 34*(4), 575–608. https://doi.org/10.1111/j.1745-9125.1996.tb01220.x

Georgia Data. (2021). *Data tables. Georgia Data.* https://georgiadata.org/data/data-tables

Georgia’s Uniform Crime Reporting Program. (2013). *Crime statistics.* Georgia Bureau of Investigation. https://gbi.georgia.gov/services/crime-statistics

Harcourt, B. E. (2015). Risk as a proxy for race: The dangers of risk assessment. *Federal Sentencing Reporter, 27*(4), 237–243. https://doi.org/10.1525/fsr.2015.27.4.237

Harm, N. J., & Phillips, S. D. (2001). You can’t go home again. *Journal of Offender Rehabilitation, 32*(3), 3–21. https://doi.org/10.1300/J076v32n03_02

Hohman, M. M., McGaffigan, R. P., & Segars, L. (2000). Predictors of successful completion of a postincarceration drug treatment program. *Journal of Addictions & Offender Counseling, 21*(1), 12–22. https://doi.org/10.1002/j.2161-1874.2000.tb00148.x

Holder, E. (2014). *Attorney General Eric Holder Speaks at the National Association of Criminal Defense Lawyers 57th Annual Meeting.* http://www.justice.gov/opa/speech/attorney-general-eric-holder-speaks-nationalassociation-criminal-defense-lawyers-57th

Huebner, B. M., DeJong, C., & Cobbina, J. (2010). Women coming home: Long-term patterns of recidivism. *Justice Quarterly, 27*(2), 225–254. https://doi.org/10.1080/07418820902870486

Huq, A. Z. (2019). Racial equity in algorithmic criminal justice. *Duke Law Journal, 68,* 1043–1134.

Kleinberg, J., Mullainathan, S., & Raghavan, M. (2016). *Inherent trade-offs in the fair determination of risk scores.* ArXiv. https://doi.org/1609.05807

Lattimore, P. K., Tueller, S., Levin-Rector, A., & Witwer, A. (2020). The prevalence of local criminal justice practices. *Federal Probation, 84*(1), 28–37.

Mayson, S. G. (2018). Bias in, bias out. *Yale Law Journal, 128*(8), 2218–2301.

Minow, M., Zittrain, J., & Bowers. (2019). *Technical flaws of pretrial risk assessments raise grave concerns.* Berkman Klein Center. https://cyber.harvard.edu/story/2019-07/technical-flaws-pretrial-risk-assessments-raise-grave-concerns

Monahan, J. (2006). A jurisprudence of risk assessment: Forecasting harm among prisoners, predators, and patients. *Virginia Law Review,* 391-435.

Porter, L. C., & DeMarco, L. M. (2019). Beyond the dichotomy: Incarceration dosage and mental health. *Criminology, 57*(1), 136–156. https://doi.org/10.1111/1745-9125.12199

Pretrial Justice Institute. (2020). *Updated position on pretrial risk assessment tools.* https://www.pretrial.org/wp-content/uploads/Risk-Statement-PJI-2020.pdf

Schnittker, J., & John, A. (2007). Enduring stigma: The long-term effects of incarceration on health. *Journal of Health and Social Behavior, 48*(2), 115–130. https://doi.org/10.1177/002214650704800202

Schumacher, M., Holländer, N., & Sauerbrei, W. (1997). Resampling and cross-validation techniques: A tool to reduce bias caused by model building? *Statistics in Medicine, 16*(24), 2813–2827. https://doi.org/10.1002/(SICI)1097-0258(19971230)16:24<2813::AID-SIM701>3.0.CO;2-Z

Scott, T., & Brown, S. L. (2018). Risks, strengths, gender, and recidivism among justice-involved youth: A meta-analysis. *Journal of Consulting and Clinical Psychology, 86*(11), 931–945. https://doi.org/10.1037/ccp0000343

Skeem, J. L., & Lowenkamp, C. T. (2016). Risk, race, and recidivism: Predictive bias and disparate impact. *Criminology, 54*(4), 680–712. https://doi.org/10.1111/1745-9125.12123

Skeem, J. L., & Lowenkamp, C. T. (2020). Using algorithms to address trade-offs inherent in predicting recidivism. *Behavioral Sciences & the Law, 38*(3), 259–278. https://doi.org/10.1002/bsl.2465

Starr, S. B. (2014). Evidence-based sentencing and the scientific rationalization of discrimination. *Stanford Law Review, 66*, 803.

Turanovic, J. J., Rodriguez, N., & Pratt, T. C. (2012). The collateral consequences of incarceration revisited: A qualitative analysis of the effects on caregivers of children of incarcerated parents. *Criminology, 50*(4), 913–959. https://doi.org/10.1111/j.1745-9125.2012.00283.x

US Census Bureau. (2021). *Accessing PUMS data.* The United States Census Bureau. https://www.census.gov/programs-surveys/acs/microdata/access.html

Vincent, G. M., & Viljoen, J. L. (2020). Racist algorithms or systemic problems? Risk assessments and racial disparities. *Criminal Justice and Behavior*, 0093854820954501. https://doi.org/10.1177/0093854820954501

Wachter, A. (2015). *Statewide risk assessment in juvenile probation. National Center for Juvenile Justice.* http://www.ncjj.org/pdf/JJGPS%20StateScan/JJGPS_StateScan_Statewide_Risk%20Assessment_2014_2.pdf

Wagner, P., & Rabuy, B. (2017). *Following the money of mass incarceration.* Prison Policy Initiative. https://www.prisonpolicy.org/reports/money.html

Walters, G. D. (2015). Recidivism and the "Worst of Both Worlds" hypothesis: Do substance misuse and crime interact or accumulate? *Criminal Justice and Behavior, 42*(4), 435–451. https://doi.org/10.1177/0093854814551018

Western, B., Kling, J. R., & Weiman, D. F. (2001). The labor market consequences of incarceration. *Crime & Delinquency, 47*(3), 410–427. https://doi.org/10.1177/0011128701047003007

Wollert, R. (2002). The importance of cross-validation in actuarial test construction. *Journal of Threat Assessment, 2*(1), 87–102. https://doi.org/10.1300/J177v02n01_06

World Prison Brief. (2018). *Highest to lowest—Prison Population Rate.* https://www.prisonstudies.org/highest-to-lowest/prison_population_rate?field_region_taxonomy_tid=All

Yukhnenko, D., Blackwood, N., & Fazel, S. (2020). Risk factors for recidivism in individuals receiving community sentences: A systematic review and meta-analysis. *CNS Spectrums, 25*(2), 252–263. https://doi.org/10.1017/S1092852919001056

Table 1

Model Performance by Dataset, Year and Gender, for training dataset and holdout sample

				False Positive Rate		
Dataset	Accuracy	Sensitivity	Specificity	White	Black	Brier Score
Train Year 1						
Female	0.7943	0.0044	0.9994	0.0000	0.0017	0.1481
Male	0.7099	0.1669	0.9553	0.0494	0.0416	0.1892
Average						0.1686
Test Year 1						
Female	0.7800	0.0144	0.9946	0.0083	0.0000	0.1585
Male	0.7025	0.1505	0.9527	0.0486	0.0464	0.1951
Average						0.1768
Train Year 2						
Female	0.7943	0.0467	0.9893	0.0098	0.0126	0.1394
Male	0.7634	0.1942	0.9689	0.0335	0.0295	0.1620
Average						0.1507
Test Year 2						
Female	0.8288	0.0687	0.9918	0.0130	0.0000	0.1263
Male	0.7482	0.1513	0.9479	0.0500	0.0534	0.1699
Average						0.1481
Train Year 3						
Female	0.8639	0.0000	1.0000	0.0000	0.0000	0.1047
Male	0.7999	0.0000	1.0000	0.0000	0.0000	0.1409
Average						0.1228
Test Year 3						
Female	0.8494	0.0000	1.0000	0.0000	0.0000	0.1193
Male	0.7935	0.0000	1.0000	0.0000	0.0000	0.1524
Average						0.1358

Table 2

Contest Score by Year and Gender

	Contest Score	
Dataset	Male	Female
Train 1	0.8045	0.8504
Test 1	0.8031	0.8345
Train 2	0.8347	0.8582
Test 2	0.8273	0.8624
Train 3	0.8591	0.8953
Test 3	0.8476	0.8807

Table 3

Confusion Matrix by Dataset, Year and Gender

		Observed			
		Male		Female	
Dataset		No	Yes	No	Yes
Train Year 1					
Prediction	No	10404	4099	1759	455
	Yes	487	821	1	2
Test Year 1					
Prediction	No	4495	1817	738	205
	Yes	223	322	4	3
Train Year 2					
Prediction	No	7753	2328	1381	347
	Yes	249	561	15	17
Test Year 2					
Prediction	No	3351	1004	606	122
	Yes	184	179	5	9
Train Year 3					
Prediction	No	6401	1601	1206	190
	Yes	0	0	0	0
Test Year 3					
Prediction	No	2805	730	519	92
	Yes	0	0	0	0

Table 4
Mean, Minimum, and Maximum Predicted Probabilities

Set	Mean	Min	Max
Train 1	0.30	0.03	0.75
Test 1	0.30	0.03	0.72
Train 2	0.26	0.00	0.83
Test 2	0.26	0.00	0.80
Train 3	0.17	0.04	0.48
Test 3	0.17	0.04	0.44

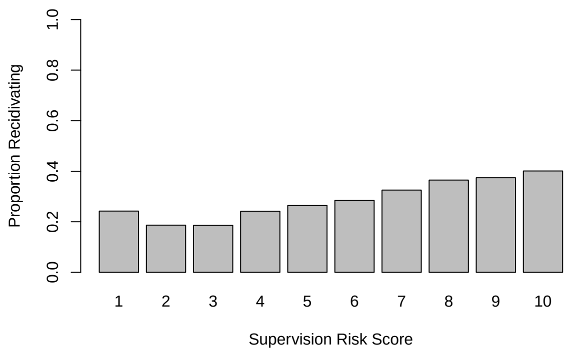

Figure 1

Proportion Recidivating at the End of Year 1 by Supervision Risk Score

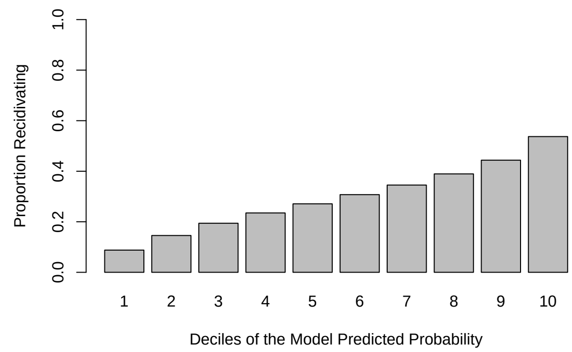

Figure 2

Proportion Recidivating at the End of Year 1 by Decile of Predicted Probability

Appendix A

Feature Importance Tables for XGBoosted Models

Table A1

Importance Matrix for Year 1 XGBoosted Gradient Tree Model: 30 Most Important Variables

Feature	Gain
Gang_Affiliated.true	0.1644442
Prior_Arrest_Episodes_Property	0.1312534
Supervision_Risk_Score_First	0.1024736
Prior_Arrest_Episodes_PPViolation.5_or_more	0.0943669
MHSA_Priors	0.0494180
Age_at_Release.48_or_older	0.0489115
Prison_Years.Less_than_1_year	0.0363765
Prior_Arrest_Episodes_Misd	0.0349866
Prior_Arrest_Episodes_Felony	0.0280256
Gender.M	0.0271737
Prison_Years.More_than_3_years	0.0249934
Prior_Conviction_Episodes_Misd	0.0231134
Age_at_Release.23_27	0.0184198
Rate_Sum_LessHalfYearWorked	0.0161371
Prison_Offense.Violent_Sex	0.0154813
Female_Dependents	0.0152849
CogED_Priors	0.0133683
Prior_Revocations_Parole.true	0.0107930
Age_at_Release.38_42	0.0107469
Male_PriorViolent	0.0101646
Rate_PUMA_95	0.0090422
Prison_Offense.Property	0.0089828
Age_at_Release.43_47	0.0082721
Rate_PUMA_14	0.0081327
Age_at_Release.33_37	0.0078163
Supervision_Level_First.Standard	0.0072990
Prison_Offense.Drug	0.0071623
Prison_Years.Greater_than_2_to_3_years	0.0058724
Prior_Conviction_Episodes_Prop	0.0054493
Age_at_Release.28_32	0.0043577

Table A2

Importance Matrix for Year 2 XGBoosted Gradient Tree Model: 30 Most Important Variables

Feature	Gain
Probability	0.4186117
Jobs_Per_Year	0.1603833
Percent_Days_Employed	0.1549017
Avg_Days_per_DrugTest	0.0796209
DrugTests_THC_Positive	0.0238482
Violations	0.0127656
Delinquency_Reports	0.0113833
Prior_Revocations_Parole.true	0.0094820
MHSA_Drug	0.0079200
MHSADrugs	0.0076962
Residence_Changes	0.0067345
Rate_PUMA_14	0.0061730
Program_Attendances	0.0061482
Prior_Arrest_Episodes_Misd	0.0054487
Supervision_Risk_Score_First	0.0053076
Program_UnexcusedAbsences	0.0049762
PUMA_164	0.0048642
MHSA_Priors	0.0045581
Prior_Arrest_Episodes_Violent	0.0044344
Race.WHITE	0.0042768
Prior_Arrest_Episodes_Felony	0.0042493
Prior_Arrest_Episodes_Drug	0.0042385
Education_Level.High_School_Diploma	0.0036929
Residence_PUMA.6	0.0033324
Age_at_Release.38_42	0.0033078
Age_at_Release.33_37	0.0030666
Prior_Conviction_Episodes_Felony	0.0029866
Residence_PUMA.14	0.0027040
PUMA_19	0.0025633
PUMA_221	0.0023009

Table A3

Importance Matrix for Year 3 XGBoosted Gradient Tree Model without Prior Probabilities: 30 Most Important Variables

Feature	Gain
Prior_Arrest_Episodes_PPViolationCharges	0.1302818
Percent_Days_Employed	0.1083164
Avg_Days_per_DrugTest	0.1017630
Jobs_Per_Year	0.0927024
Gang_Affiliated.true	0.0659487
Supervision_Risk_Score_First	0.0646153
MHSA_Priors	0.0451457
Prior_Conviction_Episodes_Misd	0.0344736
Prior_Arrest_Episodes_Felony	0.0295227
Age_at_Release.48_or_older	0.0270246
Prior_Arrest_Episodes_Misd	0.0244191
DrugTests_THC_Positive	0.0208614
Violations	0.0206476
Prior_Arrest_Episodes_Property	0.0191017
Gender.M	0.0159579
DrugTests_Meth_Positive	0.0150301
Rate_Sum_DayCommute	0.0130575
Prior_Conviction_Episodes_Felony	0.0129458
Delinquency_Reports	0.0113725
Violations_Instruction.true	0.0112837
Program_Attendances	0.0112657
Prior_Conviction_Episodes_Prop	0.0089142
Age_at_Release.43_47	0.0089018
Supervision_Risk_Score_First_isNA	0.0084530
Residence_Changes	0.0081764
Rate_PUMA_14	0.0077555
Prison_Years.More_than_3_years	0.0074178
MHSADrugs	0.0070424
Prison_Years.Less_than_1_year	0.0067171
PUMA_19	0.0064709

Table A4

Importance Matrix for Year 3 XGBoosted Gradient Tree Model with Prior Probabilities: 30 Most Important Variables

Feature	Gain
Probability2	0.4474033
Probability1	0.1500780
Jobs_Per_Year	0.0943376
Avg_Days_per_DrugTest	0.0410181
Percent_Days_Employed	0.0344985
DrugTests_Meth_Positive	0.0212334
Violations	0.0192419
Prior_Arrest_Episodes_Felony	0.0179984
Prior_Arrest_Episodes_Misd	0.0160644
DrugTests_THC_Positive	0.0107427
Rate_Sum_DayCommute	0.0100725
MHSA_Priors	0.0084051
Dependents	0.0065557
Program_Attendances	0.0063252
Prior_Conviction_Episodes_Misd	0.0061631
Rate_PUMA_95	0.0060066
Gender.M	0.0056610
Prior_Conviction_Episodes_Prop	0.0055870
Residence_Changes	0.0052577
Prior_Arrest_Episodes_GunCharges.true	0.0052062
Violations_MoveWithoutPermission.true	0.0048099
Rate_PUMA_49	0.0047176
Prior_Revocations_Parole.true	0.0046641
DrugTests_Cocaine_Positive	0.0044009
Supervision_Risk_Score_First_isNA	0.0040181
Program_UnexcusedAbsences	0.0040113
Employment_Exempt.true	0.0038456
MHSADrugs	0.0038033
CogED_Priors	0.0035876
DrugTests_Other_Positive	0.0035120

Appendix B

PUMA Data Information

Table B1
PUMA Data Tables Extracted from Census Data

Census Table Label	Construct	Number Variables Contained
AGEP	Average age	1
FES	Total number of families and employment status	10
FPARC	Total number of families	6
GASP	Average gas cost	1
GRPIP	Average gross rent cost	1
HISP	Total number of people who identify as Spanish, Hispanic, or Latino	25
INTP	Average amount of interest, dividends, and net rental income in the past 12 months	1
JWMNP	Average travel time to work	1
MV	Total number of people who moved into their house a specified number of years ago	9
OIP	Average amount of income from other sources	1
R60	Total number of families with a person over 60 years old in the household	5
RACAIAN	Total number of people who identify as American Indian	3
RACASN	Total number of people who identify as Asian	3
RACBLK	Total number of people who identify as Black	3
RACSOR	Total number of people who identify as other race	3
RACWHT	Total number of people who identify as White	3
RNTP	Average monthly rent	1
WAGP	Average wage or salary income in the past 12 months	1
WKHP	Average hours worked per week	1
CIT	Total number of citizen status	6
COW	Total number of employment	11
ESP	Total number of parent employment	10
ESR	Total number of armed forces and civilian employee	8
GCL	Total number of grandparent living with grandchildren	4
HHT	Total number family and nonfamily household	9
JWTR	Total number of transportation to work	14
LNGI	Total number of households that have a person at least 14 years old speak English	4
MAR	Total number of marriage status	6

(continued)

Census Table Label	Construct	Number Variables Contained
MLPA	Total number of people who served in the Armed Forces before and after 9/11	4
MSP	Total number of marriage status	8
NOC	Average number of children in household	1
NP	Average number of people in the household	1
NPP	Total number of grandparent headed household	4
NR	Total number of households that have nonrelatives	4
PSF	Total number of households that have subfamilies	4
R18	Total number of households that have a person under 18 years old	4
RNTM	Total number of meals included with rent	4
SEX	Total number of sex	3
SRNT	Total number of specified rental units on 10 acres of land	4
VACS	Total number of vacancy	9
VEH	Total number of vehicles	9
WIF	Total number of workers in the family	6
BUS	Total number of businesses on the property	4
GRNTP	Average gross monthly rent	1
PAP	Average PUMS supplementary security income/AFDC/other welfare income	1
SSIP	Average supplementary security income in the past 12 months	1
SSP	Average PUMS social security or railroad retirement income	1
PARTNER	Total number of unmarried partners in the household	7
HICOV	Total number of people who have health insurance coverage	3
GCR	Total number of grandparents responsible for children	4
FS	Total number of people who received food stamps in the past year	4
FINCP	Average amount of family income in the past year	1
DOUT	Total number of people who independent living difficultly	4
MARHD	Total number of people who divorced in the past year	4
MARHT	Total number of times a person married	5
MIG	Total number of people who lived in the same house one year ago	5
MULTG	Total number of multigenerational households	4
PRIVCOV	Total number of people with private health coverage	3
PUBCOV	Total number of people with public health coverage	3
RWAT	Total number of people with hot and cold running water	4

(continued)

Census Table Label	Construct	Number Variables Contained
TAXP	Total number of property tax amounts for the year	70
WKW	Total number of weeks worked in the year	8
FER	Total number of women who gave birth in the last year	4
JWDP	Total number of time of departure to work	152

Table B2
Final PUMA Variables

Characteristic	Description
Variable	PUMA_1
Census Label	Age
Description	Average age
Type	Continuous
Original Unit	Years
Variable Manipulation	Direct use
Variable	Rate_PUMA_3
Census Label	FES: N/A (GQ/vacant/not a family/same-sex married-couple families)
Description	Other types of families
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_4
Census Label	FES: Married-couple family: Husband and wife in LF
Description	Married couple with both spouses in labor force
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_5
Census Label	FES: Married-couple family: Husband in labor force, wife not in LF
Description	Married couple with husband in labor force and wife unemployed
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_6
Census Label	FES: Married-couple family: Husband not in LF, wife in LF
Description	Married couple with husband unemployed and wife in the labor force
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_7

(continued)

Characteristic	Description
Census Label	FES: Married-couple family: Neither husband nor wife in LF
Description	Married couple with both spouses unemployed
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_8
Census Label	FES: Other family: Male householder, no wife present, in LF
Description	Male headed-household, and the male is in the labor force
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_9
Census Label	FES: Other family: Male householder, no wife present, not in LF
Description	Male headed-household, and the male is unemployed
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_10
Census Label	FES: Other family: Female householder, no husband present, in LF
Description	Female headed-household, and the female in labor force
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_11
Census Label	FES: Other family: Female householder, no husband present, not in LF
Description	Female headed-household, and the female is unemplyed
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_13
Census Label	FPARC: N/A (GQ/vacant/not a family)
Description	Other types of families
Type	Continuous (Rate)
Original Unit	Person

(continued)

Characteristic	Description
Variable Manipulation	Rate
Variable	Rate_PUMA_14
Census Label	FPARC: With related children under 5 years only
Description	Family has children only under 5 years old
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_15
Census Label	FPARC: With related children 5 to 17 years only
Description	Family had children between 5 and 17 years old
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_16
Census Label	FPARC: With related children under 5 years and 5 to 17 years
Description	Family has children between 0 and 17 years old
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_17
Census Label	FPARC: No related children
Description	Family has no children
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	PUMA_18
Census Label	GASP: Gas (monthly cost, use ADJHSG to adjust GASP values 4 and over to constant dollars)
Description	Average Gas Cost
Type	Continuous
Original Unit	Dollars
Variable Manipulation	Direct use
Variable	PUMA_19
Census Label	GRPIP: Gross rent as a percentage of household income past 12 months
Description	Average Gross Rent Cost
Type	Continuous
Original Unit	Dollars

(continued)

Characteristic	Description
Variable Manipulation	Direct use
Variable	Rate_PUMA_21
Census Label	HISP: Not Spanish/Hispanic/Latino
Description	Not Spanish or Hispanic or Latino
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_22
Census Label	HISP: Mexican
Description	Mexican
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_23
Census Label	HISP: Puerto Rican
Description	Puerto Rican
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_24
Census Label	HISP: Cuban
Description	Cuban
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	PUMA_45
Census Label	INTP: Interest, dividends, and net rental income past 12 months (signed, use ADJINC to adjust to constant dollars)
Description	Interest, dividends, and net rental income past 12 months (signed, use ADJINC to adjust to constant dollars)
Type	Continuous
Original Unit	Dollars
Variable Manipulation	Direct use
Variable	PUMA_46
Census Label	JWMNP: Travel time to work
Description	Average Travel Time to Work
Type	Continuous
Original Unit	Minutes

(continued)

Characteristic	Description
Variable Manipulation	Direct use
Variable	Rate_PUMA_49
Census Label	MV: 12 months or less
Description	Person moved into their house or apartment less than 12 months ago
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	PUMA_56
Census Label	OIP: All other income past 12 months (use ADJINC to adjust to constant dollars)
Description	All other income sources in the past year
Type	Continuous
Original Unit	Dollars
Variable Manipulation	Direct use
Variable	Rate_PUMA_59
Census Label	R60: No person 60 and over
Description	No person over 60 years old lives in the household
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_64
Census Label	RACAIAN: Yes
Description	Number of People who do identify as American Indian or Alaska Native
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_67
Census Label	RACASN: Yes
Description	Number of People who identify as Asian
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_70
Census Label	RACBLK: Yes
Description	Number of People who identify as Black
Type	Continuous (Rate)
Original Unit	Person

(continued)

Characteristic	Description
Variable Manipulation	Rate
Variable	Rate_PUMA_73
Census Label	RACSOR: Yes
Description	Number of People who identify as other race
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_76
Census Label	RACWHT: Yes
Description	Number of People who identify as White
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	PUMA_77
Census Label	RNTP: Monthly rent (use ADJHSG to adjust RNTP to constant dollars)
Description	Monthly Rent
Type	Continuous
Original Unit	Dollars
Variable Manipulation	Direct use
Variable	PUMA_78
Census Label	WAGP: Wages or salary income past 12 months (use ADJINC to adjust WAGP to constant dollars)
Description	Wages and/or Salary Income
Type	Continuous
Original Unit	Dollars
Variable Manipulation	Direct use
Variable	PUMA_79
Census Label	WKHP: Usual hours worked per week past 12 months
Description	Average hours worked per week
Type	Continuous
Original Unit	Hours
Variable Manipulation	Direct use
Variable	Rate_PUMA_81
Census Label	CIT: Born in the U.S.
Description	Born in the US
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate

(continued)

Characteristic	Description
Variable	Rate_PUMA_85
Census Label	CIT: Not a citizen of the U.S.
Description	Not a U.S. citizen
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_88
Census Label	COW: Employee of a private for-profit co. or bus., or of an individual, for wages, salary, commissions
Description	Employee of a private for-profit company or business, or of an individual, for wages, salary, commissions
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_89
Census Label	COW: Employee of a private not-for-profit, tax-exempt, or charitable organization
Description	Employee of a private not-for-profit, tax-exempt, or charitable organization
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_90
Census Label	COW: Local government employee (city, county, etc.)
Description	Local government employee
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_91
Census Label	COW: State government employee
Description	State government employee
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_92
Census Label	COW: Federal government employee
Description	Federal government employee
Type	Continuous (Rate)
Original Unit	Person

(continued)

Characteristic	Description
Variable Manipulation	Rate
Variable	Rate_PUMA_93
Census Label	COW: Self-employed in own not incorporated business, professional practice, or farm
Description	Self-employed in own not incorporated business, professional practice, or farm
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_94
Census Label	COW: Self-employed in own incorporated business, professional practice or farm
Description	Self-employed in own incorporated business, professional practice or farm
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_95
Census Label	COW: Working without pay in family business or farm
Description	Works without pay in the family business or farm
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_96
Census Label	COW: Unemployed and last worked 5 years ago or earlier or never worked
Description	Unemployed at least 5 years or never worked
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_111
Census Label	ESR: Unemployed
Description	Unemployed
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_114
Census Label	ESR: Not in Labor Force
Description	Not in the labor force
Type	Continuous (Rate)

(continued)

Characteristic	Description
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_117
Census Label	GCL: Yes
Description	Grandparents live with grandchildren
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_121
Census Label	HHT: Married couple household
Description	Married couple household
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_122
Census Label	HHT: Other family household: Male householder, no spouse present
Description	Other family household: Male householder, no spouse present
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_123
Census Label	HHT: Other family household: Female householder, no spouse present
Description	Other family household: Female householder, no spouse present
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_124
Census Label	HHT: Nonfamily household: Male householder: Living alone
Description	Nonfamily household: Male householder: Living alone
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_125
Census Label	HHT: Nonfamily household: Male householder: Not living alone

(continued)

Characteristic	Description
Description	Nonfamily household: Male householder: Not living alone
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_126
Census Label	HHT: Nonfamily household: Female householder: Living alone
Description	Nonfamily household: Female householder: Living alone
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_127
Census Label	HHT: Nonfamily household: Female householder: Not living alone
Description	Nonfamily household: Female householder: Not living alone
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_130
Census Label	JWTR: Car/truck/van
Description	Car/truck/van used as transportation to get to work
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_144
Census Label	LNGI: At least one person in the household 14 and over speaks English only or speaks English ’very well’
Description	At least one person in the household 14 and over speaks English only or speaks English ’very well’
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_147
Census Label	MAR: Married
Description	Number of Married People
Type	Continuous (Rate)
Original Unit	Person

(continued)

Characteristic	Description
Variable Manipulation	Rate
Variable	Rate_PUMA_148
Census Label	MAR: Widowed
Description	Number of Widowed People
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_149
Census Label	MAR: Divorced
Description	Number of Divorced People
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_150
Census Label	MAR: Separated
Description	Number of Separated People
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_151
Census Label	MAR: Never married or under 15 years old
Description	Number of Never Married People
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_155
Census Label	MLPA: Served This Period
Description	Served in the Armed Forces after September 2001
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_159
Census Label	MSP: Now Married, Spouse Absent
Description	Married with spouse absent
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	PUMA_164

(continued)

Characteristic	Description
Census Label	NOC: Number of own children in household (unweighted)
Description	Number of children in the household
Type	Continuous
Original Unit	Person
Variable Manipulation	Direct use
Variable	PUMA_165
Census Label	NP: Number of persons in this household
Description	Number of persons in the household
Type	Continuous
Original Unit	Person
Variable Manipulation	Direct use
Variable	Rate_PUMA_169
Census Label	NPP: Grandparent headed household with no parent present
Description	Grandparent headed household with no parent present
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_173
Census Label	NR: 1 or more nonrelatives
Description	1 or more nonrelatives in the household
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_177
Census Label	PSF: 1 or more subfamilies
Description	1 or more subfamilies in the household
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_181
Census Label	R18: 1 or more persons under 18 in household
Description	1 or more persons under 18 in the household
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_188
Census Label	SEX: Female
Description	Number of Females

(continued)

Characteristic	Description
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_191
Census Label	SRNT: A single-family home on 10 or more acres.
Description	A single-family home on 10 or more acres.
Type	Continuous (Rate)
Original Unit	Building
Variable Manipulation	Rate
Variable	Rate_PUMA_204
Census Label	VEH: No vehicles
Description	No Vehicles
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_205
Census Label	VEH: 1 vehicle
Description	1 Vehicle
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_206
Census Label	VEH: 2 vehicles
Description	2 Vehicles
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_213
Census Label	WIF: No workers
Description	No workers in the family in the past 12 months
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_214
Census Label	WIF: 1 worker
Description	1 worker in the family in the past 12 months
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate

(continued)

Characteristic	Description
Variable	Rate_PUMA_215
Census Label	WIF: 2 workers
Description	2 workers in the family in the past 12 months
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_216
Census Label	WIF: 3 or more workers in family
Description	3 or more workers in the family in the past 12 months
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_219
Census Label	BUS: Yes
Description	Business is on Property
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	PUMA_221
Census Label	GRNTP: Gross rent (monthly amount)
Description	Average gross monthly rent
Type	Continuous
Original Unit	Dollars
Variable Manipulation	Direct use
Variable	PUMA_222
Census Label	PAP: PUMS SSI/AFDC/other welfare income
Description	Social Security Income/Aid to Families with Dependent Children/Other welfare income
Type	Continuous
Original Unit	Dollars
Variable Manipulation	Direct use
Variable	PUMA_223
Census Label	SSIP: Supplementary Security Income past 12 months
Description	Supplementary Security Income in the past year
Type	Continuous
Original Unit	Dollars
Variable Manipulation	Direct use
Variable	PUMA_224

(continued)

Characteristic	Description
Census Label	SSP: PUMS Social Security or Railroad Retirement Income
Description	Social Security or Railroad Retirement Income
Type	Continuous
Original Unit	Dollars
Variable Manipulation	Direct use
Variable	Rate_PUMA_228
Census Label	PARTNER: Male householder, male partner
Description	Male householder, male partner in the household
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_229
Census Label	PARTNER: Male householder, female partner
Description	Male householder, female partner in the household
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_230
Census Label	PARTNER: Female householder, female partner
Description	Female householder, female partner in the household
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_231
Census Label	PARTNER: Female householder, male partner
Description	Female householder, male partner in the household
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_233
Census Label	HICOV: With health insurance coverage
Description	Has health insurance coverage
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_237
Census Label	GCR: Yes
Description	Grandparents are responsible for children
Type	Continuous (Rate)

(continued)

Characteristic	Description
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_241
Census Label	FS: Yes
Description	Received food stamps in the past year
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	PUMA_243
Census Label	FINCP: Family income (past 12 months, use ADJINC to adjust FINCP to constant dollars)
Description	Total family income in the past year
Type	Continuous
Original Unit	Dollars
Variable Manipulation	Direct use
Variable	Rate_PUMA_246
Census Label	DOUT: Yes
Description	Has independent living difficulty
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_250
Census Label	MARHD: Yes
Description	Divorced in the past year
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_254
Census Label	MARHT: One time
Description	Married one time
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_255
Census Label	MARHT: Two Times
Description	Married two times
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate

(continued)

Characteristic	Description
Variable	Rate_PUMA_256
Census Label	MARHT: Three or more times
Description	Married three or more times
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_259
Census Label	MIG: Yes, same house (nonmovers)
Description	Lived in the same house one year ago
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_265
Census Label	MULTG: Yes, is a multigenerational household
Description	Is a Multigenerational Household
Type	Continuous (Rate)
Original Unit	Household
Variable Manipulation	Rate
Variable	Rate_PUMA_267
Census Label	PRIVCOV: With private health insurance coverage
Description	Has private health insurance
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_270
Census Label	PUBCOV: With public health coverage
Description	Has public health insurance
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_274
Census Label	RWAT: Yes
Description	Has hot and cold running water
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_PUMA_356
Census Label	FER: Yes
Description	Gave birth in the past year

(continued)

Characteristic	Description
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Rate
Variable	Rate_Sum_ThreePlusVeh
Census Label	Three or more vehicles in household
Description	Vehicles in household
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Summed category counts to rate
Variable	Rate_Sum_Less2KPropertyTax
Census Label	< 2,000 annual property tax
Description	Annual property tax
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Summed category counts to rate
Variable	Rate_Sum_2KPropertyTax
Census Label	> 2,000 annual property tax
Description	Annual property tax
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Summed category counts to rate
Variable	Rate_Sum_HalfYearWorked
Census Label	Worked at least half of year
Description	Based on Weeks Worked in the Past Year
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Summed category counts to rate
Variable	Rate_Sum_LessHalfYearWorked
Census Label	Worked less than half year
Description	Based on Weeks Worked in the Past Year
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Summed category counts to rate
Variable	Rate_Sum_OvernightCommute
Census Label	Commute overnight (5pm-5am)
Description	Time of Departure to Work (Hour and Minute)
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Summed category counts to rate

(continued)

Characteristic	Description
Variable	Rate_Sum_MorningCommute
Census Label	Commute morning (5am-9:30am)
Description	Time of Departure to Work (Hour and Minute)
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Summed category counts to rate
Variable	Rate_Sum_DayCommute
Census Label	Commute day (9:30am-5pm)
Description	Time of Departure to Work (Hour and Minute)
Type	Continuous (Rate)
Original Unit	Person
Variable Manipulation	Summed category counts to rate