University of Toronto
Department of Economics

Working Paper 684

Prison Rehabilitation Programs: Efficiency and Targeting

By William Arbour, Guy Lacroix and Steeve Marchand

January 08, 2021

Prison Rehabilitation Programs: Efficiency and Targeting*

William Arbour† Guy Lacroix‡ Steeve Marchand§

January 7, 2021

Abstract

Increasing evidence suggests that incarceration, under certain circumstances, can improve inmates' social reintegration upon release. Yet, the mechanisms through which incarceration can lead to successful rehabilitation remain largely unknown. This paper finds that participation in social rehabilitation programs while incarcerated can significantly reduce recidivism. This result is entirely driven by inmates whose risk and needs were evaluated by a widely used assessment tool identifying their criminogenic needs. For this group, we estimate that participation in these programs reduces recidivism by about 9 percentage points within three years following release. Our results suggest targeting criminogenic needs is crucial for successful rehabilitation. We also find considerable heterogeneous program treatment effects: inmates with a high overall risk score, or who exhibit procriminal attitudes, benefit little if at all from program participation. We investigate the stability of the treatment effect coefficients and conclude they unlikely suffer from an omitted variable bias.

Keywords: Incarceration, Recidivism, Rehabilitation Programs, Risk Assessment

JEL code: K42

*We wish to thank Quebec's Ministry of Public Security, in particular Bernard Chéné for providing key institutional details and for crucial assistance with the data. The views expressed in this paper are the authors' only and do not necessarily represent those from the Ministry. We also thank seminar participants at the Public Economics and Policy Seminar at HEC, Université de Lausanne for useful comments.

†Department of Economics, University of Toronto (william.arbour@mail.utoronto.ca)

‡Department of Economics, Université Laval (guy.lacroix@ecn.ulaval.ca)

§Melbourne Institute: Applied Economic & Social Research, University of Melbourne (steeve.marchand@unimelb.edu.au)

1 Introduction

Incarceration can affect inmates' criminal behavior in various ways. Prison may prove criminogenic if it develops criminal networks or expertise or lessens inmates' capacity to reintegrate into the labor market due to social stigma or human capital deterioration. There is, however, increasing evidence suggesting prison establishments that emphasize rehabilitation can reduce recidivism (Landersø, 2015; Mastrobuoni and Terlizzese, 2019; Bhuller et al., 2020; Lotti, 2020; Hjalmarsson and Lindquist, 2020). While this nascent literature provides convincing evidence that favorable prison conditions may reduce recidivism, it is still unclear which specific factors are truly beneficial. Rehabilitation programs that focus on education, enhancing job skills or providing specific psychological assistance are often advanced as potential mechanisms to explain these findings. In addition to being potentially privately beneficial, such programs could turn out to be cost-efficient from society's perspective if they reduce future costly incarcerations and reliance on social assistance programs.

Unfortunately, empirical evidence showing which types of programs work, and under which conditions, remains limited. In their meta-analysis of correctional rehabilitation programs for adult offenders, Wilson et al. (2000) and Davis et al. (2013) both conclude that, while evidence suggests that participation decreases recidivism and increases employment, the research designs are deemed too poor to provide reliable estimates of their impact. The major caveat in most studies is a failure to thoroughly consider selection into participation. Selection issues may occur if (self-)selected participants differ systematically from non-participants. Observed and unobserved differences between the two groups may partly explain differences in post-release outcomes, thus biasing the programs' causal effects.

In this paper, we use administrative data from three male provincial prisons in Quebec (Canada) to study the effects of participation in social rehabilitation programs on recidivism between 2007 and 2019. We use various econometric and machine learning techniques to compare non-participants and participants and find that the programs can substantially decrease recidivism if they are well targeted to the offender's criminogenic needs. Because our baseline estimation strategy may not fully account for selection due to unobservable factors, we follow Oster (2019)'s proposed methodology which relates selection on observable and on unobservable characteristics, and find clear evidence that our results do not suffer from an omitted variable bias.

The inmates in Quebec's provincial prisons all serve sentences of less than two years. This population is particularly relevant for the analysis of social rehabilitation programs. Since prison sentences are relatively short, participants may rapidly put to use their newly acquired skills. In addition, all inmates who serve a sentence of at least six months are evaluated using the "Level of Service/Case Management Inventory" (LS/CMI), a proprietary assessment tool widely used in Canada and in the United States. This tool is used to gauge individual risks and determine criminogenic needs. The risk assessment is based upon several aspects of the inmate's history and behavior (e.g., criminal history, alcohol and drug problems, social influences). In Quebec, the evaluation is conducted after sentencing and is used to predict recidivism as well as to direct inmates toward rehabilitation programs adapted to their needs.[1] While risk assessment is compulsory, participation in a rehabilitation program is not.

Our empirical strategy is as follows. First, we estimate baseline regressions of recidivism on participation. We find that participation decreases recidivism likelihood only for those who were evaluated by the LS/CMI, suggesting the evaluation leads to a better fit between programs and needs. Among evaluated inmates, participation is associated with a decrease of around 9.5 percentage points in the probability to recidivate within three years. Importantly, our regression results are robust to the inclusion of the LS/CMI scores, all of which strongly correlate with recidivism. Following Oster (2019), we next exploit the predictive power of these scores to consider selection on unobservable factors. We compute lower bounds (in absolute value) of our treatment effects and show that our estimates are remarkably robust to her correction, even when setting bounding parameters to exceedingly conservative values.

The rehabilitation programs we study are quite heterogeneous. Indeed, the data include programs related to job skills, education, self-development, violence control, addictions and others. We investigate their relative efficacy and find that, save for those in the *other* category, the programs all decrease recidivism substantially for evaluated inmates. In addition, we report evidence that inmates are matched to programs well suited to their criminogenic risk factors. Because the LS/CMI evaluation is specifically designed to identify such risk factors, these results suggest evaluation-based targeting is a key mechanism underlying efficient rehabilitation.

Finally, we extend our analysis on targeting by exploring heterogeneous program effects using machine learning methods. We use the random causal forest algorithm proposed by Wager and Athey (2018) to estimate individual-level treatment effects. Our findings have important policy implications regarding targeting. The component of the LS/CMI that focuses on "procriminal" attitudes is found to be strongly inversely associated with program efficiency. Furthermore, we find that the estimated effect of programs dissipates completely for individuals with the highest overall LS/CMI score, who are considered to be the most at risk. Since program participation is relatively independent of the LS/CMI scores, our findings suggest that there remain additional benefits to be gained from more focused targeting in our setting.

Contributions to the Literature

The contributions of the paper are threefold. First, it contributes to the broad literature investigating the link between incarceration and post-release outcomes. Some find that lengthy prison sentences may deplete human capital and alter inmates' capacity to reintegrate into the labor market (Lochner, 2004; Aizer and Doyle Jr, 2015; Mueller-Smith, 2015). Others argue that prison spells might act as a *school of crime* in which networks of criminals share knowledge and influence one another (Bayer et al., 2009; Stevenson, 2017). Conversely, incarceration may decrease crime through incapacitation and deterrence in

[1] In other jurisdictions, the evaluations are conducted before sentencing and are taken into account in the judges' decisions.

addition to providing a unique opportunity for intervention.[2] Recent papers have shown that incarceration in rehabilitation-friendly prisons can be beneficial. Landersø (2015) studies a reform in Denmark that lengthened sentences for violent offenders and shows that the reform increased employment and earnings. Bhuller et al. (2020) use a judge leniency design and provide causal evidence that inmates who serve longer sentences in Norway are less prone to reoffend. This effect is driven by previously unemployed offenders who are likely to gain from job training programs. However, papers focusing on the United States and using randomly assigned judges find no effects of incarceration on economic outcomes (Kling, 2006) or on recidivism (Green and Winik, 2010).

A related strand of the literature compares prisons with harsh treatment to rehabilitation-oriented ones. Lotti (2020) exploits a major reform in England and Whales which marked a shift from strict warehousing to rehabilitating of young offenders. The reform is found to have decreased recidivism substantially. Mastrobuoni and Terlizzese (2019) compare (harsh) *closed* to (rehabilitation-oriented) *open* prison regimes on recidivism. The endogeneity of assignment is accounted for by using variations in nearby prisons' overcrowding, and they find that open prisons yield lower recidivism. Tobón (2020) shows that exogenous assignment of inmates from older to newer Colombian prisons—which are less crowded and offer better conditions, services, and rehabilitation programs—substantially reduce recidivism. Finally, Hjalmarsson and Lindquist (2020) study the effects of Swedish reforms that lengthened sentences. They show that time spent in prison allows inmates with mental health issues to engage in therapy and to participate in various programs, resulting in long-lasting health benefits and reduced recidivism. These studies provide indirect evidence that rehabilitation programs are perhaps responsible for the observed link between incarceration and post-release outcomes. Our paper complements this literature by estimating the effect of several prison-based interventions and by providing evidence that they are indeed a driving force underlying rehabilitation.

The second contribution of the paper is to add to the growing literature that evaluates different types of programs aimed at offenders at varying stages of their criminal trajectory. An important strand of the literature focuses on programs aimed at young offenders. Heller et al. (2017) provide evidence that behavioral interventions may reduce recidivism among this population. They exploit three RCTs in Chicago in which at-risk youths and juvenile delinquents participated in programs and find a significant reduction in arrests and reincarcerations. Seroczynski et al. (2016) also study an RCT for a similar program and find that it decreases recidivism significantly. On the other hand, Armstrong (2003) uses an RCT to study similar interventions targeted at young offenders and finds no effect of participation.

Studies on programs targeted at adult offenders can be separated into two groups: prison based and external. Doleac et al. (2020) review three randomized controlled trials of external reentry programs and

[2] See Owens (2009); Buonanno and Raphael (2013) and Barbarino and Mastrobuoni (2014) for recent evidence documenting the importance of incapacitation effects of incarceration. See Chen and Shapiro (2007) for a review of empirical evidence of the deterrence effects of incarceration, and Mastrobuoni and Rivers (2016) for more recent evidence.

find, at best, mixed evidence of their effectiveness.[3] Blattman et al. (2017) study the effect of behavioral therapy and show that it can reduce violent crimes. Many studies have sought to estimate the effects of prison-based programs, but as mentioned earlier, this literature faces challenges due to selection issues.[4] However, recent studies have found encouraging results. Balafoutas et al. (2020) randomly asked inmates to reflect on their incarceration and show that this simple intervention increased the inmates' social aptitudes. Kuziemko (2013) exploits a 1998 policy reform which canceled parole eligibility for convicts in the state of Georgia. She presents evidence that inmates respond by decreasing their rehabilitation effort, including reducing their program participation while incarcerated, thus increasing recidivism likelihood. Macdonald (2020) exploits a reform in Arizona that eliminated judges' discretion in the decision to release inmates early. He shows that eliminating early release opportunities decreases recidivism, likely through reduced rehabilitation efforts in educational programs. Overall, convincing evidence that programs can benefit inmates exists but remains scarce. Our paper fills a gap in the literature by showing that prison-based interventions can impact adult inmates' reentry and also suggests that different types of programs (educational, vocational, and therapy) can be beneficial if well targeted.

Our third contribution adds to the growing literature on the use of inmates' risk assessment in judicial proceedings. Assessment tools used to predict recidivism are now widespread in most jurisdictions: bail and sentencing judges (Stevenson and Doleac, 2019; Albright, 2019), parole board members (Berk, 2017), and probation officers (Kopf, 2014) use the assessment scores to determine the level of risk. However, the extent to which risk assessments can be used to provide targeted rehabilitation services is less understood (Long et al., 2019). Mastrobuoni and Terlizzese (2019) argue that targeting might play an important role, as they do not find that high-risk offenders benefit from moving from closed prisons to open prisons like their lower-risk counterparts. We complement these findings by showing that risk assessments allow better targeting. Consistent with Mastrobuoni and Terlizzese (2019), we find that high-risk inmates benefit the least from rehabilitation programs. Further, we find that inmates with a high procriminal score do not benefit from programs.

The rest of the paper is structured as follows. In Section 2, we provide the institutional details of Quebec's criminal justice system and develop on the several rehabilitation programs offered. We present our results in Section 3, beginning with baseline regressions. We then show that the main results are not driven by selection and explore heterogeneity using various econometric techniques. Section 4 concludes and discusses policy implications.

[3]Other studies find mixed results. For instance, Cook et al. (2015) conduct an RCT on high-risk offenders in Wisconsin. They find a decrease in the likelihood of a rearrest but no difference in the likelihood of reincarceration.

[4]In her literature review, Doleac (2019) notes that "future work exploiting natural experiments or field experiments that avoid selection bias would be valuable for determining the power of educational programs—alone or in combination with other [prison-based] interventions—to encourage desistance from crime." Davis et al. (2013) and Visher et al. (2005) recognize the same gap in the literature.

2 Institutional Details and Data

In this section, we first provide institutional details on the prisons the data cover and on their social rehabilitation programs. We then describe the measure of inmates' risk assessment we use, discuss how we construct our recidivism variables, and present summary statistics.

2.1 Prisons in Canada

Two types of prisons shape Canada's justice system: penitentiaries, or federal prisons, administered by the Canadian government, and provincial prisons, administered by provincial governments. Adult individuals sentenced to less than two years are incarcerated in provincial prisons and thus fall under their provincial government's responsibility. Those sentenced to two years or more are incarcerated in federal prisons. The data used in this paper are provided by the province of Quebec's Ministry of Public Security and therefore comprise offenders who committed not so severe crimes, with an average sentence length of five months. Furthermore, individuals typically spend only two-third of their sentence incarcerated, so the actual time spent in incarceration is even shorter.[5] This population is highly relevant to investigate the effects of social rehabilitation programs. Because of their short sentences, inmates will leave prison in the short run and hopefully get the opportunity to rapidly reintegrate into society. Thus, the programs' effects may manifest themselves early in the observation period. Moreover, because these inmates are incarcerated for relatively minor offenses, rehabilitation into society is less likely to be out of reach.

The data come from DACOR (Administrative Correctional Files), an administrative database run by Quebec's Ministry of Public Security that oversees eighteen prisons in the province and accommodates around 5,000 inmates daily. These data detail each individual's trajectory when involved with the criminal justice system. They provide information on the crimes committed as well as characteristics of interest such as the inmate's age, number of dependents, and indigenous status. Offenders are tracked through time with unique identifiers, allowing us to monitor their comings and goings in the judicial system.

2.2 Social Rehabilitation Programs

In an effort to facilitate the social reintegration of their inmates, all provincial prisons offer a diversified set of programs. These programs' primary purposes are to turn the participants' time in prison into a unique opportunity for them to start reflecting, sharpen their sense of responsibility, and acquire skills that they will use to transition into the workforce. In this respect, prisons have established partnerships with other government bodies, such as the Ministry of Education and Higher Education and the Ministry of Labour, to offer tailored programs to match the labor market demand. Social organizations and external agencies

[5]Individuals who are granted parole spend only one-third of their sentence incarcerated, but only a small portion of individuals in our sample are granted parole. Less than 35% of individuals are eligible for a parole audience (which requires having a sentence or at least 6 months). Among those eligible, less than 30% are granted parole. This represents less than 10% of our whole sample.

can also be responsible for developing other types of programs.

Among all male provincial prisons, three of the largest closely recorded the data on inmates' program participation and agreed to share these data.[6] We were thus able to match these prisons' data on participation in programs with the DACOR files through the unique identifier. The program data comprise participants in more than 150 unique programs. With the guidance of authorities responsible for managing the programs, we devised an intuitive classification based on the criminogenic risks and needs targeted by each. We divided programs into six categories: self-development, violence, addiction, education, job skills, and *other*. Table 1 provides the different categories of programs as well as the share of observations participating in each. We also list examples of programs in each category.

Table 1: Program Classification

Category	Share of Participants	Examples of Programs
Self-Development	0.365	stress/anger management, problem solving, accountability
Violence	0.054	violence/aggressiveness management, domestic violence
Addiction	0.166	substance abuse, drugs/alcohol addiction
Education	0.267	literacy, languages, mathematics
Job skills	0.075	resume building, laundry, construction work
Other	0.074	leisure, art, spirituality
Total	1,00	
Observations	3,888	

It is not rare for participants to enroll, concurrently or consecutively, in multiple programs during the same sentence. For instance, we observe participants who completed programs in both the violence and addiction categories. In those cases we record the category in which the participant has completed most programs.[7] In the case of an equal number of programs in two or more categories, one of them is picked randomly. Additionally, all the participants become involved in the programs voluntarily. Inmates can be recommended programs or be encouraged to enroll based on an assessment made at the onset of incarceration. We specifically discuss this assessment in the next section.

2.3 Risk Evaluation

At the onset of incarceration, every inmate's physical and mental conditions are assessed by a qualified professional. The assessment's extent depends on the incarceration length. Inmates incarcerated for less than six months receive a short assessment conducted by a correctional officer within seven days before the sixth of their sentence. Those with a sentence of six months or more receive a substantially more thorough assessment and are evaluated using a comprehensive, standardized risk assessment tool. Starting in 2007, probation officers have employed the "Level of Service/Case Management Inventory" (LS/CMI,

[6]These are the *Établissement de détention de* Montreal, Quebec and St-Jerome.

[7]For instance, if a participant engages in two stress management therapies and one domestic violence program, we keep self-development as the program category for this individual The results are virtually unchanged if we keep, alternatively, either the first or last category in which the inmate has participated.

afterward) to conduct such evaluations.[8] The LS/CMI is a widely implemented proprietary assessment tool in North America (Andrews et al., 2000) and is the culmination of ulterior versions of tools under the "Level of Supervision Inventory" family.[9] Probation officers must complete the evaluation seven days before the sixth of the sentence or 45 days after sentencing, whichever comes first. If an inmate is re-incarcerated, a new evaluation is not deemed necessary if a prior evaluation is still considered appropriate and was completed less than two years prior. In the rest of this paper, we refer to *evaluated* inmates as those assessed using the thorough LS/CMI evaluation.

The LS/CMI questionnaire comprises eight components, each containing between two and nine questions, for a total of 43 questions. Completing the LS/CMI involves a few hours of work since some questions require extensive research in administrative and health files. To get a reliable sense of criminogenic characteristics absent from the files, the probation officer (the evaluator) runs an interview with the offender. During the interview, he or she gauges how the inmates' social network, recreational habits, and procriminal attitude could cause them to recidivate after release. The questionnaire gives the inmate a score for each component.

Table 2 presents the eight components and their summary statistics in the sample. For each of these categories, a higher score means a higher risk of recidivism as predicted by the evaluator. For example, the first component, "Criminal History," measures to what extent the inmate is judged to have past criminal behaviors that make him more at risk of committing crimes again. The other categories are "Education/Employment," "Family/Marital," "Leisure /Recreation," "Companions," "Alcohol/Drug," "Procriminal Attitude," and "Antisocial Pattern." The eight scores are also aggregated by simply summing them up, yielding a total score out of 43 which categorizes inmates in terms of recidivism risk: very low (0–4), low (5–10), medium (11–19), high (20–29), or very high (30–43). Appendix Table A.1 provides a reproduction of the questionnaire.

Table 2: LS/CMI Scores

LS/CMI Component	Range	Mean	SD	Observations
Criminal History	0 to 8	5.59	2.03	8,767
Education/Employment	0 to 9	4.96	2.83	8,767
Family/Marital	0 to 4	1.88	1.17	8,767
Procriminal Attitude	0 to 4	1.51	1.29	8,767
Companions	0 to 4	2.50	1.05	8,767
Leisure /Recreation	0 to 2	1.53	0.64	8,767
Alcohol/Drug	0 to 8	3.93	2.33	8,767
Antisocial Pattern	0 to 4	1.83	1.23	8,767
Total score	0 to 43	23.73	8.66	8,767

Aside from providing risk scores, the LS/CMI identifies some of the offenders' needs with respect to rehabilitation. To this end, evaluators use the scores obtained in each section to outline a tailored intervention plan. For instance, an inmate with a high score in the Alcohol/Drug Problem section might

[8]Only trained probation officers can conduct the assessment.
[9]For a broad review, we refer to Vose et al. (2008).

be recommended to a program focusing on addiction. Once the evaluator has devised an intervention plan, the case is transferred to another agent for the remaining of the sentence and the inmate decides whether or not to follow the initial recommendations.

2.4 Recidivism Measures

The very definition of recidivism is a point of debate in the literature and across jurisdictions.[10] In this study, we define recidivism as an incarceration sentence arising from another conviction during a fixed follow-up period after release. The results are virtually unchanged if we, alternatively, record any subsequent sentence (even sentences that do not include serving time) as recidivism. The results are also not sensitive to calculating the fixed follow-up period based on the initial planned released instead of the actual release.[11] To meet with the overall consensus, the main results will consider recidivism within three years following release. We vary the follow-up period from one to five years to provide robustness checks, discarding each time observations we do not observe long enough.

2.5 Descriptive Statistics

The LS/CMI evaluation and scores are key components in our analysis, as they are designed in part to predict recidivism and various criminogenic problems. Figure 1 presents the relationship between the score for each component of the LS/CMI evaluation and recidivism within three years in the sample. For each component, a higher score implies a higher predicted risk, as measured by the evaluator. The figure shows that evaluators' predictions are consistent with observed recidivism: the proportion of ex-offenders recidivating increases with the attributed score for each component. The components displaying the strongest relationships are criminal history, education and employment, alcohol and drugs, and antisocial pattern. For these components, the lowest score predicts a recidivism rate of around 20%, while the highest score predicts one of around 80%.

As mentioned above, the scores of the eight components are often aggregated into one single total score to predict recidivism. Figure 2 presents the strong relationship between this total score and recidivism within three years. For the lowest total scores, the rate of recidivism is close to zero. The recidivism rate increases steadily with the total score and comes close to one for the highest total scores. The figure also shows a kernel-smoothed density of total LS/CMI scores in the sample. Few inmates are at either extreme of possible scores. The highest peak is around 29, where close to 70% of inmates recidivate within three

[10]For instance, Kuziemko (2013) considers as a recidivist an inmate returning to prison within three years. In contrast, Bhuller et al. (2020) define recidivism as the event of being charged with at least one crime during a given period. Quebec's Ministry of Public Security and Public Safety Canada consider recidivism a new decision for another offense limited to two years following to end of the sentence. In contrast, the United States National Institute of Justice measures recidivism as any subsequent involvement with the criminal justice system within three years following release, whether or not a new sentence was issued.

[11]This is not surprising in our context, because the vast majority of inmates are released near their planned date. Less than 10% of individuals in our sample are granted parole. See Footnote 5 for further details.

years.

We now compare participants in programs with non-participants. Table 3 provides summary statistics and measures of differences between the two groups. The "Evaluation" variable indicates whether an inmate is assessed using a thorough LS/CMI evaluation. Participants are more likely to be evaluated than non-participants (66.3% versus 40.7%). The next lines show differences in average LS/CMI scores. Note that these comparisons are only made within the subsample of evaluated inmates, thus the smaller number of observations. The total LS/CMI scores are strikingly similar in the two groups. Given that the total score is extremely predictive of recidivism (see Figure 2), this similarity suggests there are little differences in participants' and non-participants' propensity to recidivate in the absence of programs. The average score indicating problems related to criminal history, which is also strongly related to recidivism, is also similar between the two groups. However, we find statistically significant differences in average scores for other components. The differences are, however, small compared with the standard deviations of scores.

We divide the types of crime committed into four categories: assault, burglary and theft, drug crimes, and *other*. Thus these four types of crimes presented in the table are mutually exclusive. Participants and non-participants tend to differ regarding these: participants are more likely to have committed assault or burglary and theft crimes than a crime in the *other* category. We also find that participants tend to be slightly younger than non-participants and the share of individuals belonging to an indigenous ethnic group is higher among participants.[12]

Table 3: Summary Statistics for Participants and Non-Participants

	Non-Participants			Participants			Diff.
	Mean	SD	Obs.	Mean	SD	Obs.	p-value
	(a)	(b)	(c)	(d)	(e)	(f)	(g)
Evaluation	0.407		15,205	0.663		3,898	0.000
LS/CMI: Total Score (0 to 43)	23.731	8.713	6,184	23.752	8.527	2,583	0.919
LS/CMI: Crim. History (0 to 8)	5.577	2.046	6,184	5.632	1.985	2,583	0.246
LS/CMI: Educ./Empl. (0 to 9)	4.918	2.844	6,184	5.069	2.786	2,583	0.023
LS/CMI: Family/Marital (0 to 4)	1.897	1.174	6,184	1.847	1.150	2,583	0.071
LS/CMI: Procrim. Attitude (0 to 4)	1.538	1.311	6,184	1.428	1.230	2,583	0.000
LS/CMI: Companions (0 to 4)	2.483	1.081	6,184	2.551	0.987	2,583	0.006
LS/CMI: Leisure/Recreation (0 to 2)	1.556	0.635	6,184	1.480	0.661	2,583	0.000
LS/CMI: Alcohol/Drug (0 to 8)	3.902	2.303	6,184	3.997	2.381	2,583	0.080
LS/CMI: Antisocial Pattern (0 to 4)	1.861	1.231	6,184	1.748	1.212	2,583	0.000
Crime: Other	0.362		15,205	0.268		3,898	0.000
Crime: Assault	0.149		15,205	0.180		3,898	0.000
Crime: Burglary and theft	0.165		15,205	0.215		3,898	0.000
Crime: Drugs	0.324		15,205	0.337		3,898	0.136
Age < 30	0.323		15,205	0.301		3,898	0.007
$30 \leq$ Age < 40	0.251		15,205	0.252		3,898	0.943
Age ≥ 40	0.426		15,205	0.447		3,898	0.014
At least one dependent	0.309		15205	0.403		3,898	0.000
Indigenous	0.025		15,205	0.104		3,898	0.000

All variables except LS/CMI scores are binary variables.
Column (g) reports p-values for tests of differences in means or proportions between columns (a) and (d).

[12]The higher share of individuals belonging to an indigenous ethnic group among participants is explained by some programs being specifically targeted for this population in one of our three prisons.

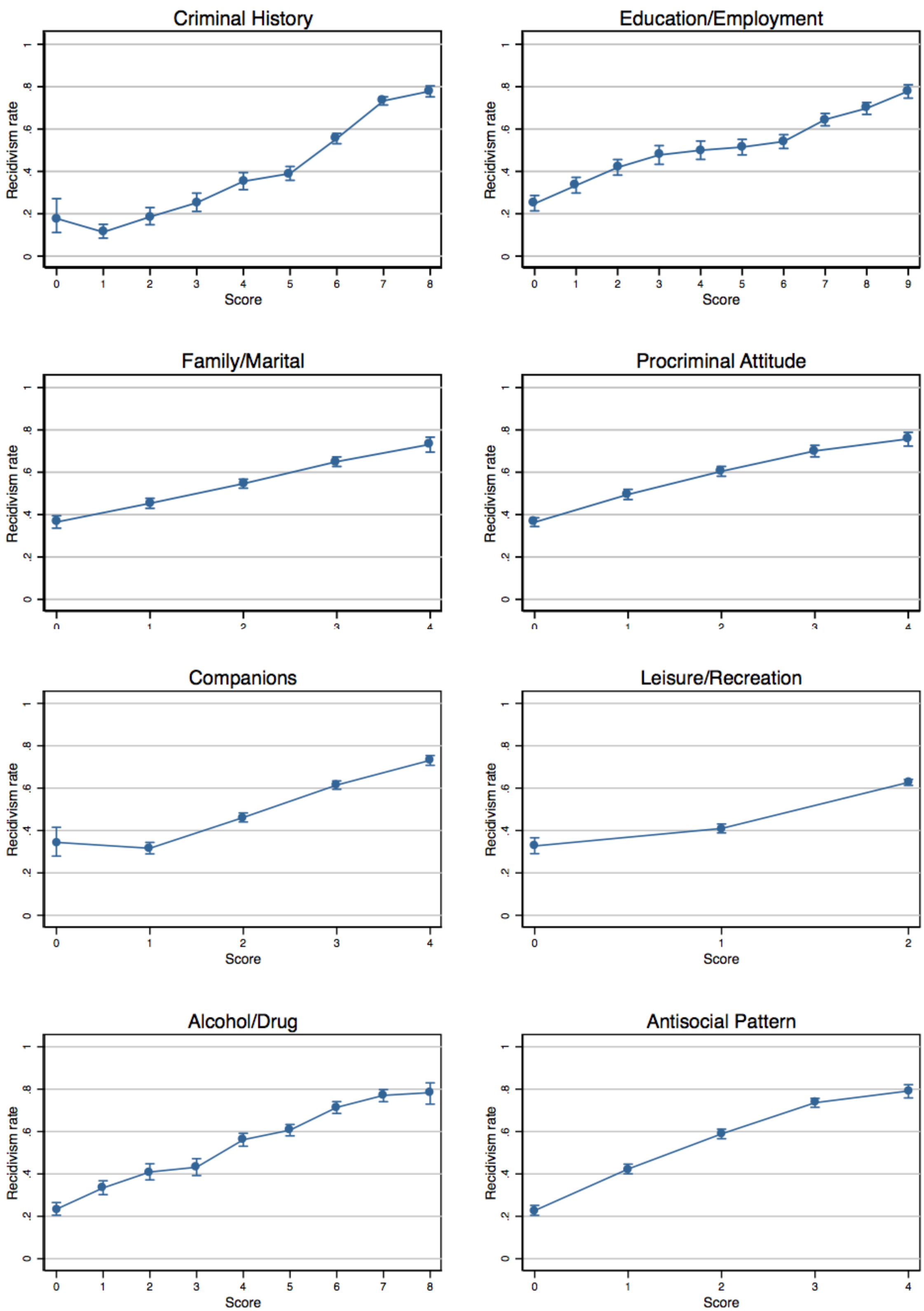


Figure 1: Relationships between LS/CMI Scores and Recidivism within Three Years

Note: 95% confidence intervals for proportions are reported.

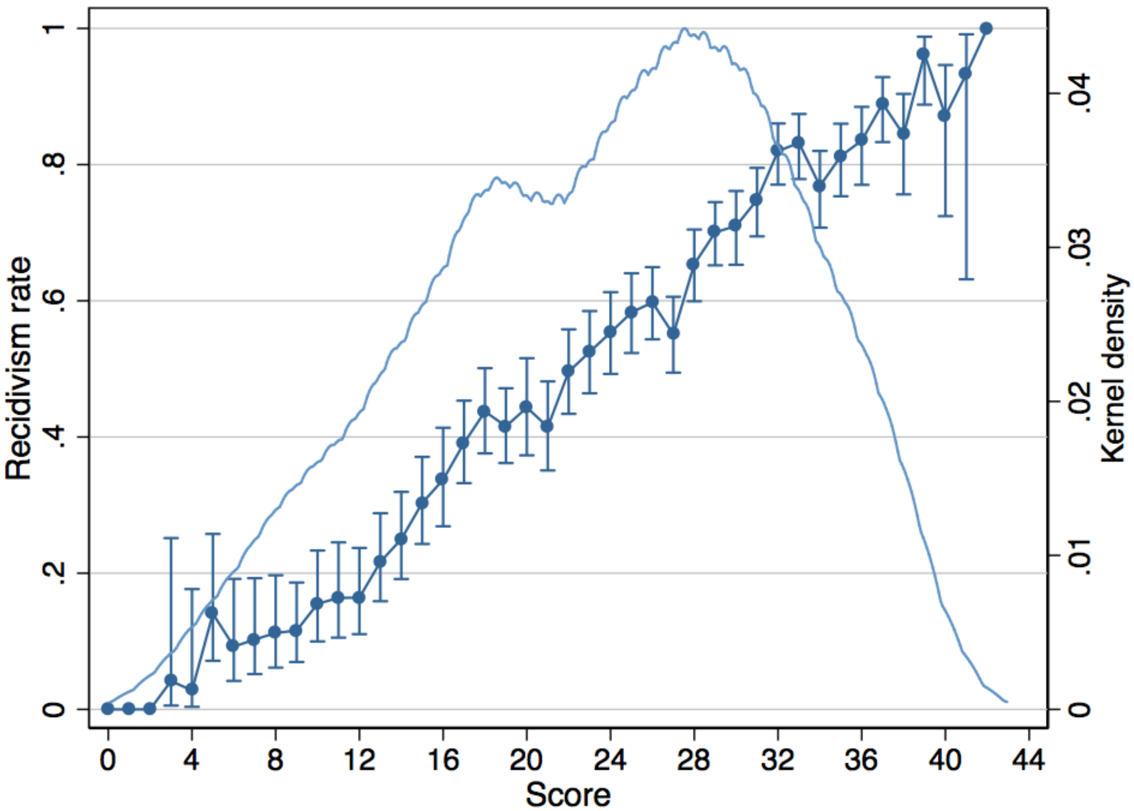


Figure 2: Relationship between Total LS/CMI Score and Recidivism within Three Years

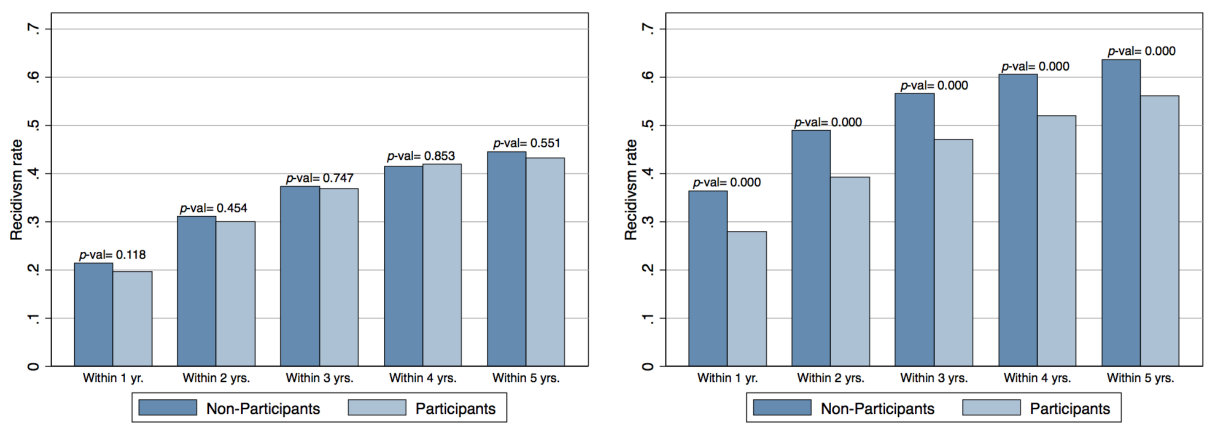
Notes: 95% confidence intervals for proportions are reported for recidivism rate. Kernel density is estimated using an Epanechnikov kernel and a Silverman bandwidth.

3 Estimation and Results

In this section, we first present our main results on the effect of participating in programs on recidivism. We then explore robustness to potential selection on unobserved characteristics using the approach from Oster (2019). Finally, we estimate heterogeneous effects of participation using modern machine learning techniques.

3.1 Effect of Participating in Programs

The histograms in Figure 3 plot the recidivism rates for participants and non-participants for several intervals. Figure 3a presents rates for inmates not evaluated by the LS/CMI, and Figure 3b shows such rates for evaluated inmates. While there is no evidence of any difference in recidivism between participants and non-participants for inmates who did not receive an evaluation, recidivism is substantially lower for participants among those who were evaluated regardless of the duration considered when generating recidivism.



(a) Prisoners Not Evaluated by LS/CMI

(b) Prisoners Evaluated by LS/CMI

Figure 3: Recidivism Rates by Intervals

Note: p-values are for tests of difference of proportions between recidivism rates for participants and non-participants.

We now turn to ordinary least squares (OLS) regressions, focusing on recidivism within three years.[13] Our main specification takes the following form:

$$y_i = \beta_0 + \mathbb{1}\left[eval_i = 0\right] \times \beta_1 T_i + \mathbb{1}\left[eval_i = 1\right]\left(\alpha + \beta_2 T_i + \mathbf{s}_i'\boldsymbol{\omega}\right) + \mathbf{x}_i'\boldsymbol{\theta} + \epsilon_i, \tag{1}$$

where y_i is an indicator equal to 1 if inmate i recidivates within three years following release and T_i is a program participation binary variable. The binary variable $eval_i$ indicates whether inmate i's risk was

[13] The main results presented below are robust to estimating the model using a probit estimation. The advantage of the linear probability model is that it simplifies the presentation of the marginal effects for interaction variables.

assessed with the LS/CMI evaluation. If they were evaluated, we include a vector of scores for components of the LS/CMI in the regression ($\mathbf{s}_i$); we set these scores to 0 for individuals not evaluated.[14] $\mathbf{x}_i$ is a vector of other controls, including fixed effects for the sentence year and for the prison, dummy variables for the category of crime, age groups, whether the inmate has dependents, and whether he is part of an indigenous ethnic group. The two coefficients of interest are β_1 and β_2, which capture the effects of participating in programs for inmates not evaluated and evaluated, respectively.

Table 4 presents the results. Column (a) simply regresses recidivism within three years on participation for all inmates and shows no evidence of any effect. Column (b) separates the effect of programs according to whether or not participants were evaluated by the LS/CMI and adds a dummy variable controlling for the effect of being evaluated (but not for the scores). Evaluated inmates are substantially more at risk of recidivating (19.2 percentage points) than those not evaluated. Consistent with the evidence presented in Figure 3, we only find a significant effect of programs for participants who were evaluated by the LS/CMI (–9.6 percentage points).

Column (c) adds as a control the total LS/CMI score, which is significantly associated with an increase in the probability to recidivate. This score being very predictive of recidivism, the R^2 increases from 0.03 to 0.13 with respect to column (b). Remarkably, this only marginally affects our estimate of the effect of programs on recidivism for evaluated participants (–9.3 percentage points), suggesting our estimates of the effect of programs are not driven by selection (we elaborate on this point in the next subsection). Unsurprisingly, the effect of the dummy *evaluation* variable drops substantially, because it now measures the effect of being evaluated for individuals who would have an LS/CMI score of zero. Column (d) controls for each component of the LS/CMI separately instead of using the total score, which increases the R^2 to 0.142 while still leaving our estimate of the effect of programs unchanged (–9.6 percentage points for evaluated participants). Finally, column (e) includes our full vector of controls. This further increases our R^2 to 0.19 but leaves our estimate of the effect of programs essentially unchanged (–9.4 percentage points).

3.2 Robustness to Unobserved Selection

Thus far, one might be concerned about unobserved selection driving the results from our main specification. The error term could contain critical unobserved characteristics, such as the level of motivation, remorse, or ability, which could be correlated with both the selection into treatment and recidivism. In this section, we adopt the approach of Oster (2019) to argue that the omitted variables bias in our setting is negligible. The idea behind her approach is to compare the change in the estimated coefficient to the change in the R^2 after adding additional control variables. The approach draws from the Altonji et al. (2005) framework, which suggests a method for formally relating selection on observables to selection on unobservable characteristics. Under the assumption that selection into programs based on observable

[14] Note that the parameter α allows for individuals not evaluated to differ from those evaluated with a score of zero.

Table 4: OLS Regressions: Recidivism within Three Years (Dependent Variable)

	(a)	(b)	(c)	(d)	(e)
Program	−0.016				
	(0.010)				
Program× Eval.=0		−0.006	−0.006	−0.006	−0.039**
		(0.017)	(0.017)	(0.017)	(0.017)
Program× Eval.=1		−0.096***	−0.093***	−0.096***	−0.094***
		(0.013)	(0.011)	(0.011)	(0.012)
Eval.		0.192***	−0.438***	−0.473***	−0.432***
		(0.009)	(0.015)	(0.018)	(0.018)
LS/CMI: Total Score			0.229***		
			(0.004)		
LS/CMI: Crim. History				0.116***	0.110***
				(0.007)	(0.007)
LS/CMI: Educ./Empl.				0.037***	0.019***
				(0.007)	(0.007)
LS/CMI: Family/Marital				0.009	0.010*
				(0.006)	(0.006)
LS/CMI: Procrim. Attitude				0.018***	0.024***
				(0.007)	(0.007)
LS/CMI: Companions				0.012*	0.011
				(0.007)	(0.007)
LS/CMI: Leisure /Recreation				0.005	0.012**
				(0.006)	(0.006)
LS/CMI: Alcohol/Drug				0.084***	0.077***
				(0.006)	(0.006)
LS/CMI: Antisocial Pattern				0.040***	0.023***
				(0.009)	(0.009)
Constant	0.455***	0.373***	0.373***	0.373***	0.486***
	(0.005)	(0.006)	(0.006)	(0.006)	(0.026)
Other controls	No	No	No	No	Yes
Observations	15,047	15,047	15,047	15,047	15,047
R^2	0.000	0.031	0.133	0.142	0.191

Each LS/CMI component is normalized by its standard deviation in the sample.
Heteroskedasticity-robust standard errors are reported in parentheses.
*** $p \leq 0.01$; ** $p \leq 0.05$; * $p \leq 0.1$
Other controls include year of sentence fixed effects and dummy variables for prisons, age groups, having at least one dependent and belonging to an indigenous ethnic group.

variables that predict recidivism is similar to selection based on unobservable variables that also predict recidivism (we clarify the meaning of *similar* below), comparing the effect of adding controls to the regression, scaled by these controls predictive value, is informative about the selection bias arising from unobservables. To illustrate the approach, consider the following regression model:

$$y = \beta_0 + \beta T + \Psi\omega^o + W_2 + \epsilon, \tag{2}$$

where β is the treatment effect we seek to estimate, ω^o is the vector of observable characteristics, and W_2 is the index of unobserved characteristics affecting both T and y. Define the observable index as $W_1 = \Psi\omega^o$. Altonji et al. (2005) and Oster (2019) relate selection on observables and unobservables using the following relationship:

$$\delta\frac{\sigma_{1T}}{\sigma_1^2} = \frac{\sigma_{2T}}{\sigma_2^2}, \tag{3}$$

where $\sigma_{jT} = cov(W_j, T)$ and $\sigma_j^2 = var(W_j)$ for $j \in \{1, 2\}$. Equation 3 implies that selection on observable characteristics is proportional to selection on unobservable characteristics with a coefficient of proportionality δ. To simplify the intuition behind the approach, it is useful to assume, first, that the observables and the unobservables are equally related to participation (i.e., $\delta = 1$) and, second, that the relative importance of each observable in explaining y is equal to its importance in explaining T (i.e., $\frac{\psi_i}{\psi_j} = \frac{\mu_i}{\mu_j}\ \forall i, j$, where μ_i is the coefficient of a regression of T on ω^o). Under these two assumptions, Oster (2019) shows that an unbiased and consistent estimator of β is given by

$$\beta^* = \tilde{\beta} - \left[\mathring{\beta} - \tilde{\beta}\right] \frac{R_{\max} - \tilde{R}}{\tilde{R} - \mathring{R}}, \tag{4}$$

where $\mathring{\beta}$ and $\mathring{R}$ are the estimate of the treatment effect and the R^2 obtained from a regression not adding the controls, and $\tilde{\beta}$ and $\tilde{R}$ are those obtained from a regression adding the controls. $R_{\max}$ is the theoretical R^2 that would be obtained could the researcher include all relevant variables explaining both participation and recidivism. It equals one only if all variables affecting recidivism also affect participation ($\epsilon = 0$). Equation (4) provides the straightforward intuition underlying the estimator. Assume, for example, an $R_{\max}$ of 1, and that adding controls increases the R^2 from $\mathring{R} = 0.25$ to $\tilde{R} = 0.5$. The added proportion of the variance explained by controls is thus 0.25, and there still remains twice this share (0.5) to be explained by unobservables (i.e., $\frac{R_{\max} - \tilde{R}}{\tilde{R} - \mathring{R}} = 2$). Under the assumption that $\delta = 1$, it follows that the correction needed to control for selection on unobservables is then twice that of the correction needed to control for selection on observables. Importantly, we do not rely on the assumptions stated above by simply computing equation (4). Oster (2019) shows that, to obtain an unbiased and consistent estimate of the treatment effect, we only need to assume that our estimated index of observable variables predicts the true direction of the relationship between this index and participation. We therefore use her *unrestricted* estimator, which only assumes that our control variables can predict the right direction of selection into programs.

The approach necessitates that the researcher sets values for $R_{\max}$ and δ. Because they need to be set arbitrarily, it is useful to think of these values as upper bounds that will provide a lower bound (in absolute value) of the treatment effect. Although the theoretical bound of 1 for $R_{\max}$ could be seen as a relevant conservative choice, Oster (2019) argues that it will often be appropriate to use a lower value since, first, measurement errors on the outcome are captured by ϵ, not W_2. Second, ϵ captures all factors arising after participation in a program occurred. This is especially relevant in our context, as the decision to recidivate will potentially be affected by numerous events happening after inmates leave prison. It thus seems more realistic to set an upper bound lower than 1 for R^2. Analyzing results from randomized experiments recently published in top economics journals, and assuming that the estimates in those papers are truly causal so that they should survive the above correction, Oster (2019) suggests setting $R_{\max} = 1.3\tilde{R}$. In our main analysis below, we set a conservative bound of $R_{\max} = 1.5\tilde{R}$. As for δ, Oster (2019) provides empirical evidence that a value of 1 will often be appropriate. We thus set δ to 1 for our main results.

Consider our *baseline* regression that separates the effects of programs for evaluated and non-evaluated individuals without controlling for observable characteristics and LS/CMI scores:

$$y_i = \beta_0 + \mathbb{1}\left[eval_i = 0\right] \times \mathring{\beta}_1 T_i + \mathbb{1}\left[eval_i = 1\right]\left(\alpha + \mathring{\beta}_2 T_i\right) + \epsilon_i. \tag{5}$$

This regression corresponds to the model estimated column (b) of Table 4. Recall from the last subsection that we only find evidence of an effect of programs for evaluated individuals, so the coefficient of interest in this regression is $\mathring{\beta}_2$. Control variables already included in this baseline regression are assumed not to be part of the *confounding set* (i.e., the set of observables proportionally related to unobservables).[15] We compare the results from this regression to those from the following *controlled* regression, which adds our full set of controls:

$$y_i = \beta_0 + \mathbb{1}\left[eval_i = 0\right] \times \tilde{\beta}_1 T_i + \mathbb{1}\left[eval_i = 1\right]\left(\alpha + \tilde{\beta}_2 T_i + \mathbf{s}_i'\boldsymbol{\omega}\right) + \mathbf{x}_i'\boldsymbol{\theta} + \epsilon_i, \tag{6}$$

where $\mathbf{s}_i'$ and $\mathbf{x}_i'$ are the vectors of LS/CMI scores and individual characteristics. This regression corresponds to the model estimated in column (e) from Table 4. Note that we label our coefficient of interest from this regression by $\tilde{\beta}_2$ and the R^2 from this regression by $\tilde{R}$.

We estimate Oster (2019)'s unrestricted estimator from these two regressions. Table 5 presents the results, varying the follow-up window within which we consider recidivism to provide additional robustness. The first row shows our estimate of $\mathring{\beta}_2$ in equation (5). Regardless of the follow-up window, we find large and significant effects of participation on recidivism and low values of R^2. Treatment effects estimated from the controlled regression (equation (6)) barely move, although adding the controls substantially increases the regressions' R^2. This indicates that the controls account for a sizable portion of the variance in recidivism, yet such essential covariates prove to be negligibly correlated with the treatment effect. The fourth row then presents Oster's unrestricted estimator, assuming $\delta = 1$. The method yields estimates largely in line with the regression estimates. In all cases, we estimate precise negative and significant effects of program participation on recidivism.

The unrestricted estimator can be seen as an upper bound (or lower bound in absolute value) of the treatment effect because it is a function of $R_{\max}$ and δ, which are set to arguably high values. Still, one might not be satisfied with this bound if they are not satisfied with these arbitrary values set by the researcher. We therefore conduct the two following exercises suggested by Oster (2019). First, we find the value δ_0 of δ that would be required to find an estimate of the treatment effect of zero conditional on setting $R_{\max} = 1.5\tilde{R}$. We find that this value ranges from 5.7 to 16.2, which are substantially above the value

[15] The reason for this choice is that we only control for evaluation and for participation for non-evaluated individuals to separate the effect of programs for both groups; these controls are not truly characteristics of evaluated individuals that we can relate to selection within this group.

considered appropriate by Oster (2019). In our context, such large values would indicate that unobservable variables should play a disproportionately important role to explain enrollment in programs relative to the observable controls, including LS/CMI risk scores that are considerably related to recidivism. Second, setting $\delta = 1$ again, we find the value of $R_{\max}$ that would be needed to estimate a treatment effect of 0. We obtain values again substantially above those considered appropriate by Oster (2019) and, in some cases, still find a negative treatment effect for $R_{\max} = 1$.

Table 5: Coefficient Stability Checks

	Recidivism within...				
	1 year (a)	2 years (b)	3 years (c)	4 years (d)	5 years (e)
Baseline effect (robust s.e.) $[\mathring{R}]$	-0.085*** (0.011) [0.024]	-0.097*** (0.014) [0.028]	-0.096*** (0.013) [0.031]	-0.085*** (0.015) [0.030]	-0.075*** (0.018) [0.032]
Controlled effect (robust s.e.) $[\tilde{R}]$	-0.080*** (0.010) [0.131]	-0.096*** (0.011) [0.175]	-0.094*** (0.012) [0.191]	-0.072*** (0.013) [0.193]	-0.063*** (0.015) [0.183]
$R_{\max} = 1.5\tilde{R}$	0.197	0.263	0.287	0.289	0.275
Oster (2019) unrestricted estimator (bootstrapped s.e.) given $R_{\max} = 1.5$ and $\delta = 1$	-0.077*** (0.011)	-0.096*** (0.011)	-0.092*** (0.013)	-0.063*** (0.015)	-0.055*** (0.017)
δ_0 for $\beta = 0$ given $R_{\max} = 1.5\tilde{R}$	10.604	16.187	13.101	5.522	5.718
$R_{\max}$ for $\beta = 0$ given $\delta = 1$	0.855	†	†	0.732	0.713

Controls in the confounding set (i.e., related to the set of proportionally related unobservables) include all the LS/CMI components scores; year of sentence fixed effects; and dummy variables for prisons, age groups, having at least one dependent, and belonging to an indigenous ethnic group. Controls in the non-confounding set (i.e., unrelated to the set of proportionally related unobservables) include the binary variable $eval_i$ and an interaction variable between participation to a program and $\mathbb{1}[eval_i = 0]$.
Heteroskedasticity-robust standard errors are reported in parentheses; R^2 are reported in brackets.
Bootstrapped standard errors are obtained with 500 replications.
*** $p \leq 0.01$; ** $p \leq 0.05$; * $p \leq 0.1$
† : β^* never reaches 0 even with $R_{\max} = 1$.

3.3 Effects by Program Types

We now estimate separate effects of different types of programs. With the help of prison experts and the administrators responsible for managing the programs, we divided the broad set of programs into six categories: self-development, job skills, educational training, addiction problems, violence issues, and *other*.[16] We define a set of six mutually exclusive dummy variables T_i^j, where $j \in \{1, ..., 6\}$, and set T_i^j equal to 1 if i spent most of his participation time in program j and to 0 otherwise. We then estimate

[16]The *other* category includes programs related to leisure, art, spirituality, and zootherapy.

the following model:

$$y_i = \beta_0 + \mathbb{1}\left[eval_i = 0\right] \sum_{j=1}^{6} \beta_1^j T_i^j + \mathbb{1}\left[eval_i = 1\right] \left(\alpha + \sum_{j=1}^{6} \beta_2^j T_i^j + \mathbf{s}_i' \boldsymbol{\omega} \right) + \mathbf{x}_i' \boldsymbol{\theta} + \epsilon_i. \tag{7}$$

The β_1^js and β_2^js are the parameters of interest that capture the effects of participating in a program of type j compared to not participating at all for participants evaluated and not evaluated, respectively. Table 6 presents the results, each column adding more controls. For inmates not evaluated, we only find a significant effect for programs related to self-development (between -6.3 and -8.7 percentage points). Furthermore, we reject the null hypothesis that all programs for non-evaluated inmates have the same effect with a p-value of less than 0.05. For evaluated inmates, we find strong significant effects for all programs, with the exception of those in the *other* category. Nevertheless, we cannot formally reject that all effects are equal at any conventional level of significance.

It may come as a surprise that such different types of programs seem to yield similar effects on recidivism for evaluated individuals. A potential explanation is that program officers can better match inmates to appropriate programs that will truly favor rehabilitation if they are evaluated. As previously mentioned, the LS/CMI evaluation is used precisely to assess inmates' needs. To empirically validate that evaluations are used to target inmates to an appropriate program, we estimate a multinomial logistic regression on all evaluated participants using our six types of programs as the dependent categorical variable and the eight LS/CMI components (divided by their standard deviation) as the explanatory variables.

Table 7 presents the relative risk ratios from the estimation and shows evidence of targeting into programs. First, an increase in the score measuring problems related to education and employment greatly increases the odds of participating in an education program or a job skills program, relative to a program in the *other* category. Second, an increase in the score measuring alcohol and drug problems significantly increases the odds of participating in an addiction-related program, although this effect is only slightly statistically significant. We find a stronger relationship between problems related to leisure and recreation and participation in an addiction-related program. This intuitively makes sense, as alcohol and drug consumption during leisure time may be included in problems related to leisure. We also find that the criminal history score is strongly related to participation in a violence-related program, though this effect relates less intuitively to targeting.

We estimate another multinomial logistic model using as explanatory variables the types of crime committed that led to the sentence. Table 8 presents the results, and we again find evidence of targeting. Having committed an assault-related crime greatly increases the odds of participating in a violence-related program, and having committed a drug-related offense significantly increases the odds of participating in an addiction-related program.

Table 6: OLS Regressions: Effects by Program Type—Recidivism within Three Years (Dependent Variable)

	(a)	(b)	(c)	(d)
Program if not Evaluated				
Other	−0.030	−0.030	−0.030	−0.079*
	(0.047)	(0.047)	(0.047)	(0.047)
Self-Development	−0.063**	−0.063**	−0.063**	−0.087***
	(0.027)	(0.027)	(0.027)	(0.026)
Violence	0.198*	0.198*	0.198*	0.155
	(0.108)	(0.108)	(0.108)	(0.106)
Addiction	−0.011	−0.011	−0.011	0.019
	(0.058)	(0.058)	(0.058)	(0.056)
Education	0.026	0.026	0.026	−0.019
	(0.031)	(0.031)	(0.031)	(0.029)
Job skills	0.073	0.073	0.073	0.026
	(0.052)	(0.052)	(0.052)	(0.050)
p-value H_0: $\beta_1^j = \beta_1^k \ \forall \ \{j,k\}$	0.033	0.033	0.034	0.047
Program if Evaluated				
Other	−0.085**	−0.048	−0.054	−0.034
	(0.041)	(0.034)	(0.034)	(0.032)
Self-Development	−0.097***	−0.087***	−0.089***	−0.103***
	(0.021)	(0.017)	(0.017)	(0.018)
Violence	−0.080	−0.151***	−0.159***	−0.165***
	(0.049)	(0.046)	(0.046)	(0.046)
Addiction	−0.096***	−0.126***	−0.136***	−0.086***
	(0.026)	(0.023)	(0.023)	(0.023)
Education	−0.111***	−0.090***	−0.087***	−0.096***
	(0.023)	(0.019)	(0.019)	(0.019)
Job skills	−0.077**	−0.064*	−0.060*	−0.075**
	(0.038)	(0.033)	(0.033)	(0.033)
p-value H_0: $\beta_2^j = \beta_2^k \ \forall \ \{j,k\}$	0.971	0.262	0.139	0.233
Controls				
Eval.	0.192***	−0.440***	−0.477***	−0.434***
	(0.009)	(0.015)	(0.018)	(0.018)
LS/CMI: Total Score		0.230***		
		(0.004)		
LS/CMI: Crim. History			0.117***	0.110***
			(0.007)	(0.007)
LS/CMI: Educ./Empl.			0.037***	0.019***
			(0.007)	(0.007)
LS/CMI: Family/Marital			0.009	0.010*
			(0.006)	(0.006)
LS/CMI: Procrim. Attitude			0.018***	0.025***
			(0.007)	(0.007)
LS/CMI: Companions			0.012*	0.011*
			(0.007)	(0.007)
LS/CMI: Leisure/Recreation			0.006	0.012**
			(0.006)	(0.006)
LS/CMI: Alcohol/Drug			0.084***	0.076***
			(0.006)	(0.006)
LS/CMI: Antisocial Pattern			0.040***	0.023***
			(0.009)	(0.009)
Constant	0.374***	0.374***	0.374***	0.488***
	(0.006)	(0.006)	(0.006)	(0.026)
Other controls	No	No	No	Yes
Observations	15,047	15,047	15,047	15,047
R^2	0.032	0.134	0.143	0.192

Heteroskedasticity-robust standard errors are reported in parentheses.
*** $p \leq 0.01$; ** $p \leq 0.05$; * $p \leq 0.1$
Other controls include year of sentence fixed effects and dummy variables for prisons, age groups, having at least one dependent, and belonging to an indigenous ethnic group.

Table 7: Relative Risk Ratios from a Multinomial Logistic Model Estimation: Effects of LS/CMI Components on Program Categories

	Self-Development (a)	Violence (b)	Addiction (c)	Education (d)	Job Skills (e)
LS/CMI: Crim. History	1.047	1.965***	1.065	1.053	0.988
	(0.106)	(0.360)	(0.118)	(0.109)	(0.128)
LS/CMI: Educ./Empl.	1.151	1.121	1.063	1.313***	1.367**
	(0.115)	(0.168)	(0.111)	(0.137)	(0.170)
LS/CMI: Family/Marital	0.951	1.139	0.925	0.912	0.943
	(0.092)	(0.158)	(0.095)	(0.090)	(0.116)
LS/CMI: Procrim. Attitude	1.153	1.087	0.889	1.093	1.381**
	(0.130)	(0.167)	(0.107)	(0.127)	(0.196)
LS/CMI: Companions	0.969	0.836	1.150	1.223*	1.062
	(0.105)	(0.123)	(0.131)	(0.138)	(0.154)
LS/CMI: Leisure/Recreation	1.060	1.399**	1.466***	0.939	0.897
	(0.097)	(0.194)	(0.147)	(0.088)	(0.099)
LS/CMI: Alcohol/Drug	0.896	0.740**	1.203*	0.789**	0.867
	(0.087)	(0.104)	(0.121)	(0.078)	(0.104)
LS/CMI: Antisocial Pattern	0.990	0.915	0.903	0.921	0.806
	(0.128)	(0.167)	(0.121)	(0.122)	(0.137)
Constant	4.018***	0.096***	0.741	2.804***	1.029
	(1.276)	(0.052)	(0.270)	(0.909)	(0.388)
Observations			2,573		

The sample consists of evaluated participants with the known program category (ten observations missing).
The reference outcome category is the "other programs" category.
Each LS/CMI component is normalized by its standard deviation in the sample.
Heteroskedasticity-robust standard errors are reported in parentheses.
Significance stars test the null hypothesis that the relative risk ratio equals one.
*** $p \leq 0.01$; ** $p \leq 0.05$; * $p \leq 0.1$

Table 8: Relative Risk Ratios from a Multinomial Logistic Model Estimation: Effects of Type of Crime Committed on Program Categories

	Self-Development (a)	Violence (b)	Addiction (c)	Education (d)	Job Skills (e)
Assault	1.373	2.312***	1.273	1.775***	1.515
	(0.280)	(0.590)	(0.301)	(0.374)	(0.394)
Burglary & Theft	1.219	1.174	1.602**	1.426*	1.237
	(0.224)	(0.298)	(0.334)	(0.274)	(0.299)
Drugs	0.929	0.392***	2.174***	1.428**	1.282
	(0.150)	(0.106)	(0.389)	(0.239)	(0.269)
Constant	4.620***	0.717**	1.457***	2.707***	0.837
	(0.531)	(0.116)	(0.197)	(0.330)	(0.129)
Observations			3,888		

The sample consists of participants with the known program category (ten observations missing)
The reference outcome category is the "other programs" category.
Heteroskedasticity-robust standard errors are reported in parentheses.
Significance stars test the null hypothesis that the relative risk ratio equals one.
*** $p \leq 0.01$; ** $p \leq 0.05$; * $p \leq 0.1$

Overall, the results from our multinomial logistic estimations suggest that efforts are made to target inmates to appropriate programs as a function of their background. This further suggests that evaluated inmates are more likely to have their needs recognized and to be assigned to an appropriate program.

3.4 Heterogeneity, or Who to Target?

The effects of programs potentially differ across inmates. Knowing which inmates benefit the most from participating is essential to target programs' resources to where they will generate the highest benefits (e.g., they may affect low- and high-risk inmates differently). Figure 4 presents simple kernel-weighted local mean smoothing curves of recidivism rates within three years given the total LS/CMI score for participants and non-participants. We discern only small differences between participants and non-participants among low-risk inmates. Potentially, the effects of programs may not fully manifest themselves through lower recidivism among populations with already low rates. The gap widens as the risk score increases up to a score of 12 and narrows as the score increases further, closing entirely at a score of 32 (at this point, inmates are deemed highly risky).

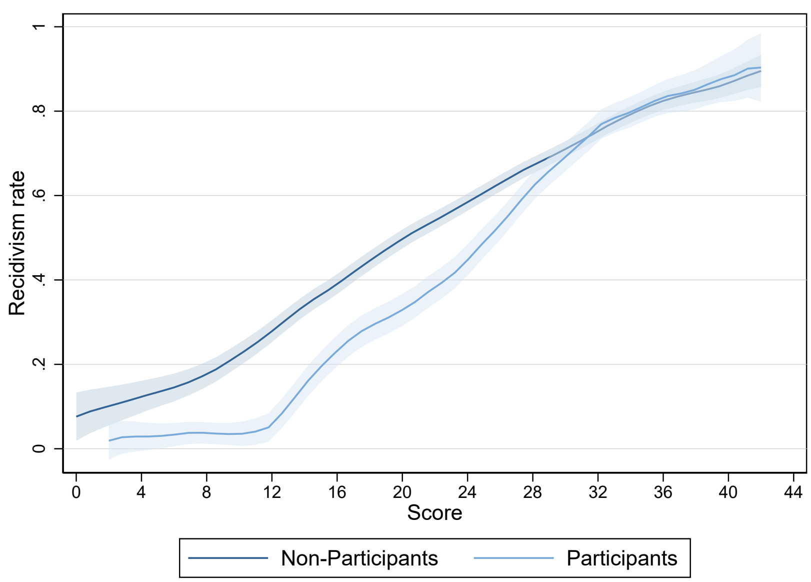


Figure 4: Smoothed Averages of Recidivism Within Three Years on Total LS/CMI Score for Participants and Non-Participants

Notes: Average recidivism rate computed by kernel-weighed local mean smoothing—Epanechnikov kernel with bandwidth of two is used for weighting. 95% confidence intervals are shown.

In light of this suggestive evidence of heterogeneous effects, we now examine heterogeneity of the effect of programs with respect to individual characteristics using recent developments in the machine learning literature. In particular, Athey and Imbens (2016), Wager and Athey (2018), and Athey et al. (2019) developed random forests methods designed to capture *honest* heterogeneity in treatment effects, circumventing overfitting issues inherent to traditional methods and thus avoiding mistaking statistical noise for heterogeneity. The method allows one to compute a treatment effect for each individual in the sample based on an array of characteristics.

The first step is to randomly divide the sample into two groups: the *training sample* and the *treatment sample*. Denote the training sample by $\mathcal{A}_1$ and the treatment sample by $\mathcal{A}_2$. The training sample is used to "grow a tree" that separates values of $\mathbf{x}_i$ into groups of observations called *nodes* or *leaves*. Each node then becomes a *parent node*. Within this parent node, the algorithm evaluates potential splits into subgroups called *children nodes*. We select nodes using recursive partitioning (Athey and Imbens, 2016) with the following splitting rule to maximize treatment effect heterogeneity:

$$C_1^*, C_2^* = \arg\max_{C_1, C_2} \frac{n_{C_1} n_{C_2}}{n_P^2} \left[\hat{\beta}(C_1, \mathcal{A}_1) - \hat{\beta}(C_2, \mathcal{A}_1)\right]^2, \tag{8}$$

where C_1 and C_2 are the two children nodes of the parent node P, n_k is the number of observations within node k, and $\hat{\beta}(C_k, \mathcal{A}_1)$ is the estimated treatment effect calculated with the data contained in node C_k. Figure A.1 provides an example of a regression tree. The nodes in blue, selected after the algorithm has gone through each value of all covariates, are called *terminal nodes*, or *terminal leaves*.

Once a tree is grown, we compute the treatment effect for each terminal node using the treatment sample. Following Wager and Athey (2018), we grow many trees, repeating these steps $B = 10,000$ times while bootstrapping the training and treatment samples at each iteration. Let us now denote the training and treatment samples in tree τ by $\mathcal{A}_1^\tau$ and $\mathcal{A}_2^\tau$, respectively. To find the individual treatment effect of inmate i, we average out the predictions over all trees. Suppose that i appears B_i times in the treatment sample. The estimated conditional average treatment effect (ATE) for i is then

$$\widehat{CATE}(\mathbf{x}_i) = \frac{1}{B_i} \sum_{\tau : i \in \mathcal{A}_2^\tau} \hat{\beta}(l(\mathbf{x}_i, \tau), \mathcal{A}_2^\tau), \tag{9}$$

where $\hat{\beta}(l(\mathbf{x}_i, \tau), \mathcal{A}_2^\tau)$ is the treatment effect in terminal node $l(\mathbf{x}_i, \tau)$ in which inmate i has landed in tree τ, calculated with data contained in $\mathcal{A}_2^\tau$.

We estimate a causal random forest on the set of inmates who were evaluated by the LS/CMI. We first include in $\mathbf{x}$ the full set of individual controls, excluding, for now, years and prisons fixed effects.[17] The outcome variable is an indicator for recidivism within three years. In Figure 5, in blue, we plot the distribution of the predicted treatment effects in the sample, and most of the distribution lies on the left of zero as expected. Thus, when computing the ATE using the method of Athey et al. (2019), we find an average decrease of 8.18 percentage points in the recidivism rate due to the program. This result is

[17]The reason for omitting years and prisons fixed effects is simple: we want to determine if there exist heterogeneous effects of programs based on characteristics that have implications for targeting at the individual level. We include years and prisons fixed effects later in the analysis to estimate robust best linear projections, although doing this does not dramatically change the results.

comparable to what we found earlier with the OLS regressions. Unlike with OLS, however, one can use random forests to compute the effect for participants and non-participants separately. We estimate that, on average, participants' recidivism likelihood decreases by 7 percentage points because of the programs. Additionally, we find that non-participants would, on average, benefit as much if they had participated (a decrease of around 8 percentage points in recidivism). These results are significant at the 0.1% level.

The interpretation of the blue density in Figure 5 has an important caveat: it is unclear whether the distribution's wideness is due to heterogeneity—what we seek to capture—or to statistical noise. We conduct a simple exercise to visualize both effects graphically. Suppose that the effect is in fact homogeneous and equals -0.082 for each individual. In this case, by estimating heterogeneous effects using a causal forest estimation, we would still find estimates varying with $\mathbf{x}_i$ because of the variance of our $\widehat{CATE}(\mathbf{x}_i)$ estimators. Fortunately, the causal random forest method described above allows to compute the individual variance associated with each individual's predicted treatment effect. We therefore use these estimates to simulate the distribution of predicted treatment effects we would expect to obtain if the treatment effect was homogeneous. Specifically, we generate the following predictions:

$$\widetilde{CATE}(\mathbf{x}_i) = \widehat{ATE} + \hat{\sigma}_i \epsilon_i, \tag{10}$$

where $\hat{\sigma}_i^2$ is the variance of $\widehat{CATE}(\mathbf{x}_i)$ and ϵ_i is a Gaussian shock with mean 0 and variance 1. We plot the distribution of $\widetilde{CATE}$, in red, in Figure 5.[18] The distribution of predicted treatment effects that we would expect to obtain just from statistical noise is less wide than our distribution of $\widehat{CATE}(\mathbf{x}_i)$, suggesting the additional spread in the latter arises from true heterogeneity in treatment effects.

To provide further evidence of heterogeneous treatment effects, we implement a test proposed by Athey and Wager (2019) and divide the sample into two groups based on whether the predicted treatment effect is below or above the median, using out-of-bag predictions to avoid selection issues.[19] The results are presented in Table 9. On the one hand, for the subset of individuals below the median, we detect a statistically significant effect of -17 percentage points. On the other hand, we estimate an effect of -5 percentage points for the subset of individuals above the median, a marginally significant result. We can reject at the 1% level that the two effects are the same. Therefore, the causal random forest successfully identifies subgroups—at least two—with differing treatment effects.

We adopt a third approach, suggested by Athey and Wager (2019), to test for heterogeneous treatment effects with respect to observables. The approach draws from the best linear predictor design suggested by Chernozhukov et al. (2018). Let $\hat{p}^{(-i)}(\mathbf{x}_i)$ be the propensity score of type $\mathbf{x}_i$ leaving out observation i. and let $\hat{m}^{(-i)}(\mathbf{x}_i)$ be the predicted outcome of the type of i, again leaving observation i out. To ease the notation, let $\hat{\beta}^{(-i)}(\mathbf{x}_i)$ be the out-of-bag predicted treatment effect of i, and denote the average predicted

[18] To make the distribution smoother, we duplicated each observation three times and added a Gaussian random shock to each.

[19] This means that individual i is left out of the calculation when determining if he falls in the below- or above-median group. Athey and Wager (2019) discuss this test thoroughly.

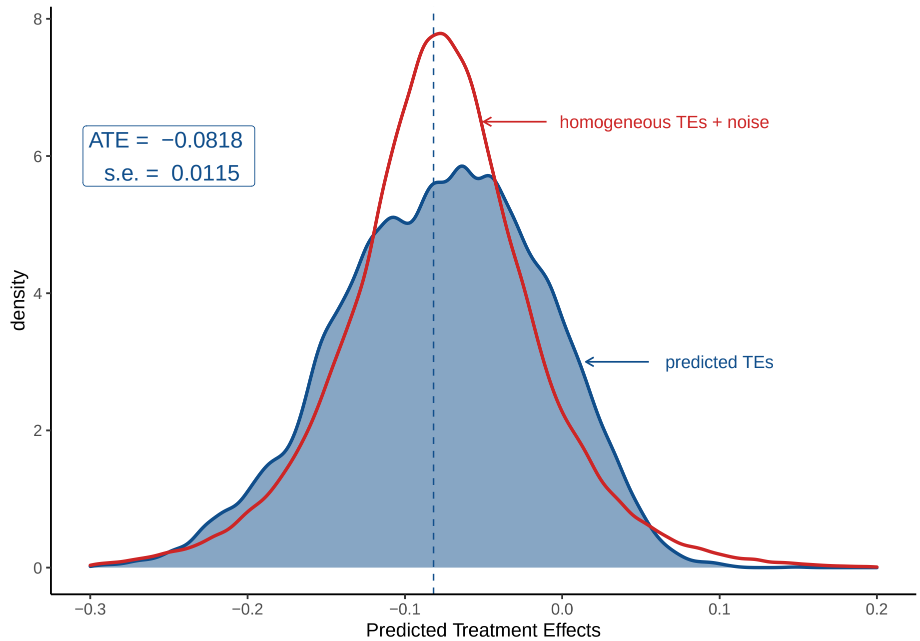


Figure 5: Density of Predicted Treatment Effects

Notes: The red line shows the distribution of the predicted treatment effects corresponding to the sum of the average treatment effect and the simulated zero-mean shocks with the estimated variances of the CATEs. The blue density shows the distribution of the estimated CATEs.

Table 9: Out-of-Bag Predictions for Effects Below and Above the Median

	Below Median	Above Median
ATE	-0.1656***	-0.0487
(s.e.)	(0.0185)	(0.0323)
N	3515	3516
p-value of the difference	0.0017	

*** $p \leq 0.01$; ** $p \leq 0.05$; * $p \leq 0.1$

treatment effect by $\bar{\beta}$. Then consider the following regression:

$$y_i - \hat{m}^{(-i)}(\mathbf{x}_i) = \alpha_0 \bar{\beta} \left[T_i - \hat{p}^{(-i)}(\mathbf{x}_i) \right] + \alpha_1 \left[\hat{\beta}^{(-i)}(\mathbf{x}_i) - \bar{\beta} \right] \left[T_i - \hat{p}^{(-i)}(\mathbf{x}_i) \right] + \epsilon_i. \tag{11}$$

The parameters α_0 and α_1 are to be estimated. α_0 captures the ATE, and if the model is well calibrated, one should estimate $\hat{\alpha}_0 = 1$. In contrast, α_1 measures the strength of the association between an individual's treatment effect with his type of ATE, meaning observable characteristics are key predictors of the treatment effect. If α_1 is statistically different from zero, one can reject the null hypothesis that there is no heterogeneity. We estimate $\hat{\alpha}_0 = 1.03$ with a standard error of 0.1488 and $\hat{\alpha}_1 = 2.06$ with a standard error of 0.16. Hence, our analysis reveals the presence of sharp heterogeneity with regards to observable characteristics. Intuitively, individuals who share an array of characteristics have similar treatment effects.

Our tests presented above suggest that programs have heterogeneous effects as a function of observable characteristics and thus targeting might be beneficial. We now seek to determine which characteristics correlate with the efficiency of programs. Semenova and Chernozhukov (2020) provide a method for estimation and inference of the best linear projections of CATEs on observable characteristics.[20] Table 10 reports the first set of results from this procedure. In each column, we specify a distinct subset of characteristics from $\mathbf{x}$. Each estimation controls for year of sentence and prison fixed effects to increase precision.[21] Two results come forth and are robust to all specifications: first, inmates with risks and needs associated with a procriminal attitude have lower treatment effects because these attitudes are deemed supportive of crime and are averse to treatment (Andrews et al., 2000). A one standard deviation increase in this component's score increases the programs' effect toward zero by 3.4 to 4.8 percentage points depending on the specification, confirming the intuition that these measurable attitude problems materialize into lower effects of programs.

Second, individuals aged from 30 to 40 respond better to rehabilitation programs compared to their older counterparts. This finding is also common to all specifications. We estimate that belonging to the 30–40 age group increases the effectiveness of programs by around 5.5 percentage points. This effect could be partly mechanical: an extensive literature shows how age is invariably negatively correlated with criminal activities (see, e.g., Landersø et al. (2017)). Then, since older inmates are less likely to reoffend in the first place, one can expect younger individuals to gain more from participating in programs. As for the youngest age group (29 and below), our point estimates suggest a similar effect of programs than that of the 30 to 40 age group, although the effect relative to the oldest group is not statistically significant.

We estimate another specification and regress the predicted treatment effects on the total LS/CMI score, separated into five groups, on top of prison and year fixed effects.[22] Doing this allows us to investigate whether such an overall measure of risk can be useful for targeting. Figure 4 suggests evidence of a higher effect for medium-risk individuals, but this figure was not based on an honest approach in measuring heterogeneous effects. The top panel of Figure 6 plots the estimate of the best linear projections of CATEs onto the total LS/CMI score. The results support the finding that medium-risk individuals benefit the most from programs: we estimate treatment effects of -10 percentage points for offenders with risk scores ranging from 8 to 23. Again, we cannot reject that the treatment effect for low-risk individuals is zero. Beyond an LS/CMI score of 24, which corresponds to high and very high risk, participating in programs does not clearly affect recidivism, but it is worth noting that the confidence intervals are large. On the bottom part of Figure 6, we present the program participation rate within each risk category along with the confidence intervals. We observe minimal differences in these rates. Thus, there seems to be room left for targeting according to this measure, as inmates who seem to benefit the most from

[20] A detailed description of the method is beyond the scope of this paper. For further details, see Semenova and Chernozhukov (2020), Corollary 4.1 and Example 2.2.

[21] Although fixed effects increase the precision of our estimates, the results do not significantly vary when they are removed.

[22] Again, the results presented below are robust to excluding these fixed effects from the estimation.

Table 10: Best Linear Projections

	Dependent Variable: Predicted Treatment Effects			
	(a)	(b)	(c)	(d)
LS/CMI-Crim. History	0.007		0.006	0.003
	(0.016)		(0.016)	(0.008)
LS/CMI: Procrim.Attitude	0.048**		0.046**	0.034**
	(0.022)		(0.021)	(0.016)
LS/CMI: Educ/Empl.	−0.014		−0.010	−0.003
	(0.016)		(0.016)	(0.006)
LS/CMI: Family/Marital	0.027*		0.024	0.019
	(0.016)		(0.015)	(0.013)
LS/CMI: Companions	−0.017		−0.017	−0.013
	(0.025)		(0.024)	(0.022)
LS/CMI: Leisure/Recreation	0.027		0.026	0.039
	(0.032)		(0.031)	(0.049)
LS/CMI: Alcohol/Drugs	0.011		0.012	0.005
	(0.024)		(0.024)	(0.010)
LS/CMI: Antisocial Pattern	−0.047		−0.044	−0.037
	(0.041)		(0.039)	(0.032)
Indigenous				−0.036
				(0.052)
At least one dependent				−0.011
				(0.014)
Age $<$ 30		−0.062	−0.034	−0.037
		(0.042)	(0.033)	(0.033)
30 $\leq$ Age $<$ 40		−0.067**	−0.055*	−0.057*
		(0.030)	(0.031)	(0.031)
Crime: Assault				0.016
				(0.036)
Crime: Burglary_Theft				−0.064
				(0.064)
Crime: Drugs				−0.044
				(0.034)
Constant	−0.156*	−0.043	−0.120	−0.092
	(0.087)	(0.085)	(0.090)	(0.094)

Heteroskedasticity-robust standard errors are reported in parentheses.
*** $p \leq 0.01$; ** $p \leq 0.05$; * $p \leq 0.1$
Each LS/CMI component is normalized by its standard deviation in the sample.
Each estimation includes year and prison fixed effects.

programs do not seem to enroll more.

So far, we have put forward two results useful for targeting: inmates with higher procriminal attitude scores and inmates with high total risk scores seem to benefit less from programs. In Figure 7, we further explore what we can learn from both scores for targeting. The left panel shows the best linear projections of the treatment effects onto the procriminal attitude score, ranging from 0 to 4. The effect of programs is significant for individuals who score between 0 and 2[23] but not for those who score higher. The participation rates shown in the bottom part of Figure 7 indicate that there seems to be targeting based on this measure. On the right panel of the Figure, we consider the subset of offenders with a total LS/CMI score of 24 and above. Even though we find no overall effect of programs for this group in our previous exercise, it is conceivable that programs work for one part of the group. However, within this group, the estimated treatment effects are close to zero and are not statistically significant, suggesting

[23]The p-value for the score of 2 is 0.0553.

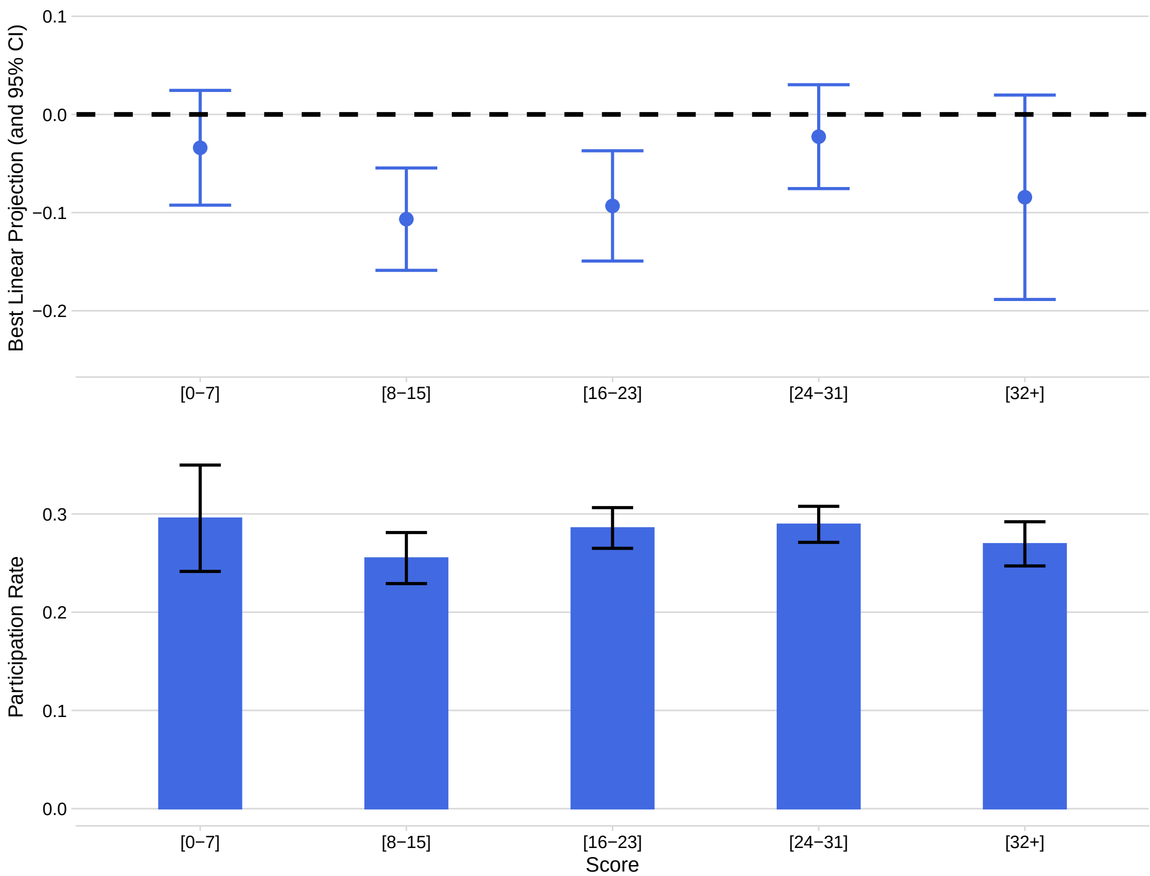


Figure 6: Participation Rates and Best Linear Projections of the Treatment Effects on the Score

Notes: On the top graph, the estimation contains year and prison fixed effects. The total score is divided into groups with at least 200 observations. A similar pattern is observed when considering more or less categories. On the top and bottom graphs, 95% confidence intervals are shown.

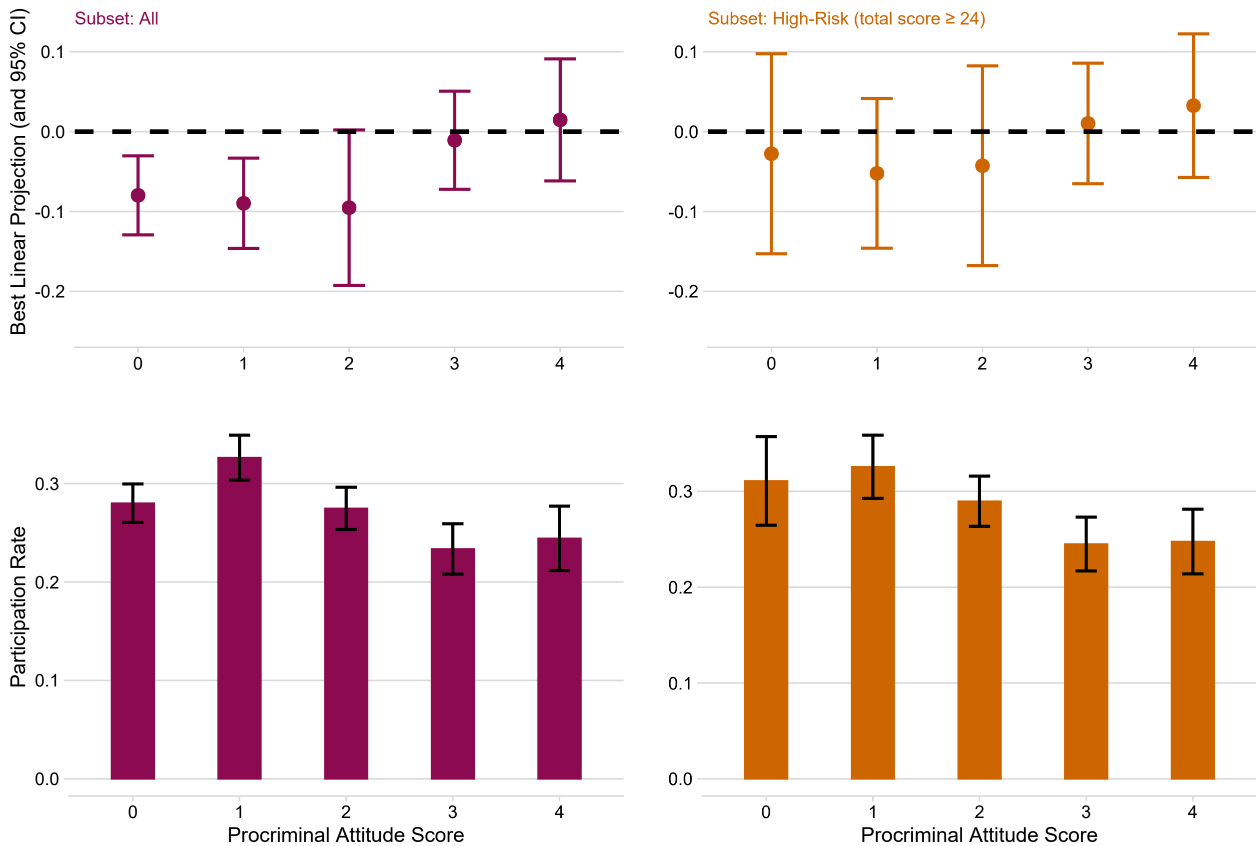


Figure 7: Participation Rates and Best Linear Projections of the Treatment Effects on Procriminal Attitude Score

Notes: Graphs on the left are for all evaluated individuals. Graphs on the right are for the subset of individuals with a total LS/CMI score of 24 or more. Top graphs: the estimation contains year and prison fixed effects. Bottom graphs: 95% confidence intervals are shown.

that the total score is a sharp discriminant factor to consider for placement in programs regardless of procriminal attitudes. However, for inmates with a score lower than 24, procriminal attitudes should be considered for targeting.

4 Discussion and Policy Implications

Our paper shows that social rehabilitation programs can significantly reduce recidivism when inmates' rehabilitation rehabilitation needs are assessed. We show that the efficiency of programs is strongly related to whether or not inmates are thoroughly evaluated to assess their criminogenic needs. We also show that such evaluations can be useful to target inmates who seem to benefit the most from programs, like those with medium overall risk measures and those with low measures of procriminal attitude.

Recent evidence suggests incarceration can favor rehabilitation (Landersø, 2015; Bhuller et al., 2020; Hjalmarsson and Lindquist, 2020) and that prison conditions may in part determine this effect (Lotti, 2020;

Mastrobuoni and Terlizzese, 2019; Tobón, 2020). Many authors have hinted at rehabilitation programs and at the quality of prison conditions in their context to explain their findings on the positive impacts of incarceration. Our paper supports their view. We find substantial effects of programs in a context where significant efforts are made to provide rehabilitation assistance.

Our finding that programs do not significantly impact inmates not undergoing a thorough risk assessment evaluation further suggests that rehabilitation programs require efforts and resources to be effective. In the United States, the average expenditure per inmate is low—–around $33,274, on average, per year in 2015 (Mai and Subramanian, 2017)—–and the literature provides little evidence of rehabilitative effects of incarceration (Kling, 2006). In Norway, Denmark and Sweden, average expenditure per inmate in 2015 were around 138,116, 79,096, and 144,139 US dollars, respectively,[24] and research has documented convincing evidence of substantial rehabilitative effects (Landersø, 2015; Bhuller et al., 2020; Hjalmarsson and Lindquist, 2020). In our setting, the cost of incarceration is middle-of-the-road: the average cost of incarceration in the provincial prisons of Quebec, where offenders are incarcerated for mostly misdemeanors and low-level crimes, is around 56,234 US dollars.[25] Though average incarceration cost is probably far from the most relevant determinant of the effect of incarceration, this hints at promising avenues for future research.

References

Aebi, M. F., M. M. Tiago, and C. Burkhardt (2016). Council of europe annual penal statistics space i: Prison populations. survey 2015. Technical report, Strasbourg: Council of Europe.

Aizer, A. and J. J. Doyle Jr (2015). Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges. *The Quarterly Journal of Economics 130*(2), 759–803.

Albright, A. (2019). If you give a judge a risk score: Evidence from kentucky bail decisions. *Th John M. Olin Center for Law, Economics, and Business Fellows' Discussion Paper Series 85*.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy 113*(1), 151–184.

Andrews, D. A., J. Bonta, and S. Wormith (2000). *Level of service/case management inventory: LS/CMI*. Multi-Health Systems Toronto, Canada.

Armstrong, T. A. (2003). The effect of moral reconation therapy on the recidivism of youthful offenders: A randomized experiment. *Criminal Justice and Behavior 30*(6), 668–687.

[24] The estimated values in euros from Aebi et al. (2016) for Norway, Denmark, and Sweden are 125,560, 71,905, and 131,035, respectively. We calculated the values in USD$ using an exchange rate of 1.1, the approximate exchange rate at the end of December 2015.

[25] This estimate comes from Segel-Brown (2018). We used an exchange rate of 1.39, which was the approximate exchange rate at the end of December 2015.

Athey, S. and G. Imbens (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences 113*(27), 7353–7360.

Athey, S., J. Tibshirani, S. Wager, et al. (2019). Generalized random forests. *The Annals of Statistics 47*(2), 1148–1178.

Athey, S. and S. Wager (2019). Estimating treatment effects with causal forests: An application. *arXiv preprint arXiv:1902.07409*.

Balafoutas, L., A. García-Gallego, N. Georgantzis, T. Jaber-Lopez, and E. Mitrokostas (2020). Rehabilitation and social behavior: Experiments in prison. *Games and Economic Behavior 119*, 148–171.

Barbarino, A. and G. Mastrobuoni (2014). The incapacitation effect of incarceration: Evidence from several Italian collective pardons. *American Economic Journal: Economic Policy 6*(1), 1–37.

Bayer, P., R. Hjalmarsson, and D. Pozen (2009). Building criminal capital behind bars: Peer effects in juvenile corrections. *The Quarterly Journal of Economics 124*(1), 105–147.

Berk, R. (2017). An impact assessment of machine learning risk forecasts on parole board decisions and recidivism. *Journal of Experimental Criminology 13*(2), 193–216.

Bhuller, M., G. B. Dahl, K. V. Løken, and M. Mogstad (2020). Incarceration, recidivism, and employment. *Journal of Political Economy 128*(4), 1269–1324.

Blattman, C., J. C. Jamison, and M. Sheridan (2017). Reducing crime and violence: Experimental evidence from cognitive behavioral therapy in Liberia. *American Economic Review 107*(4), 1165–1206.

Buonanno, P. and S. Raphael (2013). Incarceration and incapacitation: Evidence from the 2006 italian collective pardon. *American Economic Review 103*(6), 2437–65.

Chen, M. K. and J. M. Shapiro (2007). Do harsher prison conditions reduce recidivism? a discontinuity-based approach. *American Law and Economics Review 9*(1), 1–29.

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins (2018). Double/debiased machine learning for treatment and structural parameters: Double/debiased machine learning. *The Econometrics Journal 21*(1).

Cook, P. J., S. Kang, A. A. Braga, J. Ludwig, and M. E. O'Brien (2015). An experimental evaluation of a comprehensive employment-oriented prisoner re-entry program. *Journal of Quantitative Criminology 31*(3), 355–382.

Davis, L. M., R. Bozick, J. L. Steele, J. Saunders, and J. N. Miles (2013). *Evaluating the effectiveness of correctional education: A meta-analysis of programs that provide education to incarcerated adults.* Rand Corporation.

Doleac, J. L. (2019). Encouraging desistance from crime. Technical report, Mimeo, Texas A&M University.

Doleac, J. L., C. Temple, D. Pritchard, and A. Roberts (2020). Which prisoner reentry programs work? replicating and extending analyses of three rcts. *International Review of Law and Economics 62*, 105902.

Green, D. P. and D. Winik (2010). Using random judge assignments to estimate the effects of incarceration and probation on recidivism among drug offenders. *Criminology 48*(2), 357–387.

Heller, S. B., A. K. Shah, J. Guryan, J. Ludwig, S. Mullainathan, and H. A. Pollack (2017). Thinking, fast and slow? Some field experiments to reduce crime and dropout in chicago. *The Quarterly Journal of Economics 132*(1), 1–54.

Hjalmarsson, R. and M. J. Lindquist (2020). The health effects of prison.

Kling, J. R. (2006). Incarceration length, employment, and earnings. *American Economic Review 96*(3), 863–876.

Kopf, R. G. (2014). Federal supervised release and actuarial data (including age, race, and gender): The camel's nose and the use of actuarial data at sentencing. *Fed. Sentencing Reporter. 27*, 207.

Kuziemko, I. (2013). How should inmates be released from prison? An assessment of parole versus fixed-sentence regimes. *The Quarterly Journal of Economics 128*(1), 371–424.

Landersø, R. (2015). Does incarceration length affect labor market outcomes? *The Journal of Law and Economics 58*(1), 205–234.

Landersø, R., H. S. Nielsen, and M. Simonsen (2017). School starting age and the crime-age profile. *The Economic Journal 127*(602), 1096–1118.

Lochner, L. (2004). Education, work, and crime: A human capital approach. *International Economic Review 45*(3), 811–843.

Long, J. S., C. Sullivan, J. Wooldredge, A. Pompoco, and M. Lugo (2019). Matching needs to services: Prison treatment program allocations. *Criminal Justice and Behavior 46*(5), 674–696.

Lotti, G. (2020). Tough on young offenders: Harmful or helpful? *Journal of Human Resources*, 1017–9113R3.

Macdonald, D. C. (2020). Truth in sentencing, incentives and recidivism. Technical report, Mimeo, Vancouver School of Economics, University of British Columbia.

Mai, C. and R. Subramanian (2017). The price of prisons: Examining state spending trends, 2010-2015.

Mastrobuoni, G. and D. Rivers (2016). Criminal discount factors and deterrence. *Available at SSRN 2730969*.

Mastrobuoni, G. and D. Terlizzese (2019). Leave the door open? Prison conditions and recidivism. Technical report, Mimeo, University of Turin.

Mueller-Smith, M. (2015). The criminal and labor market impacts of incarceration. *Unpublished Working Paper 18*.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics 37*(2), 187–204.

Owens, E. G. (2009). More time, less crime? Estimating the incapacitative effect of sentence enhancements. *The Journal of Law and Economics 52*(3), 551–579.

Segel-Brown, B. (2018). Update on costs of incarceration. *Parliamentary Budget Officer, March 22*.

Semenova, V. and V. Chernozhukov (2020). Debiased machine learning of conditional average treatment effects and other causal functions. *The Econometrics Journal*.

Seroczynski, A., W. N. Evans, A. D. Jobst, L. Horvath, and G. Carozza (2016). Reading for life and adolescent re-arrest: Evaluating a unique juvenile diversion program. *Journal of Policy Analysis and Management 35*(3), 662–682.

Stevenson, M. (2017). Breaking bad: Mechanisms of social influence and the path to criminality in juvenile jails. *Review of Economics and Statistics 99*(5), 824–838.

Stevenson, M. T. and J. L. Doleac (2019). Algorithmic risk assessment in the hands of humans. *Available at SSRN*.

Tobón, S. (2020). Do better prisons reduce recidivism? Evidence from a prison construction program. *The Review of Economics and Statistics. (Forthcoming)*.

Visher, C. A., L. Winterfield, and M. B. Coggeshall (2005). Ex-offender employment programs and recidivism: A meta-analysis. *Journal of Experimental Criminology 1*(3), 295–316.

Vose, B., F. T. Cullen, and P. Smith (2008). The empirical status of the level of service inventory. *Federal Probation 72*, 22.

Wager, S. and S. Athey (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association 113*(523), 1228–1242.

Wilson, D. B., C. A. Gallagher, and D. L. MacKenzie (2000). A meta-analysis of corrections-based education, vocation, and work programs for adult offenders. *Journal of Research in Crime and Delinquency 37*(4), 347–368.

A Appendix

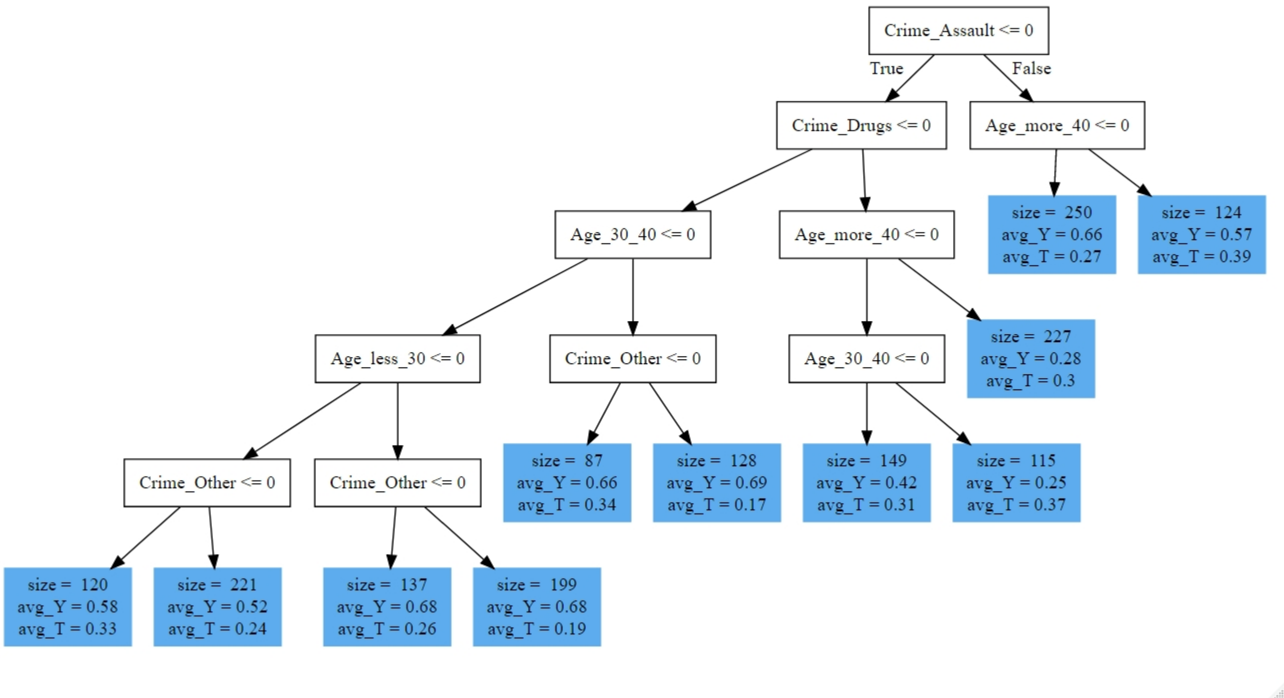


Figure A.1: Regression Tree Example

This figure is an example of a regression tree grown with only a subset of the explanatory variables. In the blue leaves (the final nodes), *size* is the total number of individuals falling in the leave, while Y and T stand for recidivism and program participation.

Table A.1: LS/CMI Risk Score Questionnaire

Item	Question
Category: Criminal History [8]	
1	Any prior youth dispositions or adult convictions?
2	Two or more prior youth/adult dispositions/convictions?
3	Three or more prior youth/adult dispositions/convictions?
4	Three of more present offenses?
5	Arrested or charged under age 16?
6	Ever incarcerated upon conviction?
7	Ever punished for institutional misconduct or a behavior report?
8	Charge laid, probation breached, or parole suspended during prior community supervision?
Category: Education/Employment [9]	
	When in the labor market (either in the community or long-term imprisonment with work opportunities):
9	Currently unemployed?
10	Frequently unemployed?
11	Never employed for a full year?
	School or when in school:
12	Less than regular grade 10 or equivalent?
13	Less than regular grade 12 or equivalent?
14	Suspended or expelled at least once?
	Classification (Education/Employment) *[In this section, the evaluator rates the inmate's behavior when in school or at work.]*
15	Participation/Performance
16	Peer interactions
17	Authority interaction
Category: Family/Marital [4]	
18	Dissatisfaction with marital or equivalent situation
19	Nonrewarding, parental
20	Nonrewarding, other relatives
21	Criminal—family/spouse
Category: Leisure/Recreation [2]	
22	Absence of recent participation in an organized activity
23	Could make a better use of time
Category: Companions [4]	
24	Some criminal acquaintances
25	Some criminal friends
26	Few anticriminal acquaintances
27	Few anticriminal friends
Category: Alcohol/Drug Problem [8]	
28	Alcohol problem, ever
29	Drub problem, ever
30	Alcohol problem, currently
31	Drug problem, currently
	If a current alcohol/drug abuse problem exists, complete the following. *[In this section, the evaluator assesses if alcohol or drugs problems contributed to other problems in the following categories.]*
32	Law violations
33	Marital/Family
34	School/Work
35	Medical or other clinical indicators?
Category: Procriminal Attitude/Orientation [4]	
36	Supportive of crime
37	Unfavorable toward convention
38	Poor, toward sentence/offense
39	Poor, toward supervision/treatment
Category: Antisocial Pattern [4]	
40	Specialized assessment for antisocial pattern. *[personality disorder, psychopathy, etc.]*
41	Early and diverse antisocial behavior *[at least two items from a list]*
42	Criminal attitude *[at least one item from a list]*
43	Pattern of generalized trouble *[at least four items from a list]*

This is a reproduction of the LS/CMI questionnaire; see Section 1 of Andrews et al. (2000).
The text in italics identifies our own additions.