

Ludwig-Maximilians-Universität München

Criminal recidivism after prison and electronic monitoring

by Rafael Di Tella and Ernesto Schargrodsky

Causal Inference for Policy Evaluation

Bakai Baiazbekov *

June 16, 2019



*B.Baiazbekov@campus.lmu.de

Contents

1	Introduction	1
2	Pure replication	3
3	Measurement and estimation analysis	3
3.1	Two - Stage Least Squares	3
3.2	Validating IV assumptions	4
4	Theory of change analysis	7
4.1	Educational outcomes after EM	7
4.2	Recidivism risk versus fixed - sentence regimes	8
5	Extensions	10
5.1	Application of Decision Trees	10
6	Discussion	13
6.1	Welfare	13
7	Conclusion	13
8	Statutory Declaration	14
9	Appendix	15

1 Introduction

All societies must decide what to do with those who commit crimes. While a standard approach for centuries has been to incarcerate offenders but since the technology doesn't stay at same place, recently was proposed to use electronic monitoring bracelets for offenders to monitor them remotely. Electronic monitoring of offenders has been in use for just over two decades and motives for using it remain diverse. Some agencies that use EM attempt to deliver humane and affordable sanctions while others seek to relieve jail crowding or to avoid the construction of new jails. As the fiscal burden of the prison population has increased and as the electronic monitoring technology became cheaper and safer with more precise, where these devices can track GPS, recognize voice with machine learning techniques and measure the trans-dermal measurement of alcohol or drug consumption. In the history, there was prisons called "panopticon", where inmates could be watched continually by guards who could not be seen.

This paper based on measurement of causal effect of criminal recidivism rate in the city of Buenos Aires. By focusing on the rearrest rates of two groups: individuals released from prison and individuals released from electronic monitoring program. Detainees are randomly assigned to different types of judges, where ideological differences across judges translate into large differences in the allocation of electronic monitoring. Taking into account this type of characteristics of judges, authors argue that there is a large, negative effect on criminal recidivism of treating individuals with electronic monitoring relative to prison.

Every year a large number of individuals are sent to prison in Argentina. By 2007, the Penitentiary Service of the Buenos Aires hosted of approximately 26,990 inmates which represented 44.5 percent of the total population under penal supervision of the whole country. The conditions of confinement are extremely poor, perhaps because the inmate population held in prisons and jails experienced a large increase from 12,223 in 1994 to 30,721 in 2005, with increase 18,498 in 11 years and without a corresponding increase in infrastructure investment. Given that prisons are costly to build and run and often involve cruel treatment of fellow citizens, possibly contributing to the conversion of inmates into "hardened" criminals, it is unsurprising that alternatives to imprisonment have been tried out. One of the more intriguing experiments in this area is the substitution of electronic monitoring for incarceration. If we look to statistics, by 2010, more than 500,000 people in the United States and Europe alone had been treated with electronic monitoring, in spite of the obvious complexity of a full cost-benefit analysis. In this paper, authors tried to contribute to an evaluation of electronic monitoring by providing the evidence on one of the estimates needed for such exercise: the difference between the recidivism rate of offenders treated with electronic monitoring and the recidivism rate of offenders released from a standard prison.

Electronic monitoring arises several debates such as that spending time under electronic monitoring rather than incarceration might make low punishment salient, implying a positive relationship between light punishment, electronic monitoring, and recidivism. On the other hand, some find a negative relationship, for instance, imprisonment might be criminogenic through harsh prison conditions or peer effects that are not present under electronic monitoring. Specifically, electronic monitoring could prevent contact with hardened criminals or reduce the perception that society

is "mean" and "deserving of the crime it receives" **Strang[2007]** [8]. In addition, electronic monitoring could differ from prison in its effect on the offender's labor market prospects or social integration, through and effect on the accumulation of skills.

At least two practical empirical problems appear when attempting to estimate a causal estimate, one of which can be called a *problem of selection*, refers to at least one potential criterion for the granting of electronic monitoring to an offender is his risk of recidivism, and the *the problem of differential risk* of the target population, refers to the possibility that electronic monitoring programs are restricted to low-risk populations(whose most serious crime is drunk driving). Thereby, the detection can fail to observe negative effect of electronic monitoring on recidivism, simply it reflects that this population is already at very low risk of crime and the control group receives very light punishment.

Authors described three important elements of the institutional and theoretical background suggesting that it is possible to use an instrumental variables approach. Initially, they described the legal process from the time individuals are arrested until they reach trial, with special attention to the role of judge. Next, they showed very large ideological differences across judges. Finally, authors presented a brief model with several of the features of Argentine setting and where ideological judges determine electronic monitoring allocation. They provided reasonable interpretation of estimate is that the causal effect of treating an apprehended offender with electronic monitoring instead of prison. The offenders are randomly matched to judges and the likelihood an offender is sent to electronic monitoring instead of prison differs substantially across judges. Another important feature is because of the usual ideological differences across judges(**Waldfogel1991**)[9] and these ideological differences become exaggerated in Argentina. The jails and prisons have frequently been denounced excessively cruel by human right organizations. This results in two stereotypes: judges who frequently allocate electronic monitoring, or liberals and the judges who never assign electronic monitoring or conservative, tough hand judges. Since the assignment of judges is exogenous to prisoners characteristics, it is possible to instrument the decision to send an offender to prison or electronic monitoring with proxy for judge's ideology. The instrumental variable results reveal a robust, negative and significant effect of electronic monitoring on later rearrest rates of between 11 and 16 percentage points or approximately half the baseline recidivism rate.

2 Pure replication

Since authors provided all data and do files, my replication was short and I decided to do the replication in R. Initially, I had troubles subsetting the data cause STATA automatically omits NA's and drops irrelevant rows, however, in R I had to do it manually. Secondly, I had an issue while reproducing the code. Everything ran smoothly, except for Column 4 of Table 5, where 15.1 SE version of STATA did not allow clustering for IV probit estimation. However, no such problem occurred in STATA 14.0 or R. Moreover, every regression estimates in all other tables and columns was clustered according to the same variable, but no such problem occurred. Below in appendix section, I attached all the tables from R, except table number 1. Since it was just descriptive statistics I just added from STATA. Moreover, if you also interested in the code, you may reach them on my github repo : bakaibaiazbekov.

3 Measurement and estimation analysis

3.1 Two - Stage Least Squares

In the general IV regression model, we do assumptions on instrument relevance and instrument exogeneity that are the same as in the simple regression model with a single endogenous regressor and only one instrument. In our case the authors took endogenous variable Electronic Monitoring and instrument court, percentage judge sent to EM and judge already used EM. Below I showed the general Instrumental Variables Regression Model and Terminology :

$$EM_i = \pi_0 + \pi_1 Z_{1i} + \pi_2 Z_{2i} + \pi_2 X_{2i} + \dots + \pi_i X_{ij} + u_i \quad (1)$$

$$Recidivism_i = \beta_0 + \beta_1 \widehat{EM}_i + \pi_2 X_{2i} + \dots + \pi_i X_{ij} + u_i \quad (2)$$

where R_i is a dummy variable that indicates whether individual i went back to detention in the Province of Buenos Aires after his release, EM_i is a dummy variable that indicates whether individual i was in the electronic monitoring group and using judge ideology as an instrument with F statistics for the joint significance is 4,36. There was 172 judges in this sample (1503), so this approach has limitations.

While computing both stages of TSLS individually is not big deal, the simple regression model with a single endogenous regressor, the more interesting part was resorting TSLS functions like `ivreg()` was more convenient when the set of potentially endogenous regressors and instruments are large.

Similarly to the simple IV regression model, the general IV model can be estimated using the two- stage least squares estimator: On the First stage, I run the OLS regressions for each of the endogenous variables, in this case *electronicmonitoring_i* on all instrumental variables, (*judgedummies_i*, *percentagejudgesenttoEM_i* , and *judgealreadyusedEM_i*), all exogenous variables and an intercept. For the second stage I regress the dependent variable, *recidivism* ,on the predicted values of all endogenous regressors, all exogenous variables and an intercept using OLS. This gave me $\widehat{\beta}_0, \dots, \widehat{\beta}_{k+r}$, the TSLS estimates of the model coefficients.

To show that our instrument is valid we have *two conditions*. In the general IV regression model, the instrument relevance and instrument exogeneity assumptions

are the same as in the simple regression model with a single endogenous regressor and only one instrument. Lets say for Z_{1i}, \dots, Z_{mi} , judge ideology, instrument to be as set of valid instruments, the following two conditions must be fulfilled: First of all, **Instrument relevance** : when there are k endogenous variables, electronic monitoring, r exogenous variables and $m \geq k$ instruments Z and the $\hat{X}_{1i}, \dots, \hat{X}_{ki}$ are the predicted values from the k population first stage regressions, it must hold that $\hat{X}_{1i}, \dots, \hat{X}_{ki}, W_{1i}, \dots, W_{ri}, 1$ are not perfectly multicollinear.

Next condition is **Instrument exogeneity**: it should hold for all m instruments and it must be uncorrelated with error term, $\rho_{Z_{1i}=0}, \dots, \rho_{Z_{mi}, u_i}$. This condition can not be tested thats why I will prove verbally. Since the assignment of judges is exogenous to prisoners characteristics (whenever a person is detained by the police, she or he is assigned to the judge who was on duty on that day, and duty turns are assigned by a lottery). Lastly, **Exclusion restriction** assumption seems to be hold, since judge's personal ideology is unlikely to affect criminal's future recidivism rate directly. It only affects y dependent variable thru endogenous variable.

There are three groups in the sample: those who were assigned to EM, those who went to prison sent by a judge who uses EM, and those sent to prison by judge who never used EM. If judges who use EM in fact select the good types(low recidivism risk) for treatment with EM, then those standing before that same judge who were not selected for EM should be bad types(high recidivism risk). In particular, their average type should be worse than the average type of the conservative judges who did zero selection. In other words, the point estimate on the dummy judge ever used EM should be positive (as the base category is those who were sent to prison by judges who never selected anyone for EM). Instead this dummy variable is insignificant, with a point estimate of -0.02 .

By the first condition I checked for Instrument Relevance, if it was a weak instrument that provides little information about the variation in X that is exploited by IV regression to estimate the effect of recidivism: the estimate of the coefficient on the endogenous regressor will be estimated inaccurately. Moreover, weak instruments cause the distribution of the estimator to deviate considerably from a normal distribution even in large samples such that the usual methods for obtaining inference about the true coefficient on X may produce wrong results.

3.2 Validating IV assumptions

The heterogeneity in views comes from a combination of ideology and practical considerations. In this case the differences across judges is over what to do with individuals typically caught in the act and with a strong presumption of guilt, before they receive final sentence. These two extreme judicial positions have been widely reported in the paper, liberals and conservatives, where liberal judges regularly assign electronic monitoring but conservative judges never do so.

A model of EM Assignment: Consider a situation with no bail where judge faces a defendant with characteristics x who must decide on pretrial detention in a prison or under EM. If judge assigns the defendant to EM but the defendant escapes with probability $p(x)$ and conditional on escaping will cause harm to others that generates expected costs for the judge $H(x)$, with $p' > 0$ and $H' > 0$. And if judge assigns the defendant to prison, there is no risk of escape but the judge experiences utility differential, where judges ideology is d , when she sends defendant to prison who

could have been treated less harshly. It can be also positive for some judges who enjoy being punitive but for many d will be negative.

$$-p(x)H(x) > d. \quad (3)$$

where judges with $d > 0$ will never use EM because they prefer punitive punishment and if $d = 0$ the use of EM declines flight risk, $p(x)$, and the expected harm given flight, $H(x)$.

The observed assignment to EM is not the product of a uniform tendency across judges. Instead, some judges in the sample tend to assign EM more frequently than others. From 293 judges acting on the 24,362 cases were analyzed, indeed, only 100 have ever used EM and other 200 judges never used EM when it was available to them. (number of EM devices <300). Hence, they constructed a measure of the intervening judge's liberal ideology, *% judge sent to EM*, defined as the proportion of alleged offenders who stood before the intervening judge that were assigned to EM. As an example in table 2 exploits the information on the intervening judge's ideology to investigate an alleged offender's tendency to receive EM in the selection sample. In column 1 of table 2 the EM assignment(its dependent binary variable, one if the alleged offender received EM or zero otherwise) is positively correlated with *% judge sent to EM*, controlling for judicial district dummies. The implied effect was large: a one standard deviation increased in judge's EM rate (0,067) would imply an increase in the likelihood of EM assignment of 0.044(0.663 x 0,067).

Next in column two of the same table includes the crime category for the most serious alleged offence and the omitted category is homicide, noting that the seriousness of the alleged offense has no predictive power and in fact those accused of homicide were more likely to receive EM than those accused of less serious crimes, such as: larceny etc.. The results are inconsistent because of flight risks. The adjusted R^2 of a baseline regression of EM assignment exclusively on judicial district fixed effects is .028. Similarly, when such a baseline regression includes both district fixed effects and the alleged offence categories, at same time it excludes *% judge sent to EM*, R^2 is 0.029. The instrument contributes to the model's fit more than the dummies capturing the seriousness of the imputed offense because the adjusted R^2 jumps from 0.028 to 0.042.

Another ambiguous column was the third column, where it includes the alleged offender's criminal history and reveals a negative and marginally significant coefficient. It was unclear why the severity of the crime, is a measure of police-reported crime that reflects the relative seriousness of individual offences, should have such different impact on judge's evaluation of the flight risk than criminal history. Moreover, the coefficient on *% judge sent to EM* is unaffected by the inclusion of the crime categories or the alleged offender's criminal history as controls.

Finally, in column five they added a proxy for the judge's ideology incorporating a time dimension: judge already used EM, a binary variable one if the judge has ever previously sent an alleged offender to EM prior to facing the particular individual and equals zero otherwise. This instrument was also marginal significant, reinforcing the idea the judges preferences affect EM assignment.

It is obvious concern with empirical strategy, that the allocation of EM to offenders is potentially nonrandom. In Particular, the concern is that judges try to assign EM to individuals who a "kind" type or have a lower risk of re offending in the future. So, in order to deal with the possibility that such unobserved traits bias

the OLS estimate, authors relied on an instrumental variable (IV) approach. Where the recidivism rate was measured by the following regression model:

$$R_i = a + bEM_i + \epsilon_i \quad (4)$$

where R_i is a binary outcome that indicates whether individual i went back to detention after his release, and EM_i is also binary variable that indicates whether individual i was in the electronic monitoring group.

Using *Judge Ideology* to Estimate the Effect of EM on Recidivism: As a benchmark it was helpful to note that in the raw data, the proportion of individuals released from prison who have returned for another crime is 22.37 percent (255/1,140). It is 13,21 percent for the group of 386 alleged offenders released after spending time under EM, for a difference of 9 percentage points. On average, the postrelease period in the sample was 2,85 years. The average yearly recidivism rate for the prison sample was 7,8(22,37/2,85) and for EM it was 4,6(13,21/2,85).

In column one of table five authors used judge dummies as an instrument. In column two of the same table they used as an instrument % judge sent to EM, the percentage of offenders whom the judge sent to electronic monitoring. The percentage were calculated using the full sample of 24,362 individuals and were restricted attention to judges with more than 10 offenders. The coefficient of the regression was still negative and significant and larger in absolute size than the OLS Estimate, implying a reduction of 59 percent of raw recidivism rate following prison.

In measurement analysis I couldn't be enough creative since authors did everything. However, I tried to play around and re-estimation was done via redefining Age and Total detention period variables in years instead of days. Moreover, when focusing on the sub sample of judges who had at least 20 cases instead of at least 10 cases that were originally specified by authors, effect of EM is still statistically significant. However, once we restrict sample to judges who at least had 30 cases, the estimation results become insignificant at 5% significance level. Next is in the original work *escape from EM* was not considered as a recidivism. But once it starts to be counted as a crime, IV estimates fail to give significant results at 5% level.

4 Theory of change analysis

4.1 Educational outcomes after EM

In theory, the difference in these two recidivism rates is ambiguous. On the one hand, specific deterrence theory suggests that spending time under electronic monitoring rather than incarceration might make low punishment salient, implying a positive relationship between light punishment, electronic monitoring, and recidivism. However, many theories show a negative relationship. For instances, imprisonment might be criminogenic through harsh prison conditions or prisoner peer effects that are not observed under electronic monitoring. In particular electronic monitoring could prevent contact with "hardened" criminals or reduce the perception that society is mean. Another also important part is electronic monitoring could differ from prison in its effect on offenders labor market prospects, social integration or educational outcomes after serving with EM. Since the predominant offenders are men at the age 25,3 years old I will focus on the effects of electronic monitoring on young offenders' educational outcomes and it is important to obtain knowledge about whether non- custodial sanctions like EM interrupt young adults education. Mainly this part is based on the literature (**Larsen, 2016**) [5], the study was based on a natural experiment exploiting a reform in Denmark in inum prison sentence of 3 months. Information on the program participation was used to estimate instrumental variable models in order to assess the causal effects of EM on young offenders educational outcomes.

The grounds for changing the Danish legislation in 2006 and giving young offenders with short - term prison sentences the opportunity to serve with electronic monitoring was actually the argument of *maintaining* labor market participation and education enrollment. The theories on deterrence argue that the severity of punishment affects the offenders subsequent criminal behaviour, because the experience of imprisonment changes the offenders perception of sanctions. (**Becker 1968**) [3] Hence, the deterrence perspective would predict higher re-offending rates among offenders serving with EM, compared to offenders serving time in prison. A less severe punishment would have a lower self - deterrent effect and offenders who experience incarceration would be more likely to desist from criminal acts. Offenders serving with EM would therefore have a higher probability of re-offending and thus be more likely to drop out of education and achieve lower educational outcomes due to subsequent imprisonment. On the other hand, theories on social learning and peer effects point to lower re-offending rates among offenders serving with EM with reference to the interactions with other criminals, the training of criminogenic skills and the hardening of offenders in prison. Within this framework, exposure to deviant peers is assumed to prompt higher levels of minor crimes due to deviant attitudes and norms, direct modeling and indirect reinforcing effects. Empirical studies have documented causal peer effects and specialization among inmates, which suggests that offenders in house arrest with electronic surveillance who have fewer interactions with other criminals and more contact with family would be less likely re-offend. (**Bayer et al. 2009**) [2] In regards to educational outcomes, this would imply that offenders serving with EM and with lower recidivism rates are less likely to drop out and have a higher chance of achieving better educational results. Moreover, in EM programs including a provision to attend school, contact with pro-social classmates, instead of criminal inmates, can increase the chances of young offenders continuing n

the school system. Finally the main purpose of this study was to examine the effects of electronic monitoring by looking at a new outcome measure, *education*, which has great significance to the young offenders future. The education is not only important to criminal outcomes, but plays a major role in determining individual life-course outcomes, such as health, income and employment.

The results of this empirical analyses showed significant differences between offenders serving with electronic monitoring and offenders serving in prison when we look at their educational outcomes after release. For offenders between 18 and 25 who fulfill the formal requirements to serve with EM, participation in the EM - program increases the probability of completing upper secondary education by 11 % points after 1.5 years and by 18 % points 3 years after release, compared to imprisonment. The analyses showed long - term effects of electronic monitoring on educational outcomes, when given to young offenders with short - term prison sanctions who are enrolled in education at the time of conviction.

4.2 Recidivism risk versus fixed - sentence regimes

Another topic that I want to discuss is about how parole aboard can help to incentivize prisoners to lower their recidivism risk and cut their sentence time and reduce jails cost. Over the past three decades, many countries didn't take into account of parole boards, which traditionally have had the foresight to release inmates before the expiration of their full sentence, in favor of fixed - sentence regimes in which the original sentence is binding. However, if the prison time lowers recidivism risk and if parole boards can accurately estimate inmates recidivism risk, then relative to fixed - sentence regime, parole can provide allocative - efficiency benefits, that costly prison space is allocated to the highest risk offenders and incentive benefits such as : prisoners know they must reduce their recidivism risk to gain an early release, so give incentives to invest in their own rehabilitation. In a quasi experiment by **Kuziemko 2013** [4], she shows that prison time reduces recidivism risk and that parole boards set prison time in an allocatively efficient manner. Estimates showed that eliminating parole for all prisoners would increase the prison population by 10 % while this means that it increases the crime rate through deleterious effects on recidivism. In her work, where I personally found reasonable, she emphasizes two key ways that release policies can affect these costs. First, policies can be more or less allocatively efficient, it requires that costly prison space be allocated to inmates with the greatest recidivism risk. Second, release policies can incentivize inmates to lower their incarceration and recidivism costs by investing in their own rehabilitation. For instance, inmates who know that they must lower their recidivism risk to gain an early release may behave better in prison, meaning that to lower incarceration costs, or take steps to prepare themselves for successful release or lowering recidivism risk. In her paper, she presents evidence that parole boards appear to perform better along these metrics than do policies in which inmates original sentences are binding. The results from the mass release experiment suggest that parole boards indeed assign longer terms to inmates with greater recidivism risk. Inmates who lost parole eligibility as a consequence of the reform accumulated more disciplinary infraction or violations. Therefore, an increase in the recidivism rate not only increases the prison population but also the overall crime rate, from her estimates is that 30 % increase in the recidivism rate could lead to a 3 % increase in the general crime rate

and a considerably higher increase for serious violent crime such as murder, that means the cost of crime may be very high.

When the criminal justice releases prisoners, it creates trade-off between two significant social costs. On the one hand, keeping individuals in prison is very costly the operating costs of the prisons are billions of dollars annually. On the other hand, the costs associated with overly lenient releases are also high given the vastly disproportionate share of crime committed by the newly released for example, past work of **Raphael and Stoll (2004)** [7] tells that individuals released over the previous 12 months account for 14 % of all murders. So, here you ask, how the choice of release policy can affect total incarceration and recidivism costs? There are two ways to approach this question: first, policies can be more or less allocatively efficient, or an allocatively efficient policy would release a prisoner at the point in time when the marginal cost of his incarceration equals the marginal expected cost of his or her recidivism risk, **Fig. 1** where the recidivism risk falls with time in prison, and thus allocative efficiency requires that lower -risk inmates leave prison sooner **Fig. 2**, given that higher -risk inmates require more time before their expected recidivism risk falls to the level of their incarceration cost.

$$\text{Marginal cost of incarceration} = \text{Marginal expected cost of recidivism risk}$$

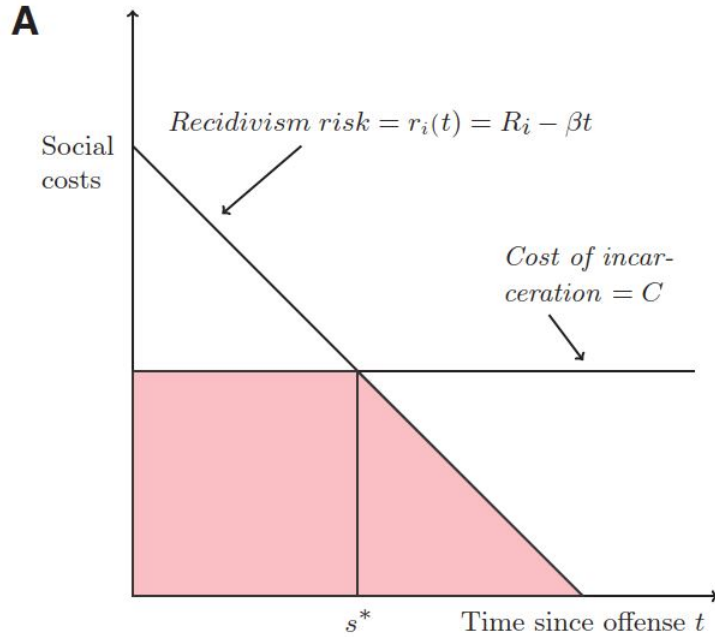


Figure 1: Total social cost at optimal time-served s^* for a linear recidivism-risk function $r_i(t) = R_i - \beta t$

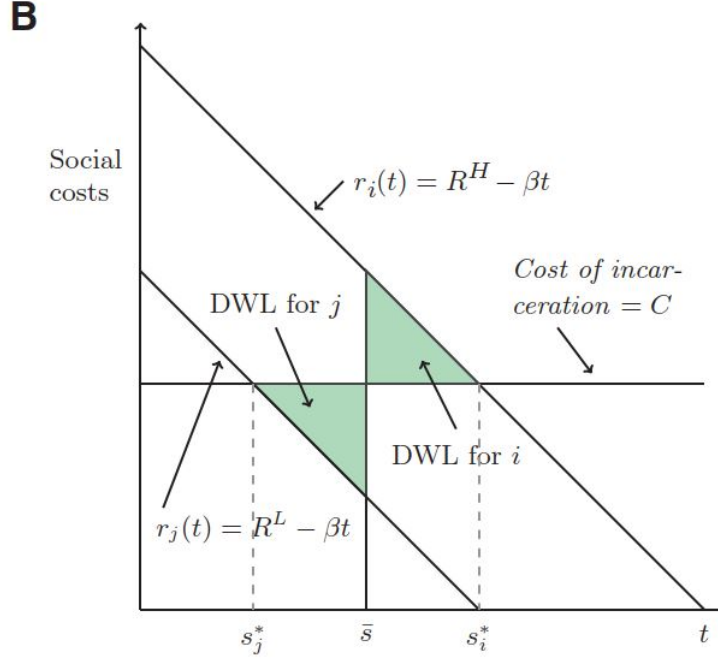


Figure 2: Dead-weight loss of a fixed release date s relative to optimal s_i^* for a high risk ($R_i = R^{High}$) inmate and optimal s_j^* for low risk ($R_j = R^{Low}$)

5 Extensions

5.1 Application of Decision Trees

In this section I decided to predict recidivism for each individual using a statistical model built from given data. Here I wanted to see if an analytical model could perform and predict recidivism to prevent the future criminal individuals or if they were assigned to prison, keep them for a longer time. **(Kuziemko 2013)**[4] In this model I used a method called classification and regression trees, CART. In our case the outcome is binary(0 for no recidivism and 1 for recidivism). I could also used a logistic regression models but they are not easily interpretable. The model coefficients in the Logit regression indicate the importance and relative effect of variables, but do not give a simple explanation of how a decision is made.

A CART model is represented by tree that splits by observation outcome [6]. The first split, as shown in the figure, explains whether year is equal to 1993, 2003, 2004, 2005, 2006, 2007. If yes, the model says to predict red and if no the model moves on to the next split. Then the second split whether or not the number of imprisonment is less than one. If no, the model moves to the last split. The last split checks whether the year of entering to penitentiary system is equal to 1997, 1998, 2000, 2001, 2002. If yes, the model predicts no recidivism, if no, the model predicts recidivism.

After cleaning and omitting all the NAs in the data, I used the method CART. This method builds what is called a tree by splitting on the values of the independent variables. To predict the recidivism for new observation, it is easy to follow the splits in the tree and at the end, I could see the most frequent outcome of recidivism in the training set that followed the same path. The plot below illustrates sample data for

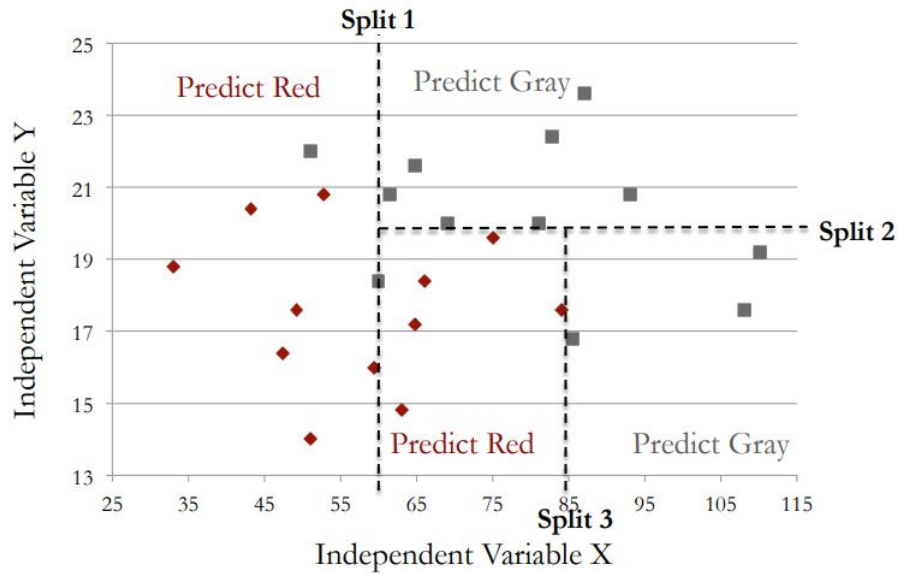


Figure 3: CART split

two independent variables x and y , and each data point is colored by the outcome variable, red or gray. CART tries to split this data into subsets so that each subset is as pure or homogenous as possible. Then the model prediction is just the majority in each subset. If a new observation fell into one of these two subsets then it would predict red, since the majority of the observations in those subsets are red. However, if a new observation fell into one of these two subsets model would predict gray, since the majority of the observations in those two subsets are gray.

From this point, I will tell about the method of CART and related method called *random forests*. To predict the outcome variable (recidivism), I used the same data from 1998 to 2007. My independent variables are electronic monitoring (Dummy 1 if the offender was assigned to EM), number of previous imprisonments (it is number of previous imprisonment that alleged offender had before), judge ever used EM (dummy 1 who ever used EM or liberal judges), age of the offender (it was in days but I could convert them to years), year (the year of entering offender to penitentiary system), family visits (dummy = 1 if the alleged offender received family visits), escapees (dummy = 1 if the alleged offender escaped) and dummies for each type of crime. Moreover, there are different ways to control how many splits are generated. One way is by setting a lower bound for the number of data points in each subset. In R programming it is called the minbucket parameter, for the minimum number of observation in each bucket or subset. The smaller minbucket is the more splits will be generated. But if it's too small, overfitting will occur. This means that CART will fit the training set almost perfectly. But this is bad because then the model will probably not perform well on test set data or new data. On the other hand, if the minbucket parameter is too large the model will be too simple and accuracy will be poor. In our case, I classified observations as either recidivism or no recidivism. Instead of just taking the majority outcome to be the prediction, I computed the percentage of data in a subset of each type of outcome. For instance, if in my subset 10 no recidivism and 2 recidivism, then the model will predict 87 % of no recidivism. And like in logistic regression I used threshold value to obtain the prediction. Since

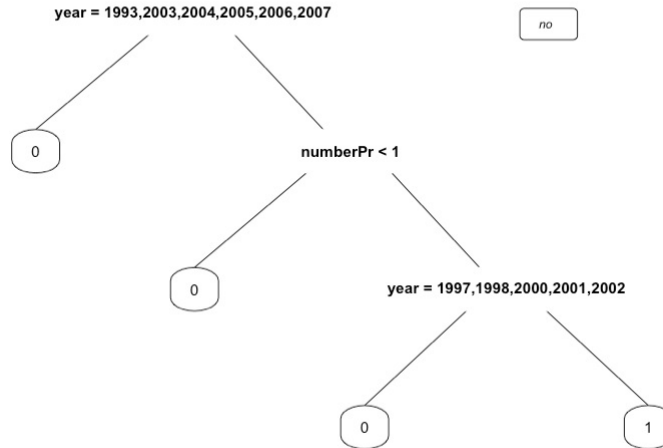


Figure 4: Decision Tree

the majority is no recidivism(0) model predict with threshold of 0.5, no recidivism. However, if I increase that threshold to 0.9 the model would predict recidivism. So by changing the threshold, it is easy to compute ROC curve and compute an AUC value to evaluate the model.

Now, I will briefly go thru on the pipeline. First I splitted the data to *Train*, *Test* and *Cross Validation* datasets. In case, I put 70 % of the data in the training set and other 30 % into testing set. After installing relevant packages into R studio, I built CART model with all X and Y variables stated previously. As a result I observed a confusion matrix on test dataset that model got correct 354 no recidivism(0) that are actually no recidivism and 5 for recidivism that are actually recidivism(1). First I calculated the diagonal and divided for total number of observations in the table. From here I got accuracy of the model was 0.79. Lastly, to evaluate the model I generated ROC curve using ROCR package. ROC gives me probability of outcomes, in this case (0,1) for each observation in the test set. Quick note about AUC. It is an area under the ROC curve of 0.8, means that randomly selected case from group with the target equals Y has a score larger than that for a randomly chosen case from the group with the target equals N in 80 % of the time.

In addition, I used a method that is similar to CART model named *random forests*. This method was designed to improve the predicition accuracy of CART and works by building a large number of CART trees, Unfortunately, this makes the method less interpretable than CART. So, there is trade off whether you need interpretability or the increase in accuracy more. To make prediction for a new observation each tree in the forest votes on the outcome and I pick the outcome that receives the majority of the votes. So here comes a question how does random forest build many CART trees? **Random Forests** only allows each tree to split on a random subset of the available independent variables, and each tree is built from what people call bagged or bootstrapped sample of the data. This means that the data used as the training data for each tree is selected randomly with replacement.

Being short, the random forest model improved accuracy a little bit over CART.

6 Discussion

6.1 Welfare

After several years, it is clear that EM has been almost desperately applied without adequate vision, planning, program integration, staff training, and concurrent research. It has punished, perhaps more humanely and cheaply, and it has been an element in the avoidance of prison crowding and prison construction, but it is not free and it is not without unintended effects.

In times with large prison populations and discussions regarding downsizing, it is important to recognise that alternative sanctions can have positive influences on young offenders future outcome, both in terms of recidivism and in regards to labour participation **Andersen 2014** [1] and educational outcomes. These findings are important in order to evaluate the returns of introducing noncustodial sanctions and to inform future policy decisions and discussion on electronic monitoring. It will, however, be important to do further qualitative and quantitative research to achieve a deeper understanding of the mechanisms behind these results: *how* the different program elements affect offenders future outcomes, *how* sanction types influence educational trajectories and *how* different outcomes like crime, education and labour market participation are interrelated.

Moreover, I personally find an interesting approach to investigate the optimal assignment of EM according to a cutoff rule. After deriving each individual probability expectation of recidivism and calculating their cost of incarceration, I expect that only those subjects, whose cost of recidivism is below the cost of incarceration, will be released under EM. Since the cutoff will induce a discontinuity in the distribution, we should only use the inmates which are below the cutoff for investigating the causal effects in future outcomes.

7 Conclusion

In conclusion, authors seeked to contribute to this debate by providing an estimate of the effect on recidivism of having a person serve time under electronic surveillance instead of prison. One of the key challenges they faced is that, ideally they tried to compare similar individuals after their release from EM and prison. This was rarely observed in practice because judicial allocation decisions are obviously influenced by the offenders perceived "meanness" and risk of recidivism. In this paper, they showed several essentials of the legal system in Argentina to study the effect of an EM program, where it is used to substitute for imprisonment of alleged offenders never receive final sentence. In fact, liberal - learning judges have allocated EM to individuals accused of very serious crimes(including homicide) and with prior records of imprisonment.

Exploiting random assignment to judges with differing inclinations to assign EM, the main IV estimates suggest that treating alleged offenders with electronic monitoring instead of prison induces a large and significant reduction in recidivism of between 11 and 16 percentage points. In this replication of paper, I replicated all

tables, checked IV validity assumptions , minor measurement analysis, described how original paper could be extended to investigate in the future, and got recidivism probabilities for each individual in the data set for the given characteristics to implement them in further analysis too.

8 Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

9 Appendix

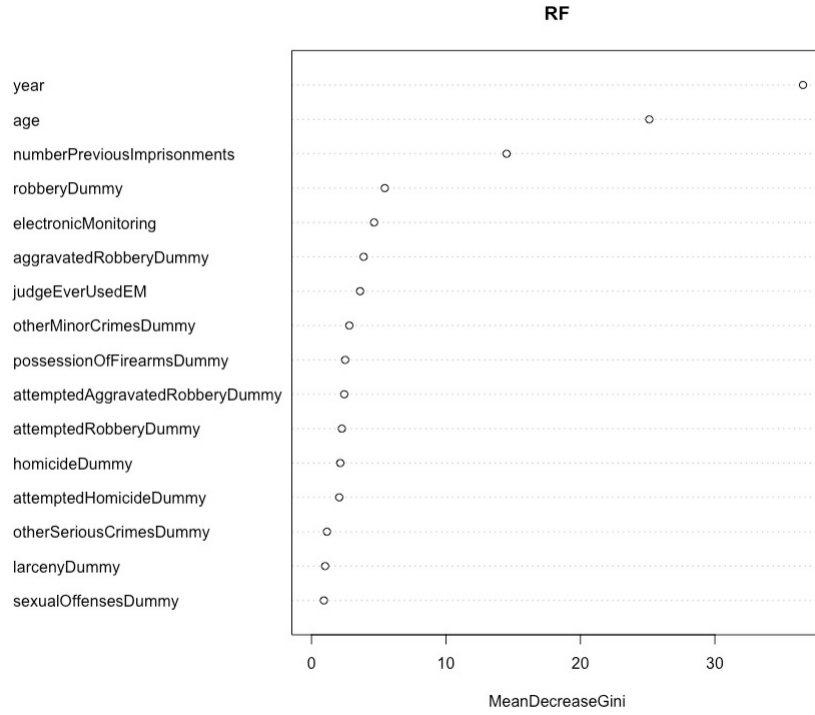


Figure 5: Feature Importance Plot

Table 1: Electronic Monitoring Assignment and Type of Crimes

	Dependent variable:				
	(I)	(II)	electronicMonitoring (III)	(IV)	(V)
percJudgeSentToEM	0.663*** (0.085)	0.660*** (0.085)	0.660*** (0.085)	0.660*** (0.085)	0.637*** (0.087)
judgeAlreadyUsedEM					0.006** (0.002)
mostSeriousCrime02 - Attempted homicide		-0.001 (0.008)	-0.001 (0.008)	-0.001 (0.008)	-0.001 (0.008)
mostSeriousCrime03 - Sexual offenses		0.001 (0.008)	0.001 (0.008)	0.0004 (0.008)	0.0003 (0.008)
mostSeriousCrime04 - Other serious crimes		-0.001 (0.007)	0.0001 (0.007)	0.0001 (0.007)	-0.00000 (0.007)
mostSeriousCrime05 - Aggravated robbery		-0.004 (0.004)	-0.003 (0.004)	-0.003 (0.004)	-0.003 (0.004)
mostSeriousCrime06 - Attempted aggravated robbery		-0.014*** (0.004)	-0.013*** (0.004)	-0.012*** (0.004)	-0.012*** (0.004)
mostSeriousCrime07 - Robbery		-0.011*** (0.004)	-0.010** (0.004)	-0.009** (0.004)	-0.009** (0.004)
mostSeriousCrime08 - Attempted robbery		-0.009* (0.005)	-0.007 (0.005)	-0.005 (0.005)	-0.006 (0.005)
mostSeriousCrime09 - Possession of firearms		-0.008 (0.005)	-0.007 (0.005)	-0.006 (0.006)	-0.007 (0.005)
mostSeriousCrime10 - Larceny / Attempted larceny		-0.012*** (0.005)	-0.010** (0.005)	-0.009* (0.005)	-0.009* (0.005)
Constant	0.003 (0.005)	0.009 (0.006)	0.010 (0.006)	-2.677*** (0.728)	-2.389*** (0.722)
Observations	24,003	24,003	24,003	23,928	23,928
R ²	0.043	0.044	0.045	0.046	0.046
Adjusted R ²	0.042	0.043	0.043	0.045	0.045
Residual Std. Error	0.122 (df = 23984)	0.122 (df = 23974)	0.122 (df = 23973)	0.122 (df = 23894)	0.122 (df = 23893)
F Statistic	59.209*** (df = 18; 23984)	39.347*** (df = 28; 23974)	38.579*** (df = 29; 23973)	34.821*** (df = 33; 23894)	34.054*** (df = 34; 23893)

Table 2:

	<i>Dependent variable:</i>				
	(I)	(II)	electronicMonitoring (III)	(IV)	(V)
percJudgeSentToEM	0.663*** (0.085)	0.660*** (0.085)	0.660*** (0.085)	0.660*** (0.085)	0.637*** (0.087)
judgeAlreadyUsedEM					0.006** (0.002)
mostSeriousCrime02 - Attempted homicide		-0.001 (0.008)	-0.001 (0.008)	-0.001 (0.008)	-0.001 (0.008)
mostSeriousCrime03 - Sexual offenses		0.001 (0.008)	0.001 (0.008)	0.0004 (0.008)	0.0003 (0.008)
mostSeriousCrime04 - Other serious crimes		-0.001 (0.007)	0.0001 (0.007)	0.0001 (0.007)	-0.00000 (0.007)
mostSeriousCrime05 - Aggravated robbery		-0.004 (0.004)	-0.003 (0.004)	-0.003 (0.004)	-0.003 (0.004)
mostSeriousCrime06 - Attempted aggravated robbery		-0.014*** (0.004)	-0.013*** (0.004)	-0.012*** (0.004)	-0.012*** (0.004)
mostSeriousCrime07 - Robbery		-0.011*** (0.004)	-0.010** (0.004)	-0.009** (0.004)	-0.009** (0.004)
mostSeriousCrime08 - Attempted robbery		-0.009* (0.005)	-0.007 (0.005)	-0.005 (0.005)	-0.006 (0.005)
mostSeriousCrime09 - Possession of firearms		-0.008 (0.005)	-0.007 (0.005)	-0.006 (0.006)	-0.007 (0.005)
mostSeriousCrime10 - Larceny / Attempted larceny		-0.012*** (0.005)	-0.010** (0.005)	-0.009* (0.005)	-0.009* (0.005)
Constant	0.003 (0.005)	0.009 (0.006)	0.010 (0.006)	-2.677*** (0.728)	-2.389*** (0.722)
Observations	24,003	24,003	24,003	23,928	23,928
R ²	0.043	0.044	0.045	0.046	0.046
Adjusted R ²	0.042	0.043	0.043	0.045	0.045
Residual Std. Error	0.122 (df = 23984)	0.122 (df = 23974)	0.122 (df = 23973)	0.122 (df = 23894)	0.122 (df = 23893)
F Statistic	59.209*** (df = 18; 23984)	39.347*** (df = 28; 23974)	38.579*** (df = 29; 23973)	34.821*** (df = 33; 23894)	34.054*** (df = 34; 23893)

Table 3: Recidivism and Electronic Monitoring: OLS Regression

	<i>Dependent variable:</i>				
	(I)	OLS (II)	recidivism probit (III)	(IV)	OLS (V)
electronicMonitoring	-0.092*** (0.021)	-0.090*** (0.024)	-0.405*** (0.107)	-0.089*** (0.024)	-0.086*** (0.023)
judgeEverUsedEM					-0.017 (0.024)
mostSeriousCrime02 - Attempted homicide		0.010 (0.064)		0.011 (0.064)	0.011 (0.063)
mostSeriousCrime03 - Sexual offenses		-0.038 (0.052)		-0.045 (0.052)	-0.029 (0.050)
mostSeriousCrime04 - Other serious crimes		0.200** (0.086)		0.202** (0.087)	0.226*** (0.088)
mostSeriousCrime05 - Aggravated robbery		0.008 (0.034)		0.010 (0.034)	0.015 (0.033)
mostSeriousCrime06 - Attempted aggravated robbery		-0.032 (0.049)		-0.033 (0.049)	-0.015 (0.048)
mostSeriousCrime07 - Robbery		0.027 (0.050)		0.025 (0.050)	0.053 (0.050)
mostSeriousCrime08 - Attempted robbery		0.046 (0.055)		0.045 (0.055)	0.064 (0.054)
mostSeriousCrime09 - Possession of firearms		0.080 (0.059)		0.087 (0.060)	0.072 (0.058)
mostSeriousCrime10 - Larceny / Attempted larceny		-0.019 (0.083)		-0.014 (0.086)	-0.006 (0.080)
Constant	0.224*** (0.012)	116.588*** (36.256)	527.711*** (113.254)	115.193*** (36.845)	1.622*** (0.305)
Observations	1,526	1,526	1,513	1,526	1,503
R ²	0.010	0.185		0.180	0.177
Adjusted R ²	0.009	0.162		0.158	0.163
Log Likelihood			-611.823		
Akaike Inf. Crit.			1,273.645		
Residual Std. Error	0.399 (df = 1524)	0.367 (df = 1484)		0.368 (df = 1485)	0.368 (df = 1477)
F Statistic	15.210*** (df = 1; 1524)	8.212*** (df = 41; 1484)		8.154*** (df = 40; 1485)	12.685*** (df = 25; 1477)

Table 4: Recidivism and Electronic Monitoring: IV Regressions

	Dependent Variable: recidivism				
	IV: court (1)	IV: percJudgeSentToEM (2)	IVs: percJudgeSentToEM, judgeAlreadyUsedEM (3)	IV probit (4)	IV: largeSampleEstimate (5)
EM1_pred	-0.112** (0.057)				
EM12_pred		-0.138 (0.093)			
EM13_pred			-0.158* (0.091)		-0.158* (0.091)
EM14_pred				-0.174* (0.095)	
Constant	92.101*** (8.341)	91.615*** (8.349)	91.534*** (8.344)	91.692*** (8.353)	91.534*** (8.344)
Observations	1,503	1,503	1,503	1,494	1,503
R ²	0.168	0.167	0.167	0.167	0.167
Adjusted R ²	0.149	0.148	0.148	0.148	0.148
Residual Std. Error	0.371 (df = 1469)	0.371 (df = 1469)	0.371 (df = 1469)	0.372 (df = 1461)	0.371 (df = 1469)
F Statistic	8.985*** (df = 33; 1469)	8.897*** (df = 33; 1469)	8.924*** (df = 33; 1469)	9.136*** (df = 32; 1461)	8.924*** (df = 33; 1469)

Table 5: Recidivism and Escape within EM

	Dependent variable:		
	(I) recidivism	(II)	(III) escapees
EMLengthTotalPrisonRatio	-0.087 (0.056)		
mostSeriousCrime02 - Attempted homicide	0.079 (0.118)	0.073 (0.124)	-0.088 (0.079)
mostSeriousCrime03 - Sexual offenses	0.135 (0.129)	0.122 (0.127)	0.209 (0.163)
mostSeriousCrime04 - Other serious crimes	-0.071 (0.080)	-0.071 (0.079)	0.172 (0.152)
mostSeriousCrime05 - Aggravated robbery	0.007 (0.062)	0.001 (0.062)	-0.004 (0.076)
mostSeriousCrime06 - Attempted aggravated robbery	0.067 (0.107)	0.053 (0.108)	-0.120 (0.104)
mostSeriousCrime07 - Robbery	0.151 (0.105)	0.144 (0.105)	-0.032 (0.113)
mostSeriousCrime08 - Attempted robbery	-0.030 (0.093)	-0.046 (0.092)	-0.020 (0.113)
mostSeriousCrime09 - Possession of firearms	-0.026 (0.094)	-0.041 (0.093)	0.062 (0.114)
mostSeriousCrime10 - Larceny / Attempted larceny	-0.142 (0.149)	-0.156 (0.142)	0.098 (0.206)
Constant	65.583*** (16.191)	65.105*** (16.205)	6.985 (20.380)
Observations	386	386	386
R ²	0.209	0.203	0.106
Adjusted R ²	0.140	0.136	0.030
Residual Std. Error	0.314 (df = 354)	0.315 (df = 355)	0.371 (df = 355)
F Statistic	3.018*** (df = 31; 354)	3.021*** (df = 30; 355)	1.401* (df = 30; 355)

Table 6: Recidivism and Electronic Monitoring: Robustness

	Dependent Variable: recidivism			
	(1)	(2)	(3)	(4)
EM161_pred	-0.243** (0.105)			
EM162_pred		-0.151 (0.117)		
EM163_pred			-0.154* (0.091)	
EM165_pred				-0.210** (0.094)
incomeProfession	0.00005* (0.00002)			
spouse		0.029 (0.025)		
familyVisits		0.048 (0.068)		
Constant	80.358*** (10.415)	91.259*** (9.856)	94.417*** (8.454)	89.300*** (8.442)
Observations	941	1,147	1,503	1,441
R ²	0.162	0.169	0.170	0.169
Adjusted R ²	0.131	0.143	0.150	0.150
Residual Std. Error	0.372 (df = 906)	0.383 (df = 1111)	0.371 (df = 1467)	0.368 (df = 1407)
F Statistic	5.156*** (df = 34; 906)	6.471*** (df = 35; 1111)	8.574*** (df = 35; 1467)	8.685*** (df = 33; 1407)

References

- [1] Lars H. Andersen and Signe H. Andersen. Effect of electronic monitoring on social welfare dependence. *Criminology & Public Policy*, 13(3):349–379, 2014.
- [2] EA Bayer, SP Smith, I Noach, O Alber, JJ Adams, R Lamed, LJW Shimon, and F Frolov. Can we crystallize a cellulosome. *Biotechnology of lignocellulose degradation and biomass utilization*, 183:205, 2009.
- [3] Gary S. Becker. Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2):169–217, 1968.
- [4] Ilyana Kuziemko. How should inmates be released from prison? an assessment of parole versus fixed sentence regimes. *Quarterly Journal of Economics*, 128(1):371–424, 2013.
- [5] Britt Østergaard Larsen. Educational outcomes after serving with electronic monitoring: Results from a natural experiment. *Journal of Quantitative Criminology*, 33(1):157–178, Mar 2017.
- [6] MIT. Mitx 15.071x - the analytics edge. 2016.
- [7] Steven Raphael, Michael A. Stoll, Mark Duggan, and Anne Morrison Piehl. The effect of prison releases on regional crime rates [with comments]. *Brookings-Wharton Papers on Urban Affairs*, pages 207–255, 2004.

- [8] Lawrence Sherman and Heather Strang. Restorative justice as evidence-based sentencing. *The Oxford Handbook of Sentencing and Corrections*, 01 2012.
- [9] Joel Waldfogel. Aggregate inter-judge disparity in federal sentencing: evidence from three districts. *Federal Sentencing Reporter*, 4:151–154, 11 1991.