Incarceration, Recidivism and Employment

Manudeep Bhuller* Gordon B. Dahl† Katrine V. Løken‡ Magne Mogstad§

September 9, 2016

Abstract: Understanding whether, and in what situations, time spent in prison is criminogenic or preventive has proven challenging due to data availability and correlated unobservables. This paper overcomes these challenges in the context of Norway's criminal justice system, offering new insights into how incarceration affects subsequent crime and employment. We construct a panel dataset containing the criminal behavior and labor market outcomes of the entire population, and exploit the random assignment of criminal cases to judges who differ systematically in their stringency in sentencing defendants to prison. Using judge stringency as an instrumental variable, we find that imprisonment discourages further criminal behavior, and that the reduction extends beyond incapacitation. Incarceration decreases the probability an individual will reoffend within 5 years by 27 percentage points, and reduces the number of offenses over this same period by 10 criminal charges. In comparison, OLS shows positive associations between incarceration and subsequent criminal behavior. This sharp contrast suggests the high rates of recidivism among ex-convicts is due to selection, and not a consequence of the experience of being in prison. Exploring factors that may explain the preventive effect of incarceration, we find the decline in crime is driven by individuals who were not working prior to incarceration. Among these individuals, imprisonment increases participation in programs directed at improving employability and reducing recidivism, and ultimately, raises employment and earnings while discouraging further criminal behavior. Contrary to the widely embraced 'nothing works' doctrine, these findings demonstrate that time spent in prison with a focus on rehabilitation can indeed be preventive.

Keywords: crime, employment, incarceration, recidivism
JEL codes: K42, J24
Acknowledgements: We thank Derek Neal and seminar participants at several universities and conferences for valuable feedback and suggestions. We are grateful to Baard Marstrand at the Norwegian Courts Administration for help in accessing the data and in understanding institutional details. The project received generous financial support from the Norwegian Research Council.

*Department of Economics, University of Chicago; Research Department, Statistics Norway. Email: bhuller@uchicago.edu

†Department of Economics, UC San Diego. E-mail: gdahl@ucsd.edu

‡Department of Economics, University of Bergen. E-mail: katrine.loken@econ.uib.no

§Department of Economics, University of Chicago; Research Department, Statistics Norway; NBER. Email: magne.mogstad@gmail.com

1 Introduction

Over the past several decades, incarceration rates have risen dramatically in many OECD countries. In the U.S., for example, the incarceration rate has increased from 220 per 100,000 residents in 1980 to over 700 per 100,000 in 2012. In Europe, the increases (and levels) tend to be smaller but still substantial, with the average incarceration rate per 100,000 residents rising from 62 in 1980 to 112 in 2010 in Western European nations.[1] These increases raise important questions about how well ex-convicts reintegrate into society after incarceration, and in particular, whether they return to a life of crime. Prison time could convince offenders that crime does not pay, or rehabilitate them by providing vocational and life skills training. Conversely, prison time could cause human capital to depreciate, expose offenders to hardened criminals, or limit opportunities due to employment discrimination or societal stigma. Indeed, the effects of incarceration could vary in magnitude and sign depending on a prisoner's background (e.g., work history), as well as prison conditions (e.g., availability of prison programs and sentence lengths).

Understanding whether, and in what situations, time spent in prison is criminogenic or preventive has proven challenging for several reasons. One problem is data availability. The ideal dataset would be a long and representative panel with individual-level information on criminal behavior and labor market outcomes. In most countries, however, the required data sources cannot be accessed and linked together. Another major challenge is the threat to identification from correlated unobservables. While ex-convicts have relatively high rates of criminal activity and weak labor market attachment, these correlations could be driven by their unobserved characteristics as opposed to the experience of being in prison.

Due to these challenges, evidence on the causal effects of incarceration is scarce. Nagin et al. (2009), in their review article, summarize the state of the literature well: "Remarkably little is known about the effects of imprisonment on reoffending. The existing research is limited in size, in quality, [and] in its insights into why a prison term might be criminogenic or preventative." Our paper overcomes both the data and the identification challenges in the context of Norway's criminal justice system, offering new insights into how imprisonment affects subsequent criminal behavior.

Our work draws on two strengths of the Norwegian environment. First, by linking several administrative data sources we are able to construct a panel dataset containing complete records of the criminal behavior and labor market outcomes of every Norwegian. Second, we

[1] These figures come from the World Prison Brief (2016). The Western European countries used to construct the population-weighted average include Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the UK.

address threats to identification by exploiting the random assignment of criminal cases to Norwegian judges who differ systematically in their stringency. Our baseline sample consists of suspected criminals who appear in court with a non-confession case.[2] Our measure of judge stringency is the average incarceration rate in other cases a judge has handled. This stringency measure serves as an instrument for incarceration since it is highly predictive of the judge's decision in the current case, but as we document, uncorrelated with observable case characteristics.

Our paper offers three sets of results. First, imprisonment discourages further criminal behavior. Using our measure of judge stringency as an instrument, we estimate that incarceration lowers the probability of reoffending within 5 years by 27 percentage points and reduces the corresponding number of criminal charges per individual by 10. These reductions are not simply due to an incapacitation effect. We find sizable decreases in reoffending probabilities and cumulative charged crimes even after defendants are released from prison.

Second, bias due to selection on unobservables, if ignored, leads to the erroneous conclusion that time spent in prison is criminogenic. Consistent with existing descriptive work, our OLS estimates show positive assocations between incarceration and subsequent criminal behavior. This is true even when we control for a rich set of demographic and crime category controls. Using the panel structure of our data reduces the estimates somewhat, but there are noticeable changes in crime and employment in the year prior to the court case, raising concerns about the validity of offender fixed effects or lagged dependent variable models. In contrast, our IV estimates, which address the issues of selection bias and reverse causality, find that incarceration is strongly preventive for many individuals, both on the extensive and intensive margins of crime.

Third, the reduction in crime is driven by individuals who were not working prior to incarceration. Among these individuals, imprisonment increases participation in programs directed at improving employability and reducing recidivism, and ultimately, raises employment and earnings while discouraging criminal behavior.[3] The effects of incarceration for this group are large and economically important. Imprisonment causes a 34 percentage point increase in participation in job training programs for the previously nonemployed, and within 5 years, their employment rate increases by 40 percentage points. At the same time, the likelihood of reoffending within 5 years is cut in half (by 46 percentage points), and the average number

[2] In Norway, both confession and non-confession cases are settled by trial. In confession cases, however, the accused admits guilt to the police/prosecutor before his case is assigned to a judge (who ultimately decides on both guilt and sentencing).

[3] Since we observe charges and not actual crimes committed, it is in theory possible that ex-convicts do not, in fact, reduce their criminal activity, but rather learn how to avoid being caught while in prison. The fact that incarceration increases formal sector employment, which is a time substitute for criminal activity, suggests this explanation is unlikely.

of criminal charges falls by 22. A very different pattern emerges for individuals who were previously attached to the labor market. Among this group, there is no significant effect of incarceration on either the probability of reoffending or the number of charged crimes. Moreover, they experience an immediate 25 percentage point drop in employment due to incarceration and this effect continues out to 5 years. This drop is driven almost entirely by defendants losing their job with their previous employer while they are in prison.

Taken together, our findings have important implications for ongoing policy debates over the growth in incarceration rates and the nature of prison. Our estimates indicate that the high rates of recidivism among ex-convicts is due to selection, and not a consequence of the experience of being in prison. Indeed, the Norwegian prison system is successful in discouraging crime and encouraging employment, largely due to changes in the behavior of individuals who were not working prior to incarceration. These individuals had no job to lose, and low levels of education and work experience. Norwegian prisons offer them access to rehabilitation programs, job training and re-entry support. Upon release, these previously unemployed individuals become more attached to the formal labor market, and find crime relatively less attractive. In contrast, for individuals with some attachment to the labor market, many of them had an actual job to lose and human capital to depreciate by going to prison. These negative effects may well offset any positive impacts of rehabilitation, and therefore help explain why incarceration does not seem to materially affect their criminal behavior or labor market outcomes.

Our paper contributes to a large literature across the social sciences on the impact of incarceration on both recidivism and future employment. Much of this literature focuses on incapacitation effects, finding reductions in crime while offenders are in prison.[4] There is less evidence on longer-term recidivism, and the findings are mixed. In terms of labor market outcomes, OLS studies usually find either negative or no effect on earnings and employment.[5] More sophisticated work uses panel data and offender fixed effects to minimize selection issues. For recidivism, there are fewer studies using this approach and the evidence is mixed, while for labor market outcomes a handful of studies find either no impact or a negative effect.[6]

[4] Recent studies in economics isolating incapacitation effects include Barbarino and Mastrobuoni (2014), Buonanno and Raphael (2013), and Owens (2009). We refer to Chalfin and McCrary (forthcoming) for a recent review of the extensive literature on criminal deterrence.

[5] For example, Bernburg et al. (2006), Gottfredson (1999), and Brennan and Mednick (1994) all reach different conclusions for recidivism. For a summary of observational research on labor market outcomes, see Western et al. (2001).

[6] See Freeman (1992) and Western and Beckett (1999) for early papers using panel data. Other evidence based on fixed effects or even study design include Grogger (1995), Kling (1999), Skardhammer and Telle (2012), and Waldfogel (1994).

More closely related to our paper, some recent work has relied on the quasi-random assignment of judges to study the effects of incarceration.[7] While each of these studies uses data from the U.S, the findings are mixed. Kling (2006) presents results suggesting that time in prison improves labor market outcomes after release, although the IV estimates based on quasi-random assignment of judges are too imprecise to draw firm conclusions. Green and Winik (2010) and Loeffler (2013) report no detectable effects of incarceration on recividism, whereas Aizer and Doyle (2015) find that juvenile incarceration results in lower high school completion rates and higher adult incarceration rates. Mueller-Smith (2015) uses data from Texas to investigate the impacts of adult incarceration and reports that incarceration increases recidivism rates, and worsens labor market outcomes.

There are several possible reasons why no consensus has emerged as to how well ex-convicts reintegrate into society. While quasi-random assignment of judges can be useful to address concerns over correlated unobservables, there remain issues that could bias the estimates. In Green and Winik (2010), for instance, the estimation sample is small and the instrument is weak, which may lead to severe bias in the IV estimates. Mueller-Smith (2015) additionally explores the importance of two other issues. He argues in his setting that standard instrumental variable estimates could be biased due to violation of the exclusion and monotonicity assumptions. To assess the relevance and validity of our instrument, we therefore perform a number of checks, all of which suggest that our instrument is strong, as good as randomly assigned, and satisfying exclusion and monotonicity.

Another possible explanation for the lack of consensus is that incarceration effects could vary depending on a prisoner's background or prison conditions. As documented later, prisoners in Norway have broadly similar observable characteristics as prisoners in many other countries. Instead, what is quite distinct, especially compared to the U.S., is the prison system. In Scandinavian countries like Norway, the prison system focuses on rehabilitation, preparing inmates for life on the outside.[8] This is done in part by investing in education and training programs, but also through extensive use of "open prisons" in which prisoners are

[7] Research designs using quasi-random assignment of cases to examiners or judges have provided important evidence in a variety of contexts in recent years. For example, Dobbie et al. (2016) and Stevenson (2016) use the detention tendencies of quasi-randomly assigned bail judges to estimate the causal effects of pre-trial detention on subsequent defendant outcomes, whereas Di Tella and Schargrodsky (2013) use a similar strategy to estimate negative causal effects on criminal recidivism of electronic monitoring relative to prison. For studies using quasi-random assignment of examiners or judges in contexts other than crime, see e.g. Autor et al. (2015), Belloni et al. (2012), Dahl et al. (2014), Dobbie and Song (2015), Doyle (2007, 2008), Doyle et al. (2012), French and Song (2014), and Maestas et al. (2013).

[8] A recent New York Times article summarizes the system's rehabilitative aims: "The goal of the Norwegian penal system is to get inmates out of it... 'Better out than in' is an unofficial motto of the Norwegian Correctional Service... It works with other government agencies to secure a home, a job and access to a supportive social network for each inmate before release."

housed in low-security surroundings and allowed frequent visits to families while electronically monitored.[9] In comparison, in many other countries rehabilitation has taken a back seat in favor of prison policies emphasizing punishment and incapacitation. In the U.S., a pivotal point was the 1974 Martinson report, concluding that "nothing works" in rehabilitating prisoners (Martinson, 1974; Lipton et al., 1975). While influential, leading criminology scholars have questioned the evidence base for this conclusion (e.g., see the review in Cullen, 2005). Our study serves as a proof-of-concept demonstrating that time spent in prison with a focus on rehabilitation can indeed be preventive.[10]

The remainder of the paper proceeds as follows. The next section provides background on the Norwegian court system, describes how criminal cases are assigned to judges, and outlines the baseline IV model. Section 3 presents our data. This section also describes similarities and differences in the criminal population and the criminal justice system of Norway versus other countries. In Section 4, we discuss our instrument and its validity. Section 5 presents our main results for recidivism, while Section 6 documents the important role of employment in reducing recidivism. Section 7 concludes.

2 Research Design

In this section, we describe our research design. We begin by reviewing key aspects of the criminal justice system in Norway, documenting how criminal court cases are assigned to judges. We then describe how our empirical model uses the random assignment of cases to judges to estimate the effects of incarceration on subsequent criminal behavior and labor market outcomes.

2.1 The Norwegian Court System

The court system in Norway consists of three levels: the district court, the court of appeals, and the supreme court. The vast majority of cases are settled at the district court level. In this paper, we focus on criminal cases tried in one of the 87 district courts in existence at

[9] Other European countries are trying out open prisons. See, for example, Mastrobuoni and Terlizzese (2015) who study an open prison in Italy (Bellato prison). They show that inmates transferred to this prison commit fewer crimes after release.

[10] The existing evidence base is scarce, and does not answer our research question of whether, and in what situations, imprisonment as compared to not being incarcerated is preventive or criminogenic. Kuziemko (2013) uses data on inmates in Georgia, and finds that access to parole boards increases participation in rehabilitation programs and reduces recidivism. There are also a few randomized controlled trials in the U.S. focusing primarily on post-release training and education programs for ex-convicts. These studies have estimated zero or small (and often imprecise due to small samples) effects on long-term labor market and recidivism outcomes (see Cook et al. 2014, Redcross et al. 2012, and Visher et al. 2005).

one time or another in Norway during the period of our study. The largest district court is located in Oslo and has around 100 judges, while the smallest courts only have a few judges.

There are two types of professional judges in district courts, regular judges and deputy judges. Regular judges are appointed civil servants, and can only be dismissed for malfeasance. One of the regular judges is appointed as chief judge to oversee the administration of the local court. In 2010 there were 370 full-time regular judges (including chief judges); their average age was 53 and 62 percent were male. Deputy judges, like regular judges, are also law school graduates, but are appointed to a court for a limited period of time which cannot exceed three years.[11] Deputy judges have a somewhat different caseload compared to regular judges, as discussed in the next subsection. Not all deputy judges become regular judges, and those that do typically need several of years of experience in other legal settings before applying for and being appointed as a regular judge.[12]

Criminal cases are classified into two broad types, confession cases and non-confession cases. Both types of cases are settled by trial (as opposed to the U.S. system which has plea bargains). In confession cases, the accused has confessed to the police/prosecutor before his case is assigned to a judge. The confession is entered into evidence, but the prosecution is not absolved of the duty to present a full case and the judge may still decide that the defendant is innocent.[13] In practice, most confession cases are relatively straightforward. To save on time and costs, they are therefore heard by a single professional judge who decides on sentencing. Non-confession cases are heard by a panel of one professional and two lay judges, or in the case of extremely serious crimes, by two professional judges and three lay judges. The lay judges are individuals chosen from the general population to serve for a limited term.[14] The professional judge presides over the case, while the lay judges participate on the questions of guilt and sentencing. As opposed to professional judges, lay judges hear only a few cases a year.[15]

One advantage of the Norwegian criminal justice system compared to some other countries

[11] Oslo is an exception, where deputy judges can serve for up to five years.

[12] In Norway, it has been a fundamental principle in the recruitment of judges that the judiciary should reflect the broadest possible professional legal background. For this reason, Norwegian judges are not recruited to internal careers in the courts as soon as they graduate from law school. The deputy judge arrangement was therefore introduced to give law graduates practical experience with the courts.

[13] These rules apply to most civil law systems, in contrast to common law systems where a majority of criminal cases are settled by confession and plea bargain rather than by a trial.

[14] The municipal council appoints lay judges for four year periods. To be a lay judge, one must satisfy certain requirements, such as not having a criminal record and not working in certain occupations (e.g., police officer, judge, lawyer, public official, government officer). In a municipal district the pool of lay judges is usually between 30-60 individuals. Lay judges are partially compensated for days absent from work if not covered by their employer; students and nonworkers get a small compensation of around $40 per day.

[15] We do not observe the identify of the lay judges in our data, but since they are randomly assigned to judges within a court, they should not create any bias in our estimates.

is that it has no plea bargaining. For example, in the U.S. criminal defendants often know their assigned judge before deciding whether to plead guilty in exchange for a reduced sentence. The fact that these pre-trial strategies are not taking place in our setting makes the interpretation of our IV estimates based on judge stringency easier to interpret (see Dobbie et al., 2016).

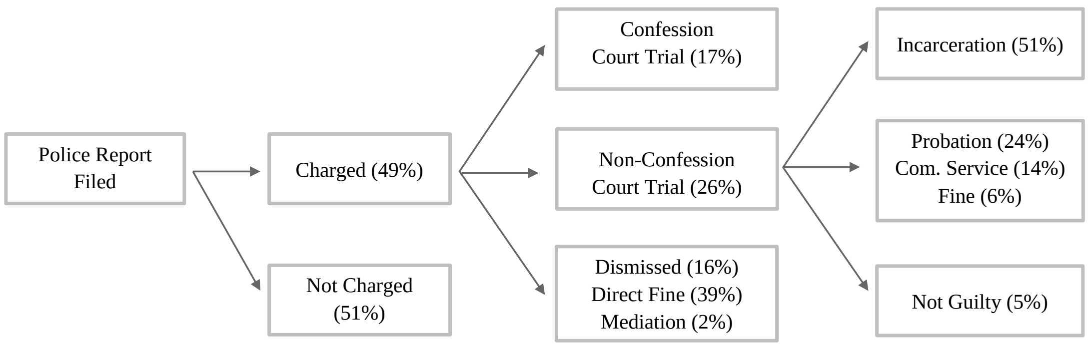


Figure 1. Processing of Suspected Crimes in Norway's Criminal Justice System.

Note: Sample consists of all criminal cases reported to the police in Norway between 2005-2009.

Figure 1 charts how suspected crimes are processed in Norway's criminal justice system. The figure reports percentages for the period 2005-2009. If the police suspect an individual of a crime, they file a formal report. A public prosecutor then decides whether the individual should be charged with a crime as well as whether the case should proceed to a court trial. As reported in the figure, about half of police reports lead to a formal criminal charge. Of these charged cases, the public prosecutor advances 43% of them to a trial. The other charged cases are either dismissed, directly assigned a fine, or sent to mediation by the public prosecutor. Around 60% of the cases that proceed to trial are non-confession cases. Once a case proceeds to trial, it is assigned to a judge. If the judge finds the accused guilty, he or she can assign a combination of possible punishments which are not necessarily mutually exclusive. In the figure, we show percentages based on the strictest penalty received, so that the percentages add up to 100%. Just over half of cases result in incarceration, with probation, community service and fines combined accounting for 44% of outcomes. In a small fraction of cases (5%), the defendant is found not guilty.

2.2 Assignment of Cases to Judges

In Norway, the law dictates that cases be assigned to judges according to the "principle of randomization" (Bohn, 2000; NOU, 2002). The goal is to treat all cases ex-ante equally and prevent outsiders from influencing the process of the criminal justice system. In practice, cases are assigned by the chief judge to other judges on a mechanical, rotating basis based on the date a case is received. Each time a new cases arrives, it is assigned to the next judge on the list, with judges rotating between criminal and civil cases.[16]

There are some special instances where the assignment of cases does not follow the principle of randomization. These include cases involving juvenile offenders, extremely serious cases which require two professional judges, and complex cases expected to take a longer time to process, all of which can be assigned to more experienced judges. The Norwegian Department of Justice provides guidelines on the types of cases that can be non-randomly assigned and the Norwegian Courts Administration has flagged such cases in our dataset. While all other cases are randomly assigned, some case types can only be assigned to regular judges, and deputy judges are assigned relatively more confession cases. This means that randomization occurs within judge type, but not necessarily across judge types. Therefore, to have a sample of *randomly assigned cases* to the *same pool of judges* we: (i) exclude the special cases described above and (ii) focus on regular judges handling non-confession cases. Further details on how we construct the estimation sample are found in the data section.

A key to our design is that not only are judges randomly assigned, but they also differ in terms of their propensity to incarcerate defendants. In our baseline specification, we measure the strictness of a judge based on their incarceration rate for other cases they have handled, including both past and future confession and non-confession cases, and not just those cases which appear in our estimation sample. Our estimation sample has 500 judges, each of whom have presided over an average of 258 randomly assigned court cases. To construct our judge stringency measure, we calculate the leave-out mean judge incarceration rate and regress it on fully interacted court and year fixed effects to account for the fact that randomization occurs within the pool of available judges. This controls for any differences over time or across judicial districts in the types of criminals or the strictness of judges. The residual from this regression is our baseline measure of judge strictness. In a number of specification checks, we show robustness of the results to how we measure judge strictness (see Section 5.3).

Table 1 verifies that judges in our baseline sample are randomly assigned to cases. The first column regresses incarceration on a variety of variables measured before the court decision. It reveals that demographic, type of crime, and past work and criminal history variables are highly predictive of whether a defendant will be incarcerated, with most being individually significant. In column 3, we examine whether our measure of judge stringency can be predicted by this same set of characteristics. This is the same type of test that would

[16]Baard Marstrand at the Norwegian Courts Administration verified that district courts are required to randomly assign cases to judges, except in a few instances which we discuss in the text. We also checked with both the Bergen District Court (the second largest court, behind Oslo) and the Nedre Telemark District Court (a medium-sized court) that they follow the principle of randomization.

be done to verify random assignment in a randomized controlled trial. There is no statistically significant relationship between the judge stringency variable and the various demographic, crime type and labor market variables. The estimates are all close to zero, with none of them being statistically significant at the 5% level. The variables are not jointly significant either (p-value=.917). This provides strong evidence that criminal court cases are randomly assigned to judges in our sample, conditional on fully interacted court and year fixed effects.

Table 1. Testing for Random Assignment of Criminal Cases to Judges.

	***Dependent Variables*:**				***Explanatory Variables*:**	
	Pr(Incarcerated)		**Judge Stringency**			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	Coefficient Estimate	Standard Error	Coefficient Estimate	Standard Error	Mean	Standard Deviation
Demographics and Type of Crime:						
Age	0.0049***	(0.0004)	-0.0000	(0.0000)	32.64	(11.35)
Female	-0.0651***	(0.0074)	-0.0011	(0.0007)	0.106	(0.308)
Foreign born	0.0084	(0.0064)	0.0007	(0.0007)	0.135	(0.342)
Married, year t-1	-0.0442***	(0.0119)	-0.0018	(0.0012)	0.111	(0.314)
Number of children, year t-1	-0.0029	(0.0033)	0.0002	(0.0004)	0.783	(1.244)
High school degree, year t-1	-0.0675***	(0.0134)	-0.0014	(0.0015)	0.172	(0.377)
Some college, year t-1	0.0060	(0.0084)	0.0003	(0.0009)	0.046	(0.209)
Violent crime	0.0795***	(0.0087)	0.0014	(0.0011)	0.256	(0.437)
Property crime	-0.0054	(0.0115)	0.0012	(0.0012)	0.139	(0.346)
Economic crime	-0.0477***	(0.0117)	0.0018	(0.0015)	0.113	(0.316)
Drug related	-0.0385***	(0.0115)	0.0000	(0.0013)	0.119	(0.324)
Drunk driving	0.0667***	(0.0132)	0.0001	(0.0014)	0.071	(0.257)
Other traffic	-0.0457***	(0.0127)	0.0003	(0.0012)	0.087	(0.281)
Missing Xs	-0.4905***	(0.1415)	-0.0102	(0.0152)	0.030	(0.170)
Past Work and Criminal History:						
Employed, year t-1	0.0104	(0.0083)	0.0001	(0.0008)	0.352	(0.478)
Ever Employed, years t-2 to t-5	-0.0055	(0.0085)	0.0001	(0.0009)	0.470	(0.499)
Charged, year t-1	0.0932***	(0.0074)	0.0004	(0.0008)	0.459	(0.498)
Ever Charged, years t-2 to t-5	0.0925***	(0.0078)	-0.0006	(0.0009)	0.627	(0.483)
F-statistic for joint test	60.61		.577			
[p-value]	[.000]		[.917]			
Number of cases	33,509				33,509	

Note: Baseline sample consisting of 33,509 non-confession criminal cases processed 2005-2009. All estimations include controls for court x court entry year FEs. Reported F-statistic refers to a joint test of the null hypothesis for all variables. The omitted category for education is "Less than high school, year t-1" and the omitted category for type of crime is "Other crimes". Standard errors are two-way clustered at judge and defendant level. **p<0.1, **p<0.05, ***p<0.01.

It is natural to ask why some judges are more likely to incarcerate than others. While we do not observe personal characteristics of judges in our data for privacy reasons, we can measure how many cases they have handled. Using an OLS regression with the same controls as in Table 1, we find no relationship between the number of cases handled and judge stringency in our baseline sample. While there may be a variety of other reasons a judge is more or less likely to incarcerate, it is important to keep in mind that as long as judges are randomly assigned, the underlying reasons should not matter for our analysis.

2.3 Baseline IV Model

Our goal is to estimate the causal effects of incarceration on subsequent criminal behavior and labor market outcomes. As we document below, some judges are systematically more stringent than others. Combined with the conditional random assignment of cases to judges, this gives rise to exogenous variation in the probability a defendant is incarcerated.

Our estimation sample consists of defendants in non-confession, randomly-assigned cases, and we normalize the data so that period zero is the time of the court decision. Our baseline empirical model can be described by the following two-equation system:

$$I_{i,0} = \gamma Z_{j(i)} + X_i'\delta + v_{i,0} \tag{1}$$

$$Y_{i,t} = \beta_t I_{i,0} + X_i'\theta_t + \eta_{i,t} \tag{2}$$

where $I_{i,0}$ is an indicator variable equal to 1 if defendant i is sentenced to prison in period zero, $Z_{j(i)}$ denotes the stringency measure for judge j assigned to defendant i's case, X_i is a vector of control variables that always includes a full set of interacted year-of-case registration by court dummy variables, and Y_{it} is the dependent variable of interest measured at some point t after individual i's court decision (e.g., cumulative criminal charges five years after the court decision). The goal is to estimate the average β_t among defendants who were sentenced to prison because they were assigned to a strict judge. To estimate this local average treatment effect (LATE), we use 2SLS with first and second stage equations given by (1) and (2). Additionally, we present estimates of the reduced form effect of judge stringency on defendant's outcomes by directly regressing Y on Z and X. To compare our results to existing research we also carry out OLS estimation of equation (2), with and without controls for pre-determined covariates and lagged dependent variables.

3 Data and Background

3.1 Data and Sample Selection

Our analysis employs several data sources that we can link through unique identifiers for each individual. Information on the court cases comes from the Norwegian Courts Administration.

The dataset contains information for all court cases over the period 2005-2014. We observe the start and end dates of every trial, various case characteristics, the verdict, and unique identifiers for both judges, defendants, and district courts. We link this information with administrative data that contain complete records for all criminal charges, including the type of crime, when it took place, and suspected offenders. This data can be additionally linked to the prison register with information on actual time spent in prison. We merge these data sets with administrative registers provided by Statistics Norway, using a rich longitudinal database that covers every resident from 1967 to 2014. For each year, it contains individual demographic information (including sex, age, and number of children), socioeconomic data (such as years of education, earnings, employment), as well as geographical and firm identifiers.

To construct our baseline sample, we exclude the non-randomly assigned cases described in Section 2.2 and focus on regular judges handling non-confession cases.[17] This yields a sample of randomly assigned cases to the same pool of judges. Excluding the non-randomly assigned cases is straightforward, as these cases are flagged in our dataset. Our baseline sample further restricts the dataset to judges who handle at least 50 randomly assigned confession or non-confession cases between the years 2005 and 2014 (i.e., at least 50 of the cases used to construct our judge stringency instrument). Since we will be including court by year of case registration fixed effects in all our estimates, we also limit the dataset to courts which have at least two regular judges in a given year. Our main estimation sample uses cases decided between 2005 and 2009 so that each defendant can be followed for up to five years after decision, while the judge stringency instrument is based on the entire period from 2005 to 2014. Appendix Table A1 shows how the various restrictions affect the number of cases, defendants, judges and courts in our sample. After applying our restrictions, the baseline estimation sample includes 33,509 cases, 23,345 unique defendants, and 500 judges.

3.2 Descriptive Statistics

We now provide some summary statistics for defendants, crime types and judges. Panel A in Table 2 shows that defendants are relatively likely to be young single men. They also tend to have little education, low earnings and high unemployment prior to the charge, with under 40% of defendants working in the prior year. Serial offenders are common, with 38% of defendants having been charged for a different crime in the previous year. Panel B reports the

[17] The reason we focus on the non-cofession cases is that our research design requires that judges differ systematically in the sentencing decisions on non-confession cases. By way of comparison, judges are fairly similar in their incarceration rates for confession cases (where the accused has admitted guilt to the police/prosecutor before his case is assigned to a judge). Replicating the IV specification of column (3) of Table 5, but using confession cases instead of non-confession cases, we estimate an incarceration effect of -.327 (s.e. = 0.315). While the magnitude of the coefficient is similar to what we will find for non-confession cases, the standard error is over three times larger.

fraction of cases by primary crime category. Around one fourth of cases involve violent crime, while property, economic, and drug crime each comprise a little more than 10 percent of crimes. Drunk driving, other traffic offenses and miscellaneous crime make up the remainder.

Table 2. Descriptive Statistics.

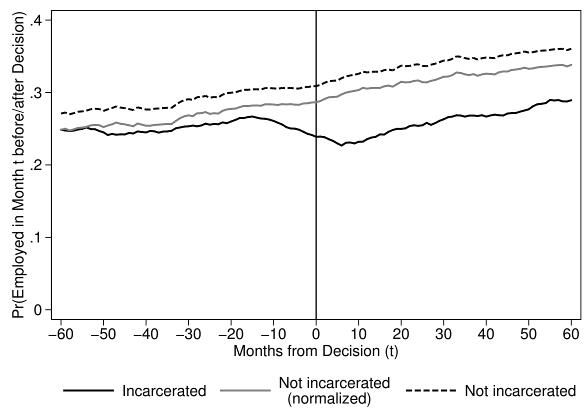
	(1)	(2)
	Mean	**Standard Deviation**
A. Defendant Characteristics:		
Demographics:		
Age	33.07	(11.78)
Female	0.119	(0.324)
Foreign born	0.148	(0.355)
Married, year t-1	0.128	(0.334)
Number of children, year t-1	0.822	(1.284)
Some college, year t-1	0.056	(0.229)
High school degree, year t-1	0.186	(0.389)
Less than high school, year t-1	0.758	(0.428)
Missing Xs	0.034	(0.181)
Past Work and Criminal History:		
Employed, year t-1	0.393	(0.488)
Ever Employed, years t-2 to t-5	0.505	(0.499)
Charged, year t-1	0.378	(0.485)
Ever Charged, years t-2 to t-5	0.572	(0.495)
Number of defendants	23,345	
B. Type of Crime:		
Violent crime	0.256	(0.437)
Property crime	0.139	(0.346)
Economic crime	0.113	(0.316)
Drug related	0.119	(0.324)
Drunk driving	0.071	(0.257)
Other traffic	0.087	(0.281)
Other crimes	0.215	(0.419)
Number of cases	33,509	

Note: Baseline sample consisting of 33,509 non-confession criminal cases processed 2005-2009 with 23,345 unique defendants.

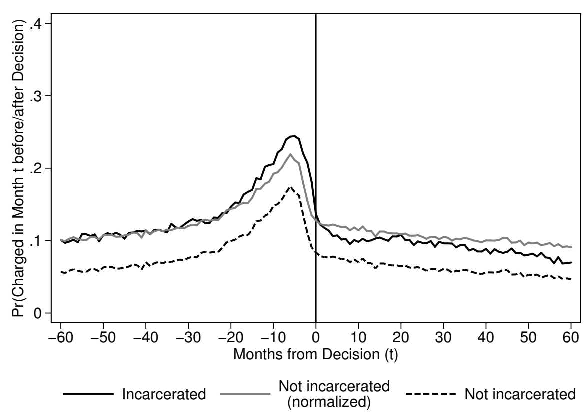
In Figure 2, we document the typical employment and crime levels for our sample over time. Panel (a) plots the probability a defendant has any paid employment in a given month during the 10 year period surrounding their court decision (normalized to period 0).[18] There are separate lines for defendants who are sentenced to incarceration versus not sentenced to incarceration. The first fact which emerges is that prior to the court decision, labor

[18]Using hours per month, full-time employment or wages instead of any employment as the outcome all reveal a similar pattern.

market participation is low for both groups, with less than 30% of defendants working in any month. Employment rates for the incarcerated group are a few precentage points lower; to ease comparison of changes over time, the graph also adjusts the non-incarcerated group's employment line to be the same as the incarcerated group's at the beginning of the sample period. Both groups have monthly employment rates which increase over time, reflecting the fact that employment rises as individuals get older.



(a) Pr(Employed in Month t)



(b) Pr(Charged in Month t)

Figure 2. Employment and Criminal Charges before and after Month of Court Decision.

Note: Baseline sample consisting of 33,509 non-confession criminal cases processed 2005-2009. Defendants are categorized in two groups, either incarcerated as shown in the solid black line or not incarcerated as shown in the dashed black line. To ease the comparison of trends, in each panel we normalize the level of the not incarcerated group's outcomes to the level of the incarcerated group's outcome in month t=-60. Outcomes for this "normalized" not incarcerated group are shown by the gray solid line. In both panels, the x-axis denotes months since court decision (normalized to period 0).

The most striking pattern in the graph is the divergence in employment between the incarcerated and non-incarcerated defendants around the time of the court decision. The positively sloped pre-trends for both groups are fairly similar up until about one year before the court decision date. However, around 12 months prior to the decision, the incarcerated line trends sharply downwards. This could be the result of incarcerated individuals being more likely to lose their jobs and turn to crime prior to the court's decision, or alternatively, incarcerated individuals being more likely to commit crime and lose their jobs as a result. Either way, the divergent trends prior to treatment suggest the two groups are not comparable. The downward trend continues until about 6 months after the decision, at which point it resumes its upward trend. Comparing the two lines reveals a sizable and stubbornly persistent drop in employment for the incarcerated group relative to the non-incarcerated.[19]

In panel (b) of Figure 2, we plot the probability an individual is charged with at least one crime in a month over time.[20] The figure reveals that both types of defendants have a high propensity to commit a crime. Five years before the court decision, defendants who will be incarcerated have a 10 percent chance of committing a crime in a month, compared to 7 percent for those who will not be incarcerated. Examining the pre-trends, there is a large jump around the court decision for both groups, since in order to have a court decision an individual must first be charged with a crime. While the two groups have similar trends for much of the pre period, they begin to diverge a little more than a year before the court decision, with the incarcerated group exceeding the non-incarcerated group by around 10 percent. That is, the incarcerated defendants get into more trouble with the police in the months leading up to their court decision. After the court decision, the probability of being charged with a crime returns to around 10 percent for both groups.[21]

In addition to describing our data, the graphs presented in Figure 2 highlight the hazards of using OLS or difference-in-differences to estimate the effects of incarceration. The incarcerated and non-incarcerated groups are not comparable in their pre-incarceration levels. Moreover, the trends in employment and criminal activity diverge before the court decision in ways that indicate there is an "Ashenfelter dip" prior to incarceration. These patterns motivate our quasi-experimental approach using the random assignment of judges.[22]

[19]There are several reasons why employment does not drop to zero after the court decision for those sentenced to prison. First, the average waiting time after a court decision before being sent to prison is around 4 months, and many prison stays are short. Second, the receipt of employment-related payments while in prison, such as vacation pay, shows up as working for pay in our dataset. Third, a small number of individuals are allowed to work outside of prison while incarcerated.

[20]Crime is serially correlated, so a single individual is often responsible for more than one charge in the figure. Later on, we will explore cumulative crime probabilities and the total number of crimes.

[21]There are several reasons why both types of defendants can be charged with crimes in the months immediately following a court decision. First, we measure when an individual was charged, not when the crime was committed. Second, individuals can commit additional crimes after their court decision before they have been imprisoned (4 month waiting time on average), as well as additional crimes while in prison.

[22]While one could omit the 12 months on either side of treatment in an attempt to avoid the Ashenfelter dip, this would assume the pre-treatment changes are caused by transitory shocks rather than a trend break (see the discussion in Ashenfelter and Card, 1985).

3.3 What Does it Mean to Be Incarcerated in Norway?

To help interpret our results, we briefly describe prison conditions in Norway.[23] Prisons emphasize rehabilitation and follow the "principle of normality" set forth by the Directorate of Norwegian Correctional Services. The principle dictates that "life inside will resemble life outside as much as possible" and that "offenders shall be placed in the lowest possible security regime." This means that low-level offenders go directly to open prisons, which have minimal security, as well as more freedoms and responsibilities. Physically, these open prisons resemble dormitories rather than rows of cells with bars on the door. More serious offenders who are at risk of violent or disruptive behaviors are sent to closed prisons, which have heightened security. The two types of prisons create a separation between minor and more hardened criminals, at least until the more hardened criminals have demonstrated good behavior.[24] While more serious offenders serve the majority of their sentence in closed prisons, they are usually transferred to open prisons for resocialization and further rehabilitation before release. Overall, one third of prison beds are in open prisons and the rest are in closed prisons.[25]

In Norway, there are a total of 61 prisons. The largest prison (in Oslo) has 392 cells, while the smallest has 13. Norway has a strict policy of one prisoner per cell and tries to place prisoners close to home so that they can maintain links with the families. This means that there is often a waiting list for (non-violent) individuals before they can serve their prison time. Sentenced individuals are released after their trial and receive a letter informing them when a cell opens up; in our data we calculate a wait time of 4 months on average.

To help with rehabilitation, all prisons offer education, mental health and training programs. In 2014, 38% and 33% of inmates in open and closed prisons, respectively, participated in some type of educational or training program. The most common programs are for high school and work-related training although inmates can also take miscellaneous courses. All inmates are involved in some type of regular daily activity, unless they have a serious mental or physical disability. If they are not enrolled in an educational or training program, they must work within prison.[26]

All inmates have the right to daily physical exercise and access to a library and newspapers. By law, all prisoners have the same rights to health and care services as the rest of the

[23] For a detailed discussion, see kriminalomsorgen.no.

[24] This separation could be important, as Bayer et al. (2009) find that inmates build "criminal capital" through interactions with other criminals.

[25] We do not observe whether an individual is assigned to an open or closed prison in our dataset.

[26] All prisoners, whether working or participating in training or education programs, receive a small stipend while in prison (around $8 per day in 2015). This stipend is not included in any of the earnings measures we use in the paper.

population. The Norwegian Directorate of Health is responsible for managing health programs for inmates. Most notably, eighteen percent of inmates participate in a drug-related program while in prison. After release, there is also an emphasis on helping offenders reintegrate into society, with access to active labor market and other programs set up to help ex-convicts find a job and access social services like housing support.[27]

3.4 Comparison to Other Countries

There are both similarities and differences in the criminal population and the criminal justice system of Norway versus the rest of the world. Along most dimensions, Norway looks broadly similar to many other Western European nations. And while it shares some commonalities with the U.S., the U.S. is an international outlier in some respects.

Incarceration rates. Appendix Figure A1 graphs Norway's incarceration rate over time. In 1980, there were an estimated 44 incarcerated individuals per 100,000 in Norway. This rate has increased gradually over time, with a rate of 72 per 100,000 in 2012. This 64% increase is not merely due to more crime being committed over time, as there has been a more modest 25% increase in crime over the same period (Lappi-Seppälä, 2012). Norway's gradual increase is mirrored in other Western European countries as well, although Norway's rate is slightly lower. In comparison, the U.S. incarceration rate has shot up dramatically, so much so that a separate scale is needed in the figure for the U.S. Not only did the U.S. start at a higher rate of 220 in 1980, but this rate reached over 700 by 2012.[28]

Comparing Norway and the U.S. to a broader set of countries, the U.S. remains an outlier, especially given how wealthy it is. This can be seen in Appendix Figure A2, which plots incarceration rates versus GDP for 160 countries with a population of greater than half a million. No other country comes close to the U.S. rate of just over 700 per 100,000, and only the six countries of Rwanda, El Salvador, Turkmenistan, Thailand, Cuba and Russia have over 400 per 100,000. In contrast, the figure shows that Norway's incarceration rate is similar to the average for other Western European countries (102 per 100,000). The U.S. is particularly an outlier after controlling for GDP per capita; while several countries have high GDP per capita (purchasing power adjusted), the U.S. incarceration rate is several multiples higher.[29]

[27] It is important to realize that the initial judge assigned to a case does not determine which prison a defendant is sent to, the type of training, educational, or work program a defendant participates in, or when a defendant is eligible for parole.

[28] Neal and Rick (2016) show that most of the growth in incarceration rates in the U.S. can be explained by changes in sentencing policy as opposed to higher crime and arrest rates.

[29] It is more difficult to compare measures of criminal activity across countries due to differences in reporting.

Inmate characteristics. Along many dimensions, the prison populations in Norway, Western Europe and the U.S. are similar.[30] In the U.S., Norway, and many of the European countries for which data is available, roughly three fourths of inmates have not completed the equivalent of high school. Five percent of prisoners in Norway are female compared to 5% in Western Europe and 7% in the U.S. In all of these countries, inmates are in their early or mid-thirties on average.

The types of offenses committed by inmates differs across countries, but perhaps less than one might expect. In terms of the fraction of prisoners who have committed a drug offense, the rates are surprisingly similar, with 24% in Norway, 22% in Western Europe and 20% in the U.S. By comparison, fourteen percent are serving a sentence for assault/battery and 4% for rape/sexual assault in Norway, respectively, compared to 11% and 7% in Western Europe and 9% and 11% in the U.S. Of course, these comparisons needed to be understood in the context of a much higher incarceration rate in the U.S. But they point to a considerable overlap in the types of crimes committed by inmates across countries.[31]

Prison expenditures, sentence lengths, and post-release support. One difference across countries is the amount of money spent on prisoners. Western European countries spend an average of $66 thousand per inmate per year, which is roughly double the average of $31 thousand for the U.S. But these averages mask substantial heterogeneity, in part due to differences in wages and labor costs.[32] For example, in Norway the cost is $118 thousand (about the same as Sweden, Denmark, and the Netherlands), in Italy $61 thousand, and in Portugal $19 thousand. In the U.S., the state of New York spends $60 thousand per prisoner, Iowa $33 thousand, and Alabama $17 thousand. And in New York City, the annual cost per inmate reaches $167 thousand.[33]

With this caveat in mind, the U.S. has more than double the number of reported assaults than either Norway or the rest of Western Europe according to the United Nations Survey on Crime Trends (787 per 100,000 in the U.S. compared to 346 in Norway and 377 in Western Europe for the year 2006; see Harrendorf et. al, 2010). But the large incarceration gap between the U.S. and other countries cannot be fully explained by differences in how many crimes are committed. Instead, at least part of the difference is due to longer mandatory sentencing policies for minor crimes (see Raphael and Stoll, 2013).

[30] For details on the U.S. criminal population, see Bureau of Justice Statistics (2015) and Raphael and Stoll (2013). For Scandinavia and other European countries, see Kristoffersen (2014) and Aebi et al. (2015).

[31] These numbers for Norway differ from our estimation sample for two reasons: we do not have illegal immigrants in our dataset, and our sample is restricted to non-confession cases which are randomly assigned. The numbers for the U.S. are the weighted average of inmates in federal and state prisons.

[32] In most countries, a majority of prison costs are due to labor expenditures; for example, in Norway, two-thirds of the prison budget is spent on labor.

[33] Cost estimates are calculated by dividing total prison budgets by the number of prisoners. The numbers for Western Europe (with Belgium and Switzerland excluded due to lack of data) are for the year 2013 and are purchasing power parity adjusted (Abei et al, 2015). The data for 40 U.S. states with available data are for 2010 (Vera Institute of Justice, 2012). New York City data are for 2012 (NYC Independent Budget Office, 2013; http://www.nytimes.com/2013/08/24/nyregion/citys-annual-cost-per-inmate-is-nearly-168000-study-

Norway is able to maintain the type of prison conditions summarized in Section 3.3 in part due to its larger prison budget. In particular, more resources can be devoted to education and training programs and overcrowding is not an issue. In contrast, while most state prison systems in the U.S. aim to provide GED test preparation, adult basic education and vocational skills training, a recent RAND report (2014) finds that funding for such initiatives is scarce. The U.S. also faces serious overcrowding issues, with Federal prisons being 39% over capacity (GAO, 2012) and over half of states at or above their operational capacity (Bureau of Justice Statistics, 2014).

Another difference between Norway (and Western Europe) versus the U.S. is sentence length. The average time spent in prison using our judge stringency instrument is estimated to be 175 days, or 5.6 months, for our Norwegian sample. Almost 90% of spells are less than 1 year. This is considerably shorter compared to the average prison time of 2.9 years for the U.S. (Pew Center, 2011), and fairly similar to the median of 6.8 months in other Western European countries (Aebi et al., 2015). Because of this disparity in sentence lengths, the average cost per prisoner spell in Norway and Europe is smaller compared to the U.S., even though the cost per prisoner per year is generally higher.

Norway has been a leader in reforming its penal system to help integrate inmates back into society upon release. While offenders in Norway may lose their job when going to prison, they are usually not asked or required to disclose their criminal record on most job applications.[34] This stands in contrast to the U.S. Moreover, while gaps will still appear on employment resumes, these will often span months rather than years due to shorter prison spells. Upon release all inmates have access to support from the Norwegian work and welfare services. This includes work training programs and help searching for a job, as well as access to a variety of social support programs such as unemployment benefits, disability insurance and social assistance.

4 Assessing the Instrument

4.1 Instrument Relevance

Figure 3 shows the identifying variation in our data, providing a graphical representation of the first stage of the IV model. In the background of this figure is a histogram that shows the distribution of our instrument (controlling for fully interacted year and court dummies). Our instrument is the average judge incarceration rate in other cases a judge has handled, including the judge's past and future cases that may fall outside of our estimation sample.

says.html?_r=0).

[34]Exceptions include occupations such as child care workers or police officers.

The mean of the instrument is 0.45 with a standard deviation of 0.08.The histogram reveals a wide spread in a judge's tendency to incarcerate. For example, a judge at the 90th percentile incarcerates about 54% of cases as compared to approximately 34% for a judge at the 10th percentile.

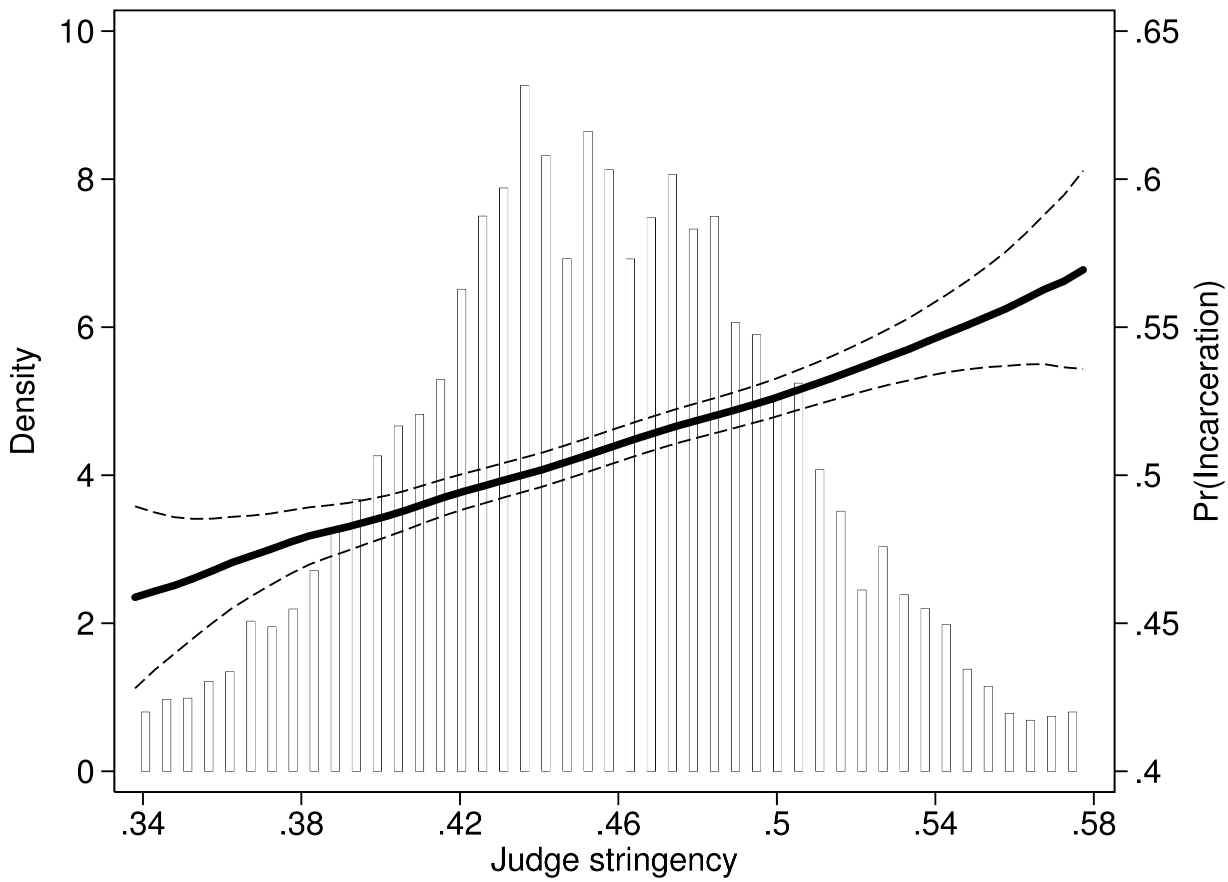


Figure 3. First Stage Graph of Incarceration on Judge Stringency.

Note: Baseline sample consisting of 33,509 non-confession criminal cases processed 2005-2009. Probability of incarceration is plotted on the right y-axis against leave-out mean judge stringency of the assigned judge shown along the x-axis. The plotted values are mean-standardized residuals from regressions on court x court entry year interacted fixed effects and all variables listed in Table 1. The solid line shows a local linear regression of incarceration on judge stringency. Dashed lines show 90% confidence intervals. The histogram shows the density of judge stringency along the left y-axis (top and bottom 2% excluded).

Table 3. First Stage Estimates of Incarceration on Judge Stringency.

	(1)	(2)	(3)	(4)	(5)	(6)
***Estimation Sample*:**	Time of Decision	Month 12 after Decision	Month 24 after Decision	Month 36 after Decision	Month 48 after Decision	Month 60 after Decision
***Dependent Variable*:**	**Pr(Incarcerated)**					
A. Court × Year of Court Case Registration Interacted Fixed Effects						
Judge Stringency	0.4921***	0.4949***	0.4895***	0.4845***	0.4834***	0.4742***
	(0.0666)	(0.0666)	(0.0666)	(0.0660)	(0.0666)	(0.0672)
F-stat. (Instrument)	53.81	55.06	53.59	53.15	51.92	49.14
B. Add Controls for Demographics and Type of Crime						
Judge Stringency	0.4816***	0.4839***	0.4763***	0.4716***	0.4715***	0.4627***
	(0.0667)	(0.0663)	(0.0664)	(0.0660)	(0.0666)	(0.0672)
F-stat. (Instrument)	51.41	52.49	50.77	50.38	49.40	46.69
C. Add Controls for Demographics, Type of Crime, Past Work and Criminal History						
Judge Stringency	0.4828***	0.4855***	0.4777***	0.4730***	0.4741***	0.4648***
	(0.0657)	(0.0654)	(0.0654)	(0.0651)	(0.0659)	(0.0664)
F-stat. (Instrument)	53.30	54.31	52.57	52.03	51.08	48.43
Dependent mean	0.5082	0.5075	0.5062	0.5051	0.5044	0.5043
Number of cases	33,509	33,231	32,723	32,257	31,767	31,287

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=33,509 at time of decision and N=31,287 in month 60 after decision). Standard errors are two-way clustered at judge and defendant level. **p<0.1, **p<0.05, ***p<0.01.

Figure 3 also plots the probability a defendant is sent to prison in the current case as a function of whether he is assigned to a strict or lenient judge. The graph is a flexible analog to the first stage in equation (1), plotting estimates from a local linear regression. The likelihood of receiving a prison sentence is monotonically increasing in the judge stringency instrument, and is close to linear. Table 3 reports first stage estimates where we regress a dummy for whether a defendant is incarcerated in the current case on our judge stringency instrument. In panel A, we include fully interacted court and year dummies but otherwise no other controls. The first column reports the first stage estimate at the time of the court decision, whereas the other columns report first stages estimates in each of the five subsequent years. These columns are identical except for the very modest impact of sample attrition (less than seven percent over five years) stemming from death or emigration of defendants.[35] The point estimate of nearly 0.5 barely moves across columns, indicating that attrition exerts a negligible impact on the first stage relationship. The estimates are highly significant, suggesting that being assigned to a judge with a 10 percentage point higher overall incarceration rate increases the probability of receiving a prison sentence by roughly 5 percentage points. [36]

4.2 Instrument Validity

Conditional Independence. For our instrument to be valid, the stringency of a judge must be uncorrelated with both defendant and case characteristics that could affect a defendant's future outcomes (controlling for fully interacted court and year dummies). As discussed in Section 2.2, Table 1 provides strong empirical support for the claim that the criminal justice system in Norway randomly assigns cases to judges within each court in a given time period.

As a second test, panels B and C of Table 3 explore what happens if a large set of control variables are added to the first stage regressions. If judges are randomly assigned, per-determined variables should not significantly change the estimates, as they should be uncorrelated with the instrument. As expected, the coefficient does not change appreciably when demographic and crime type controls are added in panel B. As shown in panel C, this coefficient stability continues to hold when we additionally condition on lagged dependent variables capturing a defendant's prior work and criminal history.

Exclusion. Conditional random assignment of cases to judges is sufficient for a causal interpretation of the reduced form impact of being assigned to a stricter judge. However, interpreting the IV estimates as measuring the causal effect of incarceration requires an exclusion restriction: the incarceration rate of the judge should affect the defendant's outcomes only through the incarceration sentencing channel, and not directly in any other way. The key challenge here is that trial decisions are multidimensional, with the judge deciding on incarceration, fines, community service, probation, and guilt. After discussing our main results, we will present empirical evidence that the exclusion restriction holds (see Section 5.4). In particular, we will show that our estimates do not change appreciably when we augment our baseline model to either control for judge stringency in other dimensions or include and instrument for other trial sentencing decisions.

[35] Another test of selective attrition is to regress the probability of dropping out of the sample on the judge stringency instrument conditional on fully interacted court and year dummies. When performing this test, we find no evidence of a significant relationship between dropping out and the instrument (see Table A3).

[36] Note the first stage coefficient need not be one, unless the following conditions hold: (i) the sample of cases used to calculate the stringency measure is exactly the same as estimation sample, (ii) there are no covariates, and (iii) there are a large number of cases per judge. In our setting, there is no reason to expect a coefficient of one. In particular, the full set of court times year dummies breaks this mechanical relationship.

Monotonicity. If the causal effect of incarceration is constant across defendants, then the instrument only needs to satisfy the conditional independence and exclusion assumptions. With heterogeneous effects, however, monotonicity must also be assumed. In our setting, the monotonicity assumption requires that defendants who are incarcerated by a lenient judge would also be incarcerated by a stricter judge, and vice versa for non-incarceration. This assumption ensures that IV identifies the LATE of incarceration; that is, the average effect among the subgroup of defendants who could have received a different incarceration decision had their case been assigned to a different judge.

One testable implication of the monotonicity assumption is that the first stage estimates should be non-negative for any subsample. For this test, we continue to construct the judge stringency variable using the full sample of available cases, but estimate the first stage on the specified subsample. Results are reported in column (1) of Appendix Table A9. In panel A, we construct a composite index of all of the characteristics found in Table 1, namely predicted probability of incarceration, using the coefficients from an OLS regression of the probability of incarceration on these variables (while conditioning on fully interacted court and year dummies). We then estimate separate first stage estimates for the four quartiles of predicted incarceration. Panel B breaks the data into six crime types. Panels C and D split the data by previous labor market attachment and by whether the defendant has previously been incarcerated, respectively. Panels E, F and G split the samples by age, education, and number of children. For all of these subsamples, the first stage estimates are large, positive and statistically different from zero, consistent with the monotonicity assumption.

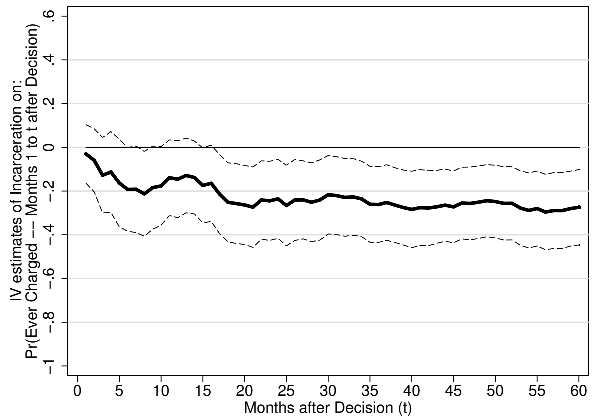
A second implication of monotonicity is that judges should be stricter for a specific case type (e.g., violent crimes) if they are stricter in other case types (e.g., all crimes except for violent crimes). To test this implication, we break the data into the same subsamples as we did for the first test, but redefine the instrument for each subsample to be the judge's incarceration rate for cases outside of the subsample. For example, for the violent crime subsample, we use a judge's incarceration rate constructed from all cases except violent crime cases. Column (2) of Appendix Table A9 lists the first stage estimates using this "reverse-sample instrument" which excludes own-type cases. The first stage estimates using this redefined instrument are all positive and statistically different from zero, suggesting that judges who are stricter for one type of case are also stricter for other case types, as monotonicity requires.

5 Effects of Incarceration on Recidivism

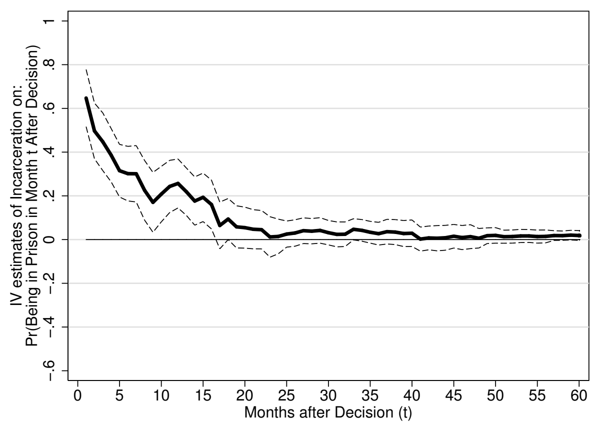
In this section, we present our main findings, showing that (i) incarceration causes a large reduction in the probability of reoffending, (ii) the drop is not due only to incapacitation, with further reductions in criminal charges after release, and (iii) the total number of charged crimes falls over time, with many individuals being diverted from a future life of crime. We then contrast these IV estimates to results from OLS, learning that the high rates of recidivism among ex-convicts is due to selection, and not a consequence of the experience of being in prison.

5.1 Main Results

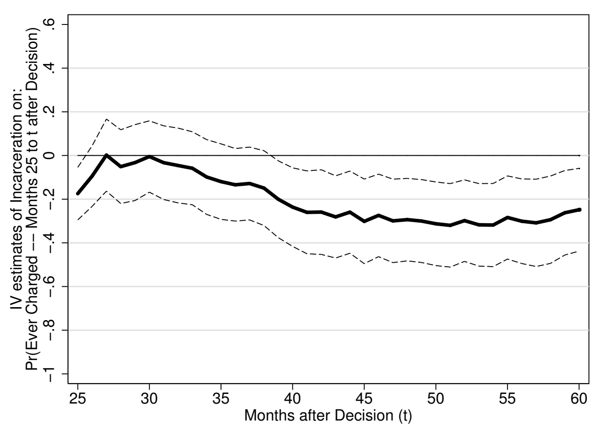
Reoffense probabilities. Panel (a) of Figure 4 graphically presents IV estimates of the effect of incarceration on the probability of reoffending. We define reoffending as the probability of being charged with at least one crime by the end of a given time period. The graph presents a series of cumulative monthly estimates from 1 month to 60 months after the court decision. For example, the estimate at month 6 uses the probability an individual has been charged with at least one crime by 6 months after the decision as the dependent variable in the second stage of the IV model. As expected, there is little effect on reoffending in the first few months after the court decision, since not much time has elapsed for the committing of new crimes. But the estimate becomes more negative over time, and at around 18 months there is a large and statistically significant reduction of over 25 percentage points in recidivism for those previously sentenced to incarceration. This negative effect persists at roughly the same level all the way to 60 months.



(a) IV Estimates: Pr(Ever Charged – Months 1 to t)



(b) IV Estimates: Pr(Being in Prison – Month t)



(c) IV Estimates: Pr(Ever Charged – Months 25 to t)

Figure 4. The Effect of Incarceration on Recidivism and Probability of Being in Prison.

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=33,509 at time of decision and N=31,287 in month 60 after decision). Panel (b) plots prison probabilities related only to the original sentence. Dashed lines show 90% confidence intervals.

Incapacitation versus post-release effects. The recidivism effect found in panel (a) of Figure 4 could simply be due to incapacitation, as individuals sentenced to prison time will be locked up and therefore have few criminal opportunities.[37] To better understand the role of incapacitation, Table 4 presents IV estimates of the effects of incarceration on prison time. We find that, on average, being incarcerated leads to a sentence of 238 days in prison. But this is sentencing time (i.e., potential prison time), not actual time served. Using the IV model, we estimate that being incarcerated leads to 175 days, or approximately 6 months, in actual prison time served. This smaller number makes sense, as Norway allows individuals to be released on parole after serving about two-thirds of their prison sentence for good behavior.[38] In column (2) of the Table 4, we also estimate the average wait time between the court decision and when individuals start serving their prison sentence. The average wait time is estimated to be a little under 4 months.

Table 4. The Effect of Incarceration on Prison Time.

	(1)	(2)	(3)
***Dependent Variable*:**	**Days of Prison Sentence (Potential Prison Time)**	**Days Spent Outside Prison Before Serving Sentence (Waiting Time)**	**Days of Prison Sentence Served (Actual Prison Time)**
RF: Judge Stringency	110.59**	53.36**	81.52**
	(50.62)	(22.93)	(35.51)
2SLS: Incarcerated	237.95***	114.81**	175.40***
	(90.53)	(44.97)	(62.64)
Dependent mean	153.37	48.38	111.14
Number of cases		31,287	

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=31,287 in month 60 after decision). Controls include all variables listed in Table 1. Days spent in pre-trial custody are included in actual prison time in column (3). Standard errors are two-way clustered at judge and defendant level. *p<0.1, **p<0.05, ***p<0.01.

In panel (b) of Figure 4, we plot a series of IV estimates for the probability of being in prison, 1 to 60 months after the court decision. The figure is similar to a survival function, in that if all treated individuals (i.e., those sentenced to prison) started out in prison in month 1, the estimates would map out 1 minus the probability of exit from prison. It is not a survival function, because not all individuals sentenced to prison begin serving their sentences immediately (due to prison capacity constraints). As expected, the probability of being in prison if an individual is sentenced to prison starts out high. This probability falls

[37]It is possible for individuals to be charged with a crime while serving prison time, as they can commit crimes while in prison. They may also have other cases working their way through the system while they are in prison.

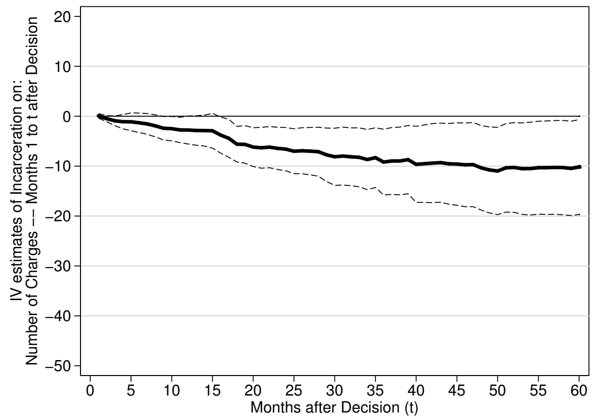
[38]The IV estimate suggests that a majority of individuals receive parole in our dataset. If an inmate commits a new offense while on parole, this counts as a new charge in our dataset.

rapidly, with fewer than 30 percent of incarcerated individuals being in prison for the original criminal charge six months after the court decision. By month 18, only around 5 percent of these individuals are still in prison, and month 24 very few are still in prison.

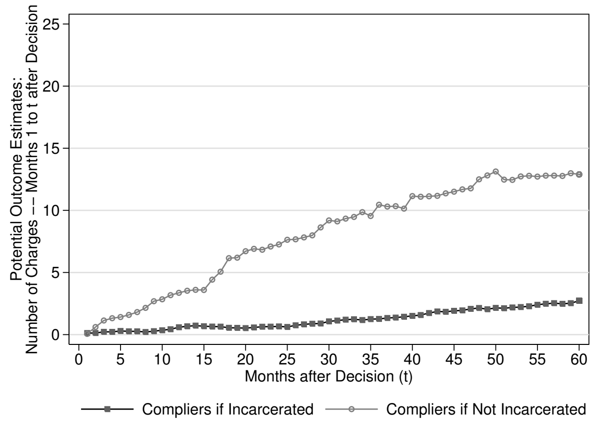
The main point to take away from panel (b) of Figure 4 is that any incapacitation effect from being incarcerated at time zero can only operate in the first two years. Using this insight, we now graph the probability of ever being charged with a crime between months 25 and 60 in panel (c) of Figure 4. By ignoring crimes committed within the first two years after the decision, we are estimating incarceration effects which cannot be attributed to the original incapacitation spell. As in panel (a) of Figure 4, it takes a few months for individuals to start being charged with a crime within this window. But within 15 months after the start of this new window (i.e., 40 months after the court decision), there is a strong and statistically significant reduction in crimes for those individuals who were previously sentenced to prison. The effect in this period is a sizable 25 percentage point reduction in reoffending at least once by month 60.

Number of crimes. A comparison of panels (a) and (c) in Figure 4 suggests that incarceration not only prevents an individual from ever committing a crime (the extensive margin), but it also prevents individuals from committing a series of future crimes (the intensive margin). In panel (a), after month 18, the probability of ever being charged with a crime is flat, suggesting that additional individuals are not being prevented from committing a crime after that time. But in panel (c), we see that the probability an individual will commit a crime between 24 and 60 months is affected by an incarceration decision at time zero. This means that many of the individuals who were prevented from committing a crime in panel (a) are also being prevented from committing another crime in panel (c).

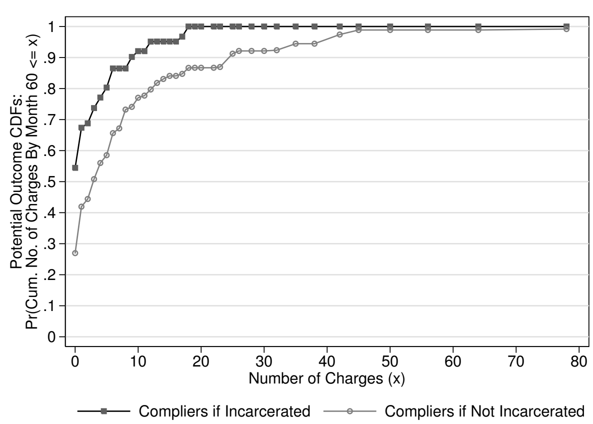
To further explore the intensive margin, panel (a) of Figure 5 plots IV estimates for the cumulative number of charges in the months after the court decision. The estimated effects generally become more negative over time. After one year, the estimated effect of an incarceration decision is around 3 fewer crimes per individual, whereas after two years, the effect is 7 fewer crimes. By four years, the effect is 10 fewer crimes per individual.



(a) IV Estimates: No. of Charges – Months 1 to t



(b) Potential Outcomes: No. of Charges in Months 1 to t



(c) Potential Outcome CDFs: No. of Charges by Month 60

Figure 5. The Effect of Incarceration on Number of Charges.

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=33,509 at time of decision and N=31,287 in month 60 after decision). Dashed lines show 90% confidence intervals.

Potential crimes. Our IV estimates represent the average causal effects for compliers who could have received a different court decision had their case been assigned to a different judge. To better understand this LATE, we follow Imbens and Rubin (1997) and Dahl et al. (2014) in decomposing the IV estimates into the average potential outcomes if the compliers would have been incarcerated and the average potential outcomes if they would not have been incarcerated. The top line in panel (b) of Figure 5 is the number of potential charges if the compliers would not have been incarcerated. The line trends upward in close to a linear fashion, with approximately 2 to 3 extra criminal charges per year and around 13 crimes on average after five years. In sharp contrast, the compliers would have been charged with far fewer crimes if incarcerated; even by month 60, they would only been charged with 3 crimes on average.

Panel (c) plots the distribution functions for cumulative potential charges as of year 5, for compliers if they would have been incarcerated and if they would not have been incarcerated. The difference between the two CDF's when the number of charges is one is around 30 percentage points, which mirrors the IV effect graphed in Figure 4, panel (a), at 5 years out. Comparing the CDF's further to the right (i.e., for a larger number of charges) makes clear that incarceration is not simply preventing low-crime individuals from committing future crime. To see this, suppose that incarceration caused individuals who would have been charged with 5 crimes or less (or some similarly small number of crimes) from being charged with any crimes, but that more hardened criminals (those charged with more than 5 crimes) were unaffected. In this case, the two lines in panel (c) would lie on top of each other starting at 5 charges. But, in fact, the two lines diverge at one charge, remain fairly parallel until around 18 charges, and do not get close to each other until around 45 charges. For instance, 12% of compliers would have been charged with more than 18 crimes if they were not incarcerated, whereas few, if any, compliers would have been charged with this many crimes if incarcerated. Taken together, the results suggest that incarceration must be preventing some individuals from being charged with a large number of crimes, and stopping some individuals from a life of crime entirely.[39]

5.2 Comparison to OLS

With few exceptions, the bulk of the research on recidivism is based on OLS regressions with controls for observable confounding factors. In Table 5, we present OLS estimates of equation (2) with and without a rich set of controls. The first OLS specification in Table 5 regresses

[39] From the graph, one cannot infer whether an individual charged with 35 crimes reduces their charges to 0 versus whether an individual charged with 35 crimes reduces their crime to 15 while the individual charged with 15 reduces their crime to 0. But the shapes of the CDF's do imply that high volume criminals must reduce their number of charged crimes.

whether an individual has reoffended (i.e., been charged with a new crime after the court decision) on whether the defendant was sentenced to prison, but includes no other control variables. The OLS estimates 0-2 years after the decision, 2-5 years after the decision, and 0-5 years after the decision are all positive and significant; for example, individuals sent to prison are 11 percentage points more likely to reoffend at least once over the next 5 years.

Table 5. The Effects of Incarceration on Recidivism.

Dependent Variable:	**Pr(Ever Charged)**			**Number of Charges**
	Months 1-24 after Decision	*Months 25-60 after Decision*	*Months 1-60 after Decision*	*Months 1-60 after Decision*
	(1)	**(2)**	**(3)**	**(4)**
OLS: Incarcerated	0.130***	0.113***	0.113***	5.092***
No controls	(0.007)	(0.007)	(0.006)	(0.315)
OLS: Incarcerated	0.126***	0.108***	0.103***	5.187***
Demographics & Type of Crime	(0.007)	(0.007)	(0.006)	(0.303)
OLS: Incarcerated	0.087***	0.068***	0.066***	4.128***
All controls	(0.006)	(0.007)	(0.006)	(0.291)
OLS: Incarcerated	0.083***	0.065***	0.066***	3.735***
Complier Re-weighted	(0.007)	(0.007)	(0.006)	(0.273)
RF: Judge Stringency	-0.103**	-0.115**	-0.127***	-4.729*
All controls	(0.049)	(0.049)	(0.045)	(2.519)
IV: Incarcerated	-0.223*	-0.248**	-0.274***	-10.176*
All controls	(0.115)	(0.115)	(0.104)	(5.759)
Dependent mean	0.57	0.56	0.70	9.91
Complier mean if not incarcerated	0.56	0.58	0.73	12.90
Number of cases		31,287		

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=31,287 in month 60 after decision). Controls include all variables listed in Table 1. RF and IV in addition also control for court x court entry year FEs. OLS standard errors are clustered at the defendant level, while RF and IV standard errors are two-way clustered at judge and defendant level. *p<0.1, **p<0.05, ***p<0.01.

In the next specification of Table 5, we add a host of defendant characteristics, including demographic variables and the type of crime they are being charged with. These controls affect the estimates only slightly. In the third specification, we additionally add lagged variables for whether defendants have previously been charged with a crime and whether they have worked in the prior year (i.e., including all of the variables listed in Table 2 as controls). This brings the coefficient down to a 6 percentage point effect.

IV estimates summarizing the results previously shown in the figures appear in the bottom half of Table 5. The divergence between the OLS estimates and the IV estimates is stark. The OLS estimates always remain positive, while the IV estimates are negative and large. One possible explanation is that the OLS estimates suffer from selection bias due to correlated unobservables. If this is the case, we can conclude that the high rates of recidivism among ex-convicts is due to selection, and not a consequence of the experience of being in prison.

Another possible explanation for the differences between the IV and OLS estimates is effect heterogeneity, so that the average causal effects for the compliers differ in sign compared to the mean impacts for the entire population. To explore this possibility, it is useful to characterize compliers by their observable characteristics. We begin by splitting our sample into eight mutually exclusive and collectively exhaustive subgroups based on prior labor market attachment and the predicted probability of incarceration (see Appendix Table A4). The predicted probability of incarceration is a composite index of all of the observable characteristics while prior employment is key source of heterogeneity in effects, as discussed in the next section. Next, we estimate the first stage equation (1) separately for each subsample, allowing us to calculate the proportion of compliers by subgroup. We then reweight the estimation sample so that the proportion of compliers in a given subgroup matches the share of the estimation sample for that subgroup. The fourth row of Table 5 presents OLS estimates based on this reweighted sample. The results suggest the differences between the IV and OLS estimates cannot be accounted for by heterogeneity in effects, at least due to observables.

5.3 Specification Checks

In Appendix Table A5, we present a variety of specification checks with respect to how the sample is selected and the instrument is constructed.[40] The first row presents our baseline results for comparison. In this specification, we include any defendant whose judge handled at least 50 cases. In the next two specifications, we instead require judges to handle at least 25 cases and at least 75 cases. These changes in the sample selection do not materially affect the estimated effects.

The next robustness check deals with individuals who appear more than once in our dataset because they are brought to trial for multiple crimes over time. Individuals appearing multiple times could be in the incarcerated group in one year and the non-incarcerated group in another year. While judges are randomly assigned for each court case, and hence the baseline estimate is still causal, the interpretation is more involved. In column (4), we limit the sample to the first court case appearing in our data, so that each individual appears only once in the estimation sample. This restriction reduces the sample of cases by a third,

[40]In Appendix Table A6, we perform the same specification checks for the subgroup that is driving the reduction in crime, namely, individuals who were not working prior to the incarceration decision.

and while the standard error rises as expected, the estimated coefficient remains remarkably similar compared to the baseline.

The final two specification checks examine how sensitive the results are to changing how the instrument is constructed. In column (5), we randomly split our sample in two. In the first subsample, the instrument is calculated using only the decisions of the judge in the second subsample, and vice versa. We then stack the two subsamples and estimate our IV model given by equations given by (1)-(2). This split-sample approach gets rid of any issue due to using leave out means to calculate the instrument. The resulting estimates (and standard errors) do not materially change. The last column shows that our findings are not sensitive to whether we calculate judge stringency based on non-confession cases only or if we include all randomly assigned cases (both confession and non-confession cases) in these calculations.

5.4 Threats to Exclusion Restriction

As discussed in Section 4.2, interpretating the IV estimates as the LATE of incarceration requires the judge stringency instrument to affect the defendant's outcomes only through the prison sentencing channel, and not directly in any other way. A potential issue is that trial decisions are multidimensional, with judges deciding on incarceration, fines, community service, probation and guilt (where the penalties are not mutually exclusive).

To make this issue precise, it is useful to extend the baseline IV model given by equations (1) and (2), distinguishing between the incarceration decision and other trial decisions:

$$I_{i,0}^{Incar} = \alpha Z_{j(i)}^{Incar} + \gamma Z_{j(i)}^{Other} + X_i^{'}\delta + v_{i,0} \tag{3}$$

$$I_{i,0}^{Other} = \zeta Z_{j(i)}^{Incar} + \lambda Z_{j(i)}^{Other} + X_i^{'}\psi + u_{i,0} \tag{4}$$

$$Y_{i,t} = \beta_t I_{i,0}^{Other} + \theta_t I_{i,0}^{Incar} + X_i^{'}\omega_t + \eta_{i,t} \tag{5}$$

where j denotes the judge that handles defendant i's case, $I_{i,0}^{Incar}$ is an indicator variable equal to 1 if defendant i is sentenced to prison in period zero, $I_{i,0}^{Other}$ is an indicator variable equal to 1 if defendant i is sentenced to fines, community service or probation, $Z_{j(i)}^{Incar}$ denotes the judge stringency instrument for the incarceration decision, $Z_{j(i)}^{Other}$ denotes the judge stringency instrument for trial decisions other than incarceration, and X_i is a vector of control variables that always includes a full set of year-of-case registration by court dummy variables. The omitted reference category is not guilty. As in the baseline model, we measure $Z_{j(i)}^{Incar}$and $Z_{j(i)}^{Other}$ as leave-out means for incarceration decision and trial decisions other than incarceration, respectively. Because cases are randomly assigned, both $Z_{j(i)}^{Incar}$and $Z_{j(i)}^{Other}$

satisfy the independence assumption (conditional on fully interacted court and year fixed effects).

There are two cases in which the baseline IV estimates based on (1)-(2) are biased because they abstract from trial decisions other than incarceration. The first case is if $Z_{j(i)}^{Incar}$correlates with $Z_{j(i)}^{Other}$, and $Z_{j(i)}^{Other}$ directly affects $Y_{i,t}$ (conditional on X_i). This would violate the exclusion restriction in the baseline IV model because $Z_{j(i)}^{Incar}$ not only affects $Y_{i,t}$ through $I_{i,0}^{Incar}$ but also through its correlation with $Z_{j(i)}^{Other}$. However, controlling for $Z_{j(i)}^{Other}$ in both (1) and (2) will eliminate this source of bias. The second case is if $Z_{j(i)}^{Incar}$ correlates with $I_{i,0}^{Other}$ conditional on $Z_{j(i)}^{Other}$, and $I_{i,0}^{Other}$ affects $Y_{i,t}$ holding $I_{i,0}^{Incar}$ fixed (conditional on X_i). In the baseline IV model, this would violate the exclusion restriction because $Z_{j(i)}^{Incar}$ not only affects $Y_{i,t}$ through $I_{i,0}^{Incar}$ but also through to its influence on $I_{i,0}^{Other}$. The augmented IV model given by (3)-(5) addresses this issue by including $I_{i,0}^{Other}$ as an additional endogeneous regressor and $Z_{j(i)}^{Other}$ as an extra instrument.

In Appendix Tables A10 and A11, we examine these two cases, finding support for the exclusion restriction. To start, we first calculate a judge's tendencies on trial decisions other than incarceration.[41] For example, we measure a judge's probation stringency as the average probation rate in the other cases a judge has handled. The top panel of Table A10 repeats our baseline specification for comparison. In panel B, we add a judge's probation stringency, community service stringency, and fine stringency as three additional controls in both the first and second stages. A decision of not guilty is the omitted category. The IV estimates for both recidivism outcomes are similar to our baseline, albeit with standard errors which are larger. To increase precision, panel C combines these three control variables into a single "probation, community service or fine" stringency variable. Again, the IV estimates for recidivism are similar to the baseline in panel A, but the standard errors are considerably smaller.

We next estimate the augmented IV model given by (3)-(5). Appendix Table A11 presents the first stage, reduced form and IV estimates. To make sure we have enough precision and avoid problems associated with weak instruments, we use a specification with three decision margins: "incarceration," "probation, community service or fine," and "not guilty." For the incarceration first stage, the judge stringency instrument for the incarceration decision has a similar coefficient as before. For the other first stage, the judge stringency instrument for the incarceration decision matters little if anything, but the other instrument

[41] While not a trial decision per se, judges could also differ in how quickly they process cases. Creating an instrument based on a judge's average processing time in other cases they have handled, and redoing the empirical tests reported below yields similar conclusions. Another decision margin of the judges is sentencing length. Empirically, however, there is little if any difference in sentence lengths across judges holding incarceration rates fixed. This is consistent with judges having discretion on the incarceration decision, but using mandatory rules or strict guidelines for sentence lengths.

is strongly significant. To formally evaluate the overall strength of the instruments we report the Sanderson-Windmeijer F-statistics, indicating that weak instruments are not an issue. Looking at the reduced form estimates, the coefficients on the judge stringency instrument for the incarceration decision are virtually unchanged compared to the baseline IV model. In contrast, we find that the judge stringency instrument for the other decisions has almost no effect on recidivism in the reduced form. Likewise, the IV estimates for incarceration in the final columns of Appendix Table A11 are similar to those from the baseline IV model which does not include and instrument for the other decision margins.

A useful byproduct of examining the threats to exclusion from trial decisions other than incarceration is that it helps with interpretation. The baseline IV model compares the potential outcomes if incarcerated to the outcomes that would have been realized if not incarcered. The augmented IV model helps to clarify what is meant by not incarcerated, distinguishing between not guilty as opposed to alternative sentences to imprisonment. The IV estimates in Appendix Table A11 suggest significant effects of being sentenced to prison compared to being found not guilty, whereas the probation, community service or fine category does not have a statistically different effect compared to not guilty.

6 Employment and Recidivism

This section explores factors that may explain the preventive effect of incarceration, showing that the decline in crime is driven by individuals who were not working prior to incarceration. Among these individuals, imprisonment increases participation in programs directed at improving employability and reducing recidivism, and ultimately, raises employment and earnings while discouraging further criminal behavior.

6.1 Recidivism as a Function of Prior Employment

To examine heterogeneity in effects by labor market attachment, we assign defendants to two similarly sized groups based on whether they were employed before the crime for which they are in court occured. We classify people as 'previously employed' if they were working in at least one of the past five years; the other individuals are defined as 'previously non-employed'.[42] We then re-estimate the IV model separately for each subgroup.

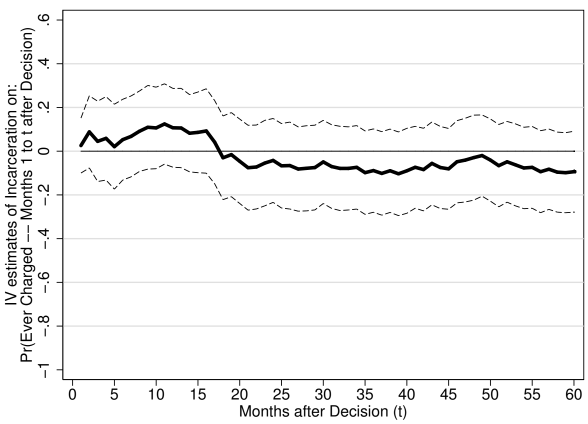
Figure 6 presents the IV estimates for the two subsamples of the effect of incarceration on the probability of reoffending. The results show that the preventive effect of incarceration is

[42] As in Kostøl and Mogstad (2014), an individual is defined as employed in a given year if his annual earnings exceed the yearly substantial gainful activity threshold (used to determine eligibility to government programs like unemployment insurance). In 2010, this amount was approximately NOK 72,900 ($12,500). Our results are not sensitive to exactly how we define employment.

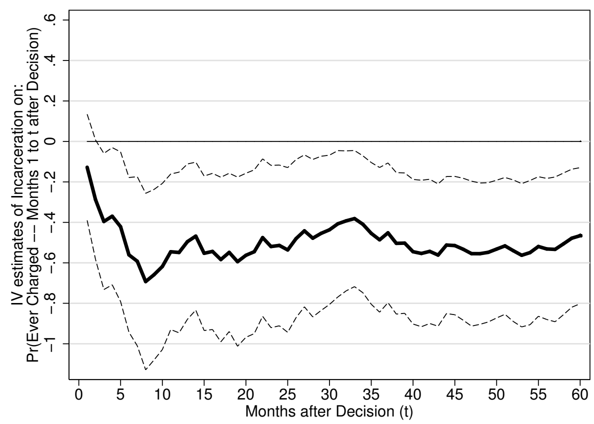
concentrated among the previously non-employed. The effects of incarceration for this group are large and economically important. In particular, the likelihood of reoffending within 5 years is cut in half due to incarceration, from 96 percent to 50 percent. Examining the results in Figure 7 reveals that incarceration not only reduces the probability of re-offending among the previously non-employed, but also the number of crimes they comitt. Five years out, this group is estimated to commit 22 fewer crimes per individual if incarcerated. By comparison, previously employed individuals experience no significant change in recidivism due to incarceration.

Column A:
Previously Employed

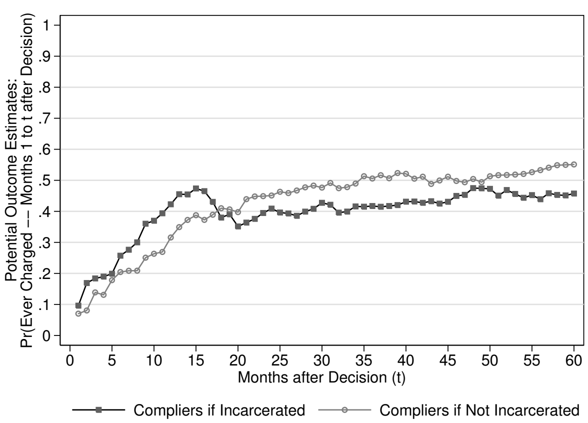
Column B:
Previously Non-employed



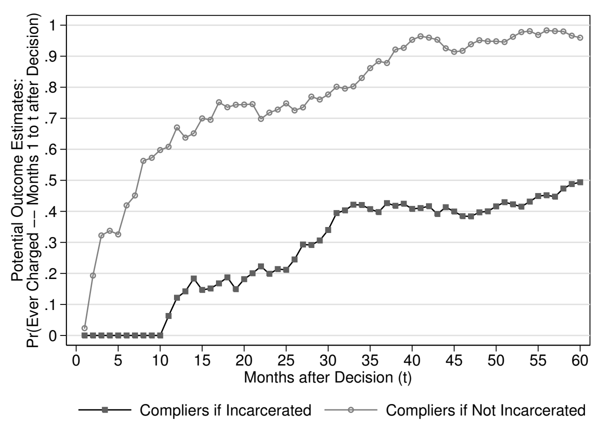
(a) IV Estimates: Pr(Ever Charged – Months 1 to *t*)



(b) IV Estimates: Pr(Ever Charged – Months 1 to *t*)



(c) Potential Outcomes: Pr(Ever Charged – Months 1 to *t*)



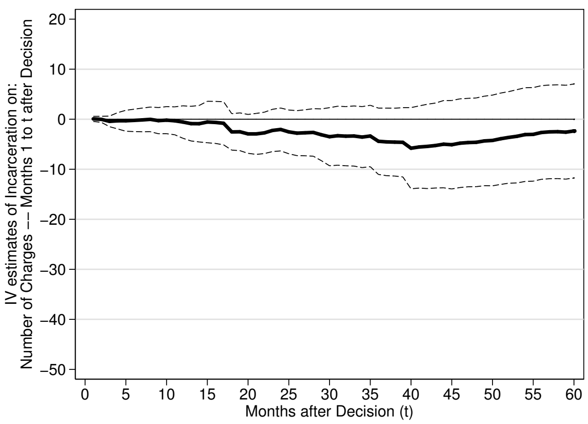
(d) Potential Outcomes: Pr(Ever Charged – Months 1 to *t*)

Figure 6. The Effect of Incarceration on Recidivism by Previous Labor Market Attachment.

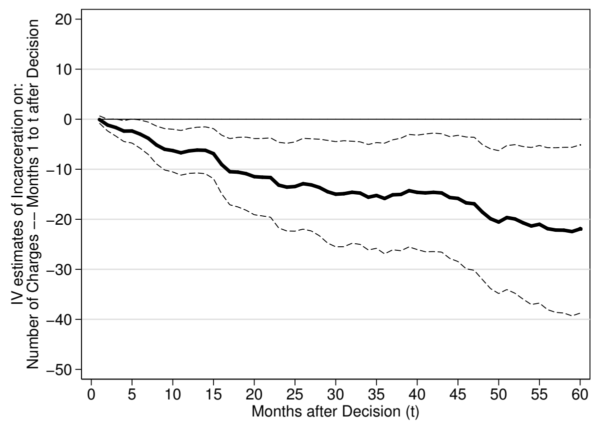
Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=33,509 at time of decision and N=31,287 in month 60 after decision). Dashed lines show 90% confidence intervals.

Column A:
Previously Employed

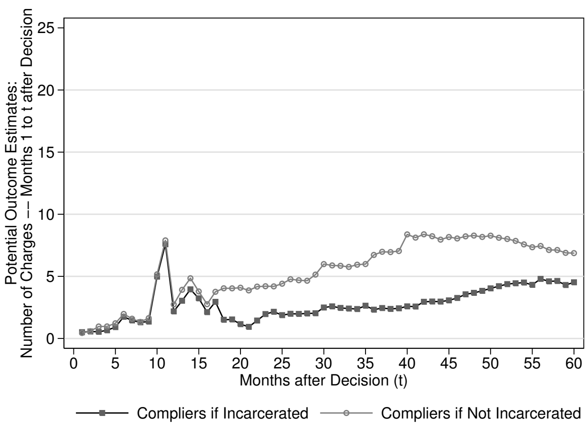
Column B:
Previously Non-employed



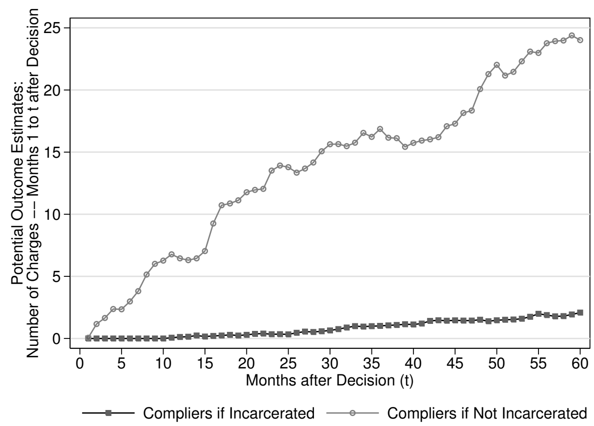
(a) IV Estimates: No. of Charges – Months 1 to t



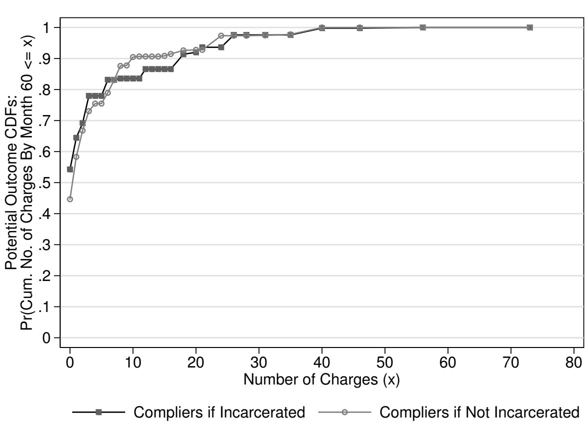
(b) IV Estimates: No. of Charges – Months 1 to t



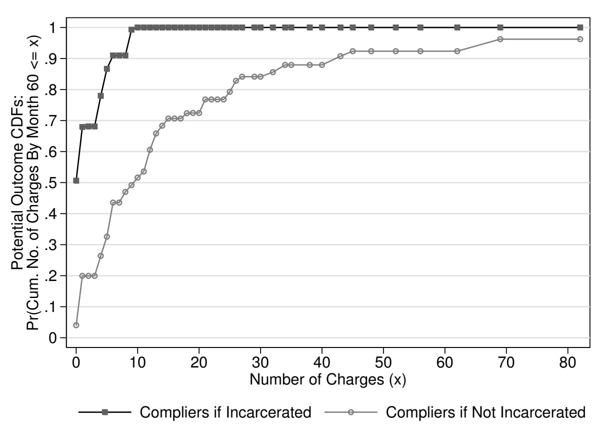
(c) Potential Outcomes: No. of Charges – Months 1 to t



(d) Potential Outcomes: No. of Charges – Months 1 to t



(e) Potential Outcome CDFs: No. of Charges by Month 60



(f) Potential Outcome CDFs: No. of Charges by Month 60

Figure 7. The Effect of Incarceration on Number of Charges by Previous Labor Market Attachment.

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=33,509 at time of decision and N=31,287 in month 60 after decision). Dashed lines show 90% confidence intervals.

A natural question is whether the heterogeneous effects are due to labor market attachment per se or to variables correlated with prior employment.To explore this, we first compare the characteristics of the previously employed and non-employed subsamples. As seen in Appendix Table A2, the two subsamples differ in characteristics other than prior employment.[43] The non-employed group is about two years younger, less likely to be married, and have lower education. They are more likely to commit property and drug-related crimes instead of economic and traffic-related offenses, including drunk driving. Both groups are charged with an equal number of violent crimes. The non-employed individuals are also 50 percent more likely to have been charged with a crime in the year before their court case.

Table 6. The Effect of Incarceration on Recidivism by Previous Labor Market Attachment.

	Sub-sample:		***Sub-sample:***	
Dependent Variable:	**Previously Employed**		**Previously Non-employed**	
A. Pr(Ever Charged)	**(1)**	**(2)**	**(3)**	**(4)**
Months 1-60 after Decision	*Baseline*	*Re-weighted*	*Baseline*	*Re-weighted*
RF: Judge Stringency	-0.053	-0.003	-0.184***	-0.143**
All controls	(0.063)	(0.068)	(0.062)	(0.069)
IV: Incarcerated	-0.094	-0.005	-0.466**	-0.332*
All controls	(0.112)	(0.116)	(0.204)	(0.188)
Dependent mean	0.61	0.67	0.79	0.76
Complier mean if not incarcerated	0.55	0.60	0.96	0.86
Dependent Variable:	**Previously Employed**		**Previously Non-employed**	
B. Number of Charges	**(1)**	**(2)**	**(3)**	**(4)**
Months 1-60 after Decision	*Baseline*	*Re-weighted*	*Baseline*	*Re-weighted*
RF: Judge Stringency	-1.320	-0.287	-8.668***	-8.511***
All controls	(3.178)	(3.778)	(3.259)	(3.086)
IV: Incarcerated	-2.350	-0.487	-21.930**	-19.790**
All controls	(5.694)	(6.406)	(10.196)	(8.853)
Dependent mean	7.08	8.60	13.05	11.85
Complier mean if not incarcerated	3.61	5.16	24.01	21.97
Number of cases	16,500		14,787	

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=31,287 in month 60 after decision). Controls include all variables listed in Table 1, besides controls for court x court entry year FEs. Standard errors are two-way clustered at judge and defendant level. In columns (2) and (4), we use propensity score re-weighting to adjust for differences in observable characteristics across sub-samples; see discussion of the re-weighting procedure in Section 6.1. *p<0.1, **p<0.05, ***p<0.01.

These comparisons make clear that the previously employed and non-employed have different characteristics. To find out whether these differences can explain the contrasting

[43] A few of the non-employed will have earned more than the minimum threshold in the year before their court case, even though by definition the five years before their crime they were non-employed. This is because the date of a court case does not line up precisely with the date of a crime.

recidivism effects, we reweight the subsamples so that they are similar based on observables. To do this, we estimate the probability of being in the previously employed group using all of the control variables listed in Table 1 (excluding the variables on past work history). Appendix Figure A3 plots the estimated propensity scores for both the previously employed and non-employed groups. There is substantial overlap for the entire range of employment probabilities. Using these propensity scores, we weight each subsample so that they have the same distribution as the opposite subsample.

Table 6 reports estimates in columns (1) and (3) for the baseline balanced sample five years after the court decision, without any reweighting. Consistent with the figures discussed above, prison time dramatically reduces the extensive and intensive recidivism margins for the previously non-employed defendants. The same is not true for the previously employed, where the effects are much smaller in absolute value, and not statistically significant. The table then reports the weighted results in columns (2) and (4). This has little effect on the estimates, indicating that differences in the observable characteristics of previously employed and previously non-employed defendants are not driving the contrasting results. Instead, it appears the differential effects are driven by labor market attachment per se or correlated unobservable characteristics.

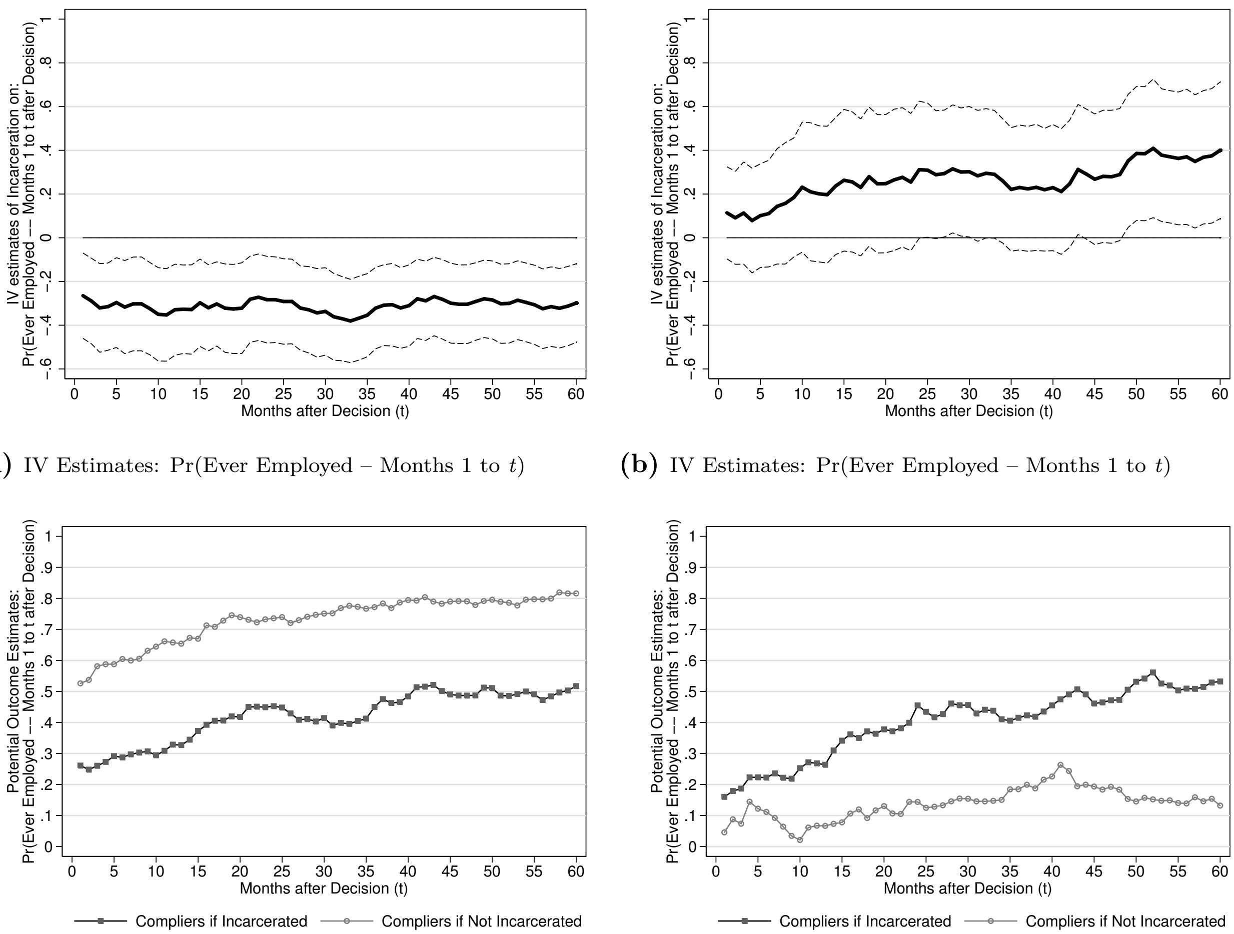
6.2 The Effect of Incarceration on Future Employment

Why are the reductions in recidivism concentrated among the group of defendants with no prior employment? To shed light on this question, we turn to an examination of the labor market consequences of incarceration depending on prior employment.

Using the previously non-employed sample, column B of Figure 8 plots the IV estimates for the probability of being ever employed by a given time period. Two years after the court decision, previously non-employed defendants experience a 30 percentage point increase in employment if incarcerated. This employment boost grows further to a nearly 40 percentage point increase within 5 years. Panel (d) decomposes the IV estimates into the potential employment rates of the compliers. This decomposition reveals that only 12 % of the previously non-employed compliers would have been employed if not incarcerated. By comparison, these compliers would experience a steady increase in employment if incarcerated, with over 50% being employed by month 60. In column A of Figure 8, a different story emerges for the previously employed. They experience an immediate 25 percentage point drop in employment due to incarceration and this effect continues out to 5 years.

Column A:
Previously Employed

Column B:
Previously Non-employed



(a) IV Estimates: Pr(Ever Employed – Months 1 to t)

(b) IV Estimates: Pr(Ever Employed – Months 1 to t)

(c) Potential Outcomes: Pr(Ever Employed – Months 1 to t)

(d) Potential Outcomes: Pr(Ever Employed – Months 1 to t)

Figure 8. The Effect of Incarceration on Future Employment by Previous Labor Market Attachment.

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=33,509 at time of decision and N=31,287 in month 60 after decision). Dashed lines show 90% confidence intervals.

Figure 9 complements the employment results by examining the effects of incarceration on the cumulative hours of work. Starting again with the previously non-employed defendants in the second column, we see a steady increase in the number of hours worked due to incarceration. The IV estimate increases modestly for the first two years, and then starts to increase at a faster rate. By month 60, incarceration increases labor supply by 2,700 hours per individual, translating into more than 550 additional hours per year. The decomposition in panels (d) and (f) help explain what is happening. If not incarcerated, few of the previously non-employed compliers would have got a job. As a result of incarceration, they get a job and continue to accumulate hours over time. Looking atpreviously employed individuals in column A of Figure 9, we see a different pattern. Incarceration has a negative effect on hours worked, consistent with the drop in employment observed for this group Interestingly, the potential employment rate of the previously employed compliers is fairly similar to that of the previously non-employed compliers if they are incarcerated (see panels (c) and (d) in Figure 9). This suggests that incarceration can take an individual who previously had almost no attachment to the labor market, and make them look like someone who also served prison time, but was previously employed.

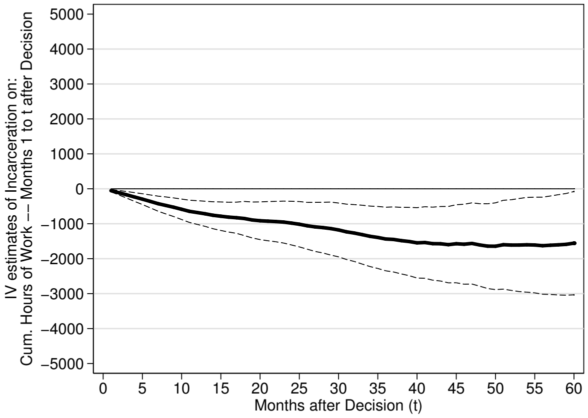
Table 7. The Effect of Incarceration on Future Employment by Previous Labor Market Attachment.

	Sub-sample:		*Sub-sample:*	
Dependent Variable:	**Previously Employed**		**Previously Non-employed**	
A. Pr(Ever Employed)	**(1)**	**(2)**	**(3)**	**(4)**
Months 1-60 after Decision	*Baseline*	*Re-weighted*	*Baseline*	*Re-weighted*
RF: Judge Stringency	-0.168***	-0.154**	0.158**	0.208**
All controls	(0.058)	(0.065)	(0.065)	(0.071)
IV: Incarcerated	-0.299***	-0.262**	0.400**	0.484**
All controls	(0.109)	(0.113)	(0.189)	(0.201)
Dependent mean	0.70	0.67	0.43	0.43
Complier mean if not incarcerated	0.82	0.75	0.13	0.13
Dependent Variable:	**Previously Employed**		**Previously Non-employed**	
B. Cumulative Hours of work	**(1)**	**(2)**	**(3)**	**(4)**
Months 1-60 after Decision	*Baseline*	*Re-weighted*	*Baseline*	*Re-weighted*
RF: Judge Stringency	-873.6*	-662.8	1081.5***	1436.2***
All controls	(506.4)	(526.5)	(366.5)	(419.2)
IV: Incarcerated	-1555.2*	-1123.4	2736.1**	3339.6***
All controls	(896.0)	(881.2)	(1203.5)	(1308.2)
Dependent mean	3810.9	3449.6	1521.4	1605.8
Complier mean if not incarcerated	4410.7	3838.3	51.4	47.8
Dependent Variable:	**Previously Employed**		**Previously Non-employed**	
C. Cumulative Earnings	**(1)**	**(2)**	**(3)**	**(4)**
Months 1-60 after Decision	*Baseline*	*Re-weighted*	*Baseline*	*Re-weighted*
RF: Judge Stringency	-161.8	-127.1	178.9**	265.0***
All controls	(145.9)	(131.0)	(74.1)	(93.2)
IV: Incarcerated	-288.1	-215.4	452.7**	616.3**
All controls	(257.8)	(220.6)	(230.8)	(276.6)
Dependent mean	835.8	713.8	257.0	277.7
Complier mean if not incarcerated	914.2	788.3	9.95	9.70
Number of cases	16,500		14,787	

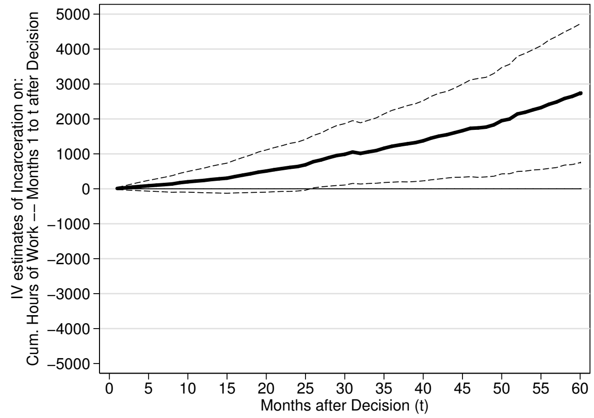
Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=31,287 in month 60 after decision). Controls include all variables listed in Table 1, besides controls for court x court entry year FEs. Standard errors are two-way clustered at judge and defendant level. In columns (2) and (4), we use propensity score re-weighting to adjust for differences in observable characteristics across sub-samples; see discussion of re-weighting in Section 6.1. *p<0.1, **p<0.05, ***p<0.01.

Column A:
Previously Employed

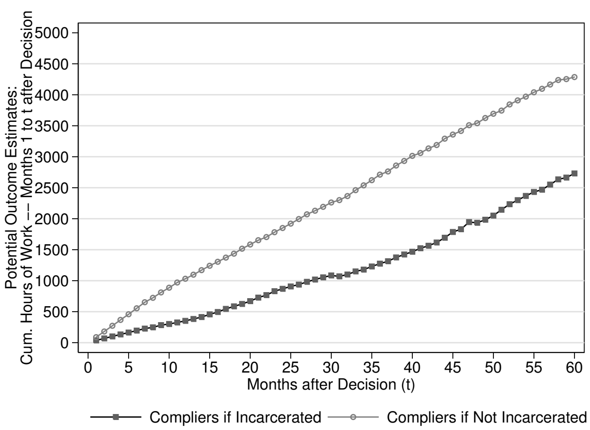
Column B:
Previously Non-employed



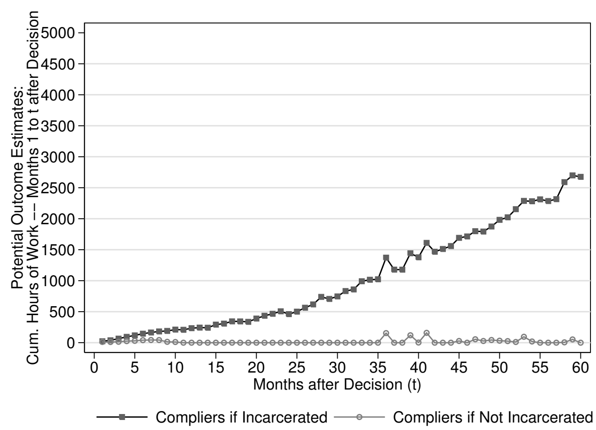
(a) IV Estimates: Hours of Work – Months 1 to t



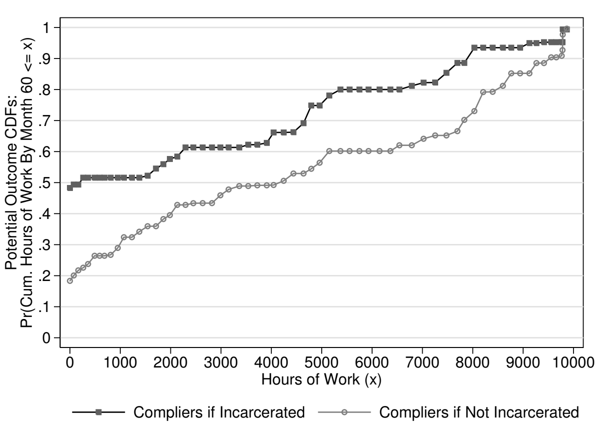
(b) IV Estimates: Hours of Work – Months 1 to t



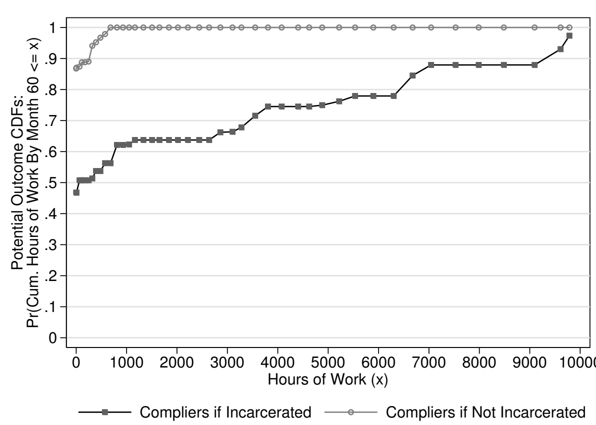
(c) Potential Outcomes: Hours of Work – Months 1 to t



(d) Potential Outcomes: Hours of Work – Months 1 to t



(e) Potential Outcome CDFs: Hours of Work by Month 60

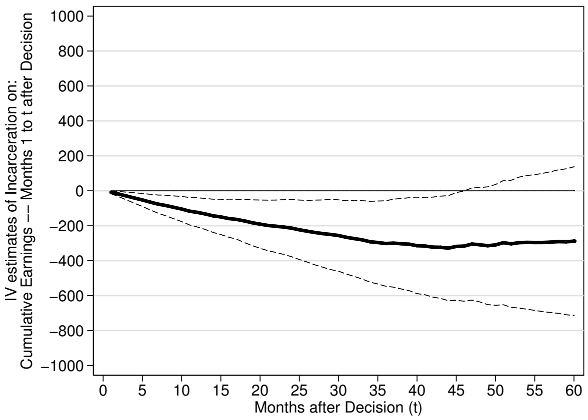


(f) Potential Outcome CDFs: Hours of Work by Month 60

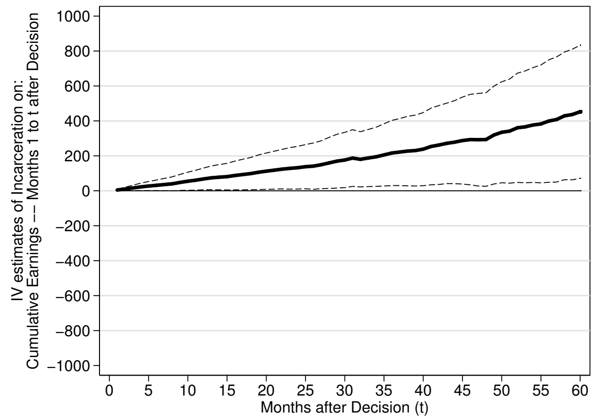
Figure 9. The Effect of Incarceration on Cumulative Hours of Work by Previous Labor Market Attachment.

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=33,509 at time of decision and N=31,287 in month 60 after decision). Dashed lines show 90% confidence intervals.

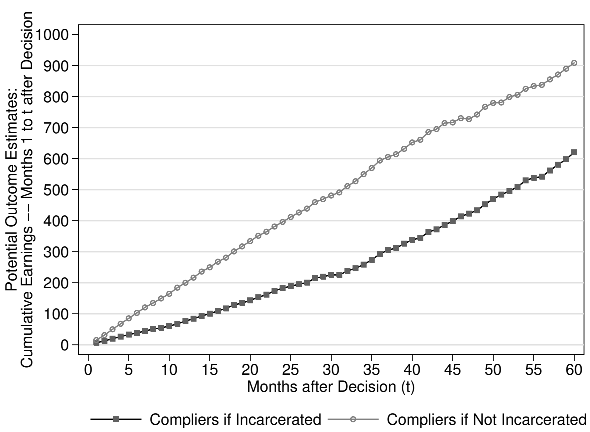
Column A: Previously Employed	**Column B: Previously Non-employed**



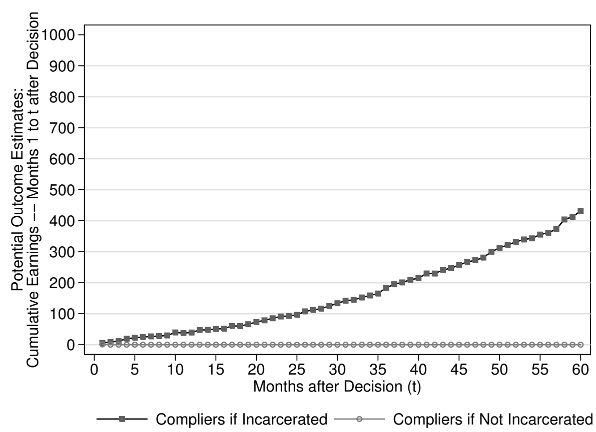
(a) IV Estimates: Cumulative Earnings – Month 1 to t



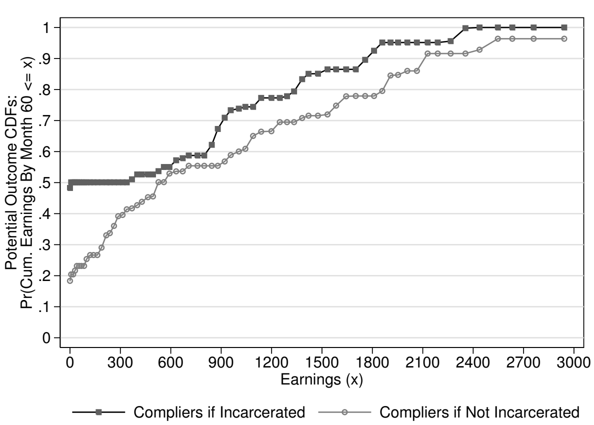
(b) IV Estimates: Cumulative Earnings – Month 1 to t



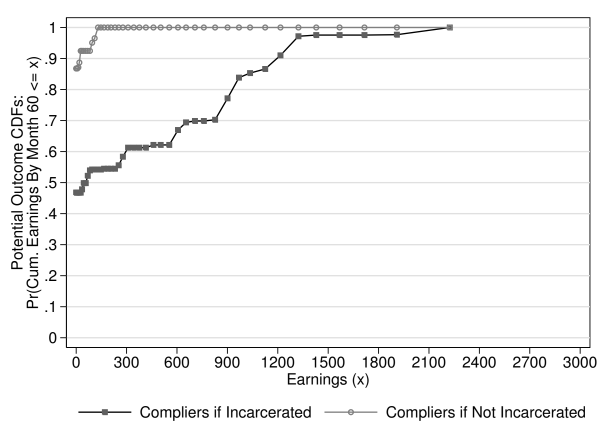
(c) Potential Outcomes: Cum. Earnings – Month 1 to t



(d) Potential Outcomes: Cum. Earnings – Month 1 to t



(e) Potential Outcome CDFs: Cum. Earnings by Month 60



(f) Potential Outcome CDFs: Cum. Earnings by Month 60

Figure 10. The Effect of Incarceration on Cumulative Earnings by Previous Labor Market Attachment.

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=33,509 at time of decision and N=31,287 in month 60 after decision). Dashed lines show 90% confidence intervals.

Figure 10 repeats the same exercise, but this time for cumulative earnings. The general patterns found for employment and hours of work are mirrored in these figures. Lastly, Table 7 shows that differences in observable characteristics other than prior employment are not driving the contrasting labor market effects for previously employed versus previously non-employed defendants.

6.3 The Role of Job Loss and Job Training Programs

The differences in labor market effects depending on prior employment are striking. For the previously employed group, the negative effects are perhaps not unexpected, as these individuals had an actual job to lose by going to prison. To test whether job loss is the explanation, we take advantage of the fact that we can link firms to workers in our data. In particular, we follow the previously employed defendants from years 2 to 5 after their court case (after virtually all incarcerated individuals should be out of prison), and track whether their first employment, if any, during this period was with the same firm as they worked in before the court case. We then run two new IV regressions, quantifying the effect of incarceration on i) the probability of being employed at a new firm and ii) the chance of being employed at the previous firm. The first column in Appendix Table A8 shows the overall employment effect in any firm, which is a 30 percentage point drop due to incarceration. As shown in the next two columns of this table, the drop in employment is almost entirely due to a reduction in the likelihood of employment at the previous firm, whereas there is only a small, and statistically insignificant, effect of incarceration on the probability of employment at a new firm.

Individuals who were not working prior to incarceration had no job to lose. However, serving time in prison could give access to educational and job training programs, both while in prison and immediately after. We collected individual-level data on participation in a variety of job training and classroom training programs. The most common job training program is on-the-job training in a regular or sheltered workplace, where the employer receives a temporary subsidy (normally up to one year) to train the individual and expose them to different jobs. Job training is specifically targeted to those who need work experience in order to find employment. It is often paired with job finding assistance, where a personal counselor helps the individual find a suitable workplace and negotiate wages and employment conditions. The classroom training programs include short skill-focused courses, vocational training and ordinary education. Classroom training is limited to 10 months for skill courses, 2 years for vocational training and 3 years for ordinary education. Forty-three percent of the previously non-employed sample participates in job training and 27% participates in classroom training. In comparison, among the previously employed, 28% participate in job training and 26% in classroom training.

Table 8 reports IV estimates for both types of training using our judge stringency instrument. We focus on the first two years after the court decision, so as to capture the training while in prison and immediately after. For the previously employed group, there are hints that participation in both job and classroom training programs increases due to incarceration, but nothing which is statistically significant. For the previously non-employed group, there is likewise no statistically significant evidence for an increase in classroom training, although the estimate is positive. Instead, what changes significantly due to incarceration is the probability that previously non-employed defendants participate in job training programs. We estimate that being incarcerated makes these individuals 34 percentage points more likely to attend a job training program. By comparison, few if any of the previously non-employed compliers would have participated in job training programs if not incarcerated.

Table 8. The Effect of Incarceration on Participation in Job Training Programs (JTP) and Classrom Training Programs (CTP).

	Sub-sample: **Previously Employed**		***Sub-sample:*** **Previously Non-employed**	
Dependent Variable:	**(1)** *Pr(Participated in Job Training Programs)*	**(2)** *Pr(Participated in Classrom Training Programs)*	**(3)** *Pr(Participated in Job Training Programs)*	**(4)** *Pr(Participated in Classrom Training Programs)*
	Months 1-24 after Decision		*Months 1-24 after Decision*	
RF: Judge Stringency	0.064	0.080	0.133**	0.056
All controls	(0.063)	(0.065)	(0.064)	(0.068)
IV: Incarcerated	0.114	0.142	0.337*	0.142
All controls	(0.113)	(0.116)	(0.181)	(0.177)
Dependent mean	0.17	0.19	0.22	0.17
Complier mean if not incarcerated	0.16	0.18	0.00	0.04
Number of cases	16,500		14,787	

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=31,287 in month 60 after decision). Control variables include all variables listed in Table 1, besides controls for court x court entry year FEs. Standard errors are two-way clustered at judge and defendant level. *p<0.1, **p<0.05, ***p<0.01.

6.4 Putting the Pieces Together

So far, we have demonstrated that the decline in crime from incarceration is driven by individuals who were not working prior to incarceration. Among these individuals, imprisonment increases participation in programs directed at improving employability and reducing recidivism, and ultimately, raises employment and earnings while discouraging further criminal behavior. A natural question is whether the people who, due to incarceration, commit fewer crimes are the same individuals as those who become more likely to participate in job training programs and work more. Or does the decline in crime occur independently of the increase in program participation and employment?We investigate this question in Appendix Table A7. In columns (2) and (3), we first break up the probability of re-offending into the probability of re-offending and employed plus the probability of re-offending and not employed. Using the IV model, we report estimates for how each of these joint probabilities are affected by incarceration. As shown in column (2), there is little change in the joint probability of re-offending and employment due to incarceration. Instead, the entire drop in recidivism appears to be driven by a reduction in the joint probability of re-offending and not employed. The only conclusion consistent with all of our estimates is that individuals who are induced to start working are the same individuals who stop committing crimes.[44]

Going a step further, in columns (4) and (5), we estimate the joint probability of re-offending, employment and job training. We find the entire drop in recidivism reported in column (1) is due to a reduction in the joint probability of being charged, not employed and not participating in a job training program. We therefore conclude that the drop in crime we find for the previously non-employed is driven by the same individuals who, due to incarceration, participate in job training and become gainfully employed.

7 Concluding Remarks

A pivotal point for prison policy was the 1974 Martinson report, which concluded that "nothing works" in rehabilitating prisoners. Around this time, incarceration rates started to rise dramatically, especially in the U.S. where they more than tripled, as an increasing emphasis was placed on punishment and incapacitation. In recent years, researchers and policymakers have questioned whether incarceration is necessarily criminogenic or whether it can instead be preventive. Our study serves as a proof-of-concept demonstrating that time spent in prison with a focus on rehabilitation can indeed be preventive. The Norwegian prison system is succesful in increasing participation in job training programs, encouraging

[44] To see this, let C denote crime, E denote employment, and I denote incarcerated. By definition, $P(C) = P(C \bigcap E) + P(C \bigcap \text{not } E)$. We estimate that $dP(C)/dI < 0$ is driven by $dP(C \bigcap \text{not } E)/dI < 0$, since $dP(C \bigcap E)/dI \approx 0$. Notice that $dP(C \bigcap \text{not } E)/dI < 0$ means that some individuals with C=1, E = 0 if I = 0 change behavior if I = 1. There are three possibilities for change: (i) C = 0, E = 0, (ii) C=1, E = 1 and (iii) C = 0, E = 1. However, (i) is inconsistent with $dP(E)/dI > 0$ and (ii) is inconsistent with $dP(C)/dI < 0$. Only (iii) is consistent with with $dP(E)/dI > 0$, $dP(C)/dI < 0$, and $dP(C \bigcap \text{not } E)/dI < 0$. Note that this is an argument about net effects; while there may be some of types (i) and (ii), they would have to be offset by even more of type (iii).

employment, and discouraging crime, largely due to changes in the behavior of individuals who were not working prior to incarceration. It is important to recognize that our results do not imply that prison is necessarily preventative in all settings. While this paper establishes an important proof of concept, evidence from other settings or populations would be useful to assess the generalizability of our findings.

References

Aebi, M., M. Tiago, and C. Burkhardt (2015). *Survey on Prison Populations (SPACE I – Prison Populations Survey 2014) Survey 2014.* Council of Europe Annual Penal Statistics.

Aizer, A. and J. J. Doyle (2015). Juvenile Incarceration, Human Capital and Future Crime: Evidence from Randomly-Assigned Judges. *The Quarterly Journal of Economics 130*(2), 759–803.

Ashenfelter, O. and D. Card (1985). Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs. *The Review of Economics and Statistics 67*(4), 648–60.

Autor, D., M. Mogstad, and A. R. Kostøl (2015). *Disability Benefits, Consumption Insurance, and Household Labor Supply.* IZA DP No. 8893.

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 Effect In Juvenile Corrections. *Quarterly Journal of Economics 124*(1).

Belloni, A., D. Chen, V. Chernozhukov, and C. Hansen (2012). Sparse Models and Methods for Optimal Instruments with an Application to Eminent Domain. *Econometrica 80*(6), 2369–2429.

Bernburg, J. G., M. D. Krohn, and C. J. Rivera (2006). Official Labeling, Criminal Embeddedness, and Subsequent Delinquency: a Longitudinal Test of Labeling Theory. *Journal of Research in Crime and Delinquency 43*(1), 67–88.

Bohn, A. (2000). *Domsstolloven, Kommentarutgave [Law of Courts, Annotated Edition].* Universitetsopplaget, Oslo (in Norwegian).

Brennan, P. A. and S. A. Mednick (1994). Learning Theory Approach to the Deterrence of Criminal Recidivism. *Journal of Abnormal Psychology 103*(3), 430–440.

Buonanno, P. and S. Raphael (2013). Incarceration and Incapacitation: Evidence from the 2006 Italian Collective Pardon. *The American Economic Review 103*(6), 2437–2465.

Bureau of Justice Statistics (2014). *Prisoners in 2013.* U.S. Department of Justice.

Bureau of Justice Statistics (2015). *Prisoners in 2014*. U.S. Department of Justice.

Chalfin, A. and J. McCrary (forthcoming). Criminal Deterrence: A Review of the Literature. *Journal of Economic Literature*.

Cook, P. J., S. Kang, A. A. Braga, J. Ludwig, and M. E. OBrien (2015). An Experimental Evaluation of a Comprehensive Employment-oriented Prisoner Re-entry Program. *Journal of Quantitative Criminology 31*(3), 355–382.

Cullen, F. T. (2005). The Twelve People Who Saved Rehabilitation: How the Science of Criminology Made a Difference-The American Society of Criminology 2004 Presidential Address. *Criminology 43*(1), 1–42.

Dahl, G. B., A. R. Kostøl, and M. Mogstad (2014). Family Welfare Cultures. *The Quarterly Journal of Economics 129*(4), 1711–1752.

Di Tella, R. and E. Schargrodsky (2013). Criminal Recidivism after Prison and Electronic Monitoring. *Journal of Political Economy 121*(1), 28–73.

Dobbie, W., J. Goldin, and C. Yang (2016). *The Effects of Pre-Trial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges*. NBER Working Paper No. 22511.

Dobbie, W. and J. Song (2015). Debt Relief and Debtor Outcomes: Measuring the Effects of Consumer Bankruptcy Protection. *The American Economic Review 105*(3), 1272–1311.

Doyle, J. J. (2007). Child Protection and Child Outcomes: Measuring the Effects of Foster Care. *The American Economic Review 97*(5), 1583–1610.

Doyle, J. J. (2008). Child Protection and Adult Crime: Using Investigator Assignment to Estimate Causal Effects of Foster Care. *Journal of Political Economy 116*(4), 746–770.

Doyle, J. J., J. A. Graves, J. Gruber, and S. Kleiner (2012). *Do High-Cost Hospitals Deliver better Care? Evidence from Ambulance Referral Patterns*. NBER Working Paper No. 17936.

Freeman, R. B. (1992). *Crime and the Economic Status of Disadvantaged Young Men*. in: George E. Peterson and Wayne Vroman, eds., Urban Labor Markets and Job Opportunities (Urban Institute Press Washington, DC), 112–152.

French, E. and J. Song (2014). The Effect of Disability Insurance Receipt on Labor Supply. *American Economic Journal: Economic Policy 6*(2), 291–337.

GAO-12-743 (2012). *Growing Inmate Crowding Negatively Affects Inmates, Staff, and Infrastructure*. United States Government Accountability Office.

Gottfredson, D. M. (1999). *Effects of Judges' Sentencing Decisions on Criminal Careers*. US Department of Justice, Office of Justice Programs, National Institute of Justice.

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.

Grogger, J. (1995). The Effect of Arrests on the Employment and Earnings of Young Men. *The Quarterly Journal of Economics 110*(1), 51–71.

Harrendorf, S., M. Heiskanen, and S. Malby (2010). *International Statistics on Crime and Justice*. European Institute for Crime Prevention and Control, affiliated with the United Nations (HEUNI).

Imbens, G. W. and D. B. Rubin (1997). Estimating Outcome Distributions for Compliers in Instrumental Variables Models. *The Review of Economic Studies 64*(4), 555–574.

Kling, J. (1999). *The Effect of Prison Sentence Length on the Subsequent Employment and Earnings of Criminal Defendants*. Princeton University, Discussion Papers in Economics, Discussion Paper No. 208.

Kling, J. R. (2006). Incarceration Length, Employment, and Earnings. *The American Economic Review 96*(3), 863–876.

Kostøl, A. R. and M. Mogstad (2014). How Financial Incentives Induce Disability Insurance Recipients to Return to Work. *The American Economic Review 104*(2), 624–655.

Kristoffersen, R. (2014). *Correctional Statistics of Denmark, Finland, Iceland, Norway and Sweden 2009 - 2013*. Correctional Service of Norway Staff Academy.

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.

Lappi-Seppälä, T. (2012). Penal Policies in the Nordic Countries 1960–2010. *Journal of Scandinavian Studies in Criminology and Crime Prevention 13*(1), 85–111.

Lipton, D., R. Martinson, and J. Wilks (1975). *The Effectiveness of Correctional Treatment: A Survey of Treatment Evaluation Studies*. New York Office of Crime Control Planning.

Loeffler, C. E. (2013). Does Imprisonment Alter the Life Course? Evidence on Crime and Employment from a Natural Experiment. *Criminology 51*(1), 137–166.

Maestas, N., K. J. Mullen, and A. Strand (2013). Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt. *The American Economic Review 103*(5), 1797–1829.

Martinson, R. (1974). What Works? - Questions and Answers About Prison Reform. *The Public Interest 35*, 22–54.

Mastrobuoni, G. and D. Terlizzese (2014). *Rehabilitation and Recidivism: Evidence from an Open Prison*. Working Paper.

Mueller-Smith, M. (2014). *The Criminal and Labor Market Impacts of Incarceration.* University of Michigan Working Paper.

Nagin, D. S., F. T. Cullen, and C. L. Jonson (2009). Imprisonment and Reoffending. *Crime and Justice 38*(1), 115–200.

Neal, D. and A. Rick (2016). The Prison Boom and Sentencing Policy. *Journal of Legal Studies 45*(1), 1–41.

NOU (2002). *Dømmes av Likemenn [Judged by Peers].* Ministry of Justice and Public Security, Norway (In Norwegian).

Owens, E. G. (2009). More Time Less Crime? Estimating the Incapacitative Effect of Sentence Enhancements. *Journal of Law and Economics 52*(3), 551–579.

Pew Center (2011). *State of Recidivism. The Revolving Door of America's Prisons.* The Pew Center on the States, Washington, DC.

RAND (2014). *How Effective Is Correctional Education, and Where Do We Go from Here?* Rand Corporation.

Raphael, S. and M. A. Stoll (2013). *Why Are So Many Americans in Prison?* Russell Sage Foundation.

Redcross, C., M. Millenky, T. Rudd, and V. Levshin (2011). More than a Job: Final Results from the Evaluation of the Center for Employment Opportunities (CEO) Transitional Jobs Program. *OPRE Report 18*.

Skardhamar, T. and K. Telle (2012). Post-Release Employment and Recidivism in Norway. *Journal of Quantitative Criminology 28*(4), 629–649.

Stevenson, M. (2016). *Distortion of Justice: How the Inability to Pay Bail Affects Case Outcomes*. Working Paper.

Vera Institute of Justice (2012). *The Price of Prisons: What Incarceration Costs Taxpayers.* Technical Report, Center on Sentencing and Corrections.

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.

Waldfogel, J. (1994). The Effect of Criminal Conviction on Income and the Trust "Reposed in the Workmen". *Journal of Human Resources 29*(1), 62–81.

Western, B. and K. Beckett (1999). How Unregulated is the US Labor Market? The Penal System as a Labor Market Institution. *American Journal of Sociology 104*(4), 1030–60.

Western, B., J. R. Kling, and D. F. Weiman (2001). The Labor Market Consequences of Incarceration. *Crime and Delinquency 47*(3), 410–427.

World Prison Brief (2016). *World Prison Population List (11th edition)*. Institute for Criminal Policy Research (Author: Roy Walmsley).

Appendix Figures and Tables

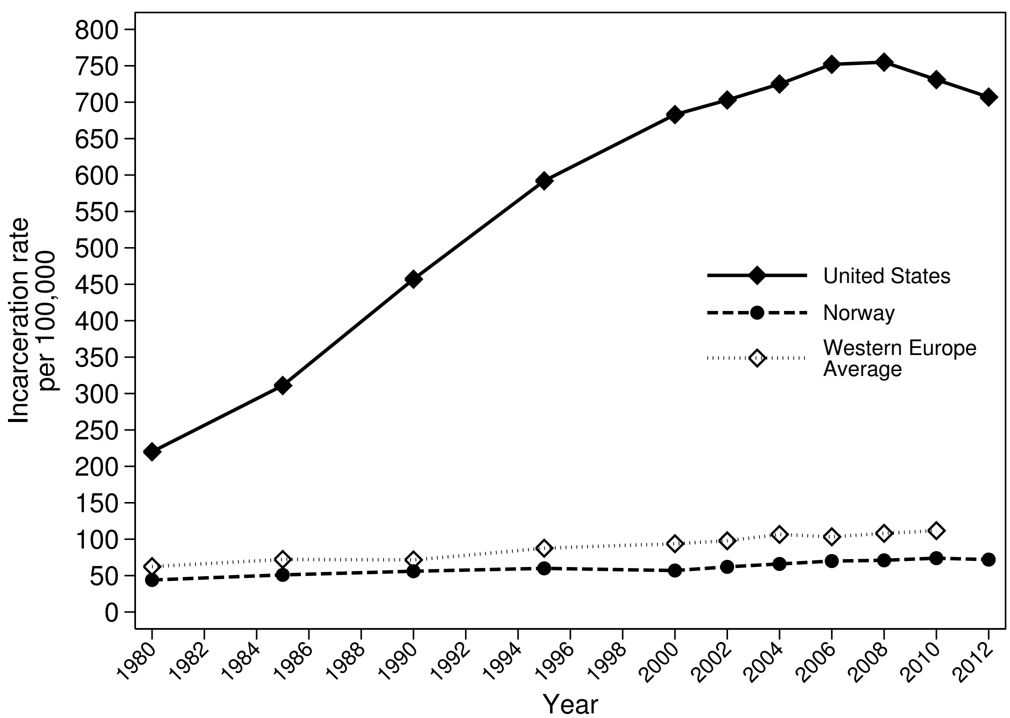


Figure A1. Incarceration Trends in Norway, Western Europe and the U.S.

Note: The Western European countries used to construct the population-weighted average include Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the UK. Source: Institute for Criminal Policy Research.

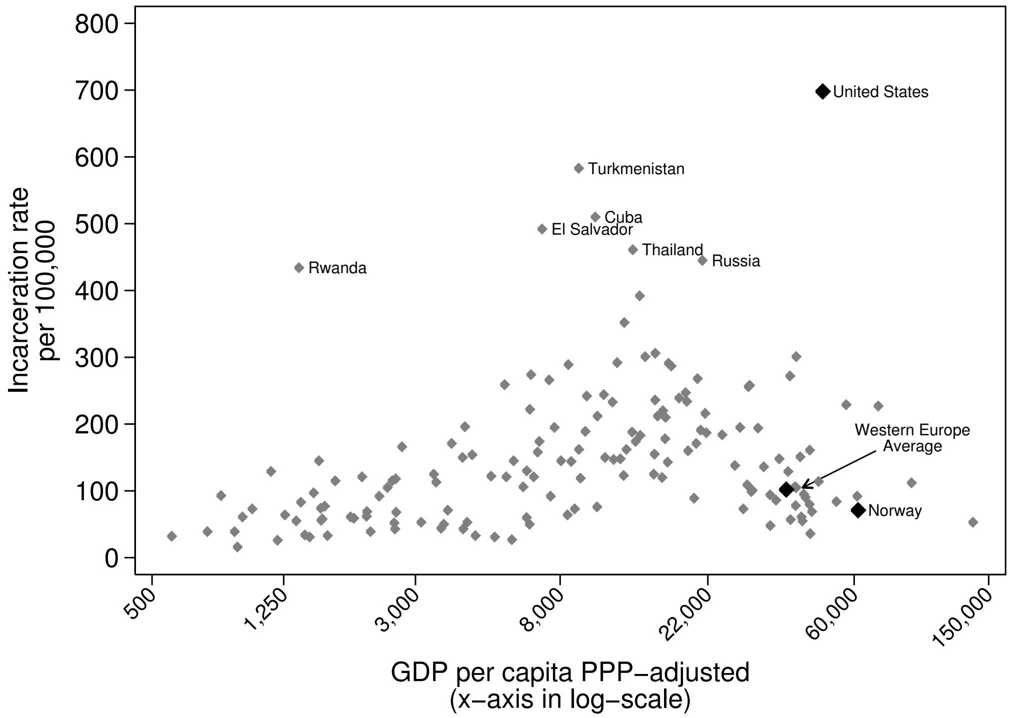


Figure A2. Incarceration Rates versus GDP per Capita.

Note: Sample consists of 160 countries with population greater than 0.5 million and with available data on incarceration and GDP. Incarceration rates and GDP are for the latest available year. GDP per capita is adjusted for purchasing power parity (PPP) and reported in 2010 US dollars. The Western European countries used to construct the population-weighted average include Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the UK. Sources: Institute for Criminal Policy Research, International Monetary Fund and the World Bank.

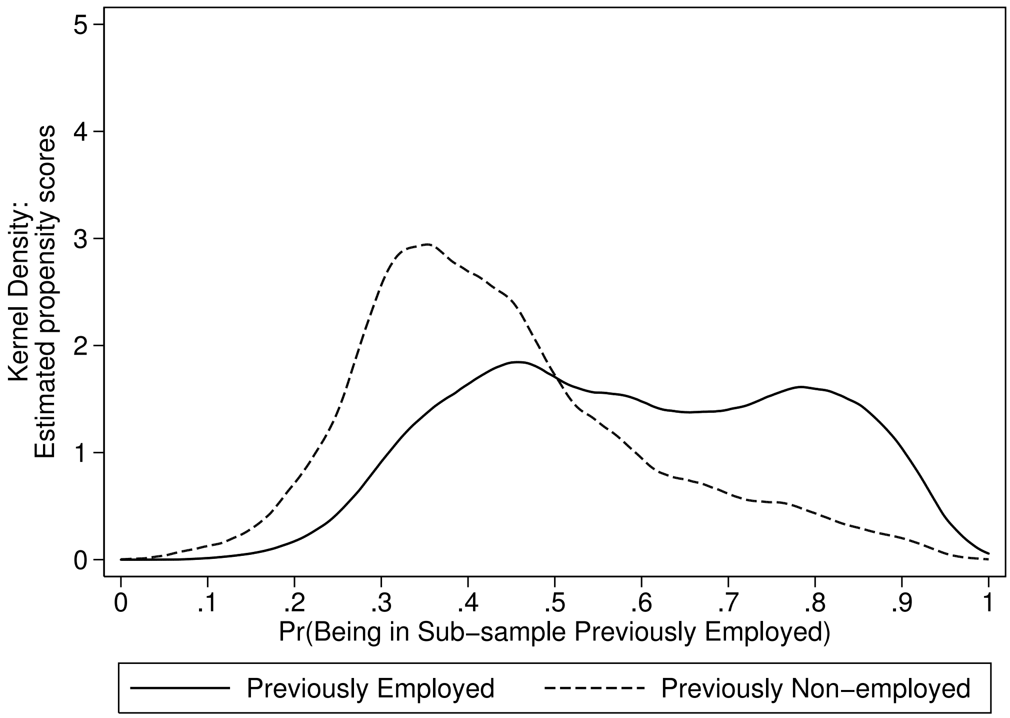


Figure A3. Propensity Score Overlap across Sub-Samples.

Note: Baseline balanced sample consisting of 33,509 non-confession criminal cases processed 2005-2009. The estimated propensity score is a composite index of all variables listed in Table 1, excluding the variables directly capturing past work history which would also fully predict the probability of being in either sub-sample with no overlap. In Tables 6-7, columns (2) and (4), we use the estimated propensity scores to adjust for differences in observable characteristics across sub-samples; see discussion of the re-weighting procedure in Section 6.1.

Table A1. Sample Restrictions.

	Sample Sizes (Remaining After Each Restriction):			
	No. of Cases (1)	**No. of Defendants (2)**	**No. of Judges (3)**	**No. of Courts (4)**
A. All Cases:	**128,804**	**83,143**	**1,059**	**87**
• drop cases with time limits (juveniles; defendant custody).	126,453	82,568	1,059	87
• drop cases with statutory sentence above 6 years.	119,860	79,593	1,059	87
• drop cases where the defendant doesn't have right to a lawyer.	101,930	68,042	1,050	87
B. Non-Confession Cases:	**76,609**	**49,989**	**1,053**	**87**
• drop cases with time limits (juveniles; defendant custody).	74,258	49,247	1,052	87
• drop cases with statutory sentence above 6 years.	68,601	46,358	1,051	87
• drop cases where the defendant doesn't have right to a lawyer.	50,671	33,182	1,041	87
C. Non-Confession Cases Assigned to Regular Judges:	**55,098**	**37,934**	**562**	**87**
• drop cases with time limits (juveniles; defendant custody).	53,294	37,238	562	87
• drop cases with statutory sentence above 6 years.	47,776	34,167	561	87
• drop cases where the defendant doesn't have right to a lawyer.	35,129	24,299	558	87
• drop courts with less than 2 regular judges stationed.	34,554	23,936	557	84
• drop judges who have handled less than 50 criminal cases.	33,509	23,345	500	83

Note: The initial sample consists of all criminal cases processed in Norwegian district courts between 2005-2009.

Table A2. Descriptive Statistics by Previous Labor Market Attachment.

	Sub-sample: Previously Employed		*Sub-sample:* Previously Non-employed	
	Mean	Standard Deviation	Mean	Standard Deviation
	(1)	(2)	(3)	(4)
A. Defendant Characteristics:				
Demographics:				
Age	34.54	(10.60)	31.23	(12.88)
Female	0.102	(0.302)	0.141	(0.348)
Foreign born	0.143	(0.350)	0.153	(0.360)
Married, year t-1	0.168	(0.374)	0.078	(0.269)
Number of children, year t-1	0.956	(1.322)	0.656	(1.215)
Some college, year t-1	0.079	(0.270)	0.025	(0.157)
High school degree, year t-1	0.267	(0.442)	0.084	(0.278)
Less than high school, year t-1	0.654	(0.474)	0.890	(0.312)
Missing Xs	0.024	(0.153)	0.046	(0.210)
Past work and criminal history:				
Employed, year t-1	0.680	(0.467)	0.037	(0.188)
Ever Employed, years t-2 to t-5	0.887	(0.317)	0.031	(0.174)
Charged, year t-1	0.314	(0.464)	0.458	(0.498)
Ever Charged, years t-2 to t-5	0.536	(0.499)	0.618	(0.486)
Number of defendants	12,925		10,420	
B. Type of Crime:				
Violent crime	0.246	(0.431)	0.267	(0.442)
Property crime	0.105	(0.306)	0.176	(0.381)
Economic crime	0.157	(0.364)	0.064	(0.246)
Drug related	0.104	(0.305)	0.136	(0.343)
Drunk driving	0.082	(0.274)	0.059	(0.236)
Other traffic	0.101	(0.301)	0.071	(0.256)
Other crimes	0.205	(0.404)	0.225	(0.418)
Number of cases	17,406		16,103	

Note: Baseline sample consisting of 33,509 non-confession criminal cases processed 2005-2009.

Table A3. Test for Selective Sample Attrition.

	(1)	(2)	(3)	(4)	(5)
***Dependent Variable*:**	Pr(Attrited by Month 12 after Decision)	Pr(Attrited by Month 24 after Decision)	Pr(Attrited by Month 36 after Decision)	Pr(Attrited by Month 48 after Decision)	Pr(Attrited by Month 60 after Decision)
RF: Judge Stringency	-0.009	-0.008	-0.009	-0.004	-0.033
	(0.009)	(0.020)	(0.022)	(0.026)	(0.026)
Dependent mean	0.008	0.024	0.037	0.052	0.066
Number of cases			33,509		

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=33,509 at time of decision and N=31,287 in month 60 after decision). Controls include all variables listed in Table 1. Standard errors are two-way clustered at judge and defendant level. Sample attrition can only occur due to either death or emigration and without such 'natural' attrition our baseline sample would be fully balanced over months 1-60 after decision. **p<0.1, **p<0.05, ***p<0.01.

Table A4. Characterization of Compliers.

		(1) Previously Employed	(2) Previously Non-Employed
1. Sub-sample: Incarceration propensity – 1^{st} quartile (lowest)			
Population share:	$Pr[X_i = x]$	0.138	0.112
Complier share:	$Pr[Complier \mid X_i = x]$	0.159	0.137
Complier cond. pop. share:	$Pr[X_i = x \mid Complier]$	0.164	0.115
Complier relative likelihood:	$\frac{Pr[X_i=x \mid Complier]}{Pr[X_i=x]}$	1.190	1.029
2. Sub-sample: Incarceration propensity – 2^{nd} quartile			
Population share:	$Pr[X_i = x]$	0.132	0.118
Complier share:	$Pr[Complier \mid X_i = x]$	0.212	0.105
Complier cond. pop. share:	$Pr[X_i = x \mid Complier]$	0.211	0.092
Complier relative likelihood:	$\frac{Pr[X_i=x \mid Complier]}{Pr[X_i=x]}$	1.594	0.785
3. Sub-sample: Incarceration propensity – 3^{rd} quartile			
Population share:	$Pr[X_i = x]$	0.123	0.127
Complier share:	$Pr[Complier \mid X_i = x]$	0.137	0.088
Complier cond. pop. share:	$Pr[X_i = x \mid Complier]$	0.126	0.084
Complier relative likelihood:	$\frac{Pr[X_i=x \mid Complier]}{Pr[X_i=x]}$	1.029	0.664
4. Sub-sample: Incarceration propensity – 4^{th} quartile (highest)			
Population share:	$Pr[X_i = x]$	0.134	0.116
Complier share:	$Pr[Complier \mid X_i = x]$	0.150	0.101
Complier cond. pop. share:	$Pr[X_i = x \mid Complier]$	0.151	0.088
Complier relative likelihood:	$\frac{Pr[X_i=x \mid Complier]}{Pr[X_i=x]}$	1.128	0.761
Number of cases		17,406	16,103

Note: Baseline sample consisting of 33,509 non-confession criminal cases processed 2005-2009. We split the sample into eight mutually exclusive and collectively exhaustive subgroups based on previous labor market attachment and quartiles of the predicted probability of incarceration which is estimated based on all variables listed in Table 1. We estimate the first stage equation (1) separately for each subgroup, which allows us to calculate the proportion of compliers by subgroup. For each subgroup, we report the population share (row 1), the complier share (row 2), and the probability of being in a subgroup conditional on being a complier (row 3). Finally, we also report the complier relative likelihood (row 4), which is the ratio of group-specific complier share to the overall complier share estimated to be 0.133 for the full baseline sample.

Table A5. Specification Checks.

	Baseline	≥ 25 cases handled by each judge	≥ 75 cases handled by each judge	Only first time court cases	Reverse-sample instrument	Non-confession sample instrument
		Sample Selection Restrictions:			***Definition of Instrument:***	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Dependent Variable:				**A. Pr(Incarcerated)**		
First stage	0.4648***	0.4636***	0.4790***	0.4758***	0.4536***	0.3713***
	(0.0664)	(0.062)	(0.0689)	(0.0791)	(0.0668)	(0.0522)
Dep. mean	0.5043	0.5043	0.5033	0.4501	0.5029	0.5053
No. of cases	31,287	32,011	30,444	20,558	30,828	28,731
Dependent Variable:				**B. Pr(Ever Charged)**		
				Months 1-60 after Decision		
RF	-0.127***	-0.106**	-0.120***	-0.134***	-0.135***	-0.084**
	(0.045)	(0.045)	(0.046)	(0.061)	(0.043)	(0.038)
IV	-0.274***	-0.229**	-0.251**	-0.281**	-0.298***	-0.226**
	(0.104)	(0.102)	(0.103)	(0.133)	(0.104)	(0.109)
Dep. mean	0.70	0.70	0.70	0.62	0.70	0.70
No. of cases	31,287	32,011	30,444	20,558	30,828	28,731
Dependent Variable:				**C. Number of Charges**		
				Months 1-60 after Decision		
RF	-4.729*	-3.706	-4.723*	-2.399	-4.769*	-3.767*
	(2.519)	(2.256)	(2.644)	(2.255)	(2.586)	(2.123)
IV	-10.176*	-7.994	-9.858*	-5.042	-10.515*	-10.145
	(5.759)	(5.072)	(5.848)	(4.782)	(6.031)	(6.173)
Dep. mean	9.91	9.93	9.90	6.96	9.91	9.91
No. of cases	31,287	32,011	30,444	20,558	30,828	28,731

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=31,287 in month 60 after decision). Controls include all variables listed in Table 1, besides controls for court x court entry year FEs. Standard errors are two-way clustered at judge and defendant level. Column (1) shows baseline estimates using leave-case-out mean judge stringency as an instrument for incarceration decision. Baseline estimates in column (1) include cases assigned to judges who have handled at least 50 cases. In columns (2) and (3), we instead require judges to handle at least 25 cases and at least 75 cases, respectively. In column (4), we restrict the estimation sample to only include first-time court cases as recorded in our dataset comprising of all criminal court cases handled 2005–2014; i.e. decisions related to all subsequent court cases are excluded from estimation sample. In column (5), we first i) randomly split the baseline estimation sample in two equal-sized and mutually exclusive sub-samples, ii) for each subsample, the instrument is then calculated using only judges' case decisions in the other subsample, and finally, iii) we stack the two subsamples to estimate our IV model given by equations given by (1)-(2). In column (6), only non-confession cases are used to construct measures of leave-out mean judge stringency. *p<0.1, **p<0.05, ***p<0.01.

Table A6. Specification Checks for the 'Previously Non-employed' Subsample.

Sub-sample: **Previously Non-employed**	Baseline	*Sample Selection Restrictions:* ≥ 25 cases handled by each judge	≥ 75 cases handled by each judge	Only first time court cases	*Definition of Instrument:* Reverse-sample instrument	Non-confession sample instrument
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Dependent Variable:				**A. Pr(Incarcerated)**		
First stage	0.3953***	0.3916***	0.4282***	0.5289***	0.4129***	0.3490***
	(0.0945)	(0.0909)	(0.0968)	(0.1224)	(0.0898)	(0.0787)
Dep. mean	0.5024	0.5021	0.5016	0.4253	0.5010	0.5036
No. of cases	14,787	15,142	14,435	8,703	14,600	13,616
Dependent Variable:				**B. Pr(Ever Charged)**		
				Months 1-60 after Decision		
RF	-0.184***	-0.179***	-0.191***	-0.281***	-0.180***	-0.148***
	(0.062)	(0.062)	(0.064)	(0.090)	(0.060)	(0.052)
IV	-0.466**	-0.456**	-0.445**	-0.531***	-0.436**	-0.424**
	(0.204)	(0.204)	(0.190)	(0.203)	(0.181)	(0.183)
Dep. mean	0.79	0.79	0.79	0.72	0.79	0.79
No. of cases	14,787	15,142	14,435	8,703	14,600	13,616
Dependent Variable:				**C. Number of Charges**		
				Months 1-60 after Decision		
RF	-8.668***	-7.049**	-9.199***	-8.105**	-8.395***	-5.067**
	(3.259)	(3.044)	(3.326)	(3.557)	(3.012)	(2.549)
IV	-21.930**	-17.999**	-21.482**	-15.323**	-20.331**	-14.516*
	(10.196)	(9.075)	(9.623)	(7.443)	(8.733)	(8.482)
Dep. mean	13.05	13.10	13.06	9.73	13.07	13.08
No. of cases	14,787	15,142	14,435	8,703	14,600	13,616

Note: Sub-sample consisting of previously non-employed defendants in non-confession criminal cases processed 2005-2009 (N=14,787 in month 60 after decision). Controls include all variables listed in Table 1, besides controls for court x court entry year FEs. Standard errors are two-way clustered at judge and defendant level. Column (1) shows baseline estimates using leave-case-out mean judge stringency as an instrument for incarceration decision. Baseline estimates in column (1) include cases assigned to judges who have handled at least 50 cases. In columns (2) and (3), we instead require judges to handle at least 25 cases and at least 75 cases, respectively. In column (4), we restrict the estimation sample to only include first-time court cases as recorded in our dataset comprising of all criminal court cases handled 2005–2014; i.e. decisions related to all subsequent court cases are excluded from estimation sample. In column (5), we first i) randomly split the baseline estimation sample in two equal-sized and mutually exclusive sub-samples, ii) for each subsample, the instrument is then calculated using only judges' case decisions in the other subsample, and finally, iii) we stack the two subsamples to estimate our IV model given by equations given by (1)-(2). In column (6), only non-confession cases are used to construct measures of leave-out mean judge stringency. *p<0.1, **p<0.05, ***p<0.01.

Table A7. Decomposing the Effect of Incarceration on Reoffending, Future Employment and Job Training Program (JTP) Participation for the 'Previously Non-employed' Subsample.

		Future Employment: *Months 1-60 after Decision*		**Future Employment & JTP Participation:** *Months 1-60 after Decision*	
	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	*Pr[Ever Charged]*	*Pr[Ever Charged ∩ Ever Employed]*	*Pr[Ever Charged ∩ Not Employed]*	*Pr[Ever Charged ∩ (Ever Employed ∪ On JTP)]*	*Pr[Ever Charged ∩ Not Employed ∩ Not on JTP]*
RF: Judge Stringency	-0.184***	0.027	-0.211***	0.020	-0.204***
All controls	(0.062)	(0.067)	(0.070)	(0.072)	(0.068)
IV: Incarcerated	-0.466**	0.069	-0.535**	0.051	-0.517**
All controls	(0.204)	(0.170)	(0.227)	(0.181)	(0.221)
Dependent mean	0.79	0.32	0.47	0.41	0.38
Complier mean if not incarcerated	0.96	0.13	0.83	0.13	0.83
Number of cases			14,787		

Note: Sub-sample consisting of previously non-employed defendants in non-confession criminal cases processed 2005-2009 (N=14,787 in month 60 after decision). Controls include all variables listed in Table 1, besides controls for court x court entry year FEs. Standard errors are two-way clustered at judge and defendant level. *p<0.1, **p<0.05, ***p<0.01.

Table A8. Decomposing the Employment Effect of Incarceration into Employment at the Previous Firm versus a New Firm for the 'Previously Employed' Subsample.

		Change in Employer: *Months 1-60 after Decision*	
	(1)	(2)	(3)
Dependent Variable:	*Pr(Ever Employed)*	*Pr(Employed at Previous Firm)*	*Pr(Employed at New Firm)*
RF: Judge Stringency	-0.168***	-0.140**	-0.028
All controls	(0.058)	(0.068)	(0.063)
IV: Incarcerated	-0.299***	-0.249**	-0.050
All controls	(0.109)	(0.122)	(0.112)
Dependent mean	0.70	0.40	0.30
Complier mean if not incarcerated	0.82	0.52	0.30
Number of cases		16,500	

Note: Sub-sample consisting of previously employed defendants in non-confession criminal cases processed 2005-2009 (N=16,500 in month 60 after decision). Controls include all variables listed in Table 1. RF and IV in addition also control for court x court entry year FEs. RF and IV standard errors are two-way clustered at judge and defendant level. *p<0.1, **p<0.05, ***p<0.01.

Table A9. Tests for the Monotonicity Assumption.

	Baseline Instrument:	Reverse-sample Instrument:
	(1)	**(2)**
	First Stage	**First Stage**
Dependent Variable:	*Pr(Incarcerated)*	*Pr(Incarcerated)*
A. INCARCERATION PROPENSITY (ALL COVARIATES):		
1. Sub-sample: Incarceration propensity – 1^{st} quartile (lowest)		
Estimate	0.4804***	0.5152***
(SE)	(0.1102)	(0.1098)
Dependent mean	0.3742	0.3746
Number of cases	7,820	7,657
2. Sub-sample: Incarceration propensity – 2^{nd} quartile		
Estimate	0.5543***	0.5974***
(SE)	(0.0964)	(0.0942)
Dependent mean	0.4690	0.4683
Number of cases	7,823	7,673
3. Sub-sample: Incarceration propensity – 3^{rd} quartile		
Estimate	0.4138***	0.4571***
(SE)	(0.1182)	(0.1122)
Dependent mean	0.5442	0.5420
Number of cases	7,822	7,679
4. Sub-sample: Incarceration propensity – 4^{th} quartile (highest)		
Estimate	0.4663***	0.4535***
(SE)	(0.0980)	(0.0999)
Dependent mean	0.6296	0.6261
Number of cases	7,822	7,648
B. TYPE OF CRIME:		
1. Sub-sample: Type of crime – Violent crimes		
Estimate	0.6331***	0.5728***
(SE)	(0.1112)	(0.1147)
Dependent mean	0.5519	0.5510
Number of cases	8,077	7,977
2. Sub-sample: Type of crime – Drug-related crimes		
Estimate	0.4392***	0.4077***
(SE)	(0.1463)	(0.1480)
Dependent mean	0.5229	0.5223
Number of cases	4,345	4,283
3. Sub-sample: Type of crime – Property crimes		
Estimate	0.4991***	0.5578***
(SE)	(0.1555)	(0.1455)
Dependent mean	0.4282	0.4270
Number of cases	3,566	3,534
4. Sub-sample: Type of crime – Economic crimes		
Estimate	0.5853***	0.5685***
(SE)	(0.1660)	(0.1625)
Dependent mean	0.4769	0.4760
Number of cases	3,674	3,649
5. Sub-sample: Type of crime – Drunk-driving & other traffic offences		
Estimate	0.2761**	0.3466***
(SE)	(0.1354)	(0.1235)
Dependent mean	0.5009	0.4998
Number of cases	4,949	4,766
6. Sub-sample: Type of crime – Other crimes		
Estimate	0.4298***	0.4640***
(SE)	(0.1212)	(0.1234)
Dependent mean	0.4927	0.4914
Number of cases	6,676	6,565
C. PREVIOUS LABOR MARKET ATTACHMENT:		
1. Sub-sample: Previously Employed		
Estimate	0.5618***	0.3831***
(SE)	(0.0817)	(0.0787)
Dependent mean	0.5059	0.5059
Number of cases	16,500	14,653

(continued on the next page)

Table A9. Tests for the Monotonicity Assumption.

	Baseline Instrument:	Reverse-sample Instrument:
	(1)	(2)
	First Stage	First Stage
Dependent Variable:	*Pr(Incarcerated)*	*Pr(Incarcerated)*
(continued from the previous page)		
C. PREVIOUS LABOR MARKET ATTACHMENT:		
2. Sub-sample: Previously Non-employed		
Estimate	0.3953***	0.3646***
(SE)	(0.0945)	(0.0792)
Dependent mean	0.5024	0.5052
Number of cases	14,787	13,992
D. PREVIOUS INCARCERATION STATUS:		
1. Sub-sample: Previously Incarcerated		
Estimate	0.4884***	0.4851***
(SE)	(0.0833)	(0.0911)
Dependent mean	0.6097	0.6085
Number of cases	15,953	15,226
2. Sub-sample: Previously Non-incarcerated		
Estimate	0.4202***	0.3562***
(SE)	(0.0808)	(0.0732)
Dependent mean	0.3945	0.3969
Number of cases	15,334	13,786
E. AGE:		
1. Sub-sample: Age at time of first offense > 30		
Estimate	0.4300***	0.3355***
(SE)	(0.0880)	(0.0868)
Dependent mean	0.5237	0.5272
Number of cases	15,403	14,248
2. Sub-sample: Age at time of first offense < 30		
Estimate	0.5544***	0.4744***
(SE)	(0.0792)	(0.0857)
Dependent mean	0.4855	0.4879
Number of cases	15,884	14,693
F. LEVEL OF EDUCATION:		
1. Sub-sample: Less than high school at time of first offense		
Estimate	0.4586***	0.2839***
(SE)	(0.0723)	(0.0797)
Dependent mean	0.5108	0.5093
Number of cases	22,533	16,631
2. Sub-sample: High school or above at time of first offense		
Estimate	0.5047***	0.4945***
(SE)	(0.1084)	(0.1078)
Dependent mean	0.4875	0.4864
Number of cases	8,754	8,523
G. NUMBER OF CHILDREN:		
1. Sub-sample: Had no children at time of first offense		
Estimate	0.4630***	0.3677***
(SE)	(0.0784)	(0.0735)
Dependent mean	0.4951	0.4986
Number of cases	19,600	15,642
2. Sub-sample: Had children at time of first offense		
Estimate	0.4834***	0.4793***
(SE)	(0.1015)	(0.0944)
Dependent mean	0.5197	0.5184
Number of cases	11,687	11,275

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=31,287 in month 60 after decision). Controls include all variables listed in Table 1, besides controls for court x court entry year FEs. Standard errors are two-way clustered at judge and defendant level. To reduce noise in judge stringency instruments, judges with less than 50 handled cases are dropped from each sample. **p<0.1, **p<0.05, ***p<0.01.

Table A10. Controlling for Judge Stringency in Decision Margins Other Than Incarceration.

	First Stage	Reduced Form		IV	
	(1)	(2)	(3)	(4)	(5)
	Decision:	*Months 1-60 after Decision:*		*Months 1-60 after Decision:*	
	Pr(Incarcerated)	**Pr(Ever Charged)**	**No. of Charges**	**Pr(Ever Charged)**	**No. of Charges**
A. Baseline Specification					
	0.4648***	-0.127***	-4.728*	-0.274***	-10.176*
	(0.066)	(0.045)	(2.519)	(0.104)	(5.759)
F-stat. (Instrument)	48.43 [0.000]				
B. Additional Controls for 'Probation Stringency', 'Community Service Stringency' and 'Fine Stringency'					
	0.3960***	-0.111*	-5.277*	-0.279*	-13.316
	(0.077)	(0.060)	(2.912)	(0.163)	(8.305)
F-stat. (Instrument)	25.90 [0.000]				
C. Additional Control for 'Probation, Community Service or Fine Stringency'					
	0.4466***	-0.127***	-4.518*	-0.285**	-10.117*
	(0.066)	(0.046)	(2.495)	(0.112)	(6.005)
F-stat. (Instrument)	45.83 [0.000]				
Number of cases			31,827		

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=31,287 in month 60 after decision). Controls include all variables listed in Table 1, besides controls for court x court entry year FEs. Standard errors are two-way clustered at judge and defendant level. *p<0.1, **p<0.05, ***p<0.01.

Table A11. IV Model with Multiple Decision Margins.

	First Stages		Reduced Form			IV	
	(1)	(2)	(3)	(4)		(5)	(6)
	Decision:		*Months 1-60 after Decision:*			*Months 1-60 after Decision:*	
	Pr(Incarcerated)	**Pr(Probation, CS or Fine)**	**Pr(Ever Charged)**	**No. of Charges**		**Pr(Ever Charged)**	**No. of Charges**
A. Baseline Specification							
Instrument:					*Decision:*		
Incarceration stringency	0.4648***		-0.127***	-4.728*	Incarcerated	-0.274***	-10.176*
	(0.066)		(0.045)	(2.519)		(0.104)	(5.759)
F-stat. (Instrument)	48.43 [0.000]						
Dependent mean	0.50		0.70	9.91		0.70	9.91
B. Specification with Three Decision Margins							
Instruments:					*Decisions:*		
Incarceration stringency	0.4466***	-0.0149	-0.127***	-4.518*	Incarcerated	-0.281**	-10.139*
	(0.066)	(0.039)	(0.046)	(2.495)		(0.110)	(5.901)
Probation, CS, or Fine	0.2134**	0.4890***	-0.000	-2.479	Probation, CS,	0.122	-0.646
stringency	(0.104)	(0.081)	(0.080)	(3.966)	or Fine	(0.196)	(9.060)
Sanderson-Windmeijer F-stat.	42.04 [0.000]	29.52 [0.000]					
Dependent mean	0.50	0.89	0.70	9.91		0.70	9.91
Number of cases				31,827			

Note: Baseline sample consisting of non-confession criminal cases processed 2005-2009 (N=31,287 in month 60 after decision). Control variables include all variables listed in Table 1, besides controls for court x court entry year FEs. Standard errors are two-way clustered at judge and defendant level. *p<0.1, **p<0.05, ***p<0.01. CS = Community Service. The omitted decision category in Panel A is "Probation, Community Service, Fine, or Not Guilty", while in Panel B the omitted category is "Not Guilty". *p<0.1, **p<0.05, ***p<0.01.