



HAL
open science

Prison, Semi-Liberty and Recidivism: Bounding Causal Effects in a Survival Model

Benjamin Monnery, François-Charles Wolff, Anaïs Henneguelle

► **To cite this version:**

Benjamin Monnery, François-Charles Wolff, Anaïs Henneguelle. Prison, Semi-Liberty and Recidivism: Bounding Causal Effects in a Survival Model. *International Review of Law and Economics*, 2020, 61, 10.1016/j.irle.2019.105884 . hal-02423761

HAL Id: hal-02423761

<https://hal.science/hal-02423761v1>

Submitted on 21 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Prison, Semi-Liberty and Recidivism: Bounding Causal Effects in a Survival Model[#]

Benjamin Monnery^{*} François-Charles Wolff^{**} Anaïs Henneguelle^{***}

First revision, *International Review of Law and Economics*

December 2019

Abstract

This paper investigates the effect of semi-liberty as an alternative to prison on recidivism in France. Our analysis is based on a unique dataset comprising 1,445 offenders, all eligible to semi-liberty. In the absence of an instrumental variable affecting access to semi-liberty but unrelated to recidivism, we turn to selection-on-observable methods as well as sensitivity analyses to bound the causal effect of interest. Our results under treatment exogeneity (Cox regressions) and conditional independence (entropy balancing matching) show that semi-liberty is associated with a reduction of 22% to 31% in offenders' hazard of recidivism in the five years after release. The estimated effects decrease, but remain negative and significant when credible confounders are introduced. Overall, our analysis lends strong support for a beneficial effect of semi-liberty compared to prison.

Keywords: Recidivism, semi-liberty, halfway houses, prison, survival analyses, sensitivity analyses

JEL Classification: K14, K42, C18

Declarations of interest: none

[#] We are indebted to one anonymous reviewer and Eric Helland, the editor, for their very helpful remarks and suggestions on a previous draft. We have also benefited from comments at the EALE 2018 conference (Milano), the AFED 2018 conference (Nancy), the JMA 2019 conference (Casablanca) and the CRED seminar (Paris). We would like to thank Annie Kensey (DAP Me5) for providing access to the raw data and Ann Thampi for research assistance. Any remaining errors are ours.

^{*} Corresponding author: EconomiX (CNRS), University Paris Nanterre, 200 avenue de la République, 92001 Nanterre Cedex, France. Email: bmonnery@parisnanterre.fr

^{**} LEMNA, University of Nantes and TEPP, France. Email: francois-charles.wolff@univ-nantes.fr

^{***} LIRIS, University of Rennes 2, France. Email: anaïs.henneguelle@univ-rennes2.fr

1. Introduction

Recidivism is a pervasive issue in many criminal justice systems worldwide. In the United States, 55% of state prisoners released in the year 2005 were reconvicted and 77% were rearrested in the next five years (Durose et al., 2014). In France, the reconviction rate is estimated at 59% within five years after prison release (Kensey and Benaouda, 2011). In light of these high recidivism rates, scholars often refer to prison as a revolving door, with a small fraction of the general population accounting for most transitions in and out of prison (Freeman, 2003)¹.

Many developed countries are increasingly turning away from prison-centered policies and adopting new strategies to prevent crime and recidivism. One of them consists of promoting alternative sanctions such as electronic monitoring, probation or semi-liberty, which are cheaper than prison and often perceived as more effective in preventing recidivism. Semi-liberty (SL hereafter) is an interesting mix of custody and liberty. Offenders under SL are free during weekdays, but have to stay at night and during weekends in dedicated correctional facilities, often called “halfway houses” or “semi-liberty centers” in France. While in the United States halfway houses are typically used to smooth the transition from prison to the community, semi-liberty can also serve as a front-door alternative to prison in France².

Semi-liberty has the potential to deliver three crime-control ingredients: i) incapacitation due to daily correctional supervision; ii) deterrence, as SL makes the threat of traditional incarceration very salient; and iii) rehabilitation, since offenders are expected to work outside or receive job training and may benefit from medical treatment or family support. While theoretically attractive, semi-liberty is (curiously) rarely mentioned in discussions on criminal justice reform, and its extent appears to decrease in several countries without clear explanation. In the United States for instance, although halfway houses receive tens of thousands of federal prison releasees annually, the trend is diminishing and residential reentry centers are being shut down (Politico, 2018).

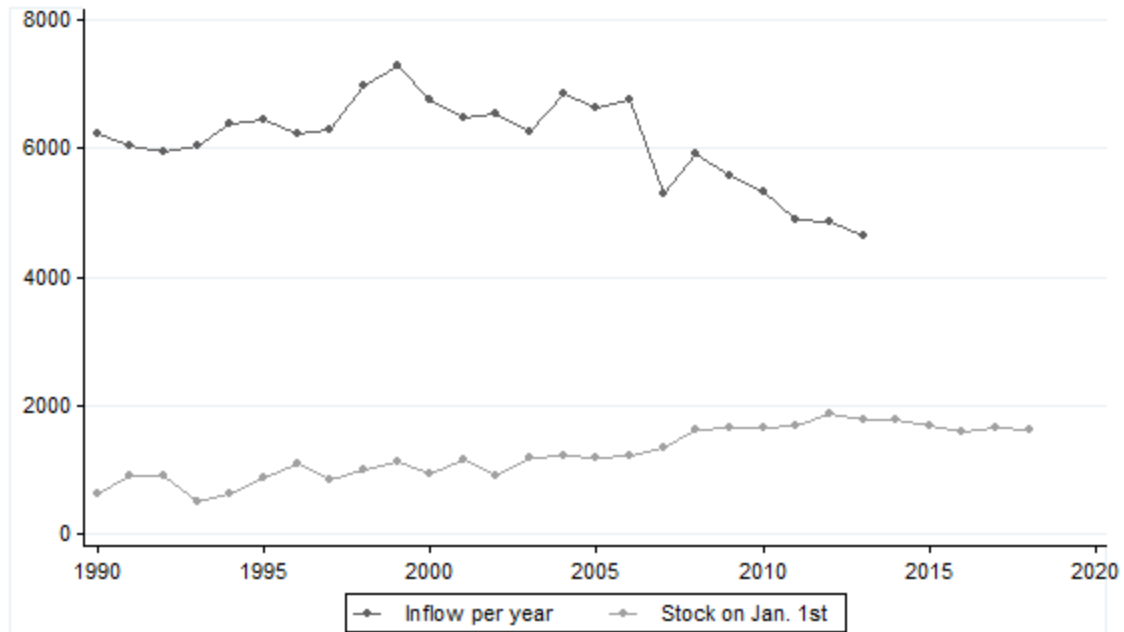
As shown in Figure 1, the annual flow of offenders benefiting from SL in France decreased from 6,000-7,000 per year between 1990 and 2006 to about 4,500 more recently. Although French prisons suffer from severe overcrowding (with around 70,000 prisoners for a 59,000 capacity in January 2019), most semi-liberty centers work under capacity with a 62% occupancy rate (DAP, 2018b). Thus, part of the pressure inside prisons due to overcrowding could be lessened by using the currently available slots in semi-liberty centers. Surprisingly, the limited appeal of semi-liberty is in stark contrast with the massive use of electronic monitoring in France (10,200 offenders on January 1 2018, for an annual

1 On recidivism and its causes (such as criminal background, but also integration into the labour market), see for instance Richey (2015), Siwach (2017) and Bhuller et al. (2019).

2 According to Lee (2019), 55% of federal prison releasees in the year 2015 went through residential reentry centers in the United States.

inflow of more than 20,000), although the two types of sanction target very similar types of offenders and share common features. One might think that so few convicts are in semi-liberty precisely because this measure has proven to fail. However, there is very limited evidence available to judges and policymakers on its empirical impact of recidivism.

Figure 1. Number of semi-liberty offenders in France



Source: data from French Prison Administration, authors' calculations.

Regarding the effects of front-door semi-liberty on future crime, there is no compelling evidence either in France or abroad, as opposed to electronic monitoring (Di Tella and Schargrodsy, 2013; Henneguëlle et al., 2016) or other probation programs (Mueller-Smith and Schnepel, 2017)³. This gap can certainly be explained by the lack of access to micro-level data, by the heterogeneity of semi-liberty programs across contexts, and by the difficulty to properly account for selection bias. Indeed, judges presumably select the best, most fit, lowest-risk offenders to enter such programs (while other offenders are incarcerated), making simple group comparisons of recidivism rates unreliable⁴. So far, the few existing studies on semi-liberty and recidivism have attempted to account for selection bias by controlling for observable differences.

3 For instance, Henneguëlle et al. (2016) find a 10% reduction of the 5-year probability of reconviction for offenders under electronic monitoring in France.

4 In the United States alone, this measure is very heterogeneous (Kilburn and Costanza, 2011). It is aimed simultaneously at former prisoners released on parole (Latessa and Allen, 1982), at convicts whose parole has gone wrong (White et al., 2011), and at other types of profiles such as convicts with psychiatric disorders (Gumrukcu, 1968; Carpenter, 1978). Moreover, existing papers use different outcomes to estimate the effect of semi-liberty: the probability of failing the measure and therefore of being re-incarcerated (Walsh and Beck,

They typically use logistic regressions (Clark, 2016), propensity score matching (Hamilton and Campbell, 2014), or a combination of the two methods (Constanza et al., 2015). However, selection on unobservables may well remain and lead to (potentially important) bias in the estimated effects. To date, the only quasi-experimental evaluation of the effect of SL on recidivism focuses on a back-door transitional program in Iowa. Using random variation in counselors' use of halfway housing versus traditional parole near the end of prisoners' sentences, Lee (2019) estimates that halfway houses significantly increase the three-years probability of return to prison for prisoners on the margin, compared to parole. This detrimental impact is partly explained by the greater supervision and detection of wrongful behavior (new crimes or technical violations) imposed in halfway houses, but it may also be related to negative interactions between inmates. While Lee (2019) offers the first clean evidence on halfway houses and recidivism, this study does not inform regarding the impact of semi-liberty used as a full substitute for incarceration.

This paper seeks to fill this gap by providing estimates of the causal effect on recidivism of semi-liberty, as a front-door alternative to incarceration. Interestingly, in France, SL is only accessible to offenders convicted of prison sentences: SL-offenders and incarcerated offenders therefore come from the same pool of prison convicts. However, selection bias is also likely to arise in this context. The decision to grant SL to prison convicts (or to incarcerate them) is possibly based on information that is observed or inferred by judges (in relation to police files or court hearings for instance), but not measured in the datasets available to researchers. Examples of such potential confounders include impulsivity and other personality traits, quality of family support, economic burden, social deprivation, drug or alcohol addictions or gang membership. Due to data restrictions, it is not possible to measure the effect of those variables on the probability of obtaining SL as well as on recidivism.

At the same time, judges themselves have certainly incomplete knowledge of those types of variables, for at least two reasons. First, probation officers who are supposed to write a social survey about all eligible convicts have only very limited time and have to deal with many people at once. Second, they rely only on self-reported data and generally do not verify convicts' statements. From an empirical perspective, a difficulty is that judges are also interested in the "rehabilitation project" of convicts (Henneguelle, 2020). This project, which does not take the form of a file but is presented orally, is not available to researchers. Its content is related to the social and economic integration project and is likely to have some influence during the hearings, as judges are keen to test the motivation of the defendant. Such confounders are expected to bias the impact of SL on recidivism (omitted variable bias).

1990), the probability of being sentenced again following a new offense (Constanza et al., 2015), or the crime rate at the local level following the opening of a new semi-liberty facility (Hyatt and Han, 2018).

Therefore, contrary to most previous studies on the effects of semi-liberty on recidivism, our empirical strategy explicitly acknowledges the potential for selection bias and carefully accounts for it. Ideally, we would like to exploit a quasi-experimental setting with similar offenders facing (randomly) different chances of serving their sentence under semi-liberty. However, we fail to identify such a situation in the French context. Alternatively, we rely on a sensitivity analysis to offer credible bounds of the causal effect of SL on recidivism in the presence of confounding factors. Statistical methods to bound causal effects in the presence of selection bias are rapidly developing for linear models (Millimet and Tchernis, 2013; Krauth, 2015; Oster, 2019). In addition to those methods, we attempt to fully exploit the time dimension of recidivism dynamics that can be precisely tracked in our data and apply sensitivity analysis in the context of survival models (Austin, 2014; Lin et al., 2013).

Our analysis is based on original survey data collected from the French Prison Administration. The dataset includes the full criminal records of a representative sample of about 8,000 inmates released in the year 2002, from their first conviction to the year 2008. We limit our attention to the 1,445 sampled offenders who either benefited from or were eligible for front-door SL, and track recidivism during a follow-up period of more than five years. Our empirical analysis seeks to disentangle the influence of selection on observable and unobservable characteristics when measuring the causal effect of SL. For that purpose, we follow the three following steps in our empirical framework. First, we report estimates from non-parametric Cox regressions under the assumption of treatment exogeneity. Second, we consider a selection-on-observables framework and implement entropy balancing matching to calculate the average treatment effect of SL. Third, we evaluate the robustness of our estimates by simulating the impact of confounding factors.

Our main results can be summarized as follows. First, we observe an unadjusted difference of 39% in terms of hazard of recidivism between SL offenders and incarcerated offenders. Second, this gap in hazard reduces in magnitude after controlling for offenders' key individual characteristics in Cox regressions (treatment exogeneity) and after balancing entropy matching (conditional independence), but remains sizeable and statistically significant with estimated reductions in hazard ranging from -22% to -31%. Third, we find that the estimated treatment effect of SL remains negative and significant under credible confounding and only turns insignificant in extreme cases of confounding.

The remainder of our paper is organized as follows. The next section presents the French institutional context and describes our data. The third section presents the various econometric results obtained by Cox regressions and matching algorithms along with sensitivity analyses to measure the robustness of our findings to the presence of unobservable confounders of different intensity. Finally, the fourth section concludes and discusses some public policy implications.

2. Institutional Context and Data

2.1. Institutional Context

In France, semi-liberty is not a proper criminal sentence but a way of serving a prison sentence. According to Article 723-15 of the Criminal Procedure Code, all offenders who are convicted of a short prison sentence (not exceeding one year⁵) and left free at trial (no bench warrant) or whose sentence remaining to be served does not exceed one year, are eligible to converting their pending incarceration into an alternative sanction either under semi-liberty, electronic monitoring or external placement. External placement is a community sanction where the convict is hosted and supervised by a third-party institution (for example, an association). All SL offenders are therefore prison convicts, whether or not this prison sentence has begun to be served⁶.

In this paper, we focus on front-door semi-liberty and do not consider semi-liberty as a back-door, early-release program. Interestingly, the decision to convert prison sentences into SL (or to proceed with incarceration) is typically made by a second, specialized judge (called *Juge de l'application des peines*) who is independent from the first judge in charge of convictions. The second judge therefore automatically receives all new cases of eligible convicts (short prison sentences without bench warrant), with no control over the docket or the initial sentence length, and must decide whether convicted offenders will be incarcerated or obtain SL as a full substitute, for example.

Selection bias is particularly likely to arise because such judicial choices (to incarcerate or not) are inherently qualitative, case-by-case decisions made by professional judges based on potentially rich information. As emphasized earlier, judges may benefit from more information to make their decisions like police investigations, police custody hearings or judicial hearings. This collection of qualitative data provides judges with private information (in the sense that this remains unobserved to researchers) to assess offenders' underlying risks of recidivism and decide on the best-suited sanction. Nevertheless, in our data, we have access to key information such as the full criminal records of convicts, offense types or age.

When semi-liberty is granted, offenders are either transferred to a semi-liberty center (a dedicated building, usually downtown) or a semi-liberty district in the premises of a traditional prison. SL-offenders must return to their room/cell every evening and weekend. However, they may work, receive training or care, meet with their family or fulfill other court-ordered obligations during the day. French halfway houses are therefore quite similar to open prisons as they may exist in Scandinavian countries or in Italy, which have proven effective in preventing recidivism (Mastrobuoni and Terlizzese,

⁵ This period was extended to 2 years by the 2009 Prison Act, but this change does not apply to our data.

⁶ This is different from the American example discussed above. In France, a convict can start SL without having spent any single day in detention (full substitution of the prison sentence).

2018). Another important feature of semi-liberty in France is the salience of incarceration. If the measure fails due to technical violations or new crimes committed while under SL, offenders are very likely to switch to a traditional prison.

2.2. Data Sources and Sample Selection

Our dataset consists of a representative sample of offenders released from French prisons and other custodial institutions (semi-liberty centers) during a period of seven months from June to December 2002 ($N = 8,419$). The survey was assembled by the French Prison Administration through a stratified sampling procedure. A few groups of special offenders were fully sampled (females, juveniles, offenders under parole), whereas other offenders were sampled based on their offense type with probability ranging from 1/1 to 1/16⁷. All our statistical analyses account for the different weights used to draw this stratified sample, making our estimates representative of the population of offenders released in France in June-December 2002. This dataset combines three key strengths for our purpose.

First, the follow-up period reaches five years or more, therefore providing credible estimates of long-term recidivism⁸. Second, our dataset records the exact dates of entry, exit and all new offenses leading to conviction, allowing us to analyze durations (calculated in days) from release to recidivism. Recidivism is established if at least one new conviction (for an offense committed after release) appears in the official, nationwide criminal records when they are retrieved, in early 2008, whatever the new sentence. Hence, we are able to study how semi-liberty affects the instantaneous hazard of recidivism and not simply the probability of recidivism in a given time window.

Third, our data collects the full criminal records of sampled offenders. It includes detailed information on their current criminal sentence (offense type, sentence length, criminal procedure, custodial facility, etc.), as well as other socio-demographic characteristics (gender, age, marital status, education, etc.). It allows us to control a large set of observable covariates that are likely to affect both judges' decisions on whether or not to incarcerate and offenders' intrinsic hazard of recidivism, thus capturing a plausibly large part of the selection bias (based on a selection on observables). A potential shortcoming is that the socio-demographic characteristics are self-assessed by offenders when entering prison (or halfway houses), meaning that they are subject to measurement errors. At the

7 The sampling was stratified with full sampling of female and juvenile releasees, and partial sampling of male releasees based on their offense type. For example, the sampling probability was 1 for offenders convicted of murder and rape, 0.5 for embezzlement and breach of trust, 0.2 for theft and 0.1 for driving under the influence.

8 The few international studies collecting recidivism data for longer periods show that most re-offenses occur in the first two years: a 5-year follow-up is likely to capture 90% of lifetime recidivism. For further details, see <https://www.bjs.gov/content/pub/pdf/18upr9yfup0514.pdf>

same time, it is important to keep in mind that judges' decision to incarcerate or divert relies only on pre-entry (and often self-declared) characteristics.

We proceed with several sample restrictions when turning to the data. First, since we focus on convicts who have either served their sentence entirely in prison or entirely in semi-liberty, we exclude from the original dataset offenders who were not legally eligible to full semi-liberty. This concerns convicts with an initial prison sentence of more than one year, offenders whose custody started before or on day of conviction (pre-trial detention, bench warrant), and offenders who obtained some other front-door alternatives (like electronic monitoring). Second, we exclude offenders who had virtually no chance in practice of obtaining SL, based on their observable characteristics. This concerns homeless people and offenders with more than five previous criminal convictions. This second restriction helps achieving balance between the two groups. Third, we exclude the few offenders whose key information was not available (in particular criminal record and sentence length)⁹.

Overall, these restrictions lead to a study sample of 1,445 offenders, all eligible for SL, and who served under two different regimes: semi-liberty ($N = 353$) or incarceration ($N = 1,092$).

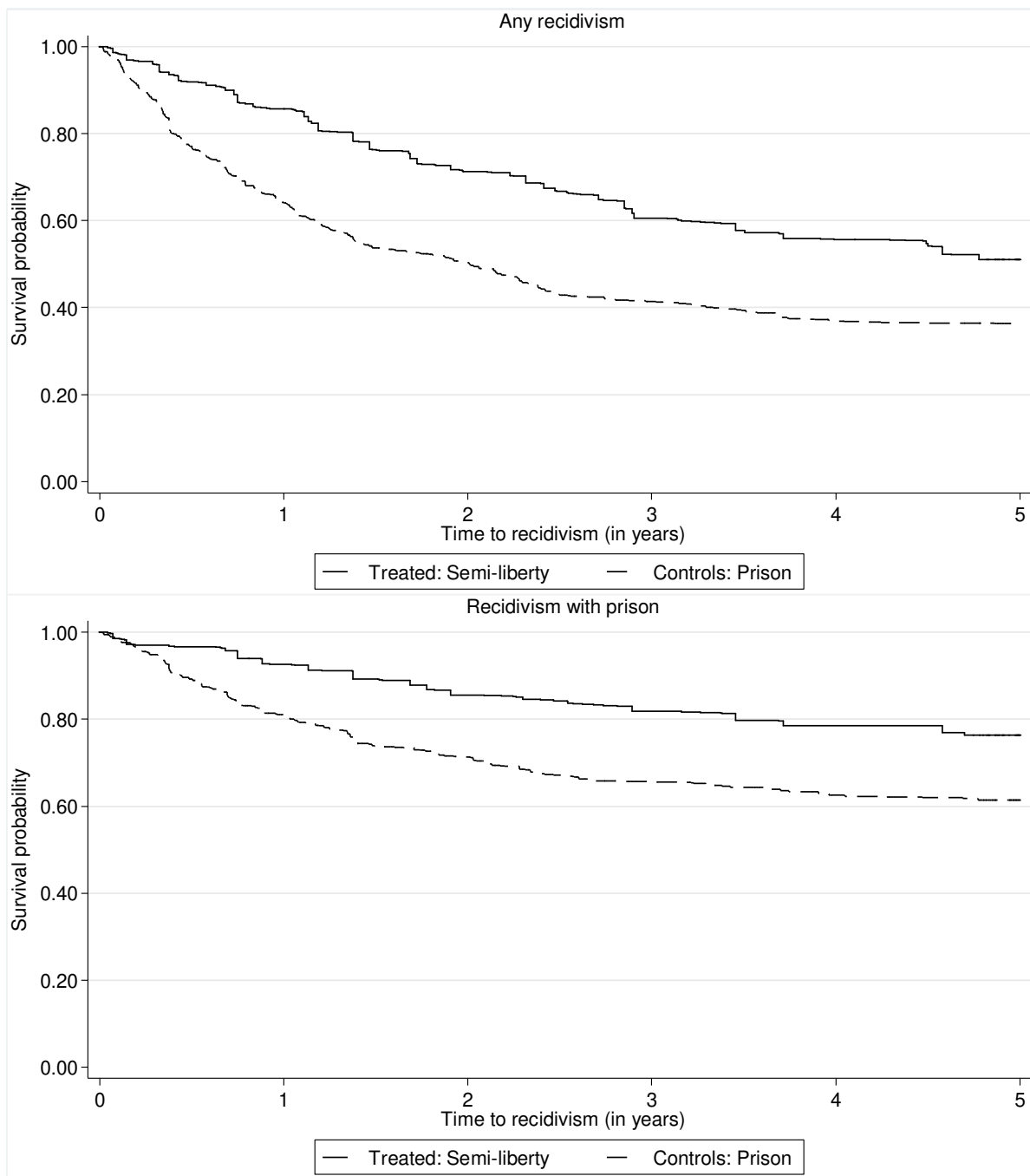
2.3. Descriptive Statistics

In this paper, we consider two different outcomes related to recidivism. First, we measure the occurrence and duration of any recidivism, defined as any new crime leading to conviction committed after the sanction under study, whatever the type of new sentence. Second, we measure the occurrence and duration of recidivism leading to prison, defined as any new crime leading to conviction to a prison sentence, excluding fully suspended prison sentences (Monnery, 2015; Henneguelle and Monnery, 2017). Durations are measured from the day of release to the day of commission of the new crime.

In our sample, the proportion of offenders having reoffended is equal to 60.7%. The data shows substantial differences between SL and non-SL offenders: recidivism is much less likely among SL offenders, with a gap of 14.6 percentage points (49.0% versus 63.6%). Results are similar when considering recidivism leading to prison, which is 14.9 percentage points lower among SL offenders compared to non-SL offenders (23.7% versus 38.6%).

⁹ For example, criminal records could not be retrieved for offenders who died during the follow-up period.

Figure 2. Survival functions for SL and non-SL offenders



Source: data from French Prison Administration, authors' calculations.

In Figure 2, we provide a graphical representation of recidivism dynamics in the two groups using the non-parametric Kaplan-Meier estimator. Panel A plots the survival rate over time, which corresponds to the share of offenders who have not yet reoffended. Five years after release, the survival rate is equal to 51.0% among SL-offenders compared to 36.4% among non-SL offenders. Figure 2 shows large differences in the prevalence of recidivism between the two groups at any point in time: 21.7 points at one year, 20.8 points at two years, and 19.2 points at three years. Panel B plots the same curve for

recidivism leading to a new prison conviction. Again, there is substantial difference in survival rates between the two groups. The prevalence in recidivism is higher among non-SL offenders: 11.8 points at one year (80.7% versus 92.5%), 14.2 points at two years (71.3% versus 85.5%) and 16.2 points at three years (65.6% versus 81.8%).

These non-parametric survival functions provide consistent evidence that SL-offenders display lower instantaneous risk of recidivism over time compared to non-SL offenders. Furthermore, Cox-regression-based tests indicate that the survival curves are significantly different between the two groups¹⁰. Interestingly, for both outcomes, most of the gap in survival rates between treated and untreated offenders emerges in the first year of follow-up. However, these differences do not imply any causal effect of SL, as both observable and unobservable characteristics may affect the decision to benefit from SL.

Table 1 reports descriptive statistics for the full sample (N=1,445), treated group (SL offenders) and control group (non-SL offenders) separately. The sample is mainly composed of male offenders (96.4%) with French citizenship (91.2%). Around one in four offenders is married and 74.3% have higher education (defined as having some high-school education or more). In around two in three cases, offenders have already experienced at least one past conviction.

When comparing the treated and untreated offenders, we observe several large and significant differences. In particular, SL-offenders are more educated on average (85.9% have higher education versus 71.5% among non-SL offenders) and they are more likely to be first-time offenders (47.7% versus 29.1%). They have received lengthier prison sentences on average, arguably so that semi-liberty has enough time to produce its rehabilitative effects¹¹. They have also committed different types of crimes, as they are more often involved in drug-related crimes (14.9% versus 8.7%), but less often in thefts and concealments (16.8% versus 28.4%).

To summarize, our descriptive analysis shows that SL-offenders are less concerned by recidivism than non-SL-offenders. However, the gap in the survival curves may be explained, either partly or fully, by differences in characteristics between treated and untreated offenders. We now turn to an econometric analysis to account for those composition effects on the hazard of recidivism and then discuss the potential influence of unobservable confounders on our results.

10 For any recidivism, the Wald statistic is equal to 14.33 with $p=0.000$. The Wald statistic is 9.44 ($p=0.002$) when considering recidivism leading to prison.

11 As SL requires some time to organize, the judges avoid SL for prison sentences of one or two months. Instead, they prefer sentences that give the convicted person time to adapt to the center and take advantage of it in terms of training or future employment.

Table 1. Description of the sample

Variables	All	Treated	Control	Difference
Female	0.036	0.033	0.037	-0.004
Foreigner	0.098	0.078	0.102	-0.024
Married	0.273	0.325	0.261	0.064
Parent	0.389	0.446	0.375	0.071
High education	0.743	0.859	0.715	0.144***
Age at entry	30.306	31.102	30.111	0.991
Sentence length	4.308	5.357	4.051	1.306***
Past convictions				
0	0.328	0.477	0.291	0.186***
1	0.244	0.203	0.254	-0.051
2	0.179	0.154	0.185	-0.031
3	0.098	0.066	0.106	-0.040
4	0.100	0.076	0.106	-0.030
5	0.051	0.024	0.058	-0.034**
Thefts – concealments	0.261	0.168	0.284	-0.116***
Degradations – economic offenses	0.098	0.099	0.097	0.002
Driving offenses	0.305	0.311	0.303	0.008
Violence – threats	0.143	0.174	0.136	0.038
Narcotic drugs	0.099	0.149	0.087	0.062**
Administrative offenses – others	0.095	0.100	0.093	0.007
Number of observations	1,445	353	1,092	

Source: data from French Prison Administration, authors' calculations.

Note: significance levels for the mean-comparison tests are 1% (***) , 5% (**) and 10% (*).

3. Econometric Results

3.1. Estimates from Cox Regressions

We rely on semi-parametric Cox regression models to provide benchmark estimates of the relationship between SL treatment and hazard of recidivism net of the influence of observable characteristics, assuming away selection on unobservables. Cox regressions write as follows:

$$h_i(t) = h_0(t) * \exp(\delta * SL_i + X_i\beta) \quad (1)$$

where $h_i(t)$ represents the hazard of recidivism at time t for individual i , $h_0(t)$ is the baseline hazard which is left unparametrized, δ measures the influence of the treatment SL_i , and X_i is a set of control variables with β being the associated vector of coefficients. We estimate separate regressions for our two different outcomes, i.e., any recidivism and recidivism with prison. Also, we consider different sets of explanatory variables. In some specifications, we account for the influence of offenders' main individual characteristics: gender, age at entry, citizenship, marital status, parenthood, educational level, current sentence length, and number of prior convictions. In other specifications, we further control the type of current offense with a distinction in six categories, administrative and other offenses being the reference category.

We report our results as odd ratios in Table 2. Without any control variables, we find that SL offenders face a 39.0% lower instantaneous hazard of recidivism compared to non-SL offenders (column 1) and a 46.4% lower instantaneous hazard of reconviction to prison (column 4). Interestingly,

the estimated hazard differentials remain large and significant after controlling for individual characteristics, albeit lower in terms of intensity: the change in estimated hazard ratios is around 20-25%. On average, SL-offenders face a 31.0% lower hazard of recidivism than observably-similar non-SL offenders (column 2) and a 34.8% lower hazard of reconviction to prison (column 5). Accounting for the type of offense has almost no impact on the estimated coefficients (columns 3 and 6).

Table 2. Cox proportional hazard estimates of recidivism

Variables	Any recidivism			Recidivism leading to prison		
	(1)	(2)	(3)	(4)	(5)	(6)
SL treatment	0.610*** (-3.79)	0.690*** (-2.59)	0.684*** (-2.77)	0.536*** (-3.07)	0.652** (-2.11)	0.665** (-2.00)
Female		0.683*** (-2.83)	0.714** (-2.42)		0.521*** (-3.24)	0.518*** (-3.17)
Foreigner		0.845 (-0.91)	0.820 (-1.05)		0.715 (-1.29)	0.694 (-1.38)
Married		0.973 (-0.19)	1.021 (0.14)		0.913 (-0.44)	0.941 (-0.29)
Parent		1.254 (1.56)	1.223 (1.41)		1.185 (0.82)	1.166 (0.72)
High education		0.750** (-2.26)	0.750** (-2.32)		0.781 (-1.50)	0.798 (-1.37)
Age at entry		0.941*** (-6.98)	0.936*** (-7.87)		0.949*** (-4.84)	0.950*** (-4.58)
Sentence length		0.950** (-2.40)	0.961* (-1.80)		0.975 (-1.00)	0.972 (-1.04)
Past convictions (reference: 0)	1	1.197 (1.17)	1.188 (1.13)		1.364 (1.49)	1.358 (1.47)
	2	1.537** (2.45)	1.431** (2.02)		1.289 (1.04)	1.261 (0.91)
	3	2.354*** (4.50)	2.223*** (4.17)		2.378*** (3.17)	2.330*** (3.08)
	4	2.348*** (4.46)	2.250*** (4.33)		3.215*** (5.03)	3.203*** (4.88)
	5	2.451*** (3.82)	2.148*** (2.98)		2.704*** (3.21)	2.639*** (3.10)
Thefts – concealments			0.743 (-1.36)			1.116 (0.40)
Degradations – economic offenses			0.914 (-0.36)			0.671 (-1.04)
Driving offenses			1.048 (0.25)			0.928 (-0.28)
Violence – threats			0.973 (-0.12)			1.232 (0.74)
Narcotic drugs			0.423*** (-2.97)			0.710 (-0.95)
Number of observations	1,445	1,445	1,445	1,445	1,445	1,445
Number of failures	813	813	813	426	426	426
Number of censored obs.	632	632	632	1019	1019	1019

Source: data from French Prison Administration, authors' calculations.

Note: estimates in exponential form (hazard ratios) from Cox proportional hazard models, with robust standard errors. Significance levels are 1% (***), 5% (**) and 10% (*).

The fact that the various individual characteristics reduce the raw hazard differentials between the two groups is consistent with judges selecting SL-offenders based on favorable observable characteristics. The various explanatory variables tend to have the expected effects. On average, female offenders experience a 32% lower hazard of recidivism and a 48% lower hazard of reconviction to prison than male offenders. Educated and older offenders also face lower hazard of recidivism. For instance, each additional year of age at entry is associated with a 5-6% reduction in the post-release hazard of reconviction. Offenders with lengthier sentences also face lower hazard and hazard significantly increases with the number of past convictions. On average, offenders with five previous convictions exhibit a hazard of recidivism that is at least twice as high as first-time convicts (the reference category). Similar findings were found for other cohorts of convicts in France (Josnin, 2013; Monnery, 2015).

3.2. Selection on Observables and Matching

Matching offers an alternative to Cox regressions in that it is more flexible and relies directly on comparisons between finely-defined pairs of observably-similar offenders. However, matching is a method which accounts for selection on observables, meaning that the distribution of potential outcomes is independent from treatment status within matched pairs (Conditional Independence Assumption). In order to build credible matches from the two groups, it is crucial to rely on information that judges use in practice to decide whether offenders are to be granted SL or not. Our detailed administrative data, with full criminal records and socio-demographic characteristics, is likely to reflect most of the critical information available to judges.

Different methods of matching exist in order to account for selection on observables (see Stuart, 2010, Austin, 2014). Matched pairs can be constructed from exact cells in terms of selected characteristics (exact matching), but this method is data-demanding and even impractical with continuous variables like age or sentence length and limited sample size. Treated and untreated offenders can also be matched by their distance on all covariates using Mahalanobis distance measures, which corresponds to nearest-neighbor matching. The propensity-score method consists of first predicting individual propensity scores based on a Logit regression, then matching treated and untreated individuals with close propensity scores (Rosenbaum and Rubin, 1983). Conditional on the propensity score, the selected covariates are expected to be balanced between treated and controls, but some imbalance can still exist if the propensity score is misspecified.

In what follows, we focus on another matching method called entropy balancing (Hainmueller, 2012). Entropy balancing matching is a data preprocessing procedure leading to perfectly balanced samples by construction. Some weights are calculated and assigned to each observation such that the covariate distributions in the reweighted dataset fulfill a set of moment conditions. Denoting by Y the

outcome under consideration, the average treatment effect on the treated in the general case is $E(Y_1|SL = 1) - E(Y_0|SL = 1)$: the second term is the counterfactual mean and remains unobserved. Under entropy balancing, the counterfactual mean is calculated as $E(Y_0|\widehat{SL} = 1) = \sum_{i|SL=0} Y_0 * w_i / \sum_{i|SL=0} w_i$ with w_i an entropy balancing weight¹². Hainmueller (2012, pp. 30-33) shows that the weights are obtained by minimizing an entropy distance metric (the directed Kullback entropy divergence) subject to both balance and normalizing constraints. A unique feature of the entropy balancing scheme is that it avoids the multiple iterations needed to find an appropriate specification of the propensity score that satisfies the balancing of the treated and control groups.

Here, data for controls (non-SL offenders) will be reweighted so as to match the first two moments, i.e. means and variance, of the treated (SL offenders). The entropy balancing method therefore ensures that all standardized biases associated with differences in characteristics between treated and control offenders go to virtually zero by design. Thus, it allows us to compare recidivism in two groups that are perfectly similar in terms of observable characteristics, both in means and variance. We consider the following covariates to reweight the control group of observations: gender, foreigner, married, parent, higher education, age at entry, sentence length, number of past convictions and type of offense.

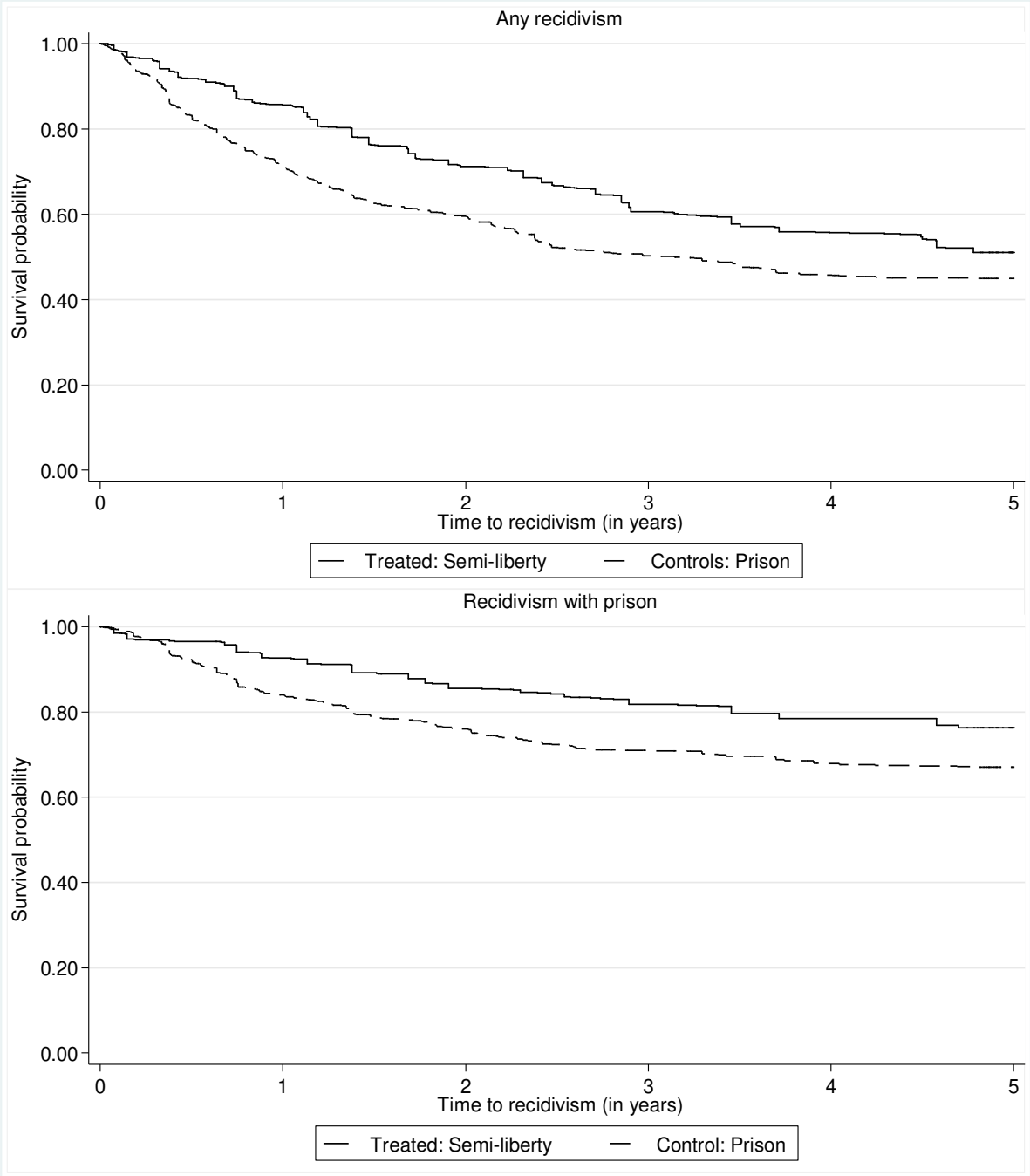
Figure 3 shows the survival functions for SL and non-SL offenders after entropy matching. Compared to the Kaplan-Meier curves presented in Figure 2, we note that the survival differentials between SL and non-SL offenders are smaller for both outcomes. However, the survival rates still remain higher among SL-offenders over the five-years period. Cox regressions on the reweighted sample show that compared to non-SL offenders, the hazard of recidivism among SL offenders is 22.2% lower for any recidivism and 34.8% lower for serious recidivism in the following five years. Still, the gap in recidivism that is observed after five years is essentially achieved within the first year after the sanction. So, even with perfectly identical characteristics on average, non SL-offenders are still more likely to be affected by recidivism.

As robustness checks, we have also implemented several other matching techniques in which the outcomes are the probability of recidivism and of recidivism leading to prison within one, three and five years, respectively. Specifically, we consider seven different techniques: nearest neighbor matching with one, two and three matches, propensity score matching with one, two and three matches (the caliper being set to 0.04), and inverse probability weighting. Overall, the various estimates of the average treatment effect are substantively similar and suggest significantly lower risk

12 The counterfactual mean under propensity score is $E(Y_0|\widehat{SL} = 1) = \sum_{i|SL=0} Y_0 * d_i / \sum_{i|SL=0} d_i$ with the weight d_i being a function of the estimated propensity score \hat{p}_i such that $d_i = \hat{p}_i / (1 - \hat{p}_i)$.

of recidivism among SL-offenders¹³. The average ATT obtained from all our matching techniques is -17.7 points at one year, -11.8 points at three years and -8.7 points at five years.

Figure 3. Marginal effect of SL on hazard rate of recidivism after entropy balancing



Source: data from French Prison Administration, authors' calculations.

13 Detailed results from other matching methods are available upon request. When estimating the propensity score, we include the following covariates : female, foreigner, married, parent, high education, age at entry, sentence length, number of past convictions, and type of offence.

3.3. Selection on Unobservables

3.3.1. Coefficient Stability

A central concern is to assess to what extent our previous results are affected by potentially important omitted variables. As the effect of SL remains rather stable (albeit slightly lower) after inclusion of a large set of controls, it could be first argued that the omitted variable bias is somewhat limited. Over the last years, a few studies have suggested that the relationship between treatment and unobservables could be inferred from the relationship between treatment and observables (see Altonji et al., 2005, 2008). However, as recently demonstrated in Oster (2019), focusing on coefficient stability is not sufficient and movements in explaining the variance of the outcome have also to be taken into account. In the case of linear models, Oster (2019) develops a new estimator based on both coefficient and R^2 movements which provides an identified set for the treatment effect.

Here, we apply this methodology specifically designed for linear models albeit we do not have any continuous outcome, so that our results should be interpreted with caution. Specifically, we turn to linear probability models and investigate to what extent unobservables may influence the effect of SL on the probability of recidivism within one, three and five years. In Oster (2019), there are two key parameters in order to gauge sensitivity to confounding : the ratio of unobservable versus observable selection into treatment (denoted by δ) and the maximum explanatory power that could be achieved if all confounders were observed (denoted by R_{max}). Different estimates for the effect of SL are obtained depending on the selected values for δ and R_{max} . While lower bounds for those parameters are straightforward to choose ($\delta = 0$ and R_{max} is set to the current R^2), the choice of the upper bounds appears more difficult.

First, regarding the possible values of R_{max} , we believe that the maximum R^2 of any fully-controlled regression between recidivism and SL should be well below 1 and probably closer to 0.5. The main arguments are the relatively hard-to-predict nature of criminal behavior, the potential for mismeasurement in recidivism (mostly type-2 errors) and the fact that many life events may occur after the treatment and thus affect recidivism. In what follows, we follow the suggestion of Oster (2019) and start with a multiplication factor of 1.3 applied to the actual R^2 obtained from a regression with extended controls¹⁴. Second, regarding the ratio of unobservable versus observable selection (δ), we note that our set of explanatory variables is very close to what judges typically observe and decide on. In particular, we control for previous convictions, age and type of offence. As a consequence, we consider values ranging between 0 and 1 for the parameter δ .

¹⁴ Oster (2019) relies on experimental data to obtain a sensible benchmark for R_{max} .

Table 3. Effect of SL using Oster (2019) sensitivity estimator

R_{max}	Any recidivism					Recidivism leading to prison				
	$\delta = 0$	$\delta = 0.25$	$\delta = 0.50$	$\delta = 0.75$	$\delta = 1.00$	$\delta = 0$	$\delta = 0.25$	$\delta = 0.50$	$\delta = 0.75$	$\delta = 1.00$
<i>Panel A. Recidivism within one year</i>										
Actual R^2	-0.162	-0.162	-0.161	-0.160	-0.159	-0.088	-0.087	-0.086	-0.086	-0.085
1.3*actual R^2	-0.162	-0.156	-0.149	-0.143	-0.136	-0.088	-0.085	-0.083	-0.080	-0.078
1.5*actual R^2	-0.162	-0.153	-0.144	-0.134	-0.124	-0.088	-0.084	-0.079	-0.075	-0.071
2*actual R^2	-0.162	-0.143	-0.123	-0.103	-0.081	-0.088	-0.079	-0.070	-0.061	-0.051
3*actual R^2	-0.162	-0.123	-0.081	-0.036	0.016	-0.088	-0.070	-0.051	-0.031	-0.008
$R^2=1$	-0.162	-0.033	0.114	0.297	0.569	-0.088	0.000	0.116	0.276	0.560
<i>Panel B. Recidivism within three years</i>										
Actual R^2	-0.123	-0.123	-0.123	-0.123	-0.123	-0.113	-0.112	-0.112	-0.111	-0.111
1.3*actual R^2	-0.123	-0.117	-0.111	-0.104	-0.097	-0.113	-0.109	-0.104	-0.100	-0.096
1.5*actual R^2	-0.123	-0.113	-0.102	-0.091	-0.080	-0.113	-0.105	-0.097	-0.089	-0.080
2*actual R^2	-0.123	-0.102	-0.080	-0.057	-0.033	-0.113	-0.097	-0.081	-0.065	-0.047
3*actual R^2	-0.123	-0.081	-0.035	0.014	0.068	-0.113	-0.081	-0.047	-0.010	0.031
$R^2=1$	-0.123	-0.036	0.062	0.175	0.315	-0.113	0.003	0.139	0.308	0.561
<i>Panel C. Recidivism within five years</i>										
Actual R^2	-0.077	-0.077	-0.077	-0.076	-0.076	-0.092	-0.090	-0.089	-0.087	-0.086
1.3*actual R^2	-0.077	-0.071	-0.066	-0.060	-0.054	-0.092	-0.086	-0.080	-0.074	-0.068
1.5*actual R^2	-0.077	-0.068	-0.058	-0.048	-0.038	-0.092	-0.083	-0.074	-0.065	-0.056
2*actual R^2	-0.077	-0.058	-0.037	-0.016	0.006	-0.092	-0.074	-0.056	-0.038	-0.018
3*actual R^2	-0.077	-0.037	0.006	0.052	0.103	-0.092	-0.056	-0.020	0.019	0.063
$R^2=1$	-0.077	-0.003	0.080	0.173	0.286	-0.092	0.041	0.191	0.375	0.644

Source: data from French Prison Administration, authors' calculations.

Note: estimates are average treatment effects (ATE) from linear probability models under different combinations of the parameters δ and R_{max} following the methodology described in Oster (2019).

We present our results for any recidivism and recidivism leading to prison in Table 3. Overall, this sensitivity analysis lead supports to the hypothesis of a crime-reducing effect of SL on any recidivism under moderate confounding. Within an horizon of one year, the point estimate of any recidivism which is -0.162 under exogeneity decreases very slightly (in absolute value) for scenarios where R_{max} is up to 2 times the actual R^2 (0.14) whatever the importance of unobservables. A very similar pattern is found for recidivism with three years and five years. It is only under strong confounding, for instance when δ is at least 0.75 and the parameter R_{max} is three times the actual R^2 that the point estimates for SL become positive in our linear probability model.

Overall, these sensitivity results fare relatively well compared to non-experimental papers surveyed by Oster (2019). Nevertheless, a potential shortcoming is that the recidivism outcome is not continuous so that we rely implicitly on a pseudo R^2 to set the different values of R_{max} . As both our outcome and the treatment are binary variables, the bivariate Probit model used in Altonji et al. (2008) could be an option to assess the role of unobservables, by setting different values for the coefficient of correlation between residuals. A difficulty here is the lack of exclusion restrictions, which would consist in the inclusion of explanatory variables having an influence on the treatment only, but not on

the outcome¹⁵. So, in the next two subsections, we turn to two complementary methods providing evidence on the influence of confounding variable within a survival regression framework.

3.3.2. Simultaneous Sensitivity Analysis

In propensity score analyses, the underlying assumption is that there is no unmeasured confounding variable (Rosenbaum and Rubin, 1984). This means that all variables expected to influence selection into treatment are indeed taken into account. As it is not possible to assess the relevance of this assumption within a selection on observables framework, we further consider the potential influence of unmeasured confounding variables.

Rosenbaum and Rubin (1984) were the first authors to propose a framework allowing the estimation of an average treatment effect on a binary outcome net of the influence of observed variables and an unobserved (binary) variable. Rosenbaum (2002, 2005) proposes a bounding approach to calculate ATE in the presence of unobserved heterogeneity (also called hidden bias) between treatment and control cases. While in a randomized experiment individuals face the same chance of being treated, they are supposed to differ in terms of unobservables and have thus different probabilities to receive the treatment. Different intensities of hidden bias are considered to assess to what extent the treatment effect remains significant. In what follows, we rely on the sensitivity analysis specifically designed for survival outcomes described in Liu et al. (2013) and Austin (2014)¹⁶.

For the presentation, let u be an unobserved confounder. The associations between the unobserved confounder u and the treatment status SL or between the unobserved confounder u and the outcome R (R corresponding to time to recidivism) are potential sources of bias which may affect the impact of the treatment status SL on the outcome R . The magnitudes of these associations are given by the odds ratios denoted by $OR_{SL,u}$ and $OR_{R,u}$, respectively. The purpose of the sensitivity analysis is to find the magnitude of the odds ratios that lead to insignificant effects of SL on R , once the influence of observable characteristics is controlled for using matching techniques. As described in Gastwirth et al. (1998), three types of sensitivity analyses must be distinguished. The primal sensitivity analysis examines the impact of $OR_{R,u}$, while the dual sensitivity analysis considers the impact of $OR_{SL,u}$. In the simultaneous sensitivity analysis, both $OR_{R,u}$ and $OR_{SL,u}$ are allowed to vary. We briefly

15 When setting the coefficient of correlation to zero (i.e the treatment is supposed to be exogenous) in the bivariate Probit model, we obtain a coefficient of -0.55 for SL in the recidivism equation (any recidivism within one year). This coefficient increases in absolute value when the coefficient of correlation increases (it is equal to -1.30 when the correlation is 0.5), and becomes positive when the correlation is -0.32 or lower.

16 For the sake of robustness, we have also implemented the sensitivity analysis described in Rosenbaum (2002). We turn to one-nearest neighbor matching and consider values ranging between 1 (no hidden bias) and 2 (the odds of being treated is doubled) for the hidden bias. We find that the negative effect of SL remains statistically significant when the odds of differential assignment due to unobserved factor is increased by 140%.

present the primal sensitivity analysis and then describe the simultaneous sensitivity analysis that we implement in our paper.

Suppose first that $OR_{SL,u}$ is bounded by Γ . Let $p^+ = \Gamma/(1 + \Gamma)$ and $p^- = 1/(1 + \Gamma)$ be the upper and lower bounds of the probability of being exposed, accounting for the influence of the unobservable confounder. Denote by T the number of discordant pairs defined as those in which the outcomes are different within the matched pairs, and a be the number of discordant pairs defined as those in which the treated had an outcome (and the untreated did not). The association between SL and u affects the lower bound and upper-bound p-values LB and UB which are given by¹⁷:

$$LB = \sum_a^T \binom{T}{a} (p^-)^a (1 - p^-)^{T-a} \quad (2)$$

$$UB = \sum_a^T \binom{T}{a} (p^+)^a (1 - p^+)^{T-a} \quad (3)$$

The purpose of the sensitivity analysis is to find how the values of Γ affect the lower bound LB and the upper-bound UB . When the upper bound UB is higher than 0.05 (or 0.1, corresponding to the 10 percent level), this means that an unmeasured confounder increasing the odds of treatment by Γ renders the effect of semi-liberty on recidivism insignificant. The dual sensitivity analysis is symmetric to the primal sensitivity analysis except that the parameter of interest is now $OR_{R,u}$, which is bounded by Δ . In the simultaneous sensitivity analysis, both $OR_{SL,u}$ and $OR_{R,u}$ (bounded by Γ and Δ , respectively) are allowed to vary. So the objective is to find combinations of values for Γ and Δ such that the impact of the treatment on the outcome net of the influence of both observable characteristics and the unobservable confounder ceases to be significant. Let $p_\Gamma = \frac{\Gamma}{1+\Gamma}$ and $p_\Delta = \frac{\Delta}{1+\Delta}$. Then, the upper bound of the probability of being treated given the unobservable confounder u is:

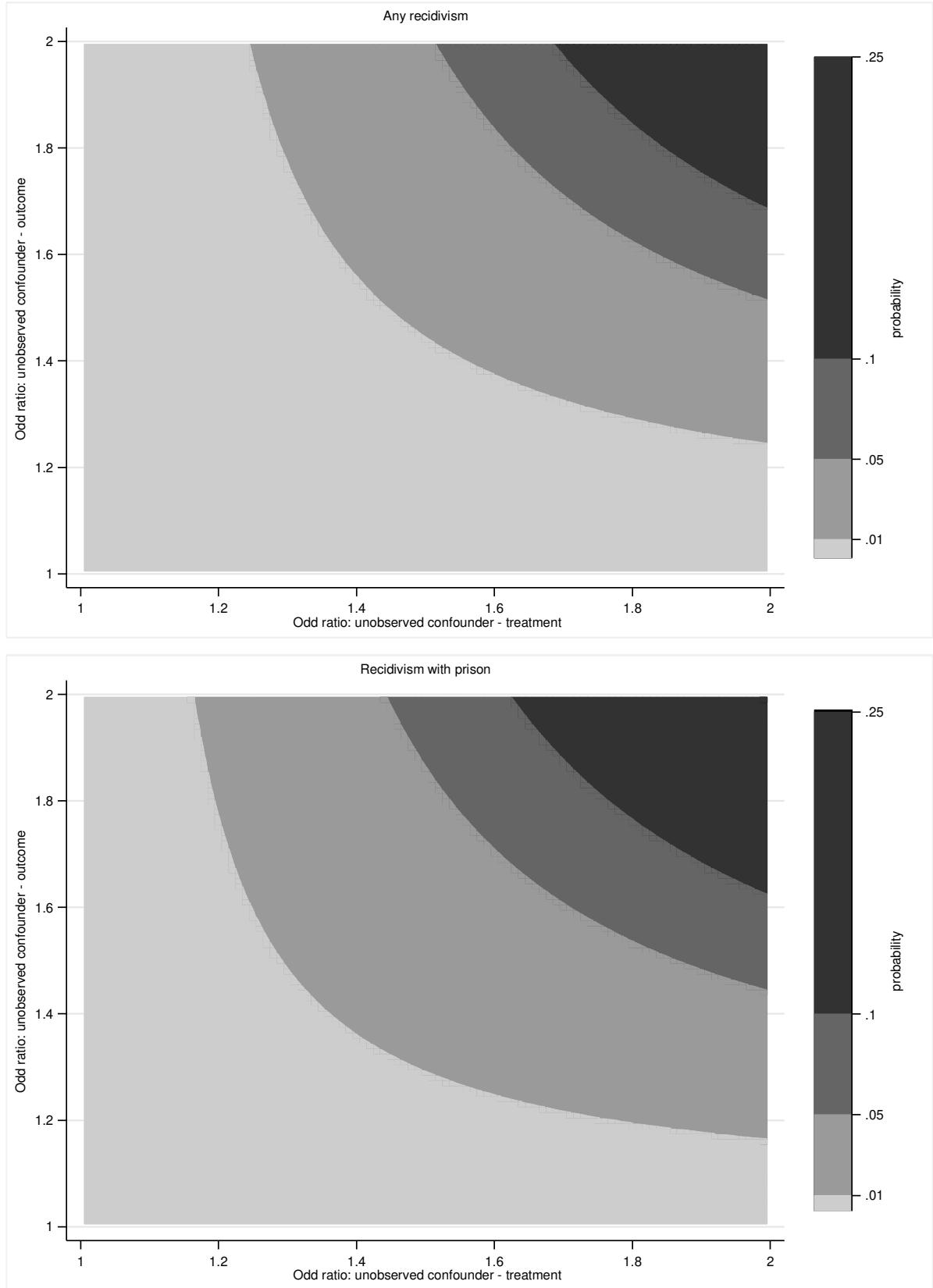
$$p^+ = p_\Gamma * p_\Delta + (1 - p_\Gamma) * (1 - p_\Delta) \quad (4)$$

If the upper bound p^+ exceeds 0.05 for low values of (Γ, Δ) , then this means that the causal effect of the treatment on the outcome is very sensitive to the assumption of no unobservable confounder. In Figure 6, we present results from a simultaneous sensitivity analysis to unmeasured confounders¹⁸.

17 For further details, see Liu et al. (2013). This corresponds to the McNemar (1947) exact test.

18 When turning to the data, we rely on propensity score matching and match each treated individual with exactly one control individual.

Figure 6. Effect of unobserved confounder in a simultaneous sensitivity analysis



Source: data from French Prison Administration, authors' calculations.

Suppose for instance that $\Gamma = 1.5$ and $\Delta = 1.5$, corresponding to a scenario in which an unmeasured confounder increases the odds of treatment by 50% and the odds of outcome by 50%. In such case, we find that the effect of semi-liberty on the risk of any recidivism would remain statistically significant (at the 5 percent level) if one were able to take such unmeasured confounder into account. The upper part of Figure 6 suggests that the influence of the unmeasured confounder must be very high. With $\Gamma = 1.7$ and $\Delta = 1.7$, the treatment effect on any recidivism is no longer significant at the 10 percent level. The sensitivity to unmeasured confounder is slightly more significant when considering recidivism with prison. With $\Gamma = 1.4$ and $\Delta = 1.4$, the effect of semi-liberty on recidivism remains significant. However, when the unmeasured confounder increases the odds of treatment (or the odds of outcome) by at least 70%, then the treatment effect is no longer statistically significant at the 10 percent level.

3.3.3. Simulations of Confounder

Another sensitivity test recently proposed by Lin et al. (2013) is directly based on Cox regressions instead of matching. This alternative has the advantage of exploiting richer information on durations to recidivism and the full study sample compared to the exact test used by Austin (2014), which only uses ranks of durations within matched pairs. Specifically, we estimate an augmented version of the non-parametric Cox model given by Equation (1):

$$h_i(t) = h_0(t) * \exp(\delta * SL_i + \alpha * \tilde{C}_i + X_i\beta) \quad (5)$$

where the new variable \tilde{C} is the unobservable confounder that we simulate. We consider different credible parameters to simulate \tilde{C}_i , allowing associations between \tilde{C}_i and both treatment SL_i and hazard of recidivism $h_i(t)$.

The method consists of successively generating different values of \tilde{C}_i with varying associations with SL and h , and examine how the inclusion of \tilde{C} in Equation (5) affects δ . More precisely, we follow three steps that are repeated sequentially for different parameters (μ, α) . First, we draw random values of \tilde{C}_i from normal distributions. For the treated group (SL offenders), we always use a standardized normal distribution $\tilde{C}_i|X \sim N(0; 1)$, where X is our vector of observable control variables. For the control group (incarcerated offenders), we shift the mean value of our normal distribution such that $\tilde{C}_i|X \sim N(\mu; 1)$, with $\mu > 0$, in order to generate mean differences in the unobservable variable \tilde{C} between the two groups (independently from differences in observables). We successively consider values for $\mu \in [0, \dots, 0.5]$ by 0.1 increments, so that $\mu = 0$ represents no systematic difference in \tilde{C} between the two groups (balanced) and $\mu = 0.5$ yields a gap equal to 50 % of a standard deviation.

Second, we set a credible constraint on the parameter α indicating the direct effect of \tilde{C} on h to estimate Equation (5), using the generated values \tilde{C} . We successively consider $\alpha \in [0, \dots, 0.3]$ by 0.1

increments, where $\alpha = 0$ implies no confounding and $\alpha = 0.3$ a very large confounding, and then keep the estimated coefficient δ from each regression. Third, we repeat step 1 (random draw) and step 2 (constrained survival regression) 500 times for each of the 24 sets of parameters (μ, α) and compute the mean of all δ to obtain an estimate of the average treatment effect, ATE , for the 24 sets of parameters. Standard errors for the treatment effects, se_{ATE} , are computed as the square root of a sum of the within-imputation variance (se_W^2) and the between-imputation variance (se_B^2)¹⁹.

To select the range values for α and μ , we consider the two following extreme scenarios. In the first one, no selection on unobservables occurs. This may happen either because the vector of control variables X captures all the information used by judges to make SL decisions (so that \tilde{C} is balanced between the two groups, i.e., $\mu = 0$) or because the unobservable variables that judges additionally use are not related to hazard of recidivism ($\alpha = 0$). In this scenario, the set of low-range parameters $(\underline{\mu}, \underline{\alpha})$ is simply $(0,0)$. In the second scenario, very strong confounding occurs. We assume that the confounder variable \tilde{C} has similar characteristics as the strongest observable predictor of treatment and hazard, namely the number of past convictions. Past convictions are a direct measure of past recidivism, which explains their potential in predicting future recidivism. Empirically, past convictions have long been found to be the single best predictor of recidivism in the literature.

Likewise, in our sample the number of past convictions strongly predicts offenders' hazard of recidivism after release (see Table 2). A Cox regression of the hazard of any recidivism on the same set of control variables X as in Column 3 of Table 2, but with past convictions as a standardized variable (instead of a set of dummy variables), yields a coefficient of 0.288 (with a t-value of 5.10) corresponding to a hazard ratio of 1.33. On average, a one-standard deviation increase in past convictions is associated with a 33% increase in hazard of recidivism among observably-similar offenders. We take this estimated coefficient of 0.288 for standardized past convictions as the largest credible magnitude for the direct effect of \tilde{C} on h (thus $\bar{\alpha} = 0.3$). It seems implausible to imagine a variable (observed by judges) whose effect on hazard of recidivism net of the influence of controls would be larger than that of past convictions.

Regarding the credible maximum magnitude for μ (which is the difference in the average between treated and untreated offenders), past convictions may again serve as a useful benchmark since this variable is well-coded in judicial files and very relevant both in the law and in practice when making SL decisions. In our sample, SL offenders accumulate an average of 1.13 past convictions compared to 1.65 for non-SL offenders. This difference of about 0.5 past convictions is very significant

19 Formally, $se_W^2 = \frac{1}{500} \sum_{r=1}^{500} se_r^2$, $se_B^2 = \frac{1}{500-1} \sum_{r=1}^{500} (\beta_r - ATE)^2$ and $se_{ATE} = \sqrt{se_W^2 + \left(1 + \frac{1}{500}\right) * se_B^2}$. For further details, see Lin et al. (2013).

($F = 13.22$, $p\text{-value} = 0.000$) and corresponds to 34% of a standard deviation ($sd = 1.52$)²⁰. In what follows, we allow for more selection on \tilde{C} than what may be used by judges on past convictions since we consider values for μ up to 0.5 of a standard deviation²¹. For this scenario of very strong confounding, the set of high-range parameters $(\bar{\mu}, \bar{\alpha})$ is therefore (0.5,0.3).

The results of our simulations for the two outcomes (any recidivism and recidivism leading to prison) are reported in Table 4. For each combination (μ, α) , we report the simulated effect of SL on hazard of recidivism, standard errors as well as 95% confidence intervals (CI) for δ . Regarding the first outcome (any recidivism), our simulations consistently yield treatment effects with confidence intervals in the negative domain, except for the least credible parameters. The results allow us to sign with strong confidence the effect of semi-liberty on hazard of recidivism after accounting for selection on unobservables. Under credible assumptions on selection on unobservables, *SL* has a significant negative effect on offenders' hazard of recidivism after release. It is only when we consider a scenario with very large confounding (for instance $\mu = 0.5$ and $\alpha = 0.3$) that our estimates no longer allow us to reject a small criminogenic effect of *SL*.

Table 4 shows that the simulation results for recidivism with prison are less conclusive. Except for the extreme cases where no confounding exists, all the confidence intervals for our treatment effects include zero, thus excluding the possibility to sign with strong confidence the causal effect of semi-liberty on "serious" recidivism. However, the standard error associated with the ATE remains large. This low precision of our simulations is presumably explained by the relatively low number of offenders who are reconvicted to prison in the follow-up period (29% overall).

20 Instead of unconditional differences in means, we may want to use differences conditional on X , i.e., differences in residuals of past convictions when regressed on other control variables. The result of this procedure is virtually identical to unconditional differences, with an estimated gap between treated and controls of 33% of a standard deviation.

21 While we find clear evidence of selection into SL related to past convictions, we fail to find such evidence for age at entry, although age is also known as one of the best predictors of recidivism (see Table 2). The difference in age is only one year (with an average age of 31.1 among SL offenders versus 30.1 among non-SL offenders) and not significant ($F = 1.27$, $p\text{-value} = 0.261$).

Table 4. Confidence intervals for the effect of SL with a confounder

A. Any recidivism

μ	0	0.1	0.2	0.3	0.4	0.5
α	ATE (SE) [Lower ; Upper]	ATE (SE) [Lower ; Upper]	ATE (SE) [Lower ; Upper]	ATE (SE) [Lower ; Upper]	ATE (SE) [Lower ; Upper]	ATE (SE) [Lower ; Upper]
0.0	-0.381 (0.137) [-0.650;-0.112]	-0.381 (0.137) [-0.650;-0.112]	-0.381 (0.137) [-0.650;-0.112]	-0.381 (0.137) [-0.650;-0.112]	-0.381 (0.137) [-0.650;-0.112]	-0.381 (0.137) [-0.650;-0.112]
0.1	-0.381 (0.138) [-0.652;-0.110]	-0.371 (0.138) [-0.642;-0.100]	-0.361 (0.138) [-0.632;-0.090]	-0.351 (0.138) [-0.622;-0.080]	-0.341 (0.138) [-0.612;-0.070]	-0.331 (0.138) [-0.602;-0.060]
0.2	-0.382 (0.141) [-0.659;-0.105]	-0.362 (0.141) [-0.639;-0.085]	-0.342 (0.141) [-0.619;-0.065]	-0.322 (0.141) [-0.599;-0.045]	-0.302 (0.141) [-0.579;-0.025]	-0.282 (0.141) [-0.559;-0.005]
0.3	-0.383 (0.146) [-0.670;-0.096]	-0.353 (0.146) [-0.640;-0.066]	-0.323 (0.146) [-0.610;-0.036]	-0.293 (0.146) [-0.580;-0.006]	-0.263 (0.146) [-0.550;0.024]	-0.233 (0.146) [-0.520;0.054]

B. Recidivism leading to prison

μ	0	0.1	0.2	0.3	0.4	0.5
α	ATE (SE) [Lower;Upper]	ATE (SE) [Lower;Upper]	ATE (SE) [Lower;Upper]	ATE (SE) [Lower;Upper]	ATE (SE) [Lower;Upper]	ATE (SE) [Lower;Upper]
0.0	-0.407 (0.204) [-0.807;-0.008]	-0.407 (0.204) [-0.807;-0.008]	-0.407 (0.204) [-0.807;-0.008]	-0.407 (0.204) [-0.807;-0.008]	-0.407 (0.204) [-0.807;-0.008]	-0.407 (0.204) [-0.807;-0.008]
0.1	-0.408 (0.205) [-0.810;-0.006]	-0.398 (0.205) [-0.800;0.004]	-0.388 (0.205) [-0.790;0.014]	-0.378 (0.205) [-0.780;0.024]	-0.368 (0.205) [-0.770;0.034]	-0.358 (0.205) [-0.760;0.044]
0.2	-0.408 (0.208) [-0.816;-0.000]	-0.388 (0.208) [-0.796;0.020]	-0.368 (0.208) [-0.776;0.040]	-0.348 (0.208) [-0.756;0.060]	-0.328 (0.208) [-0.736;0.080]	-0.308 (0.208) [-0.716;0.100]
0.3	-0.409 (0.213) [-0.827;0.009]	-0.379 (0.213) [-0.797;0.039]	-0.349 (0.213) [-0.767;0.069]	-0.319 (0.213) [-0.737;0.099]	-0.289 (0.213) [-0.707;0.129]	-0.259 (0.213) [-0.677;0.159]

Source: data from French Prison Administration, authors' calculations.

Note: ATE = average treatment effect, SE = standard error. The lower and upper bounds refer to 95% confidence intervals.

4. Conclusion

Semi-liberty is a common alternative to prison, available in many developed countries as a front-door or back-door strategy to limit incarceration. While attractive, this community sanction has not received much empirical scrutiny up until now. In this paper, we investigate the causal effect of semi-liberty on recidivism using a sample of criminal convicts eligible for SL in France. As we have no natural experiment in our data, we propose an econometric framework in which we carefully investigate the effect of selection into SL. On the one hand, we consider a selection on observables setting using Cox regressions and entropy balancing matching. On the other hand, we examine the possibility that our results may be biased due to some selection on unobservables by simulating the impact of a confounding factor.

Our results allow us to sign with strong confidence the effect of semi-liberty on recidivism as a front-door alternative to prison. Under exogeneity, SL leads to a reduction of 31% in offenders' hazard of recidivism in the subsequent five years after controlling for individual characteristics, and 22% after entropy balancing matching. Furthermore, we find that the impact of SL tends to decrease, but it remains nonetheless negative and significant when credible values of confounding factors are introduced. Specifically, the effect of SL on hazard of recidivism loses its statistical significance only when the confounder is as correlated with the decision of judges and with recidivism as the most discriminant observable characteristic, i.e., past convictions. When considering recidivism leading to prison, our results are qualitatively similar, albeit they appear more subject to the influence of confounding factors. Nevertheless, we believe that our estimates are less precise for serious recidivism due to the limited number of offenders in this situation.

Overall, we provide strong support for a beneficial effect of semi-liberty on recidivism compared to prison. Thus, our results suggest that semi-liberty has the potential to provide an effective alternative to incarceration. This finding is particularly relevant in the current context where prisons in France as well as in other countries are severely overcrowded and halfway houses work under capacity. Judges could opt more often for SL for prison convicts at the margin with no adverse consequences on recidivism. With respect to our empirical findings, the question as to why judges do not currently rely more often on SL remains a puzzle. The lack of empirical studies based on detailed individual data providing evidence on the short and long-term consequences on SL may be an answer to this puzzle. Another explanation could be that many short prison sentences cannot really be substituted by SL decisions if too short stays in semi-liberty centers are unable to provide positive effects for offenders.

Finally, a few limitations must be kept in mind when interpreting our results. First, we would like to have exogenous variations in the intensity of SL decisions but the French context does not provide such empirical setting actually. This precludes us to assess the robustness of our findings using instrumental variable techniques. Second, we have access to a limited sample of offenders, which is likely to explain the inconclusive effects (due to large standard errors) observed in the presence of large confounders when considering serious recidivism. As a consequence, we are not able to account for the possibility of heterogeneous treatment effects

on specific groups of convicts. Third, during sentencing, judges may have other choices than prison and semi-liberty. While we have only focused on those two possibilities due to data restrictions, it would be of interest to compare SL to other alternative sentencing options. Whatsoever, at that stage, our estimates provide the first empirical evidence of the positive effect of SL on recidivism in France. Further investigating the short and long-term consequences of semi-liberty and extending the comparison with other types of non-custodial sanctions are the next steps on our research agenda once detailed administrative data will become available.

References

- Altonji, J. G., T. E. Elder, and C. R. Taber (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy* 113 (1), 151-84.
- Altonji, J. G., T. E. Elder, and C. R. Taber (2008). Using Selection on Observed Variables to Assess Bias from Unobservables When Evaluating Swan-Ganz Catheterization. *American Economic Review* 98 (2), 345-50.
- Austin, P. C. (2014). The use of propensity score methods with survival or time-to-event outcomes: reporting measures of effect similar to those used in randomized. *Statistics in Medicine* 33, 1242-1258.
- Bhuller, M., G. B. Dahl, K. V. Loken and M. Mogstad (2019). Incarceration, recidivism and employment. Forthcoming in *Journal of Political Economy*.
- Carpenter, M. D. (1978). A follow-up study of Berkeley house: A psychiatric halfway house. *International Journal of Social Psychiatry* 23 (2), 120-131.
- Clark, V. A. (2016). Predicting two types of recidivism among newly released prisoners: First addresses as "launch pads" for recidivism or reentry success. *Crime & Delinquency* 62 (10), 1364-1400.
- Constanza, S., S. Cox, and M. Kilburn (2015). The impact of halfway houses on parole success and recidivism. *Journal of Sociological Research* 6 (2), 39-55.
- DAP (2018a). Recueil des coûts JDD. Technical report, Direction de l'Administration Pénitentiaire (DAP Me5).
- DAP (2018b). Statistiques mensuelles des personnes écrouées et détenues en France. Technical report, Direction de l'Administration Pénitentiaire (DAP Me5).
- Di Tella, R. and E. Schargrotsky (2013). Criminal recidivism after prison and electronic monitoring. *Journal of Political Economy* 121 (1), 28-73.
- Dum, C. P., K. M. Socia, and J. Rydberg (2017). Public support for emergency shelter housing interventions concerning stigmatized populations. *Criminology & Public Policy* 16 (3), 835-877.
- Durose, M. R., A. D. Cooper, and H. N. Snyder (2014). Recidivism of Prisoners Released in 30 States in 2005 : Patterns from 2005 to 2010. U.S. Department of Justice, Bureau of Justice Statistics.
- Freeman R. (2013). Can We Close the Revolving Door?: Recidivism vs Employment of Ex-Offenders in the U.S., *Urban Institute Roundtable: Employment Dimensions of Prisoner Reentry and Work*. New York University Law School.
- Garland, B., E. Wodahl, and C. Saxon (2017). What influences public support of transitional housing facilities for offenders during reentry? *Criminal Justice Policy Review* 28 (1), 18-40.
- Gumrukcu, P. (1968). The efficacy of a psychiatric halfway house: A three-year study of a therapeutic residence. *Sociological Quarterly* 9 (3), 374-386.
- Hainmueller, J. (2012). Entropy balancing: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis* 20 (1), 25-46.
- Hamilton, Z. K. and C. M. Campbell (2014). Uncommonly observed: the impact of New Jersey's halfway house system. *Criminal Justice and Behavior* 41 (11), 1354-1375.
- Hartmann, D. J. and P. Friday (1994). Residential probation: A seven-year follow up study of halfway house discharges. *Journal of Criminal Justice* 22 (6), 503-515.
- Henneguelle, A. (2020). L'institution judiciaire au delà de la sanction : "l'activation" dans l'aménagement des peines en France. *La Nouvelle Revue du Travail* 16.
- Henneguelle, A. and B. Monnery (2017). Prison, peines alternatives et récidive. *Revue Française d'Economie* 32 (1), 169-207.

- Henneguelle, A., B. Monnery, and A. Kensey (2016). Better at home than in prison? The effects of electronic monitoring on recidivism in France. *Journal of Law and Economics* 59 (3), 629-667.
- Hyatt, J. M. and S. Han (2018). Expanding the focus of correctional evaluations beyond recidivism: the impact of halfway houses on public safety. *Journal of Experimental Criminology* 14 (2), 187-211.
- Josnin, R. (2013). La récidive plus fréquente et plus rapide chez les jeunes condamnés. *France, portrait social 2013 INSEE*, 126-141.
- Kensey, A. and A. Benaouda (2011). Les risques de récidive des sortants de prison : une nouvelle évaluation. *Cahier d'études pénitentiaires et criminologiques* 36.
- Kilburn, M. and S. Costanza (2011). *Salvation city: Halfway House Stories*. Amherst, NY:Teneo Press.
- Krauth, B. (2015). Bounding a linear causal effect using relative correlation restrictions. *Journal of Econometric Methods* 5 (1), 117-141.
- Latessa, E. and H. E. Allen (1982). Halfway houses and parole: a national assessment. *Journal of Criminal Justice* 10 (2), 153-163.
- Lee, L. M. (2019). Halfway Home? Residential Housing and Recidivism. *Working paper*.
- Lin, N. X., S. Logan, and W. E. Henley (2013). Bias and sensitivity analysis when estimating treatment effects from the Cox model with omitted covariates. *Biometrics* 69 (4), 850-860.
- Liu, W., Kuramoto, S.J., and E.A. Stuart (2013). An introduction to sensitivity analysis for unobserved confounding in nonexperimental prevention research, *Prevention Science* 14(6), 570-580.
- Lochner, L., and E. Moretti (2004). The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports. *The American Economic Review*, 94 (1), 155-189.
- Mastrobuoni, G. and D. Terlizzese (2018). Harsh or Humane? Detention conditions and recidivism. *EIEF Working Paper*.
- McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 12 (2), 153-157.
- Millimet, D. and R. Tchernis (2013). Estimation of treatment effects without an exclusion restriction: With an application to the analysis of the school breakfast program. *Journal of Applied Econometrics* 28 (6), 982-1017.
- Monnery, B. (2015). The determinants of recidivism among ex-prisoners: a survival analysis on French data. *European Journal of Law and Economics* 39 (1), 37-56.
- Mueller-Smith, M. and K. Schnepel (2017). Diversion in the criminal justice system: Regression discontinuity evidence on court deferrals. *Working paper*.
- Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics* 37 (2), 187-204.
- Politico (2018). President Trump Says He Wants to Reform Prisons. His Attorney General Has Other Ideas. By Justin George. *Politico Magazine*.
- Richey, J. (2015), Shackled labor markets: Bounding the causal effects of criminal convictions in the U.S.. *International Review of Law and Economics* 41, 17-24.
- Rosenbaum, P. R. (2002). *Observational studies*. New York, NY : Springer, 375 p.
- Rosenbaum, P. R. (2005). Sensitivity Analysis in Observational Studies. In Everitt, B. S. and D. C. Howell (eds.) *Encyclopedia of Statistics in Behavioral Science*. Chichester, UK: Wiley, 1809-1814.
- Rosenbaum, P. R. and D. B. Rubin (1983). The central role of the propensity score in observational studies for causal effects, *Biometrika* 70 (1), 41-55.

Rosenbaum, P. R. and D. B. Rubin (1984). Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association* 79 (387), 516-24.

Siwach, G. (2017), Criminal background checks and recidivism: Bounding the causal impact. *International Review of Law and Economics* 52, 74-85.

Stuart, E.A. (2010). Matching methods for causal inference: A review and a look forward, *Statistical Science* 25, 1-21.

Walsh, C. L. and S. Beck (1990). Predictors of recidivism among halfway house residents. *American Journal of Criminal Justice* 15 (1), 137-156.

White, M. D., J. Mellow, K. Englander, and M. Ruffinengo (2011). Halfway back: an alternative to revocation for technical parole violators. *Criminal Justice Policy Review* 22 (2), 140-166.