https://economix.fr/

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

Document de Travail
Working Paper
2019-20

Anaïs Henneguelle
Benjamin Monnery
François-Charles Wolff

EconomiX - UMR7235
Université Paris Nanterre
Bâtiment G - Maurice Allais, 200, Avenue de la République
92001 Nanterre cedex

Email : secretariat@economix.fr

UMR 7235

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

Benjamin Monnery[1] **François-Charles Wolff**[2] **Anaïs Henneguelle**[3]

August 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 (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

We thank Annie Kensey (DAP Me5) for providing access to the raw data and Ann Thampi for research assistance. We have benefited from comments and suggestions at the EALE 2018 conference (Milano), the AFED 2018 conference (Nancy) and the JMA 2019 conference (Casablanca). Any remaining errors are ours.

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

2 LEMNA, University of Nantes, France. Email: francois-charles.wolff@univ-nantes.fr

3 LiRIS, University of Rennes 2, France. Email: anais.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)[4].

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 France[5].

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 work outside or receive job training, medical treatment, family support, etc. While theoretically attractive, semi-liberty is 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).

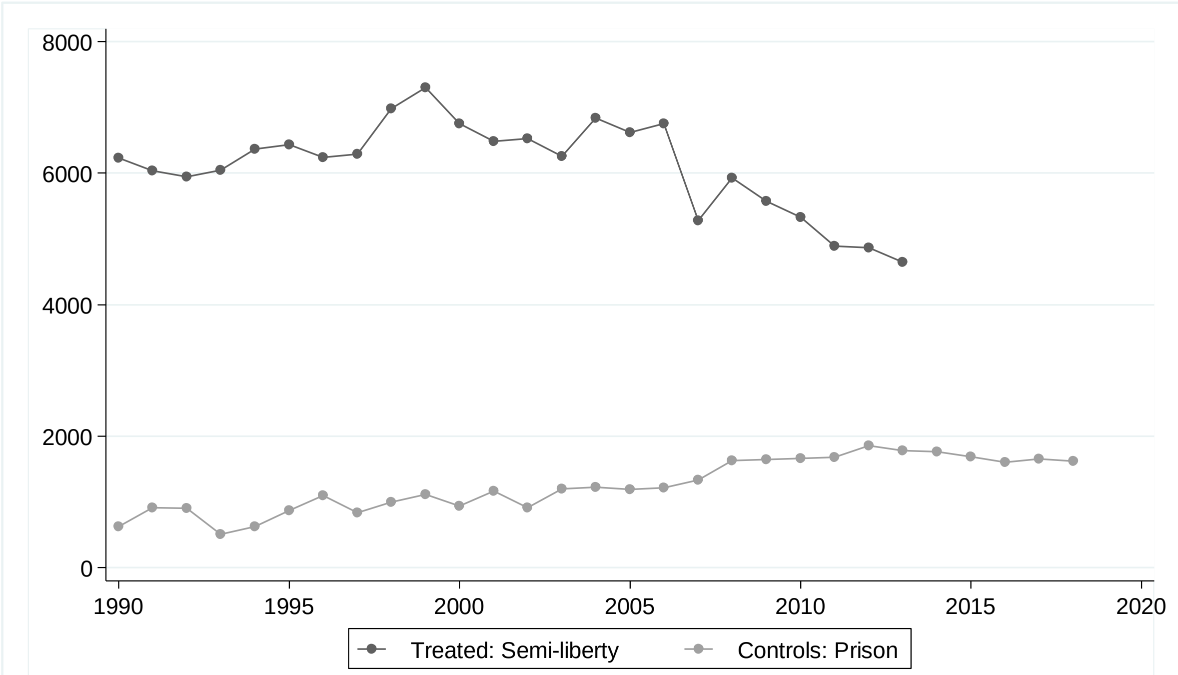
As shown in Figure 1, the annual flow of offenders benefiting from SL in France decreased from 6,000-7,000 per year between 1990 and 2006 to about 4,500 more recently. Although French prisons suffer from severe overcrowding (around 70,000 prisoners for a 59,000 capacity in January 2019), most semi-liberty 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

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

5 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 sanctions target 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 Schargrodsky, 2013; Henneguelle et al., 2016) and other probation programs (Mueller-Smith and Schnepel, 2017)[6]. 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 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 unreliable[7]. So far,

6 In France, Henneguelle et al. (2016) find a 10% reduction of the 5-year probability of reconviction thanks to electronic monitoring.

7 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, 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).

the few existing studies on semi-liberty and recidivism have attempted to account for selection bias by controlling for observable differences.

They typically use logistic regressions (Clark, 2016), propensity score matching (Hamilton and Campbell, 2014), or a combination of the two (Constanza et al., 2015). However, selection on unobservables may well remain and lead to bias in the estimated effects. 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-year 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.

Contrary to most previous attempts to study 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 where, presumably, similar offenders face 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 cautious sensitivity analysis to offer credible bounds of the causal effects 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 (Oster, 2019; Millimet and Tchernis, 2013; Krauth, 2015). Here, we 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 5 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 different matching algorithms to calculate the average treatment effect. Third, we evaluate the robustness of our estimates by accounting for 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 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 semi-liberty 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 the data used. The third section presents the various econometric results obtained by Cox regressions and matching algorithms along with some sensitivity analyses. 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 year[8]) 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 served[9].

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

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

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

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. Even though we have access to some information, judges may benefit from more information to make their decisions: police investigations, police custody hearings, judicial hearings, etc. This collection of qualitative data provides judges with private information (in the sense of "not observed by researchers") to assess offenders' underlying risks of recidivism and decide on the best-suited sanction. Note, however, that 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, 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 (e.g. females, juveniles, offenders under parole), whereas other offenders were sampled based on their offense type with probability ranging from 1/1 to 1/16[10]. All our statistical analyses account for the different weights

10 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, sampling probability was 1 for offenders convicted of murder and rape, 0.5 for embezzlement and breach of trust, 0.2 for theft, 0.1 for driving under the influence, etc.

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 recidivism[11]. 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). Although some of the socio-demographic characteristics are self-assessed by offenders and therefore potentially subject to measurement errors, having key control variables available strongly increases the reliability of our empirical analysis[12].

We proceed with several sample restrictions. 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 are not legally eligible to full semi-liberty. This concerns convicts with an initial prison sentence of more than a 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 5 previous criminal convictions. This second restriction helps achieve balance between the two groups. Third, we exclude the few offenders whose key information (in particular criminal record and sentence length) was not available[13]. Overall, these restrictions lead to a study sample of

---

11 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

12 Socio-demographic characteristics are self-assessed by offenders when they enter prison (or halfway houses), hence they can include errors and only describe offenders' socio-demographics at entry. However, since judges' decision to incarcerate or divert relies on pre-entry (and often self-declared) characteristics, this last point is actually helpful in our analysis.

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

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%).

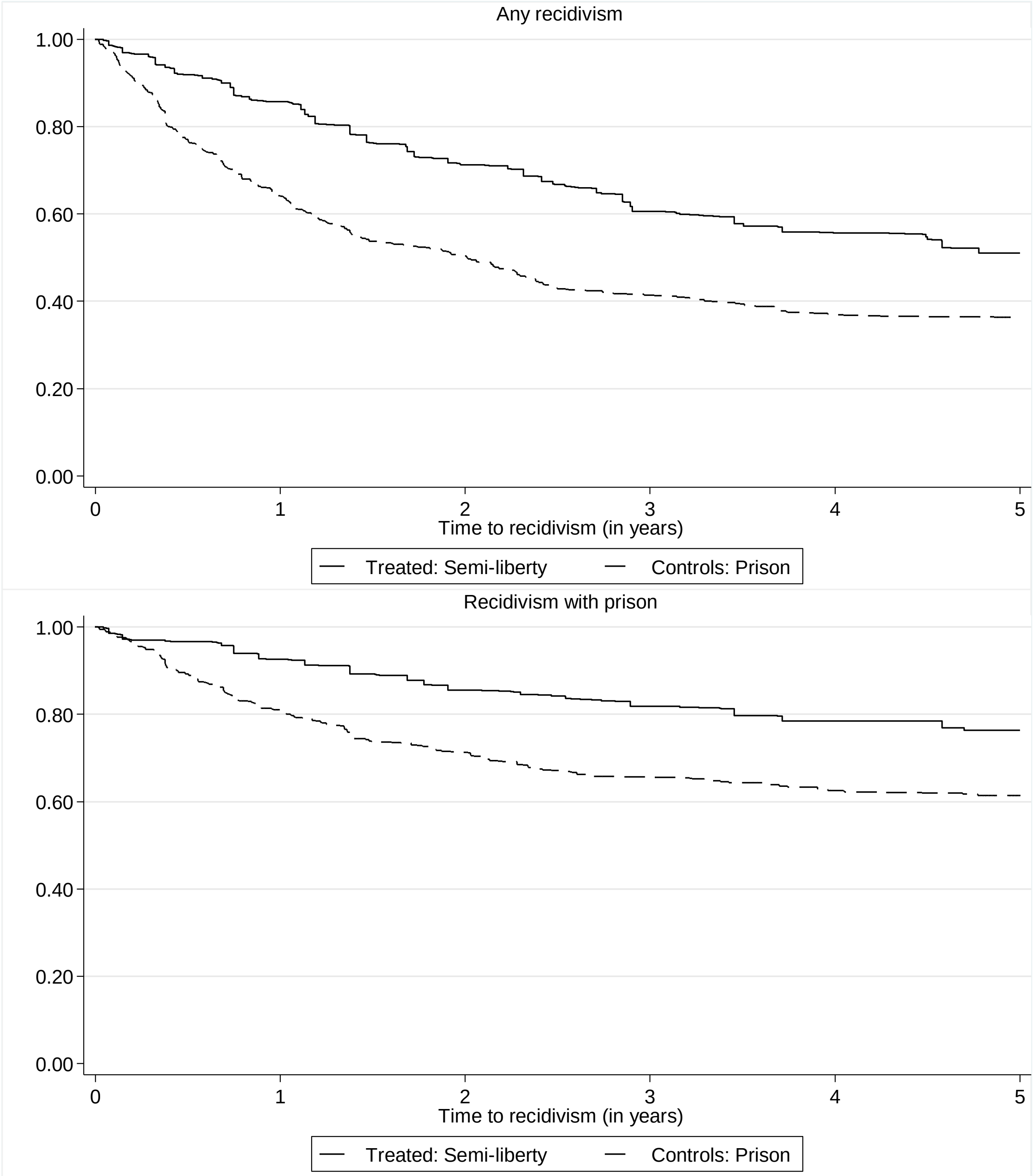
In Figure 2, we provide a graphical representation of recidivism dynamics in the two groups using the non-parametric Kaplan-Meier estimator. Panel A plots the survival rate over time, i.e., the share of offenders who have not yet reoffended. Five years after release, the survival rate is equal to 51.0% 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 groups[14]. 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.

---

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

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



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

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 effects[15]. 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%).

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% (*).

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.

15 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.

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:

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

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 in Table 2. Without any control variables, we find that SL offenders face a 39.0% lower instantaneous hazard of recidivism than non-SL offenders (corresponding to $1 - exp(-0.495)$, column 1) and a 46.4% lower instantaneous hazard of reconviction to prison ($1 - exp(-0.623)$, column 4). Interestingly, the estimated hazard differentials remain large and significant after controlling for individual characteristics, albeit lower in terms of intensity: the drop in estimated coefficients is around 25-30%. 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).

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 (see in particular Josnin, 2013; Monnery, 2015).

Table 2. Cox proportional hazard estimates of recidivism

Variables		Any recidivism			Recidivism leading to prison		
		(1)	(2)	(3)	(4)	(5)	(6)
SL treatment		-0.495***	-0.371***	-0.380***	-0.623***	-0.427**	-0.408**
		(-3.79)	(-2.59)	(-2.77)	(-3.07)	(-2.11)	(-2.00)
Female			-0.381***	-0.337**		-0.653***	-0.658***
			(-2.83)	(-2.42)		(-3.24)	(-3.17)
Foreigner			-0.169	-0.198		-0.336	-0.365
			(-0.91)	(-1.05)		(-1.29)	(-1.38)
Married			-0.027	0.020		-0.091	-0.061
			(-0.19)	(0.14)		(-0.44)	(-0.29)
Parent			0.226	0.202		0.170	0.154
			(1.56)	(1.41)		(0.82)	(0.72)
High education			-0.288**	-0.288**		-0.247	-0.226
			(-2.26)	(-2.32)		(-1.50)	(-1.37)
Age at entry			-0.061***	-0.067***		-0.052***	-0.051***
			(-6.98)	(-7.87)		(-4.84)	(-4.58)
Sentence length			-0.051**	-0.040*		-0.026	-0.028
			(-2.40)	(-1.80)		(-1.00)	(-1.04)
Past convictions	1		0.180	0.173		0.310	0.306
(reference: 0)			(1.17)	(1.13)		(1.49)	(1.47)
	2		0.430**	0.358**		0.254	0.232
			(2.45)	(2.02)		(1.04)	(0.91)
	3		0.856***	0.799***		0.866***	0.846***
			(4.50)	(4.17)		(3.17)	(3.08)
	4		0.853***	0.811***		1.168***	1.164***
			(4.46)	(4.33)		(5.03)	(4.88)
	5		0.896***	0.765***		0.995***	0.970***
			(3.82)	(2.98)		(3.21)	(3.10)
Thefts – concealments				-0.297			0.110
				(-1.36)			(0.40)
Degradations – economic offenses				-0.090			-0.399
				(-0.36)			(-1.04)
Driving offenses				0.047			-0.075
				(0.25)			(-0.28)
Violence – threats				-0.028			0.209
				(-0.12)			(0.74)
Narcotic drugs				-0.861***			-0.342
				(-2.97)			(-0.95)
Number of observations		1,445	1,445	1,445	1,445	1,445	1,445

Source: data from French Prison Administration, authors' calculations.
Note: estimates from Cox proportional hazard models, with robust standard errors. Significance levels are 1% (***), 5% (**) and 10% (*).

4.2. 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 (criminal propensities with and without semi-liberty) is independent from treatment status within matched pairs[16]. 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 Mahalabonis distance measures, which corresponds to nearest-neighbor matching. The propensity-score method consists of predicting individual propensity scores based on a Logit regression and then match treated and untreated individuals with close scores (Rosenbaum and Rubin, 1983). Finally, entropy matching is a data preprocessing procedure leading to balanced samples by construction (Hainmueller, 2012). Some weights are calculated and assigned to each observation such that the covariate distributions in the reweighted dataset fulfill a set of moment conditions. Here, data for controls (non-SL offenders) will be reweighted so as to match the moments of the treated (SL offenders).

In what follows, we implement different matching techniques in order to assess the robustness of our empirical findings. Before comparing outcomes, it is of interest to understand the role of individual characteristics on the SL decision. We estimate Logit regressions to explain SL and include many individual characteristics, as well as types of crime, as explanatory variables. According to Table 3 (column 1), we observe that many covariates are significantly correlated with the SL decision. On average, the probability of SL is significantly higher among offenders with children (at the 10 percent level) and is positively correlated with both education and sentence length. Being a parent and being educated are often perceived as preventive factors in crime (Lochner and Moretti, 2004). The fact that lengthier sentences are associated with increased propensity is explained by the fact that judges typically consider semi-liberty as a sanction requiring time to organize and yield its rehabilitative benefits (Henneguelle, 2017). As expected, access to SL is less likely among offenders with several past convictions than first-time offenders, while the type of offense has no influence[17].

16 This assumption is known as the Conditional Independence Assumption.

17 In column (2), we have introduced a set of interaction terms. We find that sentences related to violence and threats crossed with a high number of past convictions reduce the probability of SL.

Table 3. Logit estimates of semi-liberty

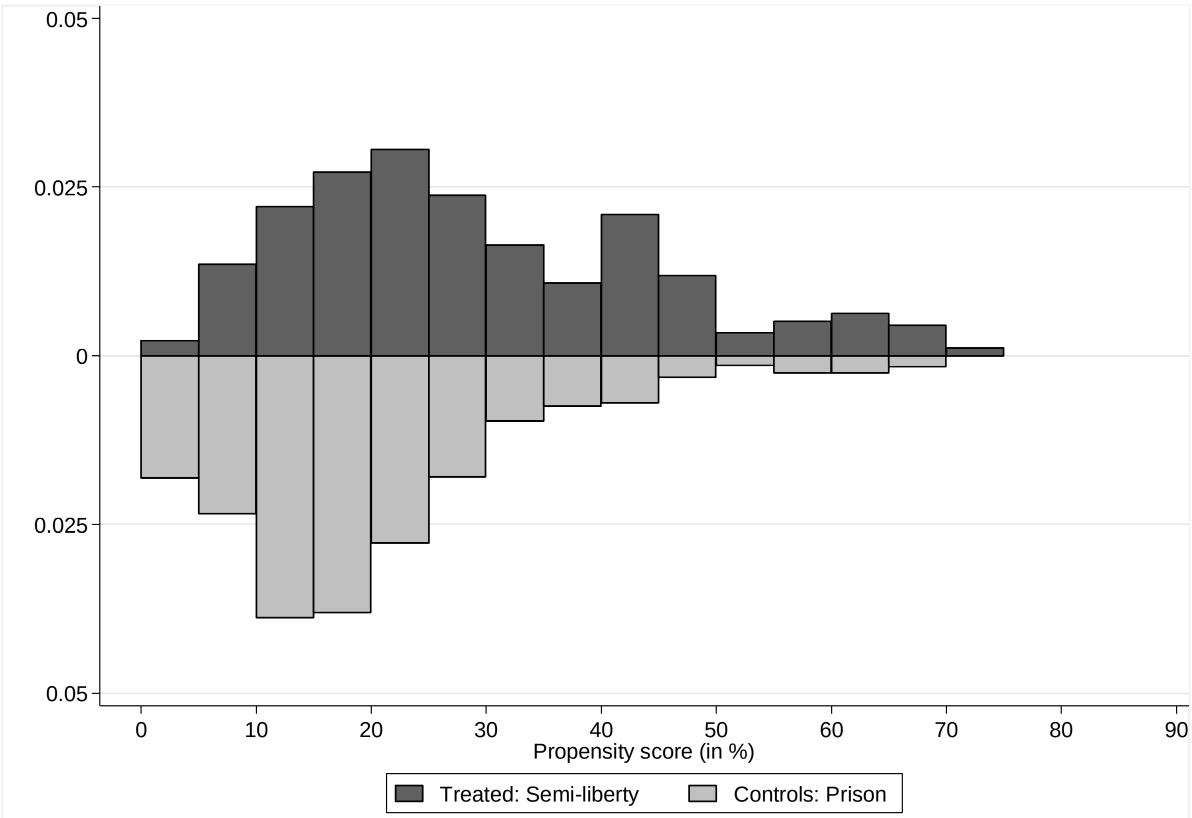
Variables		(1)	(2)
Female		-0.206	-0.240
		(-0.80)	(-0.91)
Foreigner		-0.319	-0.325
		(-0.84)	(-0.85)
Married		0.101	0.129
		(0.43)	(0.54)
Parent		0.462*	0.356
		(1.68)	(1.33)
High education		0.939***	0.785**
		(3.27)	(2.51)
Age at entry		-0.004	-0.004
		(-0.40)	(-0.34)
Sentence length		0.145***	0.146***
		(4.45)	(4.46)
Past convictions	1	-0.661**	-0.668**
(reference: 0)		(-2.35)	(-2.38)
	2	-0.801**	-0.800**
		(-2.44)	(-2.44)
	3	-0.883**	-0.852**
		(-2.26)	(-2.19)
	4	-0.860**	-0.856**
		(-2.08)	(-2.09)
	5	-1.466***	-1.872
		(-2.89)	(-0.76)
Thefts – concealments		-0.453	-1.690**
		(-1.14)	(-2.03)
Degradations – economic offenses		-0.145	-0.157
		(-0.34)	(-0.37)
Driving offenses		0.245	0.291
		(0.68)	(0.83)
Violence – threats		0.208	0.197
		(0.52)	(0.50)
Narcotic drugs		0.334	0.334
		(0.88)	(0.89)
Violence – threats x 5+ past convictions			-2.112**
			(-2.01)
Thefts – concealments x high education			1.170
			(1.59)
Age at entry x 5+ past convictions			0.032
			(0.45)
Thefts – concealments x Parent			0.532
			(0.87)
Constant		-2.410***	-2.281***
		(-4.50)	(-4.23)
Number of observations		1,445	1,445
Log likelihood		-2104.1	-2087.7

Source: data from French Prison Administration, authors' calculations.
Note: estimates from Logit models, with robust standard errors. Significance levels are 1% (***), 5% (**) and 10% (*).

The propensity score model is expected to balance (up to an acceptable level) the selected covariates between treated and control offenders. When it is not the case, then the inclusion of interaction terms among the list of explanatory variables will reduce the imbalance. In column 2 of Table 3, we report our preferred specification which includes four additional crossed terms[18]. While in the initial balance, the standardized difference in control and treated averages exceed 0.1 (in absolute value) for many covariates, in particular being married, having children, possessing higher education or number of past convictions, no standardized difference exceeds 0.1 after matching.

We report the distribution of the propensity score values for the two groups in Figure 3. For SL offenders, the distribution is bimodal with the most frequent predicted scores being around 0.25 and 0.45. There are very few SL offenders with a high propensity score. For non-SL offenders, the modal values are in the 10-20 percent interval and most offenders have an estimated propensity score below 40%. Although the two distributions strongly overlap (except for values exceeding 75% in the treated group), Figure 3 shows that alternative methods of matching like entropy balancing may be useful to assess the robustness of our results.

Figure 3. Distribution of estimated propensity scores for SL and non-SL offenders



Source: data from French Prison Administration, authors' calculations.
Note: estimates of the Logit model used to calculate the propensity score are reported in column 2 of Table 3.

18 Those interaction terms were selected "endogenously" in order to reduce the imbalance the most. We proceed in an iterative way until all variables are balanced.

We begin with the very intuitive entropy reweighting procedure such that all standardized biases associated with differences in characteristics between treated and control offenders go to zero by design (Hainmueller, 2012). Thus, this method allows us to compare recidivism in two groups that are on average perfectly similar in terms of observable characteristics. 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[19]. Figure 4 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-year 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 post-sanction.

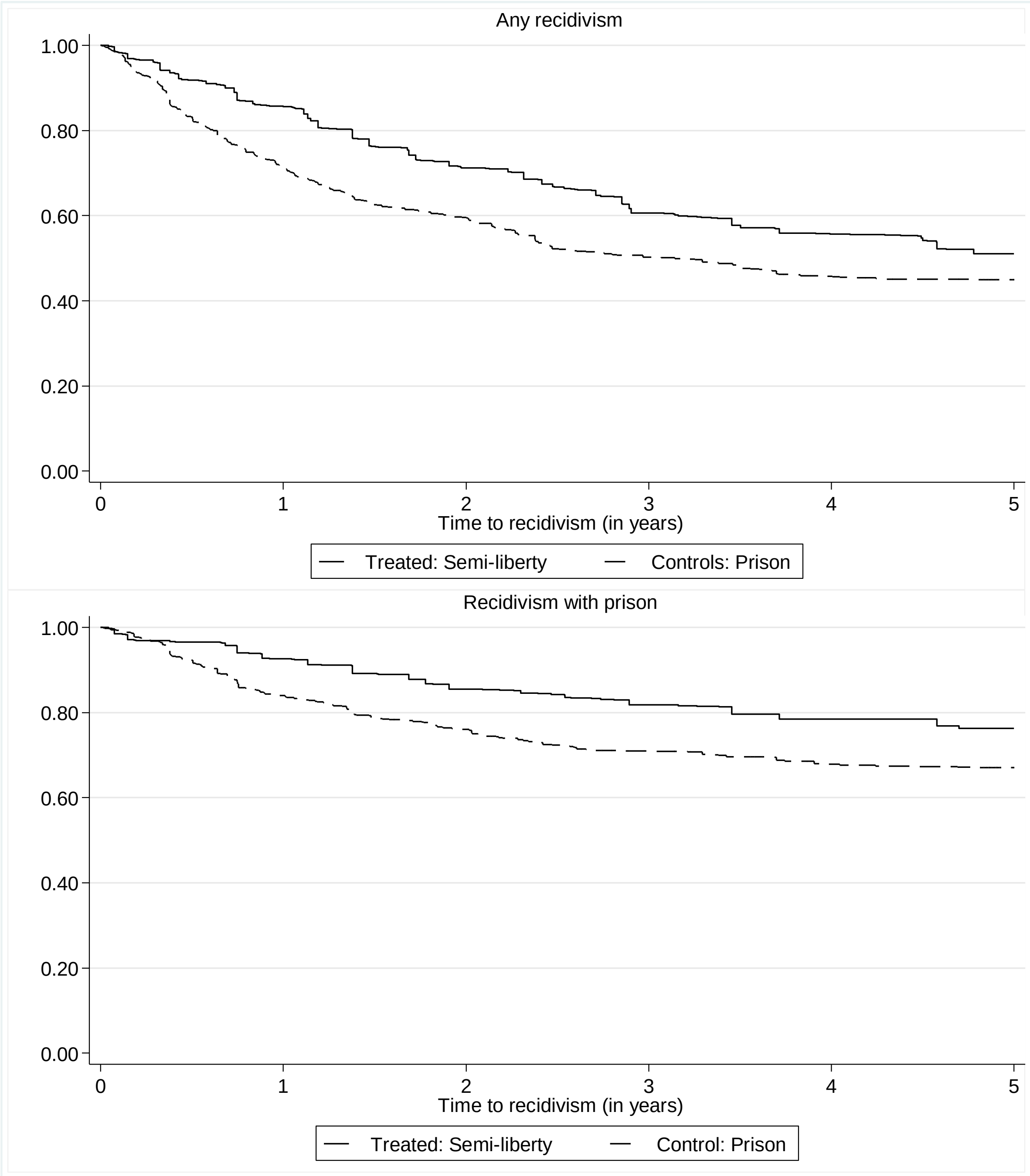
Figure 5 shows the average treatment effect (ATT) obtained from the different matching techniques, in which the outcomes are the probability of recidivism and of recidivism leading to prison within one, three and five years, respectively. Under exogeneity, results from Probit regressions lead to an ATT equal to -16.7 points at one year, -14.8 points at three years and -9.5 points at five years for any recidivism. After matching, we find rather comparable values, albeit slightly slower (in absolute value), for the three- and five-year horizon. The average ATT obtained from 8 different matching techniques is -17.7 points at one year, -11.8 points at three years and -8.7 points at five years. However, there are differences in the magnitude of the ATT depending on the underlying method. At one year for instance, the marginal effect ranges between -14.6 points (entropy balancing) and -21.3 points (propensity score matching with one neighbor and a caliper equal to 0.04), but most estimates are in a range of 16-18 points (in absolute value)[20].

When considering recidivism leading to prison, we note that the selected horizon (one-, three- or five-year) has less influence on the outcome. Under exogeneity, the ATT is -8.1 points at one year, -11.6 points at three years and -10.4 points at five years. After matching, the effect of SL on recidivism ranges between 8-10 points, whatever the horizon (in absolute value).

19 Here, the treated and control samples are balanced on their first two moments (means and variance).

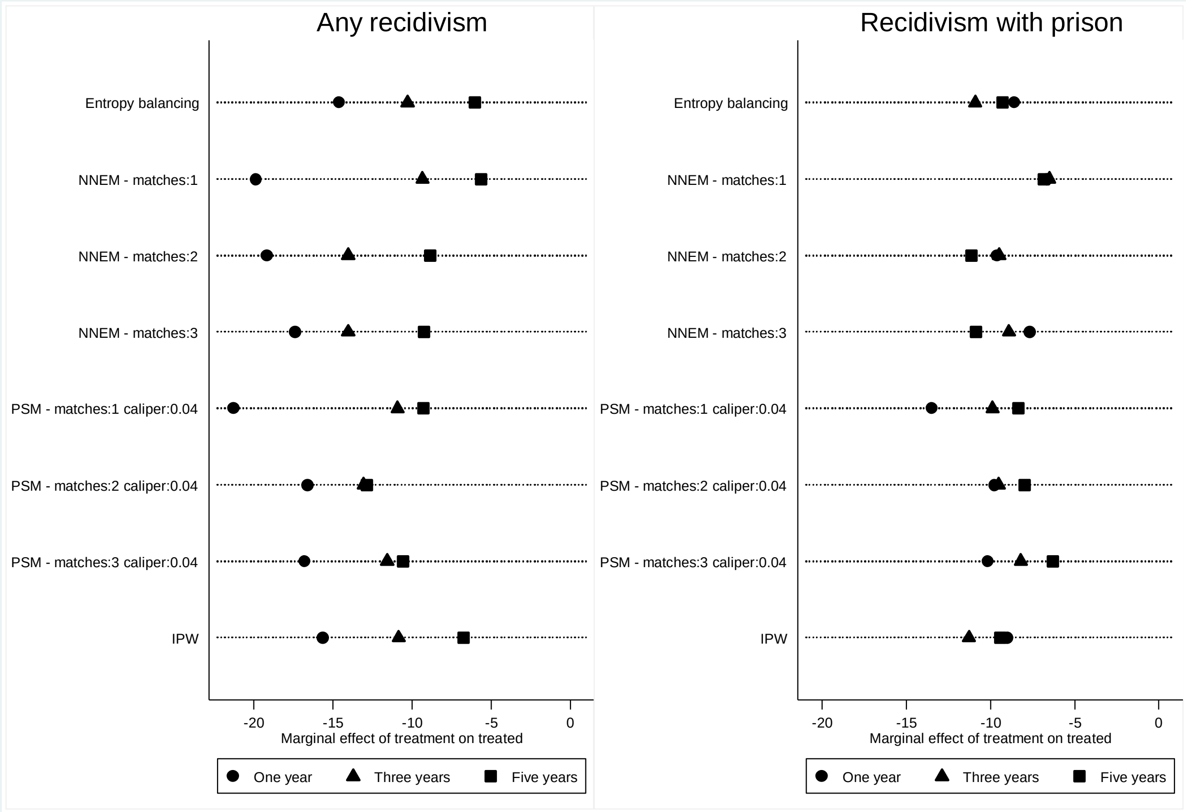
20 Both for nearest neighbor matching and propensity score matching, we find that the number of matches taken into consideration has little influence on the results (except with one match).

Figure 4. Marginal effect of SL on hazard rate of recidivism after entropy balancing



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

Figure 5. Marginal effect of SL on recidivism using different matching methods



Source: data from French Prison Administration, authors' calculations.
Note: NNEM = nearest neighbor matching, PSM = propensity score matching.

3.3. Unmeasured Confounding Variables

We consider two complementary methods to assess the influence of confounding variables on our results. First, we implement a simultaneous sensitivity analysis which relies on a one per one matching (Austin, 2014). Second, we extend the semi-parametric Cox regressions by adding a confounder in the list of covariates under some parametric assumptions.

3.3.1. Results from Simultaneous Sensitivity Analysis

In propensity score analyses, the underlying assumption is that there is no unmeasured confounding variable (see 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. In what follows, we rely on the sensitivity analysis described in Liu et al. (2013) and Austin (2014).

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 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 by[21]:

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

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

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:

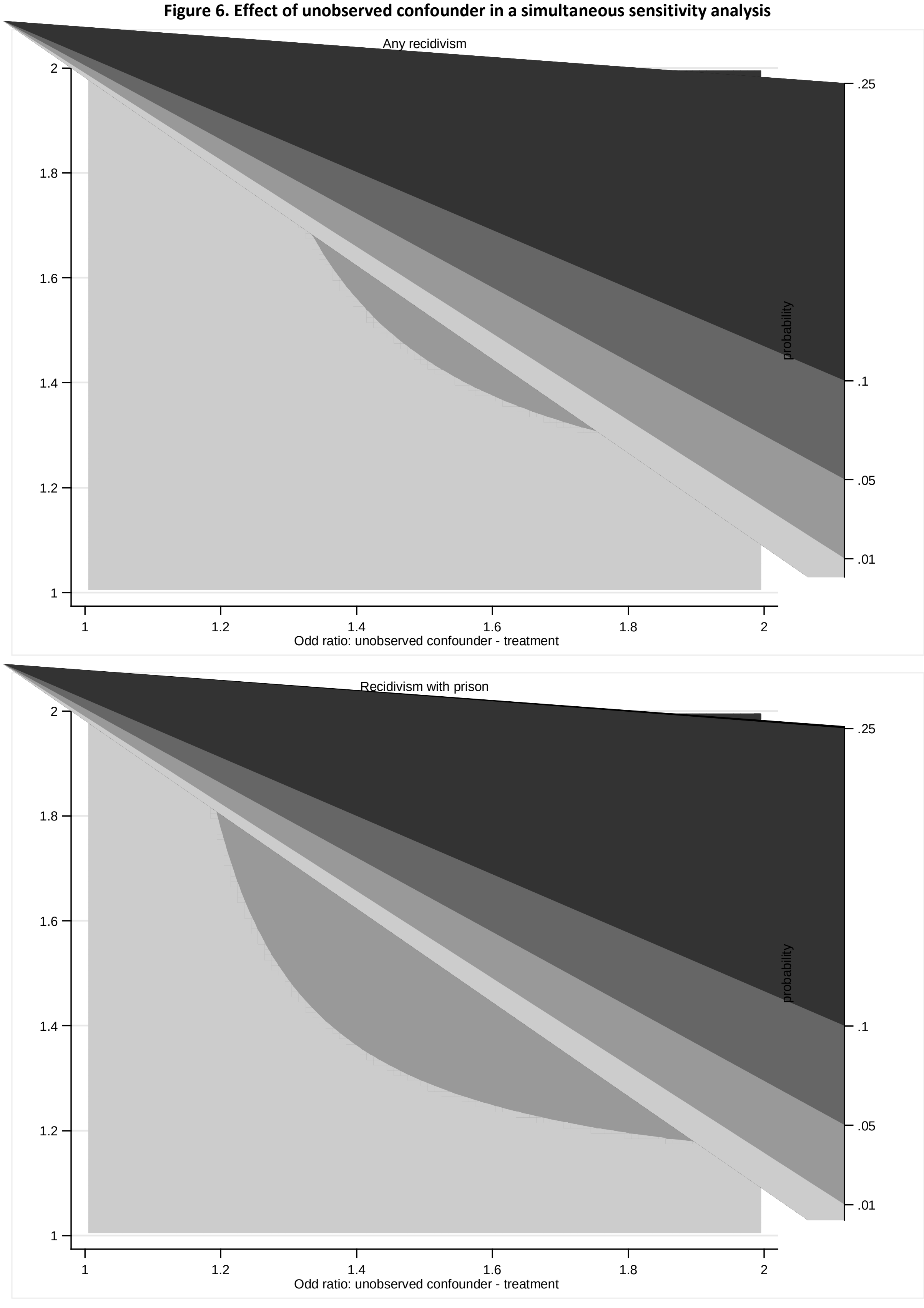
$$p^{+} = p_{\Gamma} * p_{\Delta} + (1-p_{\Gamma}) * (1-p_{\Delta}) \tag{4}$$

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.

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

In Figure 6, we present results from a simultaneous sensitivity analysis to unmeasured confounders. 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.

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



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

3.3.2 Simulations of Confounder

A complementary sensitivity test recently proposed by Lin et al. (2013) is 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):

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

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 $h_i(t)$.

To illustrate our approach intuitively, suppose that the confounder $\tilde{C}_i$ represents offenders' impulsivity (or any other credible confounder, such as severity of drug addiction, economic burden, social deprivation, etc.). While judges are likely to get a sense of impulsivity by hearing the offender or reading elements of the current case (police hearing, medical/psychological evaluation, etc.) or previous cases, such information cannot be observed or even credibly approximated with the data available to us. However, impulsivity is very likely to affect both judges' decisions on whether or not to incarcerate and offenders' hazard of recidivism after release. Because of these correlations, impulsivity is a realistic unobservable confounder that is expected to contaminate our coefficient of interest δ.

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 $\tilde{C}_i$ 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 $\tilde{C}\big|X\sim N(0;1)$, 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 $\tilde{C}\big|X\sim N(\mu;1)$, 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 $\tilde{C}$. As explained below, 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 store 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 set of parameters (μ, α) and compute the mean of all δ to obtain estimates 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)[22].

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., μ = 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 $\left(\underline{\mu}, \underline{\alpha}\right)$ 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 past convictions, age, offense type, sentence length and other socio-demographic characteristics) would be larger than that of past convictions (i.e., past recidivism).

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

22 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-value = 0.000) and corresponds to 34% of a standard deviation (sd = 1.52)[23]. 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 deviation[24]. 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*.

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).

23 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.

24 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-value = 0.261).

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

A. Any recidivism

μ / α	0	0.1	0.2	0.3	0.4	0.5
	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 / α	0	0.1	0.2	0.3	0.4	0.5
	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 different matching algorithms. On the other hand, we examine the possibility that our results are 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 remain negative and significant when credible values of confounding factors are introduced. 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, 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. Ideally, 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. Also, 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.

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 is on our research agenda once more recent detailed administrative data becomes available.

References

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. Schargrodsky (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. (2017). *Comment sortir de prison ? Le rôle des statistiques pénitentiaires dans la compréhension des comportements de récidive*. Ph Dissertation, ENS Paris-Saclay.

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.

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.