EIEF WORKING PAPER SERIES

IEF

EINAUDI INSTITUTE FOR ECONOMICS AND FINANCE

EIEF Working Paper 26/03

February 2026

Women Behind Bars: Do Single-Gender Prisons Reduce Recidivism?

By

Francesca Calamunci
(University of Catania & IZA)

Gianmarco Daniele
(University of Milan & Bocconi)

Giovanni Mastrobuoni
(Collegio Carlo Alberto & University of Torino & CEPR)

Daniele Terlizzese
(EIEF)

Women Behind Bars: Do Single-Gender Prisons Reduce Recidivism? *

Francesca Calamunci† Gianmarco Daniele‡
Giovanni Mastrobuoni § Daniele Terlizzese¶

February 13, 2026

Abstract

In 143 countries incarcerated women serve their sentence in a typically small separate section within prisons that mainly house male inmates, while 79 countries have prisons exclusively dedicated to women. Exploiting data from Italy, where both prison types coexist, and a quasi-random, proximity-based prison assignment rule, we find that women-only prisons lower three-year recidivism by about 14 percentage points. Given that proximity-based assignment rules are common in international correctional standards, our empirical strategy can be applied in other countries and to evaluate the causal effects of other (binary) prison characteristics. We use policy-relevant treatment effects to identify the optimal location of an additional women-only prison. As for the mechanisms, likely drivers are the presence of a large enough number of women in a given facility and a female prison director. We also find that a synthetic (latent) measure of prison quality causally reduces women recidivism.

JEL: K14, K42, H54

Keywords: Women, gender, prisons, recidivism, optimal policy-relevant treatment effects

*We thank Raffaella Calandra, Ida del Grosso and Neris Cimini, of the Italian Ministry of the Interior, for providing key data and for discussions about the prison system. Thanks also to the NGO Antigone and to Alessio Scandurra, for data access and discussions about the topic. Thanks also to Gemma Dipoppa, Tom Kirchmaier, Stephen Machin, Emily Owens, Paolo Pinotti, Adam Soliman and seminar participants at EIEF, 14th and 16th Petralia Workshop, University of Siena, Bolzano Applied Microeconomics Workshop, NBER Summer Institute 2025, 16th Transatlantic Workshop on the Economics of Crime, University of Palermo, Inequality in Rome Tre Seminar, LEDa Seminar Dauphine University.

†University of Catania & IZA; francesca.calamunci@unict.it

‡University of Milan & Bocconi (Baffi Carefin), Italy; gianmarco.daniele@unibocconi.it

§Corresponding author: Collegio Carlo Alberto & University of Torino & CEPR; giovanni.mastrobuoni@carloalberto.org

¶EIEF; daniele.terlizzese@eief.it

1 Introduction

Around the world, convicted women serve their sentences in different institutional settings. One possibility are mixed-gender prisons, i.e., facilities that predominantly house male convicts and have a separate female section, with virtually no interaction between the two genders.[1] Another possibility are women-only prisons, i.e., facilities entirely dedicated to women. Currently, in 79 countries, representing 14% of the world's population, there are exclusively mixed-gender prisons. In 56 countries, with 30% of the world population, there are exclusively women-only prisons. In 64 countries, with 54% of the world population, there are both types of prison.[2]

It is not clear why some countries choose one model and others choose a different one. In some cases, it might have been the legacy of historical developments. For example, in the U.S., the so-called reformatory model, more attuned to women's needs, replaced at the beginning of the twentieth century the so-called custodial model, in the wake of social feminism and the purity movement, and left both cultural and infrastructural inheritances favoring the women-only model.[3]

In some other cases, cost considerations might have played a role. Given that women represent only about 7% of the global prison population, it would often appear less costly to ring-fence a few women in an existing building, already occupied by men, rather than to build from scratch a separate facility.

However, we lack a balanced assessment of the potential benefits of these different models, for example in terms of rehabilitation. It is becoming increasingly important to have a better understanding of their consequences, given that, unfortunately, the number of female inmates is on the rise: while the world male prison population witnessed a 22% increase from 2000 to 2022, the number of incarcerated women grew by nearly 60% during the same period, surpassing 740,000 (Penal Reform International, 2023). In fact, the widespread concerns raised by NGOs and international organizations have brought attention to the difficult conditions often experienced by female prisoners, highlighting the

[1] The expression mixed-gender prison is sometimes referred to the few correctional facilities in which male and female prisoners share some common spaces and possibly rehabilitation treatments (an example is Spain's Teixeiro Prison). Our use of the expression is different.

[2] The distribution across the world is shown in Appendix Figure A1. For 12 countries, representing 2% of the world population, we were unable to find information on the type of female prisons.

[3] See Rafter (1983).

need for improved standards.[4] Depending on their gender composition, different prisons might be differently equipped to address women's specific needs, such as healthcare, tailored care for pregnant women, education policies, training and rehabilitation programs crafted with a specific focus on women's needs (World Health Organization, 2009), thereby differently affecting the conditions experienced during the time served by the women inmates. In turn, this might impact their behavior after release and in particular have an effect on female recidivism and the societal costs it entails.

However, women recidivism has received relatively little attention in the economic literature and has been explored through descriptive studies (Morash et al., 2017; Cobbina et al., 2012; Mathiassen, 2017). To our knowledge, there are no previous studies on the causal effects of prisons that differ by gender composition.

In this paper, we focus on the relative benefits of the two prison types and in particular we estimate the causal effect on female recidivism of serving a given sentence in a women-only rather than in a mixed-gender one.

To obtain this result, cross-country comparisons are hardly useful, due to multifaceted and difficult-to-control variations in prison systems between countries, compounded with distinct characteristics of the justice system. The analysis within the same country is also not free from difficulties, mainly because the allocation of inmates across the two types of prisons is typically non-random, and might be correlated with perceived recidivism, thus contaminating with selection the estimated coefficient.

We conduct our analysis using data from Italy, whose prison system contains both mixed-gender and women-only prisons, albeit with a large prevalence of the first kind: 52 prisons housing male inmates with a separate small female section, and 4 relatively large women-only prisons spread across the country, two in the South (Trani and Pozzuoli), one in Central Italy (Rome) and one in the North (Venice). As a background, it is useful to recall that in Italy, similarly to many other countries, the prison system is characterized by widespread poor conditions and overcrowding (Council of Europe, 2022).

We address the selection problem by leveraging institutional details of the assignment of inmates to prisons that allow us to mimic a quasi-random allocation of women between prisons of the two kinds. In Italy, the decision concerning the prison where a convicted woman will serve her sentence is taken by the Prison Administration Department (DAP), balancing two criteria. On the one hand, the four women-only prisons are considered

[4] Penal Reform International: "Addressing the 105,000 increase in the global female prison population, ten years after the Bangkok Rules were adopted" The challenging conditions faced by female convicts were acknowledged as early as 1833, when Auburn's prison chaplain noted that enduring a prolonged period as a female convict in often overcrowded small corners of predominantely male prisons would be worse than death, contrasting it with the relatively tolerable conditions for male convicts in the same prison (Rafter, 1983).

to be more attuned to the women needs and are in principle the preferred choice. On the other hand, the allocation tries to minimize the distance between the prison and the residence of the convict, to reduce the disruption of family and social links. Interestingly, this rule is present in many other democratic countries.[5]

When the (closest) mixed-gender prison is closer to the residence of the convict than the (closest) women-only prison, there is a tension between these two criteria. If the difference between the two distances is not too negative (i.e., when the women-only prison is not much more distant from the residence of the convict), the first consideration tends to prevail and the assigned prison is often a women-only one. If, instead, the mixed-gender prison is much closer to the residence of the convict, the second consideration tends to prevail, and the assigned prison is often a mixed-gender one.

As a result, the more negative is the difference between the two distances (mixed-gender to residence minus women-only to residence) the lower is the probability of being assigned to a women-only prison. If the (closest) women-only prison is in fact closer to the residence of the convict than the (closest) mixed-gender one, there is no conflict between the two criteria and the women-only prison tends to be chosen.[6]

The first criterion is likely to imply a selection problem, as inmates who are perceived to be more (or less) likely to recidivate might be more (or less) likely to be assigned to prisons with better environments and treatment opportunities. However, the second criterion is arguably exogenous, being only dependent on the difference between two geographical distances.

We therefore use the difference between the two distances (each from the inmate's residence) as an instrument for the likelihood of being assigned to a women-only prison.[7] In this way, we retain only the variability of the assignment that is attributable to an exogenous factor, thus getting close to random assignment. This instrumental variable (IV) approach enables the identification of the causal effect on recidivism stemming from the gender composition of prisons.[8]

A nice feature of our setting is that we can actually test the exclusion restriction, i.e.,

[5] The rule is contained in the Penitentiary code (Ordinamento Penitenziario), Art. 14, and, for example, is similar to Title 18 of the U.S. Code §3621 - Imprisonment of a convicted person.

[6] In these cases, the probability of being assigned to a women-only prison is, however, not yet 1, as other considerations, like, for example, the capacity of the prison or the need to sever the social links of high-security convicts, might play a role. However, we see that the probability is close to 0.8 and is fairly constant, regardless of the relative distance.

[7] In the Appendix Figure A2 we show the distribution of these distances across Italy.

[8] This instrument is similar in spirit to those used to instrument schooling choices (see, among others, Card, 2001, 1995), though with the advantage of exploiting relative differences across locations. As mentioned later, this allows us to control for the distance betweeen the inmate's place of residence and the prison the inmate ends up serving time.

the assumption that the effect of the instrument on recidivism is only the one mediated by the assignment to the two types of prisons, and not a direct one. Indeed, when the difference between the two distances is positive, we expect no effect on the assignment, once we control for the actual difference between the residence and the assigned prison. Hence, if the exclusion restriction is satisfied, we should observe no correlation between the instrument, on its positive range, and recidivism.

To sum up, our main contribution is to explore the role of women-only prisons as a determinant of women's recidivism. To do so, we propose a novel instrumental variable strategy that leverages a default preference of the Prison Administration for keeping inmates as close as possible to their place of residence.

Given that similar proximity-based assignment rules are common in international correctional standards, including binding EU regulations and UN guidelines (United Nations, 2010), our empirical strategy can be applied in other countries and, most importantly, can be used to evaluate the causal effects of other (binary) prison characteristics, besides the gender composition which is the focus of this study. Furthermore, our approach suggests the possibility of constructing a test, based on a revealed preference argument, to infer what drives a Prison Administration in allocating inmates to prisons. These claims will be justified in Section 2.2.

Our analysis is based on individual inmate data kindly provided by the DAP, covering the period 2012-2022. We complement these data with annual surveys conducted by the Italian NGO Antigone,[9] which provide detailed information on conditions in all Italian prisons.

We find a remarkable reduction in recidivism (measured over the three years following the release from prison), with Two-Stage Least Squares (2SLS) estimates ranging from 8 to 16 percentage points (14 percentage points in our preferred, most restricted specification), among inmates quasi-randomly assigned to women-only prisons, compared to otherwise similar inmates serving the same sentence in a mixed-gender prison. This effect is present and significant for both Italian and foreign inmates, and it is robust to various specifications and robustness tests detailed below.

In our setting (with heterogenous treatment effects and a continuous instrument) the 2SLS estimates are a variance-weighted average of covariate-specific Local Average Treatment Effects (LATEs). From these, we can also recover the so-called Marginal Treatment Effects (MTEs). These give the treatment effect for individuals at a given

[9] Antigone (https://www.antigone.it/) is a leading Italian non-governmental organization specializing in justice, human rights and prison-related matters. Each year, their independent representatives gain access to Italian prisons, documenting the conditions of inmates and identifying key challenges within the prison system.

point of the distribution of the unobserved “resistance to treatment”, which captures idiosyncratic motives that might offset the assignment to a women-only prison. We find that inmates are indeed selected on unobservables to maximize treatment effects, but this selection is of limited importance and the MTE curve is fairly flat, making it possible to interpret the 2SLS estimates as a good approximation of the Average Treatment Effect.

Leveraging on these results, we conduct two simple policy experiments: in the first, we compute “optimal” policy-relevant treatment effects (Heckman and Vytlacil, 2001b), that is, we identify the location of a single, new women-only prison that would generate the maximal effect in curbing women recidivism, given the local number of inmates, the geographic distribution of their residences and the individual MTE.

In the second, we verify the possibility to repurpose some of the existing mixed-gender prisons so as to move to a system with only single-gender prisons. In both cases, we compute the implied overall reduction in recidivism.

We have no conclusive and sharp evidence to identify the exact mechanism behind our results.

However, a large enough number of women inmates in a given facility, as well as the gender of the prison director, appear to be relevant drivers of our results, making it easier to meet their needs and design more effective and targeted rehabilitation initiatives. Indeed, we observe larger effects when the control group consists solely of women's sections within men's prisons that house a smaller share (or a smaller number) of women, which arguably lowers their bargaining position with respect to the male counterparts. This also aligns with qualitative evidence (Palmisano, 2015) suggesting a minimum-scale effect: too few women might not justify activating programs tailored to their specific needs. We also observe larger effects when the control group only includes women's sections within men's prisons directed by a man.

Women-only prisons typically exhibit better conditions in terms of improved health services, enhanced social and working spaces, better employment opportunities, and greater autonomy in organizing inmates' time and movements within the facility. It is difficult to disentangle the contributions of these different features to the estimated effect of recidivism. Rather, we are likely dealing with a bundled treatment, with the overall effect arising from the simultaneous improvement in various dimensions.

As a first attempt to deal with this, we select some qualitative and quantitative features of the prison facilities. We then repeat our analysis on the effect of two kinds of prisons on recidivism, with an identification strategy that again leverages proximity-based assignment rules, but instead of contrasting women-only and mixed-gender prisons, we use factor analysis to combine these different features into a single latent factor and we split the prisons into two groups, according to the latter. We find that the quality of the facilities and of the life inside a prison, as captured by the latent factor, has a causal effect on recidivism.

A considerable body of research has investigated the effects of imprisonment (Aizer and Doyle, 2015; Bhuller et al., 2020; Rose and Shem-Tov, 2021) sometimes with a specific focus on deterrence (Abrams, 2012; Helland and Tabarrok, 2007; Drago et al., 2009; Katz et al., 2003; Kuziemko, 2013; Vollaard, 2013) and incapacitation effects (Barbarino and Mastrobuoni, 2014; Buonanno and Raphael, 2013; Owens, 2009; Levitt, 1996).[10] Researchers have explored the influence of alternatives to traditional incarceration policies on recidivism and examined various preventive measures, such as the use of electronic monitoring or diversion (Di Tella and Schargrodsky, 2013; Mueller-Smith and T. Schnepel, 2021), rehabilitative practices (Lotti, 2022; Heller et al., 2017; Alsan et al., 2025; Arbour et al., 2024), peer effects (Bayer et al., 2009; Drago and Galbiati, 2012), better prison conditions (Drago et al., 2011), the contrast between high and minimum security levels in prisons (Chen and Shapiro, 2007), the comparison between new and old facilities (Tobón, 2022), and the distinction between open and closed prisons (Mastrobuoni and Terlizzese, 2022). Collectively, these studies indicate a consistent trend: less severe prison conditions - whether achieved through electronic monitoring, lower security level placement, or alternative approaches such as open cells — are associated with lower rates of recidivism.

2 Institutional Context

Historically, the approach towards women imprisonment in Italy has been characterized by the separation between sexes in prison and by the protection of the "incarcerated mother" (Agnella, 2023). This approach reflected a gender identity that is narrowly defined by motherhood and traditionally feminine activities, a view that traces back to early modern institutions designed for women, which aimed to preserve or rehabilitate their virtue, often related to sexual deviance. Despite reforms aimed at social reintegration and ensuring formal equality of treatment, the contemporary Italian female prison system still echoes elements of its past. According to Italian NGOs working on this topic,[11] the treatment and rehabilitation opportunities for women are limited, with a focus on stereotypically female jobs and insufficient rehabilitative, educational and vocational training. Women face inadequate healthcare, limited employment opportunities, and poor support for reintegration, exacerbating the challenges after release.

The prison system includes only four women-only facilities. The women who are not

[10] For an overview see Doleac (2023).

[11] See "Dalla parte di Antigone: Primo Rapporto sulle donne detenute in Italia (2023)" (in Italian)

housed there are assigned to one of the 52 mixed-gender prisons. The heavy reliance on the latter has been criticized for neglecting gender-specific needs, treating their issues as secondary among the various prison management challenges. A 2008 directive by the Prison Administration[12] has sought to address the absence of female-specific treatments, introducing standards that recognize and protect gender differences to mitigate the marginalization of women in prison (for example, with respect to psychological, social, and emotional needs, maternity, relationships with children, emotional well-being, specific health needs and specialized healthcare services, like gynecology, mammography,...)

2.1 The Allocation Rule

The default rule for allocating women inmates is to place them in the closest facility that can accommodate females, which can be either a women-only prison or a mixed-gender one.[13] This aims to minimize the disruption experienced by women inmates, allowing for easier access by family and legal representation, and contributing to better overall welfare.

However, this default rule is not absolute and the Department of Prison Administration (DAP) office tasked with the assignment retains some discretion. Convicted women can be assigned to facilities further away from the closest one due to security requirements,[14] capacity constraints, or specific individual needs that cannot be met by the nearest institution. Most importantly for our purposes, there is widespread awareness, in the office of the DAP tasked with the assignment, that women-only facilities offer better prison conditions and rehabilitating opportunities.[15]

The assignment therefore can be represented as a multi-stage decision tree:

0 Is the (closest) women-only prison closer to j's residence than the (closest) mixed-gender prison?

1 if **yes**, is j a security risk inmate, or are there other reasons that prompt suspension of the default rule?

- **1.1** if *no*, then assign j to the (closest) women-only prison.
- **1.2** if *yes*, then assign j to a prison different from the closest.

2 if **no**, is j a security risk inmate, or are there other reasons that prompt suspension of the default rule?

[12] *Circolare n.0308268 del 17 settembre 2008.*

[13] *Dispositivo dell'art. 14 Legge sull'ordinamento penitenziario, Titolo I, Capo III.*

[14] As a rule, members of organized crime are kept away from their familiar environment, to disrupt their criminal network, and therefore are not sent to the nearest prison.

[15] The presence of such awareness was established through interviews with DAP officials.

2.1 if *no*, does the benefit of assigning j to a women-only prison outweigh the loss of assigning her to a prison farther away from her residence?

2.1.1 if ***yes*** then assign j to the (closest) women-only prison.

2.1.2 if ***no*** then assign j to the (closest) mixed-gender prison.

2.2 if *yes*, then assign j to a prison different from the closest. This might be a women-only prison, or a mixed-gender one.

At node 1.2 of the decision tree the default rule is suspended, and closeness to j's residence is no longer a factor driving the prison assignment. In principle j could be sent to another women-only prison, but in the data we see that less than 2% of the assignments to a women-only prison involve a prison that is not the closest women-only prison. Also at node 2.2 the default rule is suspended, and the assignment could be to a women-only prison, if the latter satisfies the requirements that drive the decision when the default rule is suspended.[16] Whether or not this is the case, however, the comparison between the distances of the two closest types of prison from j's residence is no longer a relevant factor in the decision. In view of these remarks, we slightly simplify the assignment rule by assuming that at node 1.2 the assignment is indeed to a mixed-gender prison. At node 2.2, while we do not specify the trade-offs involved in the assignment decision, we just assume that they do not involve the relative distances of j's residence from the two types of (closest) prisons.

We then formalize the assignment decision as follows.

Let $A(j) \in \{W, M\}$ denote the assignment of j to a women-only (W) or mixed-gender (M) prison

Let $B_{ji}, i \in W, M$ be the perceived benefit of sending a specific inmate j to a women-only or mixed-gender prison, with $B_{jW} > B_{jM}$ for all j, though not necessarily equal for all inmates. Let $D_{ji}, i \in W, M$ be the distance between j's municipality of residence and the (closest) women-only and mixed-gender prison. Finally, let α_j denote the probability that j is either a security risk inmate or there are other reasons that might justify a deviation from the default rule, and let δ_j denote the probability (independent of relative distances D_{ji}) that a women-only prison is assigned in decision node 2.2.

Following the (simplified) decision tree, we then have that the probability of assigning j to a women-only prison is:

$$\mathbb{P}(A(j) = W) = \mathbb{1}(D_{jM} - D_{jW} \geq 0)(1 - \alpha_j) + \\ (1 - \mathbb{1}(D_{jM} - D_{jW} \geq 0))[(1 - \alpha_j)\mathbb{P}(D_{jM} - D_{jW} > \omega(B_{jM} - B_{jW})) + \alpha_j \delta_j], \tag{1}$$

[16] Given what we just noted about the data, if this were the case it would be almost surely the women-only prison closest to j's residence.

where ω is a positive scaling factor to measure the benefit of assignment to either type of prison in the same units as the distance.

To simplify the notation, let ΔD_j denote the difference $D_{jM} - D_{jW}$ and ΔB_j denote the difference $\omega(B_{jM} - B_{jW})$.

Equation (1) then implies that, as long as $\Delta D_j < 0$, that is, as long as the (closest) mixed-gender prison is closer to j' residence than the (closest) women only prison, the probability that j is assigned to a women-only prison is an increasing function of ΔD_j. Indeed, recalling that ΔB_j is strictly negative, the closer ΔD_j gets to 0 the higher is $\mathbb{P}(\Delta D_j > \Delta B_j)$.

Instead, when $\Delta D_j > 0$, the probability that j is assigned to a women-only prison does not depend on the actual value of ΔD_j.

This implies that the probability that j is assigned to a women-only prison is an increasing function of ΔD_j in its negative range and becomes a constant in the positive range.

These two features are important for our empirical strategy.

Most importantly, while ΔB_j (and possibly α_j and δ_j) could reflect selection in the assignment rule, ΔD_j is plausibly independent of any characteristic of j that could be correlated with her recidivism, once we control for a few observables of a geographic nature (more on this point in Section 4.1). Hence ΔD_j is a valid instrument for assignment to a women-only prison.

Second, the fact that for positive values of ΔD_j the assignment probability is independent of the actual value of ΔD_j implies that we can test whether, in that range, the instrument has no effect on recidivism. In other words, we can test whether the instrument affects recidivism only through its effect on the probability of assignment, and not directly, that is, we can test the exclusion restriction, which is in general an untestable assumption.

2.2 Leveraging the Proximity-Based Assignment

Before moving to the empirical specification and to the results of our analysis, it is useful to highlight that the proximity-based assignment rule, which is one of the criteria underlying the allocation rule to the two kinds of prison that we just described, can be used more broadly to construct an empirical strategy to test for the causal effect of a wide range of (binary) prison characteristics and to make inference about the unobserved policy preference (or lack thereof) concerning these characteristics.

To see that this is the case, consider again the multistage decision tree for the assignment of a given inmate to one of two types of prisons, denoted, with a label intentionally kept uninformative, as type 1 ($T1$) and type 2 ($T2$).

If it were the case that the Prison Administration (or, more generally, the authority responsible for the allocation of inmates to available correctional facilities) has a preference for, say, type 1 prisons, then we could derive for the probability of assigning inmate j to $T1$ an equation similar to (1):

$$\mathbb{P}(A(j) = T1) = \mathbb{1}(D_{jT2} - D_{jT1} \geq 0)[(1 - \alpha_j) + \alpha_j \mu_1] \quad (2)$$
$$(1 - \mathbb{1}(D_{jT2} - D_{jT1} \geq 0))[(1 - \alpha_j)\mathbb{P}(D_{jT2} - D_{jT1} > \omega(B_{jT2} - B_{jT1})) + \alpha_j \mu_1],$$

where we amended the decision tree assuming that at nodes 1.2 and 2.2, when the proximity-based assignment rule is suspended, the probability of assignment to a $T1$ prison reflects the share of that kind of prison among all prisons (μ_1) and does not depend on the relative distance ($\Delta D_j = D_{jT2} - D_{jT1}$).[17] The other symbols in the equation are a mutuated by (1) and are easily interpreted.

From (2) we would then conclude, as in Section 2.1, that the difference between distances ΔD_j is a valid instrument for assignment to type 1 prisons. It would also continue to be true that for positive values of ΔD_j the assignment probability is independent of the actual value of ΔD_j, delivering a test for the exclusion restriction.

It might, however, be the case that the Prison Administration does not have a systematic preference for either type of prisons. In this case the decison tree would be simplified, eliminating the node 2.1 (as well as nodes 2.1.1 and 2.1.2), so that the probability of assigning inmate j to $T1$ becomes:

$$\mathbb{P}(A(j) = T1) = \mathbb{1}(D_{jT2} - D_{jT1} \geq 0)(1 - \alpha_j) + \alpha_j \mu_1. \quad (3)$$

In this case the probability of an inmate to be assigned to a $T1$ prison takes the form of a step function, with a lower value ($\alpha_j \mu_1$) in the negative range of ΔD_j, and a higher value ($1 - \alpha_j + \alpha_j \mu_1$) in the positive range.

This would then suggest that a more suitable empirical strategy to assess the causal effect of assignment to $T1$ would be a regression discontinuity design, with the discontinuity taking place at 0.

Interestingly, granted the presence of a preference for proximity in the assignment rule of inmates to prisons, running a regression of the probability of being assigned to $T1$ on ΔD_j would signal, in a revealed preference spirit, whether the Prison Administration also attaches value to $T1$ prisons. If this is not the case, the regression should yield an approximate step function, with a positive discontinuity at 0, while a preference for $T1$ should result in an increasing function over the negative range of Δ_j, and a flat portion over the positive range.[18]

[17] Since type 1 and type 2 prisons are generic labels, we cannot invoke, as we did previously, the negligible share of one of the two kind of prisons.

[18] We will revisit this issue in Section 6.

3 Empirical Specification

3.1 Data

The Department of Prison Administration of the Italian Ministry of Justice granted us access to a large amount of confidential information on the universe of female prisoners who spent any time in prison between 2012 and 2022. The information includes 10,222 incarcerations.

We complement these data with yearly data on the prison population recorded by the Ministry of Justice[19] and with annual surveys on prison conditions for the years 2017, 2018, and 2019, collected in all Italian prisons by Antigone, an Italian NGO.[20]

We measure recidivism, similarly to Mastrobuoni and Terlizzese (2022), as reincarceration within three years from release from prison (we also look at shorter time windows, that is, 1 and 2 years). The data allows us to control for several individual characteristics, like age, marital status, region of residence, region or country of birth, education, sentence length, and sentence status, and the type of crime committed. We can also control for the occupancy rate of different prisons, at yearly frequency.

Antigone's surveys of prison conditions allow us to describe the main differences between women-only and mixed-gender prisons. Some differences stand out (see Table 1). Women-only prisons house on average many more female inmates, 247 against 60. In mixed-gender prisons, it is considerably less likely that inmates are allowed to freely move within the prison (26.8% versus 82.4%). The other notable differences are that inmates in women-only prisons are more likely to be allowed to work outside of the prison, are more likely to have a woman as prison director (96.9% against 64.0%) and are considerably more likely to have access to hot water, showers, obstetricians, and gynecologists. In women only prisons there is also a lower share of self-harm episodes per year (14% against 28.3% in mixed-gender prisons).

It should be noted that the unconditional rate of recidivism is in fact slightly larger among inmates who spent their sentence in a women-only prison, compared to inmates in mixed-gender prisons (about 12% vs. 11.4%). However, women-only prisons tend to be localized in or near large cities much more frequently than mixed-gender ones. Given that recidivism tends to be more prevalent in large cities (e.g., Staton-Tindall et al. (2015)), the near equality of the unconditional rates of recidivism is likely to mask important differences between the two types of prison.

[19]We collect yearly statistics from *Statistiche sulla popolazione detenuta*, reports for the years 2012-2022.

[20]The survey conducted in 2019 is the only one with a specific focus on women conditions.

3.2 Identification

We aim to estimate the causal effect on recidivism of serving the same sentence in a women-only prison, instead of a mixed-gender prison. In line with most of the recent literature, we use a linear probability model of 3-year recidivism by inmate j (R_j), depending on whether the sentence was served in a women-only prison, captured by a dummy (W_j), on fixed effects for the region of birth ($\gamma_{b(j)}$), the region of residence ($\tau_{r(j)}$) and several controls (X_j):

$$R_j = \beta_0 + \beta_1 W_j + \psi' X_j + \tau_{r(j)} + \gamma_{b(j)} + \eta_j. \tag{4}$$

Given the potential endogeneity of W_j, as argued in the previous section, we exploit the prison administration's decision rule (1), instrumenting the women-only prison dummy W_j using ΔD_j, the difference (in Km) between the (minimal) distances from the residence of inmate j of mixed-gender and women-only prisons. In our main specification, to calculate ΔD_j we use the prison of exit, which is more likely to be the one in which the largest part of the sentence was served, whereas the prison of entry might reflect only a temporary assignment. Nevertheless, most of the times prison of entry and prison of exit coincide. For robustness, we also calculate ΔD_j according to the prison of entry.

The first stage and the reduced form equations are the following:

$$\begin{pmatrix} W_j \\ R_j \end{pmatrix} = \phi + \theta \Delta D_j + \varphi X_j + \lambda_{r(j)} + \xi_{b(j)} + \epsilon_j, \tag{5}$$

where $\phi, \theta, \lambda_{r(j)}, \xi_{b(j)}$ and ϵ_j are column vectors of size 2 and φ is a $2 \times n$ matrix, where n is the number of variables included in the vector X .

Under the assumption that the relative distance from the two types of prison is conditionally exogenous, we can identify the local average treatment effect: $\hat{\beta}_1 = \hat{\theta}_2 / \hat{\theta}_1$.

Given that the distances between the municipality of residence and the two types of prison could be correlated with the characteristics of the municipality or with the ease with which the inmate can receive support from the local environment (factors that, in turn, might affect her post-release recidivism), we control for the distance between the inmate's place of residence and the actual prison, for whether the inmate resides in a metropolitan area[21], for some sociodemographic features of the inmate's residence (size of the population, share of foreign born, share of college education and population density - and their quadradic versions), for region of birth and region of residence fixed effects (the latter denoted $\xi_{b(j)}$ and $\lambda_{r(j)}$, respectively). We also control for the occupancy rate, the

[21]We classify metropolitan areas those with population exceeding 500,000 inhabitants, including the following largest cities in Italy: Rome, Milan, Naples, Turin, Palermo, Genova.

inmate's age, age squared, nationality, educational status, marital status, types of crime committed, sentence status, incarceration length, and the closing year of incarceration.[22]

A distinctive feature of our setting is that we can do better than buttressing the conditional independence assumption by including a large set of controls. Given the prison administration's decision rule (1), we know that the instrument should only be relevant in the range $\Delta D_j < 0$. This can be tested directly. Moreover, past that threshold, the instrument should not affect the probability of assignment to a women-only prison. A significant reduced-form effect past the 0 threshold would imply a violation of the exclusion restriction. Thus, in summary, we test for and estimate the following coefficients for the different samples:

Relevance: $\theta_1 > 0$ when $\Delta D_j < 0$

Decision Rule: $\theta_1 = 0$ when $\Delta D_j \geq 0$

Reduced-form Effect: $\theta_2 < 0$ when $\Delta D_j < 0$

Exclusion Restriction: $\theta_2 = 0$ when $\Delta D_j \geq 0$

Thus, we can test for something that is usually untestable in instrumental variable settings, the exclusion restriction of the instrument. It is worth noticing that in our setup the monotonicity assumption for a LATE interpretation of the 2SLS parameters should hold as well. There is no reason why inmates should become more likely to be sent to a specific prison as the distance from that prison increases.

3.3 Marginal and Average Treatment Effect

As mentioned previously, with heterogeneous treatment effects and a continuous instrument, our Two-Stage Least Squares (2SLS) estimate of equation 4 provides a variance-weighted average of covariate-specific Local Average Treatment Effects (LATEs). The latter measure the causal effect on recidivism of being assigned to a women-only prison for various groups of 'complier' inmates, that is inmates who are assigned to a women-only prison because of the relative distance from their place of residence. The interpretation of the overall 2SLS effect, being an aggregation of LATEs, is a little involved and might hide

[22]We calculate the occupancy rate as the ratio between the number of inmates and the prison's regulatory capacity. Education status is categorized as follows: without schooling, elementary school, lower secondary and higher secondary (including high school and university degrees). Marital status is defined within three categories: married or married in common law, divorced / separated, or widow, and single. We specify the types of crime as drug-related crimes, property crimes, violent crimes, and other crimes. Lastly, we define sentence status as follows: definitive sentence, under appeal, awaiting trial, mixed status, and other statuses.

potentially interesting patterns of selection induced by changes between specific pairs of values of the instrument.

However, with binary endogenous assignment (to a women-only or a mixed-gender prison) and a continuous instrumental variable, we can recover a richer array of treatment effects - the so-called marginal treatment effects (MTEs) - that better characterize the selection induced by the instrument and can be aggregated to obtain the average treatment effect (ATE), something that in general would not be possible starting from the LATE estimates.[23]

Specifically, we can estimate a propensity score, as a function of covariates and of the relative distance (our instrument), to characterize the *observed* inducement into treatment, which in our case is serving the sentence in a women-only prison. The actual assignment then results from balancing out the observed inducement with the *unobserved resistance to treatment*, which captures idiosyncratic (and unobserved by us) motives that might militate against being assigned to a women-only prison. The various LATE estimates (aggregated by the 2SLS procedure) refer to the average effect for compliers who fall in different ranges of the distribution of the unobserved resistance to treatment, while each MTE estimate is the limit as one of these ranges shrinks to a single point. Each MTE estimate then gives the treatment effect for individuals at a given point of the distribution of the unobserved resistance to treatment, in turn identified by a specific value of the propensity score.[24]

Since the percentiles of the propensity score are uniformly distributed, a simple average of the MTEs yields the average treatment effect (ATE), that is the average effect in the inmate population, the policy-relevant parameter if one were to decide to treat the entire inmate population. In Section 5.1, based on this framework, we estimate "optimal" policy relevant treatment effects to identify the ideal location of a new women-only prison.

4 Results

4.1 Balance Tests

We start by looking at whether the individual characteristics of inmates assigned to women-only or mixed-gender prisons differ. Since women-only prisons and their inmates

[23]MTE were first introduced by Björklund and Moffitt (1987) and Heckman and Vytlacil (1999) and later extended in a series of papers (see, for example, Heckman et al. (2006)). For an introduction to marginal treatment effects see Cornelissen et al. (2016).

[24]Carneiro et al. (2011) has an application of MTE to estimate returns to college education where, similar to our case, distance is used as an instrument.

tend to be located in larger cities, with features (e.g., size, density, etc.) that may correlate with crime, we control for such features. Thus, we test the statistical significance of the *conditional* differences among the characteristics of the two groups of prisoners. In other words, we provide support for the assumption of conditional independence of our instrument.

Figure 1, left panel, reports the coefficients of a linear probability model where the dependent variable is a dummy W_j equal to 1 if j is assigned to a women-only prison (0 otherwise), controlling for region of the prison and of the inmate's birth fixed effects, as well as for various characteristics of the municipality of residence (for continuous variables we add a quadratic polynomial): population, population density, a dummy for municipalities with more than 500,000 inhabitants, the share of foreigh born residents and the share with university degree. In line with the potential endogeneity of W_j, we find that several variables display a statistically significant coefficient and an F-test for joint significance of 187, with a zero p-value. In the right panel, we repeat a similar analysis, considering as the dependent variable a dummy equal to 1 if $\Delta D_j < 0$. We see that in this case the observable differences between inmates, depending on whether the relative distance of the two types of prisons from their residence is positive or negative, are almost always statistically insignificant (the joint significance test has a p-value of 0.32, well above the conventional thresholds). The lack of correlation between ΔD_j and inmates' characteristics provides reassurance regarding the conditional exogeneity of our instrument.

Table 2 shows the descriptive statistics that divide inmates according to their relative distance, $\Delta D_j < 0$ and $\Delta D_j \geq 0$. The two samples are generally comparable with some notable exceptions. As mentioned earlier, women with a positive D are more likely resident in metropolitan areas, where women-only prisons are located (22% for $\Delta D_j < 0$; 65% for $\Delta D_j \geq 0$). In turn, the share of Italians is much lower for $\Delta D_j \geq 0$ as metropolitan areas are more likely to attract migrants (73% for $\Delta D_j < 0$; 48% for $\Delta D_j \geq 0$). For this reason, we always control for whether inmates reside in cities with more than 500,000 inhabitants.

Women closer to women-only prisons are also three years younger, 22p.p more likely to be single and 8p.p. more likely to have committed a property crime. The two samples show a marked difference in the three-year recidivism with lower rates for women closer to mixed-gender prisons (10.02% for $\Delta D_j < 0$; 15.14% for $\Delta D_j \geq 0$): this difference might be driven by such imbalances in terms of citizenship, demographics, and type of crime.

4.2 Non-parametric Evidence

Figure 2 shows the fraction of inmates assigned to a women-only prison as a function of ΔD_j, without any controls. We see that the fraction increases with ΔD_j, as long as $\Delta D_j < 0$, while it is broadly constant in the range $\Delta D_j \geq 0$. In plain English, inmates whose residence is only slightly nearer to a mixed-gender prison than to a women-only prison are more likely to be assigned to the latter than inmates whose residence is much closer to a mixed-gender prison than to a women-only one. Once the women-only prison is closer, the probability of being assigned there is no longer significantly affected by the relative distance (in line with the allocation rule described in Section 2.1).

The figure is a graphical representation of the first stage of our Instrumental Variable (IV) strategy. The increasing part, for values of $\Delta D_j < 0$, highlights the relevance of the instrument. The broadly flat part, for values of $\Delta D_j \geq 0$, shows that the instrument does not predict assignment to a women-only prison for inmates whose residence is closer to that kind of prison, in line with the decision rule described in Section 2.1.

In Figure 3, the blue line shows the decreasing relationship between our instrument (in the negative range) and recidivism, again without controls. This is a graphical representation of the reduced form in our IV strategy. The fairly flat green line shown in the figure illustrates the lack of relationship between the instrument and a measure of predicted recidivism. The latter is obtained from a logistic regression model, in which recidivism is projected on all the observables considered in Figure 2.[25] The flat line, implying the absence of a correlation between ΔD_j and (a function of) these observables, confirms that the relationship between the instrument and (actual) recidivism, shown by the blue line, does not capture the effect of other observables. Therefore, it is reassuring that the exogeneity of the instrument is present.

4.3 Instrumental Variable Regressions

Table 3 presents the estimates of equation 5, with the first stage in panel A, the reduced form in panel B and the IV estimate in panel C. The bottom part of the table reports the Kleibergen-Paap statistics (KP F-Stat) for weak identification and the fitted values of recidivism, under the two alternative assumptions that *all* the inmates are either in a mixed-gender ($\hat{R}_M$) or in a women-only prison ($\hat{R}_W$), leaving all other covariates unchanged. In the first 4 columns, the estimates are restricted to the sample of inmates whose residence is closer to a mixed-gender prison than to a women-only one ($\Delta D_j < 0$), in the following 4 columns the sample only includes inmates whose residence is closer to a women-only prison ($\Delta D_j \geq 0$).

[25] Appendix Figure A3 shows the high correlation between recidivism and predicted recidivism.

In column 1 we include no controls nor fixed effects, with the exception of the dummy indicating whether or not the residence of the inmate is within a metropolitan area.[26] Hence the coefficient of interest (the effect on recidivism of being assigned to a women-only prison; β_1 in the notation of equation 4) is identified by the variability that results from all possible reciprocal positions of the inmates residence, the (closest) mixed-gender prison and the (closest) women-only prison, provided that we only compare inmates whose residence is equally within (or equally outside) a metropolitan area (though not necessarily the same). We see that both the first stage and the reduced form are highly significant, confirming the relevance of the instrument, with a large value of KP F-Stat, and the presence of a reduced-form effect (in the notation of equation 5, $\theta_1 > 0$ and $\theta_2 < 0$ when $\Delta D_j < 0$). The overall effect is about -8p.p. Hence, the fitted value of recidivism would drop from 13.8%, if all inmates were housed in mixed-gender prisons, to 5.9%, if they were all moved to women-only prisons.

In column 2 we include region of birth and region of residence fixed effects. Including the latter greatly restricts the variability of the instrument used to identify the causal effect. In particular, since different regions might be more or less conducive to recidivism,[27] and, at the same time, might be differentially endowed with women-only prisons,[28] comparing inmates who live in the same region allows us to clean the effect of being assigned to a women-only prison from the possible (positive) correlation between recidivism and presence of women-only prisons. Since this correlation would partially offset the pure effect of the assignment, we see that the latter increases substantially (in absolute value), with an estimated value of -16.1p.p. In this case, therefore, the fitted value of recidivism would drop from 17.7%, if all inmates were housed in mixed-gender prisons, to 1.6%, if they were all moved to women-only prisons.

In columns 3 and 4 we also include the controls listed in Section 3.2, so we further restrict the variability identifying the coefficient of interest, as we compare inmates who differ in terms of the relative distance to the two kinds of prisons, live in the same region, and also have the same values for all the included controls (in column 4 we add the quadratic version of the continuous controls). Both the first stage and the reduced form estimates remain significant, with little change in value, and the overall effect becomes slightly less negative, to about -14p.p.

As mentioned, we can directly test the exclusion restriction, namely the requirement that the instrument affects recidivism only through its effect on prison assignment.

[26]This dummy will always be included in all the specifications discussed in this and in later Sections.

[27]For example, regions might differ for the presence of organized crime or for labor opportunities available to released inmates.

[28]For example, it might be the case that women-only prisons were created precisely in those regions where there is more recidivism.

Columns 5 to 8 replicate the same specifications of columns 1 to 4 on the sample of inmates whose residence is closer to a women-only prison ($\Delta D_j \geq 0$). The assignment rule described in Section 2.1 implies that positive values of the variable ΔD_j should have no effect on the prison assignment. Therefore, under the validity of the exclusion restriction, positive values of the variable ΔD_j should also have no effect on recidivism. Hence, in columns 5 to 8, we should observe both an insignificant first stage and an insignificant reduced form. This is precisely what Table 3 shows.

4.4 Robustness Tests

As mentioned before, we compute the instrument ΔD_j on the basis of the prison of exit. In Table A1 we verify that our results are robust to this choice by computing the instrument on the basis of the prison of entry. The results are very similar to the ones in Table 3, with the range of IV estimates from 7 to 15p.p.

In Table A2 we compute recidivism as re-incarceration within a shorter window, namely 1-year and 2-year after release. We show that the effect of spending the sentence in a women-only prison is already apparent in the shorter time window, with the IV estimate ranging between 6 and 13 percentage points.

A further robustness check on the definition of recidivism is presented in Table A3. Instead of including all re-incarcerations after release, irrespective of whether the latter occurred after a definitive sentence,[29] we define recidivism as reincarceration after release from a definitive sentence. While this reduces the sample size and yields somewhat less precise estimates, the effect of assignment to a women-only prison are broadly similar to those observed in the main specification.

Finally, we consider two dimensions of potential heterogeneity in our results: foreign status and the type of crime. Our sample includes foreign inmates, whose recidivism might respond differently to the prison assignment. Table A4 replicates the analysis on the sample that only includes Italian inmates. The IV estimates are somewhat larger (in absolute value) than those in Table 3, particularly in the more restricted specifications, and remain highly significant.[30] Table A5 replicates the analysis distinguishing between violent and property crimes and shows that for both types of crime the assignment to a women-only prison reduces recidivism. Depending on the specification, the effect is somewhat stronger for one or the other of the two types of crime, and is in the same ballpark of what found for the entire sample.

[29]The Italian system has two levels of appeal, but a condemned inmate can be in prison while waiting for one of the appeals or because she decided not to appeal.

[30]We do not include a separate test for foreign inmates due to the small sample size.

4.5 Results for Marginal and Average Treatment Effects

We estimate the propensity to be assigned to a women-only prison using a probit model. The controls are the same used in Table 3, although we get similar results when we do not use regional fixed effects. Figure A4 shows the support of the histogram of the propensity score for women-only (treated) and mixed-gender (control) prisons and highlights two important facts: the support in the two samples is common and covers the entire range $[0, 1]$. These imply that we can use the semiparametric local instrumental variables MTE specification (see Heckman and Vytlacil, 2001a). This allows us to compare the semiparametric and more data-intensive model to the more parsimonious parametric one.

Figure 4 shows that the semiparametric local linear as well as the linear MTE curves are slightly upward sloping, which means that inmates with a larger unobserved resistance to end up in a women-only prison show smaller reductions in recidivism.

Inmates are thus selected on the basis of unobservables to maximize treatment effects. However, the slope is nearly flat and Table 4 shows that ATEs are fairly close to the main local average treatment effects estimated in Table 3. This means, importantly, that scaling up the treatment to the entire prison population, thus increasing the number of women-only prisons by a factor of about 4 (right now about one-fifth of women are housed in women-only facilities), would drastically reduce recidivism.

5 Two Policy Experiments

In this section, we present two simple policy experiments, aimed at leveraging the results obtained on the effect on recidivism of women-only assignment to explore possible targeted reorganizations of the current Italian prison system. The first experiment takes advantage of the MTEs estimated above and involves the construction of a single, new women-only prison, strategically located to serve areas with high concentrations of incarcerated women who, in addition, have larger (in absolute value) MTE. The second experiment explores the possibility to repurpose, within a given region, some of the existing mixed-gender prisons to cater exclusively for women, by reallocating the current male population. This approach would transform selected mixed-gender prisons into women-only facilities.

5.1 Optimal Policy Relevant Treatment Effects

Policy makers might want to know in which municipality, out of the 7,896 existing in Italy, a new women-only prison should be built to achieve the maximal reduction in women recidivism. For each municipality, we simulate the opening of one new women-only prison (the policy) and estimate the corresponding Policy Relevant Treatment Effects (PRTEs), that is, the marginal treatment effects for those inmates whose probability of ending up in a women-only prison changes in response to the policy (see Heckman and Vytlacil, 2001b). The simulations are based on the entire 9 years of data, and the corresponding distribution of inmates across municipalities.[31] This allows us to identify the optimal location ℓ of an additional new women-only prison, as the one that minimizes overall predicted recidivism (evaluated at the average value of the controls X):

$$\min_{\ell} E(\widehat{R}_i^{\ell} - \widehat{R}_i | X = \overline{x}) = \min_{\ell} \int_0^1 MTE(\overline{x}, u_D)\omega^{\ell}(\overline{x}, u_D) du_D,$$

where the weights are $\omega^{\ell}(\overline{x}, u_{iD}) = F_{P|\overline{X}}(u_{iD}|\overline{x}) - F_{P^{\ell}|\overline{X}}(u_{iD}|\overline{x})$, with F_P and $F_{P^{\ell}}$ denoting the distribution function of the unobserved resistance to the treatment, conditional on the *status quo* and the addition of a new women-only prison in location ℓ, respectively. In words, the weights reflect the change in the distribution of the unobserved resistance to end up in a women-only prison as the propensity scores change when a prison is build in municipality ℓ.[32] Notice that opening a new women-only prison can only increase the propensity scores, which implies that all weights are larger than zero.[33]

Figure 5 shows the distribution of the PRTEs in the left panel and the prison size requirement (the change in the propensity score times the average yearly number of prison inmates) in the right panel. Based on the current distribution of women-only prisons (Rome, Venice, Trani and Naples), the largest demand for a new women-only prison is in the North-West, in the municipality of Zinasco, province of Pavia, 34 km south of Milan, 100 km East of Turin, and 75 km North of Genoa. The prison in the optimal location would accomodate around 250 women and lead to a reduction in recidivism of about 30 inmates each year (250 times the average PRTE, -0.12). Given an estimate of the prison cost per inmate of about 50,000 euro (see Barbarino and Mastrobuoni, 2014), this would lead each year to about 1.5 million euros of savings.

5.2 Re-allocation to Existing Prisons

A bolder policy change would ban mixed-gender prison altogether. Here we propose a way to achieve such change without building new prisons.

The official prison capacity for women inmates in Italy, besides the four existing women-only prisons, is of almost 1600 places, and is present in almost all the Italian

[31]We are implicitly assuming that such distribution does not change much over time, an assumption which is supported by the data.

[32]In line with the literature, we are assuming policy invariance, meaning that the MTEs do not depend on ℓ, thus do not change with the policy (see Heckman and Vytlacil, 2007).

[33]In line with the evidence on the first stage, we set propensity scores when the difference in distance is below -110 km to zero.

regions. Setting aside possible technical difficulties in repurposing existing buildings, here we conduct the following thought experiment: within each region (except the two smaller regions which have no official capacity for women), we imagine to repurpose one or more of the existing mixed-gender prisons, transforming them into women-only facilities, so as to create enough places to host all the official capacity for women currently housed in mixed-gender prisons. In this way, all the women inmates envisaged in the official capacity of the Italian prison system could be housed in women-only prisons.

Clearly, it is only by chance that in a region there is one prison that has (or several ones that in total have) a number of places exactly equal to the total number of places for women inmates in the region, so that the women reallocated would free just enough places to house the males that were originally in the repurposed prison.[34] Nevertheless we verified that in each region the reallocation can be done in such a way that the excess number of places envisaged for male inmates almost never exceeds 3 percent of the regional residual official prison capacity for males.[35] We believe that, given the limited size of the deficit, and assuming that there could be some reorganization and rationalization of the physical space, the reallocation could be easily accommodated.

Having checked that, at least on the basis of the current distribution of the official prison capacity, the Italian prison system could become a women-only one, we can easily compute the reduction in recidivism that this would entail, given our estimates. The shift to a women-only prison would reduce the yearly number of recidivating women by a minimum of 126 to a maximum of 256 (corresponding to the range of our estimates of the effect on recidivism of women-only prison).

6 Mechanisms

As outlined in the introduction, in women-only correctional facilities it might be easier to tailor prison structures to better suit the unique needs of female inmates. Indeed, women-only prisons differ from mixed-gender ones along several dimensions (as documented in Table 1). They typically offer enhanced healthcare services, improved living areas, and a more comprehensive array of educational and vocational opportunities for inmates.

In addition, the kind of interactions and peer effects among inmates and with the prison personnel (correctional officers, but also a wide range of other staff who support the daily operations and rehabilitation programs within the prison) are likely different in a facility that only hosts women.

[34] In the current regional distribution of prison capacity, this coincidence only happens once.

[35] In fact, in most cases the male deficit is smaller than 1 percent. Only in one case, in the sparsely populated region of Sardinia, it reaches 11 percent.

As a result, we deal in this paper with a bundled treatment, and it is therefore difficult to single out one (or more) of these aspects as being the main one(s) responsible for the effects on recidivism that our analysis identifies. In fact, the ultimate driver of our results might be the complex interaction among these various aspects, so that considering them one at a time would not be able to reveal their true role.

We try to tackle this difficulty in two ways. First, focusing more directly on the specific nature of women-only prisons, we will look at factors that might enable, or prevent, the design and organization of facilities tailored to women needs.

Secondly, broadening the scope of our analysis, we will select some qualitative and quantitative features of the prison facilities that we conjecture might be drivers of our results, and we repeat our analysis on the effect of two kinds of prisons on recidivism, but instead of contrasting women-only and mixed-gender prisons, we use factor analysis to combine these different features and we split the prisons into two groups, according to the (latent) factor that summarizes them. The idea is to verify whether these features, irrespective of whether the prison is a women-only or a mixed-gender one, are associated with a reduction in recidivism. This approach leaves unanswered the question of the ultimate mechanism driving our results, as we do not know whether women-only prisons were established in facilities with those characteristics, or rather, the facilities acquired those characteristics because they were hosting only women. However, it has the advantage of potentially identifying proximate causes of a reduction in recidivism, that might be implemented irrespective of the gender composition of prisons.

Role of Section Size and Female Directorship As to the first approach, one plausible conjecture is that tailoring prison structures to better suit the specific needs of female inmates is both a more pressing concern and easier to implement when the presence of women inmates within a given facility is sufficiently sizeable.[36]

To test this conjecture, in Table A6 we compute again the effect on recidivism of being assigned to a women-only prison (as in Table 3) but, in columns 1 and 2, respectively, we restrict the control group to mixed-gender prisons with a relatively high, and relatively low, proportion of women inmates (above and below the median of the distribution). Consistently with the conjecture, the effect on recidivism of the assignment to women-only prisons is more than 3 times larger, in absolute value (although not precisely estimated) when the control group only includes prisons with relatively fewer women.[37] Similarly,

[36]According to a report from the Italian Ministry of Interior (Palmisano, 2015), the limited number of women inmates in mixed-gender facilities often leads to neglect and a lack of engagement in activities tailored to their needs, posing challenges for implementing targeted interventions.

[37]For the sake of space, we only report the most restricted specification with all the controls and the fixed effects. The appropriate comparison, therefore, is with column 4 of Table 3.

in columns 3 and 4 we consider the possible role of absolute, instead of relative, size. In column 3 the control group only includes mixed-gender prisons with relatively larger (above the median) female sections, while in column 4 it includes only mixed-gender prisons in which the size of the female section is below the median. Also in this case, the estimated coefficient is about 2 times larger, in absolute value, when the control group is the one with smaller sized female sections.

In columns 5 and 6, we consider another factor, not necessarily alternative, that might facilitate (or hinder) the organization of correctional facilities better suited to the needs of women inmates: the gender of the prison director. The plausible conjecture, in this case, is that a woman director is more attuned to other women needs. All prison directors in women-only prisons are women. In column 5, we use as a control group only mixed-gender prisons with a male director, while in column 6, we focus on mixed-gender prison with a female director. Again, the results lend support to the conjecture. The effect on recidivism of being assigned to a women-only prison is more than 2 times stronger when the control group only includes mixed-gender prisons directed by a men. This finding echoes research in political economy that highlights how the presence of women leaders can improve the welfare of the female population (Chattopadhyay and Duflo, 2004; Pande and Ford, 2009; Bochenkova et al., 2023). However, it is important to note that this result should not immediately advocate for replacing male directors with female ones, as mixed-gender prisons predominantly house male inmates, and we have so far no evidence of the potential impacts of having a woman director on the male population.

Role of Prison Quality Moving now to our second approach, we collected data from Antigone surveys on a number of features of prison facilities and their organization, such as, for example, the availability of working facilities and opportunities to carry out work outside the prison walls or inside for external firms, the presence and quality of sanitation and health services, the available space for each inmate, the opening of cells during the day (beyond the minimum required by law), the availability of spaces for social activities, etc. (see Figure A5 for the complete list of features considered).

Most of these features have an intuitively clear effect on the quality of life of inmates and, therefore, might affect their perception of imprisonment and their behavior after release. Previous research (see Mastrobuoni and Terlizzese, 2022) has shown that some of these features facilitate a smoother return to normal life after serving a sentence, thus reducing recidivism.

Taking some combination of these features as representative of an overall quality of prisons and prison life, we want to test whether this quality is the proximate cause of our results.

To this effect, we use factor analysis to extract a latent factor from these features, which we interpret as a single synthetic quality indicator. The estimated factor can be expressed as a linear combination of the observed features, with coefficients proportional to the estimated factor loadings. The square dots in the upper panel of Figure A5 show, for each of the features considered, the corresponding coefficient in the said linear combination. The bottom panel shows the corresponding factor loadings (the bivariate coefficients).

A caveat about this approach concerns the limitations of the data we were able to access. First, we observe the characteristics of the prisons hosting female inmates only for the period 2017–2019, whereas ideally we would need to measure these features as they prevailed at the time when the inmates included in our sample were serving their sentences. Secondly, in the case of mixed-gender prisons, most features are an average for the entire facility, rather than just for the section that hosts women. Thirdly, several features are missing for some of the facilities, and we suspect that the presence of missing data is not random, being potentially correlated with either weaknesses in the governance of a given facility, or with the attempt to hide unsatisfactory aspects of the prison or of the prison life.[38] All this means that the data we could access are affected by a non negligible measurement error, which is likely to attenuate the statistical significance of our estimates and could partially explain why in some cases the relationship between a feature and the latent factor is somewhat different from what one might intuitively expect.

Returning to Figure A5 we see that, in general, features clearly indicative of better quality contribute positively to the latent factor, while features suggestive of poorer quality contribute negatively. In general, we therefore believe that the estimated coefficients broadly support our interpretation of the latent factor as a synthetic indicator of overall prison quality.

We use the synthetic quality indicator to split the prisons in two quality groups: those in which the value of the latent factor is above the median (high quality), and those in which it is below (low quality). Figure A5 shows the regression coefficients of a linear probability model in which such indicator D_Q is regressed on the whole vector of observed features (upper panel), as well as the bivariate counterparts (lower panel). The resulting coefficients display a pattern similar to the one estimated for the continuous measure of the latent factor: variables that are most strongly associated with higher quality in the continuous specification also show the largest positive effects in the dichotomous one, and likewise for negative associations. This reassures us that the binary version retains the

[38] To partially control for this, we control for the sum of the number of missing variables for each given facility.

essential structure of the continuous latent measure.

On average, the synthetic quality indicator is about 1.5 standard deviations higher in women-only prisons than in mixed-gender prisons, confirming that the former exhibit systematically better conditions in the dimensions considered.

Following the general framework presented in Section 2.2, we then compute, for each female inmate, the distance from her residence to the closest prison belonging to the low quality class and to the closest prison belonging to the high quality class, and we take the difference between these two distances. A negative difference then means that the closest high quality prison is farther from the inmate's residence than the closest low quality prison. We use this distance difference as an instrument for exposure to higher-quality prison conditions in our estimation of the effect of prison quality on recidivism.

Figures A6 and A7 display the first stage and the reduced form. The first stage plots the fraction of inmates assigned to the high quality class against the distance difference, while the reduced form shows the corresponding fraction of recidivating inmates.

Given the underlying preference of the prison administration for assigning inmates to facilities closer to their residence — a preference which, as we noted before, is present in most countries — two scenarios are possible according to the framework presented in Section 2.2. If the administration has no clear preference for the features that define the high quality class, we should observe a fairly flat pattern at a low level for negative differences, and an equally flat pattern at a higher level for positive differences, as in equation (3). In this case, absent other considerations, inmates would tend not to be assigned to the high quality class when it is farther away from their residence, while they would predominantly be assigned to it when it is closer, resulting in a discrete jump at zero. Alternatively, if the prison administration trades off proximity to residence with some desirable features of the high quality class (as in our earlier analysis of mixed-gender vs. female-only prisons), we should observe an increasing pattern to the left of zero and a flat pattern to the right, as in equation (2).

The first-stage results in Figure A6 show a combination of these two possibilities, with a statistically significant and increasing relationship for negative distances, a non negligible jump that is spread across a relatively narrow range of small positive distance differences (0-30 km), and a flatter relationship, at a high level, for larger positive differences.

We interpret this as implying that, by and large, the combination of features that we interpret as a synthetic indicator of prison quality is also heeded by the Prison Administration when deciding about the assignment of inmates to alternative facilities, trading off proximity with quality.

At the same time, the presence of a jump suggests that this preference is less unambiguously and universally held than the one regarding the gender composition of the prison. Some of the officials who happen to be responsible, at each given moment, for the inmates assignment, might be indifferent about the various features aggregated by our latent indicator, being guided mainly by proximity considerations. This would explain a sizable jump at 0. The fact that the jump is not fully concentrated at 0 might be explained by the measurement error that affects our data. In particular, some prisons that we classify as high quality might be low quality in the perception of Prison Administration officials (as we measure with errors the various features, or they might weight the different features differently form our estimated linear combination, or they might pay attention to features we do not observe) and this might lead us to record in bins associated to larger differences observations that in fact belong to bin associated to smaller ones.[39]

An alternative, possibly complementary way to rationalize the stepwise jump in the range (0-30) is to suppose that there are some constraints in implementing the assignment to the closest prison (irrespective to its quality), due, for example, to temporary overcrowding of the destination prison, or bureaucratic procedures. Overcoming these constraints involves some cost, which might not be worth paying when the alternative to the closest prison is another prison that is also fairly close to the inmate's residence. Hence, the full jump would not occur entirely at 0 and would be spread over some distance range.

Turning to Figure A7, which offers a graphical representation of the reduced form, we see that overall recidivism falls as the negative difference between the two distances gets smaller (in absolute value), i.e. as the fraction of inmates sent to an high quality prison, rather than to a closer low quality prison, increases. There is also some evidence that recidivism falls in the neighborhood of 0, although statistical power becomes an issue when restricting the data to inmates with small differences in distance.

The insights we get from Figures A6 and A7 are confirmed by the regressions presented in Table A7. Panel A shows the results of the first stage regression. Given the stepwise jump in the range (0-30) the first stage is now significant both for negative and for positive values of the distance differences, differently from what we showed for gender based prisons. Accordingly, we do not restrict the sample to negative differences, also in

[39]Suppose, for example, that the distances between the residence of inmate j and prisons P_1, P_2 and P_3 are, respectively, 10, 20 and 15 km. Suppose moreover that we classify P_1 and P_3 as high quality, and P_2 as low quality, and that the inmate is assigned to P_1. Accordingly, we include the observation for inmate j in the difference bin (10-20), as the distance from the closest high quality prison (P_1) is 10 km smaller than the distance from the closest low quality prison (P_2). If however, P_3 were, in the perception of the Prison Administrators doing the assignment, a low quality prison, the correct recording of inmate j observation should have been in the bin (0,10).

panels B and C, but the results would be similar if we had just considered the negative range of the instrument.

The first stage confirms that the probability of being assigned to a prison of higher quality (more precisely, belonging to the upper half of the latent factor distribution) increases the closer to the inmate residence the latter is relative to the (closer) lower quality facility. This suggests that the latent factor captures features that the prison administration considers relevant when deciding about prison assignments.

Panels B and C of Table A7 present the reduced form and instrumental-variable estimates of the effect of assignment to a higher-quality prison on recidivism. Focusing on the IV estimates, the table shows a significant and sizable negative effect of exposure to a higher quality facility on recidivism, as long as we include fixed effects of the region of birth and the region of residence. The size of the estimated coefficient is only slightly smaller (in absolute terms) when we include other controls in the regression, and remain significant.

Nevertheless, in view of the limitations of the data we have available, we view these results more as suggestive of a promising avenue for further analysis, conditional on the availability of better and more comprehensive measures of prison characteristics, rather than a definitive identification of the mechanism underlying our results.

Moreover, as already mentioned, even sharper results on a causal role for prison quality on recidivism of women inmates would not yet clinch the analysis of the mechanisms at play. While women-only prison indeed display higher quality, it is not clear whether this is a consequence of their being women-only or the reflection of women-only prisons being located from the start in facilities with a higher quality. However, we believe that the foregoing analysis, in addition to its intrinsic interest, is illustrative of the wide applicability of our IV approach in studying the causal effect of assignment to different facilities, whenever the latter also reflects an underlying preference for proximity.

7 Conclusions

Despite a nearly 60% increase in the global population of incarcerated women over the past two decades, the issues surrounding women's incarceration remain largely neglected. In particular, little is known about the possible effects of serving a sentence in a separate section within prisons that mostly house male inmates, as is the case in 143 countries, as opposed to serving the same sentence in a dedicated, women-only prison, as is the case in 79 countries.

In this study, we show that, once we correct for the possible selection and other confounding factors, serving a sentence in a women-only prison, rather than in a mixed-gender one, substantially reduces recidivism in the three years after release.

Our results are based on Italian data, where both women-only and mixed-gender prisons coexist and where the institutional details of the process allocating women convict to prisons allow us to mimic a quasi-random assignment, and thus identify the causal effect on recidivism of the two types of prison.

Our empirical approach—leveraging a proximity-based assignment rule aligned with international correctional standards—offers broader applicability across countries and can be used to identify the causal impact of various prison characteristics on recidivism and other correctional outcomes, beyond gender composition. This framework opens a promising avenue for future research in the field of prison studies.

Women-only prisons display better features along several dimensions, including health, social and work services. At the same time, being women-only likely affects the kind of interactions and peer effects among inmates and with the prison personnel. It is difficult to single out one of these dimensions, or more likely a combination thereof, as being responsible for our results. We address this difficulty in two ways.

First, we show that both the number (in relative or absolute terms) of women inmates within a given facility, and the gender of the director, are relevant factors affecting the difference in recidivism that we estimate. This suggests that the design and organization of correctional facilities customized around women needs and fostering rehabilitation initiatives specifically targeted at women inmates might require a minimum scale and an appropriately attuned sensibility.

Secondly, as an example of the broader applicability of our empirical strategy, we use factor analysis to create a synthetic indicator of the quality of a prison and we use it to split the universe of prisons in two classes, high and low quality (irrespective of gender composition). We then use the same IV approach, leveraging the proximity-based assignment rule, to identify a causal effect on recidivism of prison quality. Since the synthetic quality indicator is on average higher for women-only prisons, we conclude that prison quality is a proximate cause of our main result. This of course still fails to pin down the precise mechanism, as we do not know why women-only prisons show a better quality.

A more in-depth analysis of the mechanisms underlying the causal effect identified in this study, as well as the collection of more accurate, time-varying data on the characteristics of the various correctional facilities, remain important topics for future research.

Our results are relevant for governments and international organizations interested in designing policies to reduce crime with a specific focus on women's incarceration. With this purpose, we also conduct two simple policy experiments to show how to leverage the results of the analysis to guide the reorganization of a prison system.

References

Abrams, D. S. (2012). Estimating the deterrent effect of incarceration using sentencing enhancements. *American Economic Journal: Applied Economics 4* (4), 32–56.

Agnella, C. (2023). Breve storia della detenzione femminile. Technical report, Antigone - Approfondimenti.

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

Alsan, M., A. Barnett, P. Hull, and C. S. Yang (2025). "something works" in u.s. jails: Misconduct and recidivism effects of the ignite program*. *The Quarterly Journal of Economics 140*, 1367–1415.

Arbour, W., G. Lacroix, and S. Marchand (2024). Prison rehabilitation programs and recidivism: evidence from variations in availability. *Journal of Human Resources*.

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

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

Bhuller, M., G. B. Dahl, K. V. Løken, and M. Mogstad (2020). Incarceration, recidivism, and employment. *American Economic Review 110* (6), 1486–1518.

Björklund, A. and R. Moffitt (1987). The estimation of wage gains and welfare gains in self-selection models. *The Review of Economics and Statistics 69* (1), 42–49.

Bochenkova, A., P. Buonanno, and S. Galletta (2023). Fighting violence against women: The role of female political representation. *Journal of Development Economics 164*, 103140.

Buonanno, P. and S. Raphael (2013). Incarceration and incapacitation: Evidence from the 2006 italian collective pardon. *American Economic Review 103* (6), 2437–2465.

Card, D. (1995). Using geographic variation in college proximity to estimate the return to schooling.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica 69*(5), 1127–1160.

Carneiro, P., J. J. Heckman, and E. J. Vytlacil (2011). Estimating marginal returns to education. *American Economic Review 101*(6), 2754–2781.

Chattopadhyay, R. and E. Duflo (2004). Women as Policy Makers: Evidence from a Randomized Policy Experiment in India. *Econometrica 72*(5), 1409–1443.

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

Cobbina, J. E., B. M. Huebner, and M. T. Berg (2012). Men, women, and postrelease offending: An examination of the nature of the link between relational ties and recidivism. *Crime & Delinquency 58*(3), 331–361.

Cornelissen, T., C. Dustmann, A. Raute, and U. Schönberg (2016). From late to mte: Alternative methods for the evaluation of policy interventions. *Labour Economics 41*, 47–60.

Council of Europe, C. (2022). Report to the italian government on the periodic visit to italy carried out by the european committee for the prevention of torture and inhuman or degrading treatment or punishment (cpt). Technical report, Council of Europe anti-torture Committee (CPT).

Di Tella, R. and E. Schargrodsky (2013). Criminal recidivism after prison and electronic monitoring. *Journal of political Economy 121*(1), 28–73.

Doleac, J. L. (2023, June). Encouraging desistance from crime. *Journal of Economic Literature 61*(2), 383–427.

Drago, F. and R. Galbiati (2012). Indirect effects of a policy altering criminal behavior: Evidence from the italian prison experiment. *American Economic Journal: Applied Economics 4*(2), 199–218.

Drago, F., R. Galbiati, and P. Vertova (2009). The deterrent effects of prison: Evidence from a natural experiment. *Journal of political Economy 117*(2), 257–280.

Drago, F., R. Galbiati, and P. Vertova (2011). Prison conditions and recidivism. *American law and economics review 13*(1), 103–130.

Heckman, J. and E. Vytlacil (2001a). Local instrumental variables. In *Nonlinear Statistical Modeling: Proceedings of the Thirteenth International Symposium in Economic*

Theory and Econometrics: Essays in the Honor of Takeshi Amemiya, pp. 1–46. New York: Cambridge University Press.

Heckman, J. J., S. Urzua, and E. Vytlacil (2006). Understanding instrumental variables in models with essential heterogeneity. *The Review of Economics and Statistics 88*(3), 389–432.

Heckman, J. J. and E. Vytlacil (2001b, May). Policy-relevant treatment effects. *American Economic Review 91*(2), 107–111.

Heckman, J. J. and E. J. Vytlacil (1999). Local instrumental variables and latent variable models for identifying and bounding treatment effects. *Proceedings of the National Academy of Sciences 96*(8), 4730–4734.

Heckman, J. J. and E. J. Vytlacil (2007). Econometric evaluation of social programs, part ii: Using the marginal treatment effect to organize alternative econometric estimators to evaluate social programs, and to forecast their effects in new environments. *Handbook of econometrics 6*, 4875–5143.

Helland, E. and A. Tabarrok (2007). Does three strikes deter? a nonparametric estimation. *Journal of Human Resources 42*(2), 309–330.

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

Katz, L., S. D. Levitt, and E. Shustorovich (2003). Prison conditions, capital punishment, and deterrence. *American Law and Economics Review 5*(2), 318–343.

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

Levitt, S. D. (1996). The effect of prison population size on crime rates: Evidence from prison overcrowding litigation. *The quarterly journal of economics 111*(2), 319–351.

Lotti, G. (2022). Tough on young offenders: Harmful or helpful? *Journal of Human Resources 57*(4), 1276–1310.

Mastrobuoni, G. and D. Terlizzese (2022). Leave the door open? prison conditions and recidivism. *American Economic Journal: Applied Economics 14*(4), 200–233.

Mathiassen, C. (2017). Being a woman in mixed-gender prisons. *Scandinavian Penal History, Culture and Prison Practice: Embraced By the Welfare State?*, 377–403.

Morash, M., D. A. Kashy, M. N. Bohmert, J. E. Cobbina, and S. W. Smith (2017). Women at the nexus of correctional and social policies: Implications for recidivism risk. *British Journal of Criminology 57*(2), 441–462.

Mueller-Smith, M. and K. T. Schnepel (2021). Diversion in the criminal justice system. *The Review of Economic Studies 88*(2), 883–936.

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

Palmisano, R. (2015). Donne e carcere - tema per stati generali dell'esecuzione penale -. Technical report, Italian Ministry of Interior, Ufficio Studi, ricerche, legislazione e rapporti internazionali.

Pande, R. and D. Ford (2009). Gender Quotas and Female Leadership: A Review. *Monthly Labor Review 42*.

Penal Reform International, P. (2023). Global prison trends 2023. Technical report, Global Prison Trends series.

Rafter, N. H. (1983). Prisons for women, 1790-1980. *Crime and Justice 5*, 129–181.

Rose, E. K. and Y. Shem-Tov (2021). How does incarceration affect reoffending? estimating the dose-response function. *Journal of Political Economy 129*(12), 3302–3356.

Staton-Tindall, M., K. L. Harp, E. Winston, J. M. Webster, and K. Pangburn (2015). Factors associated with recidivism among corrections-based treatment participants in rural and urban areas. *Journal of Substance Abuse Treatment 56*, 16–22.

Tobón, S. (2022). Do better prisons reduce recidivism? evidence from a prison construction program. *Review of Economics and Statistics 104*(6), 1256–1272.

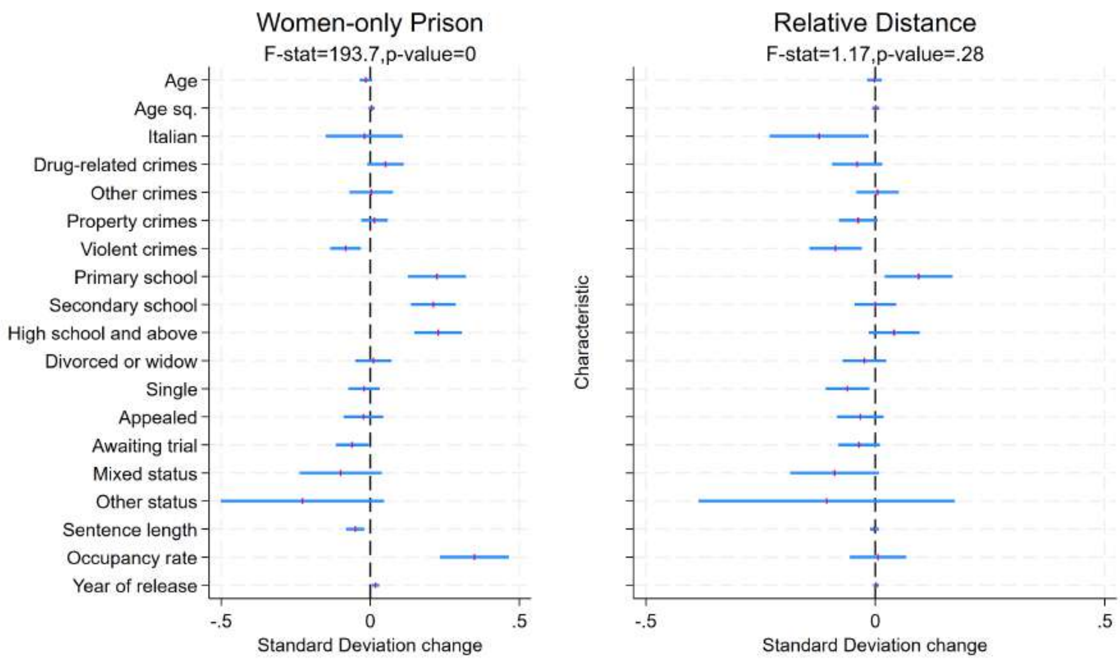
United Nations, U. (2010). United nations rules for the treatment of women prisoners and non-custodial measures for women offenders (the bangkok rules. Technical report, The United Nations General Assembly.

Vollaard, B. (2013). Preventing crime through selective incapacitation. *The Economic Journal 123*(567), 262–284.

World Health Organization, W. (2009). Women's health in prison correcting gender inequity in prison health 2009. Technical report, WHO Regional Office for Europe.

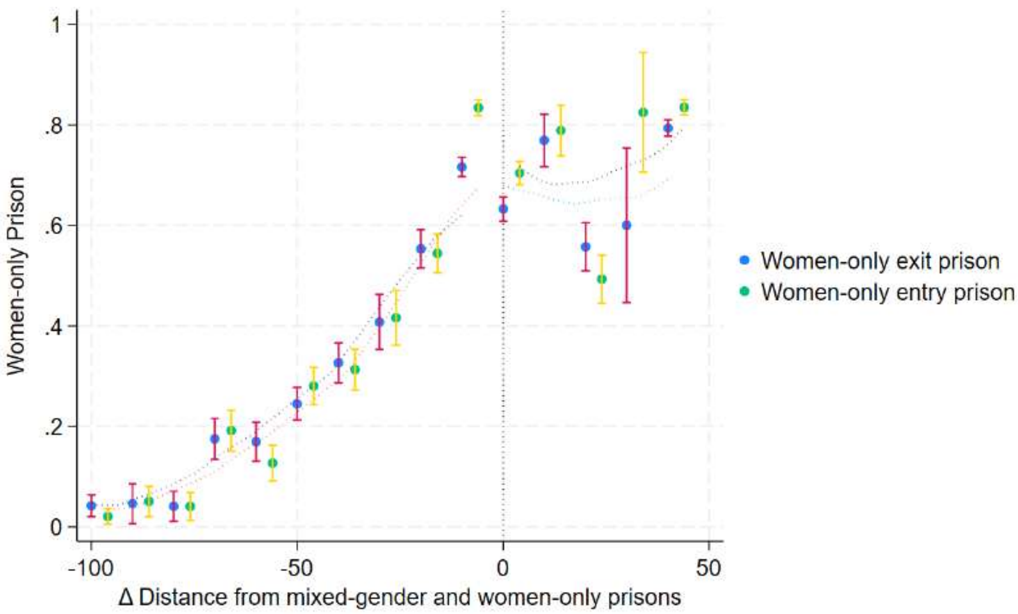
8 Figures and Tables

Figure 1: Balance Across Prisons and Relative Distance



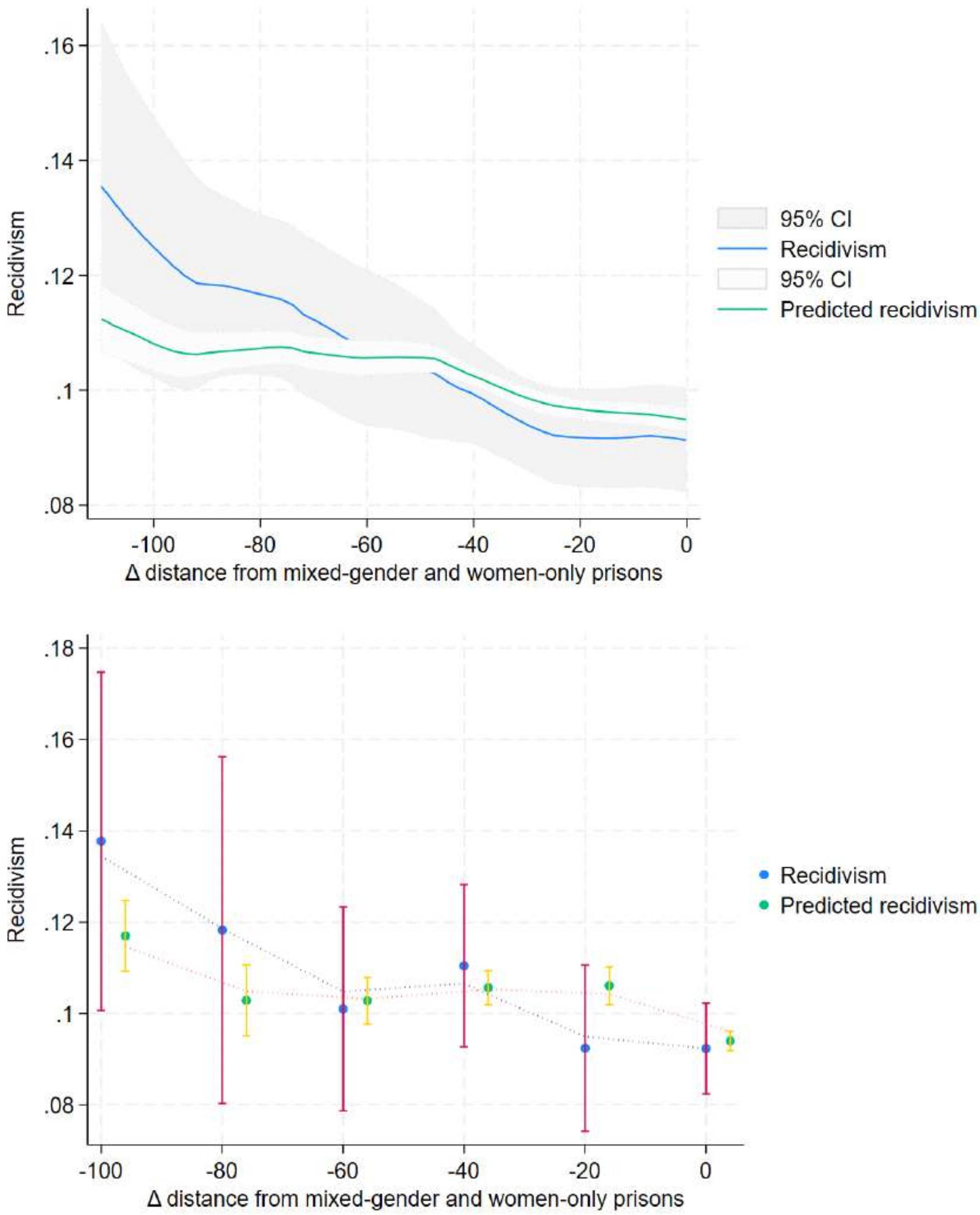
Notes - The figure reports the coefficients of a linear probability model and the corresponding 95% confidence intervals, controlling for region of the prison and of the inmate's birth fixed effects, and various controls for the municipality of residence and their square when continuous: population, population density, a dummy for municipalities with more than 500,000 inhabitants, share of foreigh born, and share with university degree. In the left panel, the dependent variable is the dummy *Women-only Prisons*, while in the right panel, it is the *Prison relative distance*. For both specifications, the joint significance test is reported.

Figure 2: Women-only Prison and Difference in Distance: First Stage



Notes - The figure shows the estimated probability to be assigned to a women-only prison of entry or exit and the corresponding 95% confidence intervals. Distance is divided into groups of 10 km each.

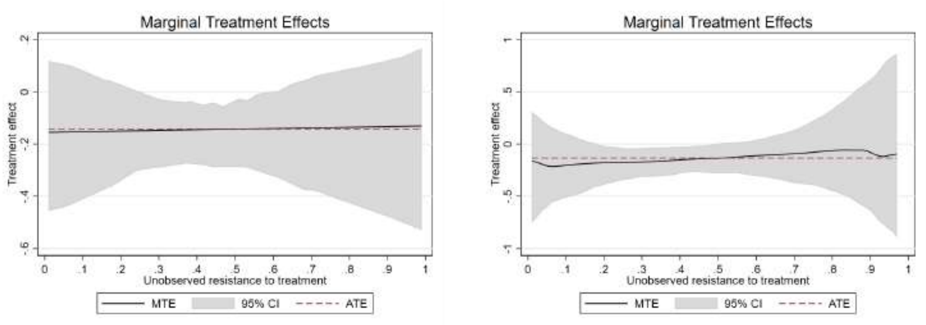
Figure 3: Difference in Distance against Recidivism and Predicted Recidivism.



Notes - The figure shows the recidivism rate and the predicted recidivism rate (and the corresponding 95% confidence intervals) against the relative distance between the place of residence of inmates and the two types of prison (mixed-gender vs. women-only). The upper panel plots the local linear regression, while in the lower panel averages are computed every 20 km.

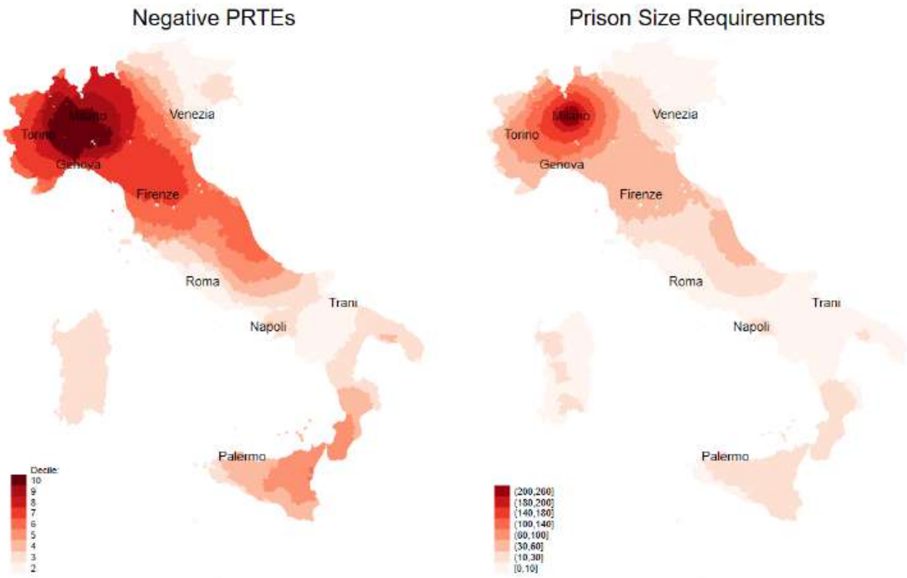
Figure 4: Marginal Treatment Effects

Linear Semi-parametric



Notes - The Figure shows the MTE curves and the Average Treatment Effects (ATE) for the semiparametric and the linear specification. The propensity score is estimated using a probit model (see Figure A4 for the common support). The dependent variable is whether the inmates recidivates within 3 years. The endogenous variable, *Women-only,* is a dummy equal to one for inmates in one of the four women-only prisons. The instrumental variable, *Prisons Relative distance* is the difference in the distance in km between the municipality of residence of the inmate and the nearest two types of prisons (Mixed-gender vs. Women-only). The controls are the same ones used in Column 3 of Table 3. The semiparametric figure uses the Epanechnikov kernel with a bandwith of the local polynomial smooth for recidivism of 0.1. Standard errors are estimated using 200 bootstrap replications.

Figure 5: Policy Relevant Treatment Effects



Notes - The maps shows the deciles of the absolute value of the sum of PRTEs (larger deciles indicate larger reductions in recidivism) and of the prison size requirements (the changes in the propensity score times the average number of yearly inmates). The PRTE are based on the linear specification of the MTEs (see Figure 4). The propensity score is estimated using a probit model (see Figure A4 for the common support). The controls are the same ones used in Table 3.

Table 1: Women-only vs Mixed-gender Prisons

	Women-only	Mixed	Difference	Std. Err.
Health				
Heating	0.934	0.870	0.064	0.007
Hot Water	0.874	0.386	0.488	0.009
Separate WC	0.714	0.941	-0.226	0.010
Shower	0.539	0.335	0.204	0.012
Presence of medical services	0.903	0.894	0.009	0.008
Gynecologist	0.962	0.842	0.120	0.007
Obstetrician	0.560	0.419	0.141	0.015
School Spaces	1.000	0.987	0.013	0.001
Spaces for work activities	1.000	0.842	0.158	0.004
Non-catholic Religious Spaces	0.288	0.221	0.067	0.011
Gym	0.707	0.668	0.040	0.011
Outdoor sports field	0.781	0.689	0.093	0.010
Hours Spent Outside the Cell	5.545	4.014	1.531	0.030
Autonomous movement in jail	0.824	0.268	0.556	0.010
Remote Meetings	0.548	0.495	0.053	0.012
Web Access	0.199	0.157	0.042	0.009
Cells with 3 m^2 per inmate	0.577	0.684	-0.107	0.011
Cells open 8 hours per day	1.000	0.545	0.455	0.006
Work & Other				
Share Working Women Inside	0.243	0.325	-0.082	0.004
Share Working Women Outside	0.058	0.037	0.021	0.002
Share Training Women	0.123	0.168	-0.045	0.006
Share Woman in School	0.275	0.279	-0.004	0.007
Number of Women	247.481	59.765	187.717	2.205
Female Occupancy Share	0.847	0.972	-0.125	0.009
Share of Foreigners	0.338	0.321	0.017	0.005
Share on Parole	0.018	0.012	0.005	0.000
Share of Disciplinary Measures	0.069	0.093	-0.024	0.003
Self-Harm	0.140	0.283	-0.143	0.008
Female Director	0.969	0.640	0.329	0.007
%Positive description	0.369	0.332	0.037	0.004
Construction year	1714	1970	-256	6.183

Notes - Prison population-weighted mean differences between the two types of prisons. Data come from the Antigone surveys. *Positive description* accounts for the fraction of positive words relative to the total number of positive and negative words used by Antigone's representatives to describe prison conditions.

Table 2: Summary Statistics

Sample	Full		$\Delta D_i < 0$		$\Delta D_i \geq 0$	
	Mean	SD	Mean	SD	Mean	SD
Recidivates	11.77	32.22	10.02	30.03	15.14	35.84
Women-only (Exit)	0.57	0.50	0.48	0.50	0.75	0.43
Relative distance (Entry)	-10.51	43.51	-35.38	30.25	37.39	17.06
Relative distance (Exit)	-9.94	42.67	-34.51	29.13	37.39	17.06
Year of release	2015.23	2.54	2015.25	2.55	2015.19	2.51
Age (10 yrs.)	39.08	23.13	40.05	27.09	37.23	12.13
Italian	0.64	0.48	0.73	0.45	0.48	0.50
Drug-related crimes	0.38	0.48	0.40	0.49	0.33	0.47
Other crimes	0.13	0.34	0.13	0.34	0.13	0.33
Property crimes	0.41	0.49	0.38	0.49	0.46	0.50
Violent crimes	0.33	0.47	0.33	0.47	0.33	0.47
Metropolitan area	0.37	0.48	0.22	0.41	0.65	0.48
Primary school	0.16	0.37	0.18	0.39	0.11	0.32
Secondary school	0.33	0.47	0.34	0.47	0.32	0.47
High school and above	0.18	0.38	0.17	0.37	0.19	0.39
Divorced or Widow	0.16	0.37	0.18	0.38	0.12	0.33
Single	0.45	0.50	0.37	0.48	0.59	0.49
Distance from prison	57.94	111.72	60.55	107.00	52.91	120.15
Appealed	0.13	0.34	0.13	0.34	0.14	0.35
Awaiting trial	0.41	0.49	0.44	0.50	0.37	0.48
Mixed status	0.03	0.17	0.03	0.17	0.03	0.18
Sentence	1.34	1.93	1.43	2.00	1.19	1.78
Other status	0.00	0.06	0.00	0.06	0.00	0.05
Occupancy rate	1.37	0.55	1.40	0.62	1.32	0.38
Population (/10k)	75.94	104.16	25.78	37.80	172.56	121.72
Share of Foreign Born (in %)	6.06	3.77	4.91	3.84	8.27	2.40
Share of College Education (in %)	12.11	5.06	10.39	3.94	15.43	5.34
Population Density	2.44	2.74	2.89	3.25	1.58	0.74
Observations	10130		6668		3462	

Notes - Means and standard deviations are reported for the full sample and by the sign of the change in relative distance (ΔD_i). Distances are in kilometers. "Women-only (Exit)" is an indicator for release from a women-only institution. Local characteristics refer to the municipality of residence.

Table 3: Effect of Women-only Prison on Recidivism.

	$\Delta D_i < 0$				$\Delta D_i \geq 0$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: First Stage - Women-only (Exit)								
Relative distance (Exit)	0.007***	0.005***	0.004***	0.004***	-0.004	-0.003	-0.006	-0.005
	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.005)	(0.004)	(0.003)
Mean women-only	0.479				0.747			
Panel B: Reduced Form - Recidivates (in %)								
Relative distance (Exit)	-0.056***	-0.075***	-0.062***	-0.062***	-0.001	-0.136	-0.053	-0.051
	(0.019)	(0.021)	(0.019)	(0.019)	(0.122)	(0.138)	(0.116)	(0.118)
Mean recidivism (%)	10.02				15.14			
Panel C: IV-Recidivates (in %)								
Women-only (Exit)	-7.955***	-16.084***	-14.327***	-14.117***				
	(2.922)	(4.823)	(4.838)	(4.802)				
Region of birth FEs		✓	✓	✓		✓	✓	✓
Region of resid. FEs		✓	✓	✓		✓	✓	✓
Controls			✓	✓			✓	✓
Controls (quadratic)				✓				✓
Observations	6,668				3,462			
KP F-Stat	186.1	56.15	60.45	71.29				
$\widehat{R}_M$ (%)	13.83	17.72	16.88	16.78				
$\widehat{R}_F$ (%)	5.871	1.633	2.550	2.659				

Notes - The dependent variable is whether the inmates recidivate within 3 years. The endogenous variable, *Women-only,* is a dummy equal to one for inmates in one of the four women-only prisons. The instrumental variable, *Prisons Relative distance* is the difference in the distance in km between the municipality of residence of the inmate and the nearest two types of prisons (Mixed-gender vs. Women-only). The measures are calculated according to the exit prison. Columns 1-4 report results for $\Delta D_i < 0$, while columns 5-8 for $\Delta D_i \geq 0$. We report the Kleibergen-Paap for weak identification. Controls include age, age squared, nationality, education status, family status, types of crime, sentence status, incarceration length, closing year of incarceration, metropolitan area, true distance, occupancy rate, as well as features of the inmate's residence (size of the population, share of foreign born, share of college education and population density - and their quadradic versions). Municipality of residence clustered standard errors in parentheses. * p< 0.10, ** p< 0.05, *** p< 0.01.

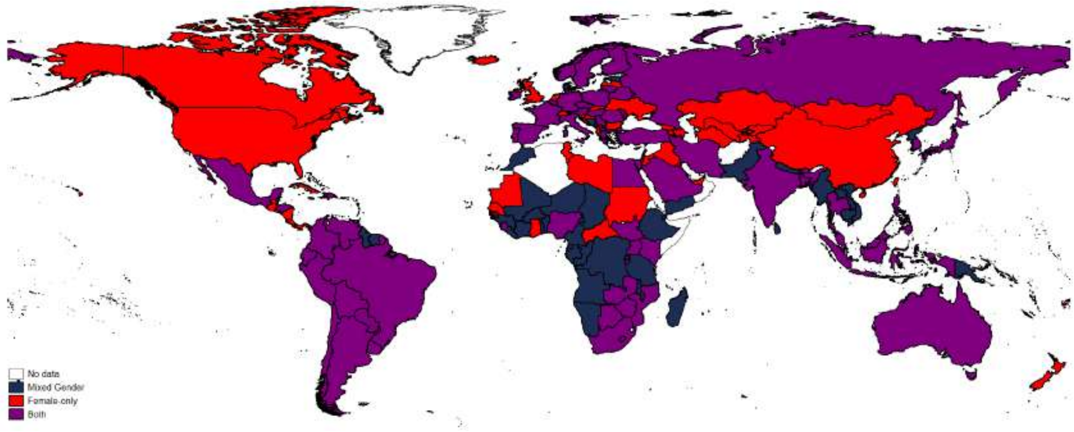
Table 4: Marginal Treatment Effects

	(1)	(2)	(3)
ATE	-0.155*	-0.120*	-0.121
	(0.0796)	(0.0727)	(0.0974)
Polynomial	3	1	-
Semiparametric			√
Observations	6,673	6,673	6,673

Notes - The Table shows Average Treatment Effects (ATE) computed using Marginal Treatment Effects estimated with parametric and sempiparametric Local Instrumental Variable specifications. The propensity score is estimated using a probit model (see Figure A4 for the common support). The dependent variable is whether the inmates recidivades within 3 years. The endogenous variable, *Women-only,* is a dummy equal to one for inmates in one of the four women-only prisons. The instrumental variable, *Prisons Relative distance* is the difference in the distance in km between the municipality of residence of the inmate and the nearest two types of prisons (Mixed-gender vs. Women-only). The controls are same ones used in Column 3 of Table 3. Standard errors are estimated using 200 bootstrap replications.

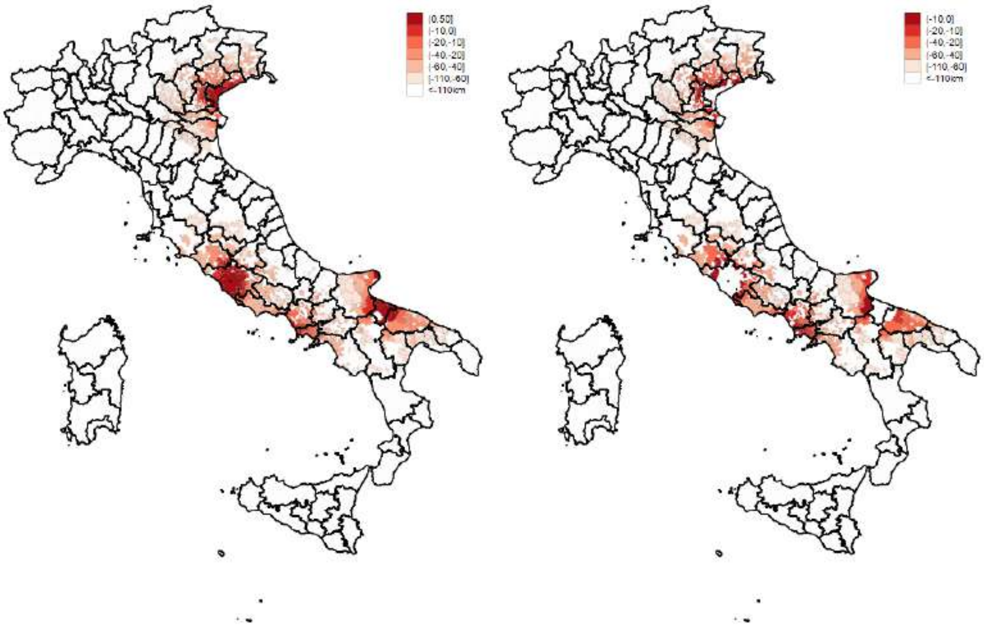
A Appendix

Figure A1: Types of Prisons Across the World



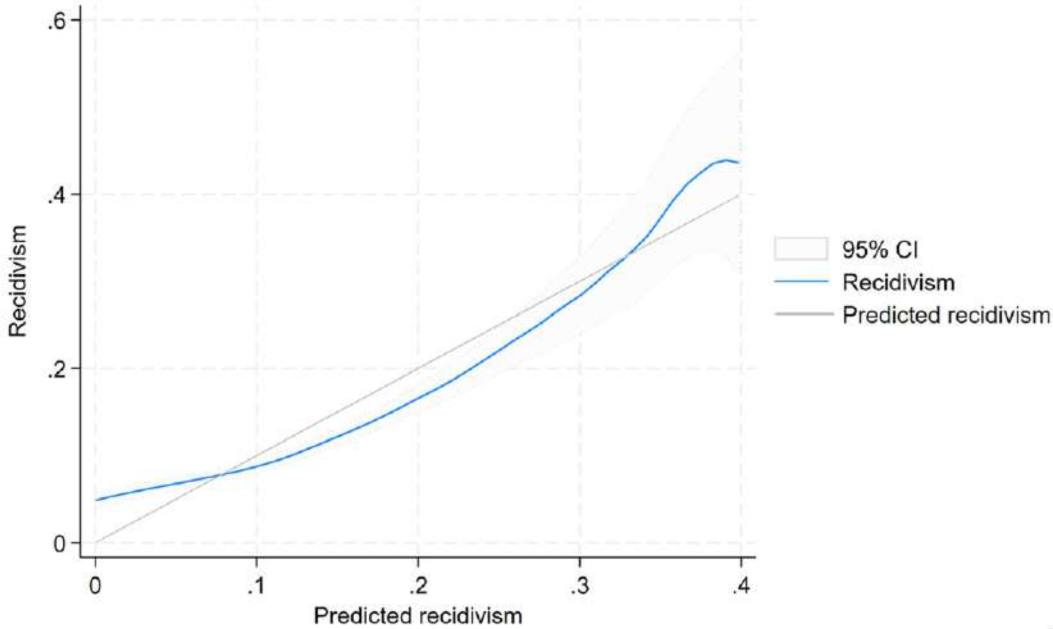
Notes: The map classifies country prison systems by the availability of facilities for women: mixed systems (women housed within predominantly male prisons), dual systems (both mixed and women-only prisons), and women-only systems (separate female institutions only).

Figure A2: Difference in Distance from the Nearest Mixed-gender and Women-only Prison



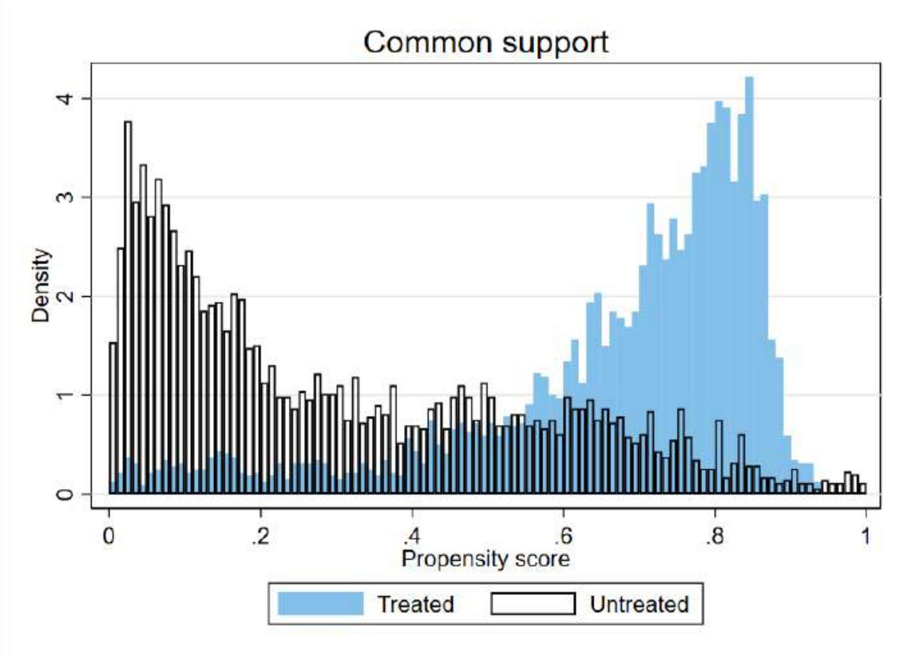
Notes - The figure plots the difference in distance (in km) between the inmate's municipality of residence and the nearest mixed-gender and women-only prison. The left map shows all distances, while the right one focuses on negative differences, where the mixed-gender one is closer. White refers to inmates housed in prison with a relative distance greater than 110 km.

Figure A3: Predicted Recidivism



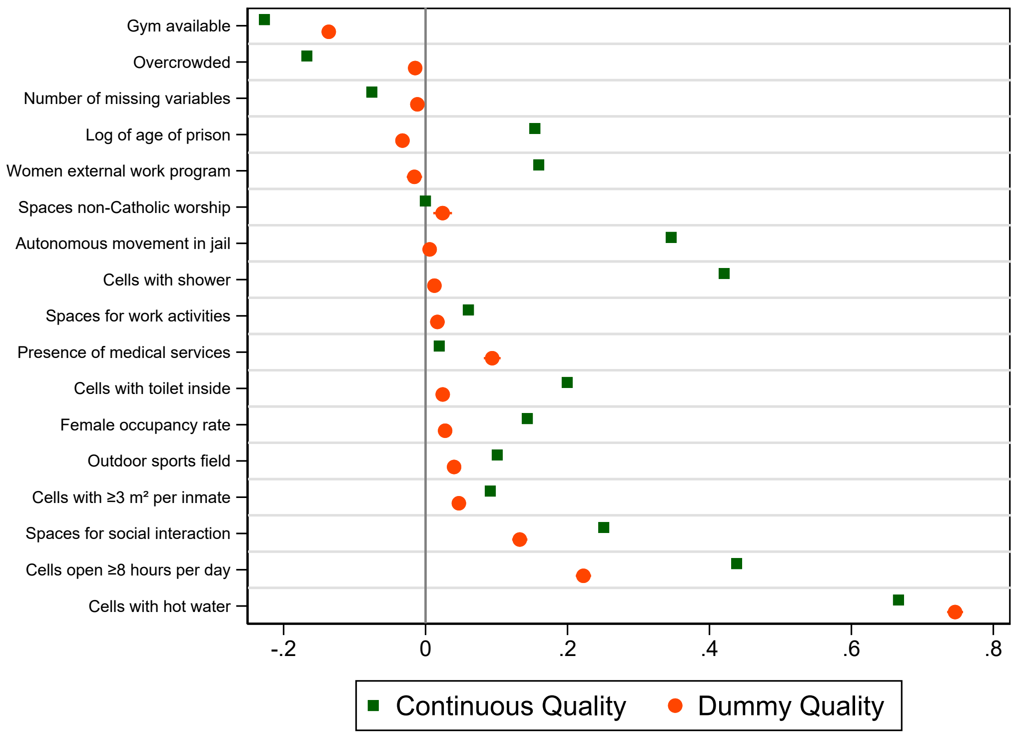
Notes - The figure plots local linear regressions for recidivism against predicted recidivism.

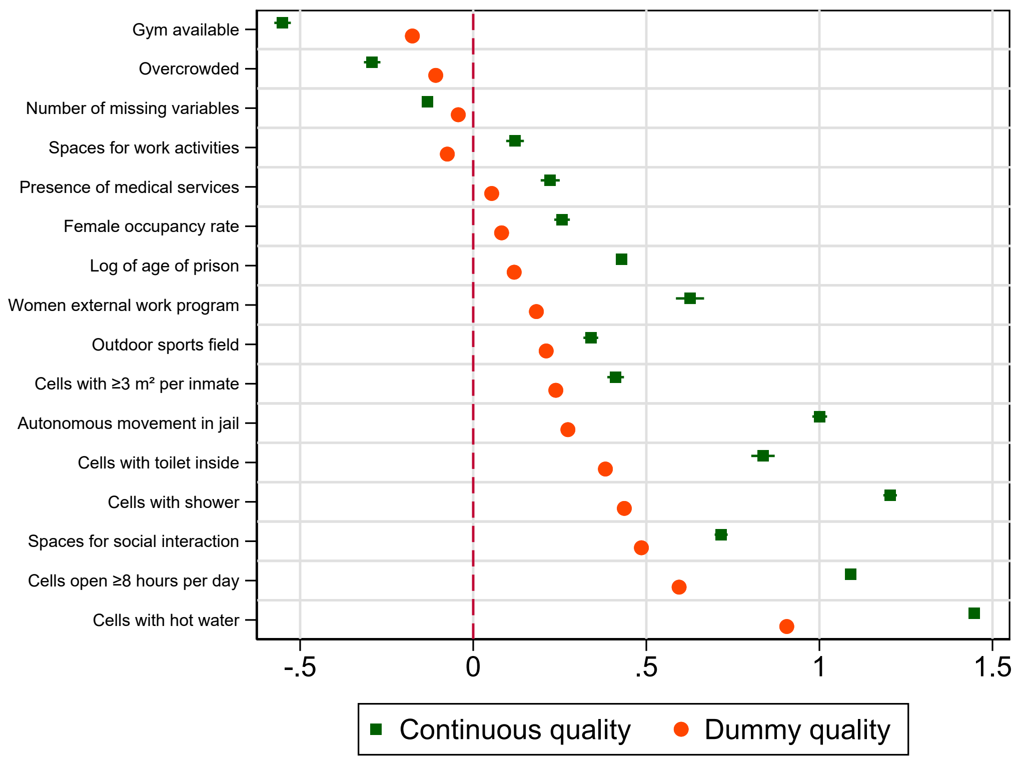
Figure A4: Propensity Score



Notes - The figures shows the histograms of the propensity score for inmates sent to the women-only prisons (treated) and mixed-gender prisons (untreated). The propensity score is estimated with a probit model that contain the variables included in column 3 of Table 3.

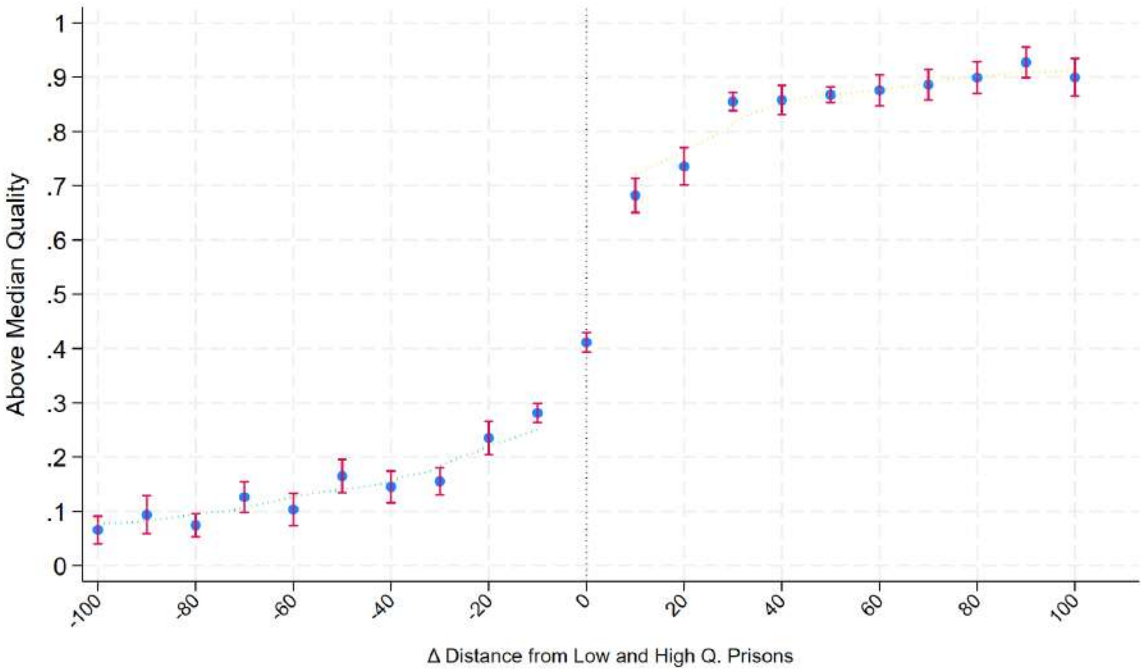
Figure A5: Quality Index and Prison Characteristics





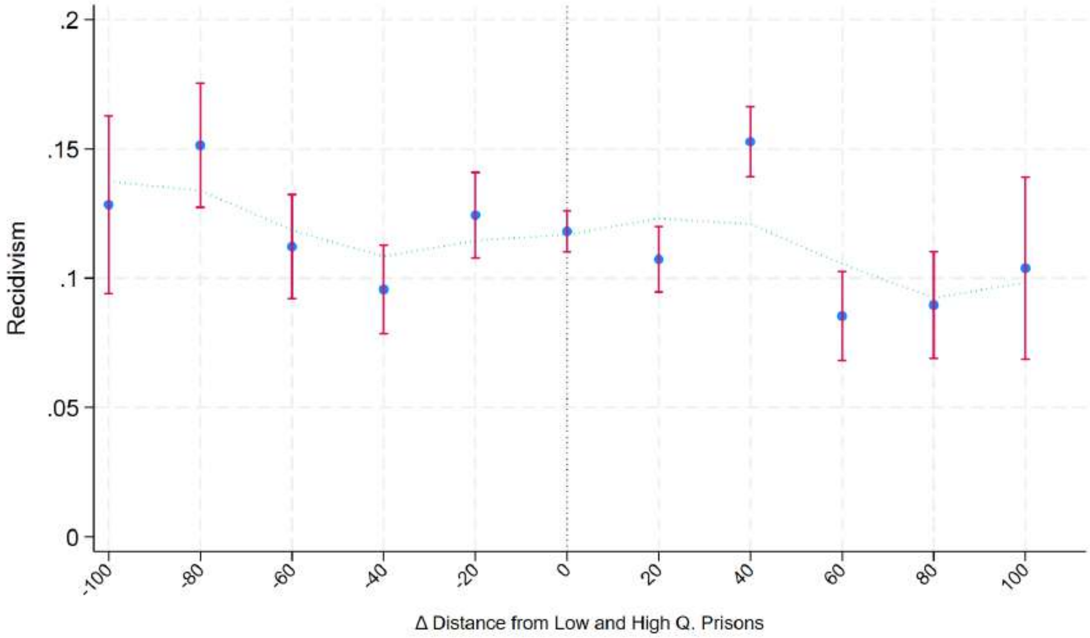
Notes: The figure displays the contribution of each feature of a prison facility to both the continuous and binary measures of prison quality, as captured by the latent factor. The top panel reports conditional correlations obtained by including all covariates jointly, while the bottom panel shows bivariate correlations, i.e., one covariate at a time.

Figure A6: Difference in Distance and Prison Quality



Notes - The figure shows the likelihood of being assigned to an above-median quality prison (and the corresponding 95% confidence intervals), against the relative distance between the place of residence of inmates and the two types of prisons (i.e. below and above median quality). Prisons are classified based on the latent factor capturing prison quality.

Figure A7: Difference in Distance and Recidivism



Notes - The figure shows the recidivism rate (and the corresponding 95% confidence intervals) against the relative distance between the place of residence of inmates and the two types of prison (i.e. below and above median quality).

Table A1: Effect of Women-only Prison on Recidivism (entry).

	$\Delta D_i < 0$		
	(1)	(2)	(3)
	Panel A: First Stage - Women-only (Entry)		
Relative distance (Entry)	0.008***	0.005***	0.004***
	(0.001)	(0.001)	(0.001)
Mean women-only		0.479	
	Panel B: Reduced Form - Recidivates		
Relative distance (Entry)	-0.053***	-0.074***	-0.059***
	(0.018)	(0.021)	(0.019)
Mean recidivism (in %)		10.02	
	Panel C: IV- Recidivates		
Women-only (Entry)	-6.680***	-14.502***	-13.373***
	(2.424)	(4.534)	(4.799)
Region of birth FEs		✓	✓
Region of resid. FEs		✓	✓
Controls			✓
Observations		6,668	
KP F-Stat	182.3	43.10	59.12
$\widehat{R}_M$ (in %)	13.51	17.59	17
$\widehat{R}_F$ (in %)	6.825	3.087	3.626

Notes - The dependent variable is whether the inmates recidivate within 3 years. The endogenous variable, *Women-only,* is a dummy equal to one for inmates in one of the four women-only prisons. The instrumental variable, *Prisons Relative distance* is the difference in the distance in km between the municipality of residence of the inmate and the nearest two types of prisons (Mixed-gender vs. Women-only). The measures are calculated according to the entry prison. We report the Kleibergen-Paap F statistic for weak identification. Controls include age, age squared, nationality, education status, family status, types of crime, sentence status, incarceration length, closing year of incarceration, metropolitan area, true distance, occupancy rate, as well as features of the inmate's residence (size of the population, share of foreign born, share of college education and population density - and their quadradic versions). Municipality of residence clustered standard errors in parentheses. * p< 0.10, ** p< 0.05, *** p< 0.01.

Table A2: Effect of Women-only Prison on Recidivism (different time windows).

	Recidivism (in %)					
	1 year		2 years		3 years	
	Panel A - Reduced Form					
Relative distance (Exit)	-0.050***	-0.058***	-0.044**	-0.054***	-0.056***	-0.062***
	(0.016)	(0.018)	(0.017)	(0.018)	(0.019)	(0.019)
Mean recidivism (in %)	6.479		8.623		10.02	
	Panel B - IV					
Women-only (Exit)	-7.081***	-13.071***	-6.331**	-12.357***	-7.955***	-14.117***
	(2.490)	(4.224)	(2.606)	(4.430)	(2.922)	(4.802)
Observations	6668					
KP F-Stat	186.1	71.29	186.1	71.29	186.1	71.29
$\widehat{R}_M$ (in %)	9.868	12.74	11.65	14.54	13.83	16.78
$\widehat{R}_F$ (in %)	2.788	-0.335	5.323	2.182	5.871	2.659
Region of birth FEs		✓		✓		✓
Region of resid. FEs		✓		✓		✓
Controls		✓		✓		✓

Notes - The table reports our main estimates differentiating by the lenght of the recidivism window (1, 2 or 3 years). The dependent variable is whether the inmates recidivate within either 1, 2 or 3 years. The endogenous variable, *Women-only,* is a dummy equal to one for inmates in one of the four women-only prisons. The instrumental variable, *Prisons Relative distance* is the difference in the distance in km between the municipality of residence of the inmate and the nearest two types of prisons (Mixed-gender vs. Women-only). The measures are calculated according to the exit prison. We report the Kleibergen-Paap F statistic for weak identification. Controls include age, age squared, nationality, education status, family status, types of crime, sentence status, incarceration length, closing year of incarceration, metropolitan area, true distance, occupancy rate, as well as features of the inmate's residence (size of the population, share of foreign born, share of college education and population density - and their quadradic versions). Municipality of residence clustered standard errors in parentheses. * p< 0.10, ** p< 0.05, *** p< 0.01.

Table A3: Effect of Women-only Prison on Recidivism (final sentence).

	(1)	(2)	(3)
Women-only (Exit)	-4.331	-13.154***	-11.362**
	(3.757)	(4.796)	(4.975)
Observations	2,668	2,668	2,668
KP F-Stat	203.8	83.14	90.20
$\widehat{R}_M$ (in %)	11.21	15.41	14.56
$\widehat{R}_F$ (in %)	6.878	2.258	3.196
Region of birth FEs		✓	✓
Region of resid. FEs		✓	✓
Controls			✓

Notes - The table reports our main estimates restricting recidivism to be computed as a new imprisonment after a previous final sentence. The dependent variable is whether the inmates recidivate within 3 years. The endogenous variable, *Women-only,* is a dummy equal to one for inmates in one of the four women-only prisons. The instrumental variable, *Prisons Relative distance* is the difference in the distance in km between the municipality of residence of the inmate and the nearest two types of prisons (Mixed-gender vs. Women-only). The measures are calculated according to the exit prison. We report the Kleibergen-Paap F statistic for weak identification. Controls include age, age squared, nationality, education status, family status, types of crime, sentence status, incarceration length, closing year of incarceration, metropolitan area, true distance, occupancy rate, as well as features of the inmate's residence (size of the population, share of foreign born, share of college education and population density - and their quadratic versions). Municipality of residence clustered standard errors in parentheses. * p< 0.10, ** p< 0.05, *** p< 0.01.

Table A4: Effect of Women-only Prison on Recidivism (Italians).

	(1)	(2)	(3)
Women-only (Exit)	-10.617***	-22.487***	-21.554***
	(2.955)	(4.771)	(6.039)
Observations	4,842	4,842	4,842
KP F-Stat	235.5	62.54	50.70
$\widehat{R}_M$ (in %)	15.96	21.98	21.51
$\widehat{R}_F$ (in %)	5.345	-0.502	-0.0425
Region of birth FEs		✓	✓
Region of resid. FEs		✓	✓
Controls			✓

Notes - The table reports our main estimates restricting the sample to Italians. The dependent variable is whether the inmates recidivate within 3 years. The endogenous variable, *Women-only,* is a dummy equal to one for inmates in one of the four women-only prisons. The instrumental variable, *Prisons Relative distance* is the difference in the distance in km between the municipality of residence of the inmate and the nearest two types of prisons (Mixed-gender vs. Women-only). The measures are calculated according to the exit prison. We report the Kleibergen-Paap F statistic for weak identification. Controls include age, age squared, nationality, education status, family status, types of crime, sentence status, incarceration length, closing year of incarceration, metropolitan area, true distance, occupancy rate, as well as features of the inmate's residence (size of the population, share of foreign born, share of college education and population density - and their quadradic versions). Municipality of residence clustered standard errors in parentheses. * p< 0.10, ** p< 0.05, *** p< 0.01.

Table A5: Effect of Women-only Prison on Recidivism by Type of Crime.

	Property Crimes		Violent Crimes	
	(1)	(2)	(3)	(4)
	Panel A: First Stage- Women-only (Exit)			
Relative distance (Exit)	0.007***	0.004***	0.006***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Mean women-only	0.499		0.420	
	Panel B: Reduced Form- Recidivates			
Relative distance (Exit)	-0.060**	-0.081***	-0.071***	-0.061**
	(0.022)	(0.021)	(0.026)	(0.032)
Mean recidivism (in %)	10.15		10.40	
	Panel C: IV- Recidivates			
Women-only (Exit)	-8.597***	-19.388***	-11.542**	-15.168**
	(3.324)	(5.901)	(4.512)	(8.399)
Observations	4,976		2,173	
Controls		✓		✓
Region of birth FEs		✓		✓
Region of resid. FEs		✓		✓
KP F-Stat	168.9	54.96	141.8	45.20
$\widehat{R}_M$ (in %)	14.44	19.82	15.24	16.77
$\widehat{R}_F$ (in %)	5.842	0.435	3.702	1.598

Notes - The table reports our main estimates differentiating between property and violent crimes. The dependent variable is whether the inmates recidivate within 3 years. The endogenous variable, *Women-only,* is a dummy equal to one for inmates in one of the four women-only prisons. The instrumental variable, *Prisons Relative distance* is the difference in the distance in km between the municipality of residence of the inmate and the nearest two types of prisons (Mixed-gender vs. Women-only). The measures are calculated according to the exit prison. We report the Kleibergen-Paap F statistic for weak identification. Controls include age, age squared, nationality, education status, family status, types of crime, sentence status, incarceration length, closing year of incarceration, metropolitan area, true distance, occupancy rate, as well as features of the inmate's residence (size of the population, share of foreign born, share of college education and population density - and their quadradic versions). Municipality of residence clustered standard errors in parentheses. * p< 0.10, ** p< 0.05, *** p< 0.01.

Table A7: Effect of Prison Quality on Recidivism.

	(1)	(2)	(3)	(4)
	Panel A: First Stage: Prison Type			
Relative distance (10km)	0.058***	0.047***	0.034***	0.033***
	(0.003)	(0.002)	(0.002)	(0.002)
Mean median quality	0.509			
	Panel B: Reduced Form: Recidivates (in %)			
Relative distance (10km)	-0.122	-0.209***	-0.152**	-0.137**
	(0.087)	(0.070)	(0.060)	(0.061)
Mean recidivism	11.95			
	Panel C: IV: Recidivates (in %)			
Above median quality	-2.103	-6.334***	-4.515**	-4.084**
	(1.502)	(2.138)	(1.770)	(1.790)
Region of birth FEs (20)		✓	✓	✓
Region of resid. FEs (20)		✓	✓	✓
Controls			✓	✓
Controls (quadratic)				✓
Observations	18,061			
KP F-Stat	429.8	204.1	292.5	332.2

Notes - The dependent variable is whether the inmates recidivate within 3 years. The endogenous variable, *Prison Quality*, is a binary variable equal to one for prisons where the quality indicator (latent factor) is above the median. The instrumental variable, *Prisons Relative distance*, is the difference in the distance in km between the municipality of residence of the inmate and the nearest two types of prisons, i.e. below/above the median of quality. We report the Kleibergen-Paap F statistic for weak identification. Controls include age, age squared, nationality, education status, family status, types of crime, sentence status, incarceration length, closing year of incarceration, metropolitan area, true distance,, as well as features of the inmate's residence (size of the population, share of foreign born, share of college education and population density – and their quadratic versions). Municipality of residence clustered standard errors in parentheses. * $p< 0.10$, ** $p< 0.05$, *** $p< 0.01$.

Table A6: Effect of Women-only Prison on Recidivism (different control groups)

	(1)	(2)	(3)	(4)	(5)	(6)
Women-only (Exit)	-7.142*	-26.173	-9.815**	-17.143	-14.168**	-6.101
	(3.954)	(15.962)	(4.522)	(11.443)	(5.747)	(6.088)
Observations	5,020	4,840	4,975	4,885	5,858	4,002
KP F-Stat	181.4	8.432	134.6	13.37	46.14	71.71
Mean recidivism	10.26	9.587	10.13	9.724	10.11	9.670
Control group	> median female share	< median female share	> median size	< median size	male director	female director
Controls	✓	✓	✓	✓	✓	✓
Region of birth FEs	✓	✓	✓	✓	✓	✓
Region of resid. FEs	✓	✓	✓	✓	✓	✓

Notes: The table reports the main estimates under alternative control group definitions. In columns 1-2, the control group is restricted to mixed-gender prisons with an above- and below-median female share, respectively. Columns 3-4 instead restrict the control group to mixed-gender prisons with an above- and below-median absolute number of female inmates. Columns 5-6 use control groups defined by the gender of the prison director (male vs. female). The dependent variable is whether the inmates recidivate within 3 years. The endogenous variable, *Women-only,* is a dummy equal to one for inmates in one of the four women-only prisons. The instrumental variable, *Prisons Relative distance* is the difference in the distance in km between the municipality of residence of the inmate and the nearest two types of prisons (Mixed-gender vs. Women-only). The measures are calculated according to the exit prison. We report the Kleibergen-Paap F statistic for weak identification. Controls include age, age squared, nationality, education status, family status, types of crime, sentence status, incarceration length, closing year of incarceration, metropolitan area, true distance, occupancy rate, as well as features of the inmate's residence (size of the population, share of foreign born, share of college education and population density - and their quadradic versions). Municipality of residence clustered standard errors in parentheses. * p< 0.10, ** p< 0.05, *** p< 0.01.