(Managing) Violence Across UK Prisons*

Rocco D'Este† Rocco Macchiavello‡ Ottavia Pesenti§

August 2025

Abstract

We apply the organizational-economics lenses of persistent performance differences and management practices to study variation in violence across prisons in England and Wales. Assembling previously unused administrative records, we (i) focus on three consistently recorded outcomes (assaults on staff, assaults among prisoners, and self-harm), (ii) validate them as proxies for distinct facets of the detention experience (anger, fear, despair) using prisoner survey microdata, and (iii) draw on systematic and detailed reports from the national prison inspectorate to measure management practices. We document large differences in violence across prisons that remain unexplained even after conditioning for detailed prison and prisoners characteristics. The dispersion in the unexplained variation in violence across establishments increases, and its persistence decreases, as endogenous responses to violence are controlled for. Still, variation in management practices recorded in inspection reports accounts for a significant share of the unexplained variation. Exploiting the asynchronous rollout for identification, we show that inspections are associated with a 0.1-0.2 SD persistent reduction in all types of violence. The reduction in violence is particularly pronounced when the prison management focuses on implementing fewer, tailored recommendations. The evidence suggests that strengthening management frameworks that promote the diffusion of adequate practices across establishments might offer a scalable, cost-effective path to improving conditions.

*Preliminary and Incomplete: Comments Welcome. We are grateful to Sophie Riley and the Research Data Thematics team at HMIP for patiently sharing data and comments. We thank Giovanni Mastrobuoni, Rosie Meek, Paolo Pinotti, Imran Rasul, and seminar and workshops participants at Bocconi, IFS, LSE, Royal Holloway for suggestions. Andrea Melillo provided extraordinary research assistance.

†University of Sussex. email: R.D-Este@sussex.ac.uk

‡LSE. email: R.Macchiavello@lse.ac.uk

§LSE. email: o.pesenti@lse.ac.uk

"No one truly knows a nation until one has been inside its jails. A nation should not be judged by how it treats its highest citizens, but its lowest ones."*

Nelson Mandela, *Long Walk to Freedom*

1 Introduction

At the time of writing, about 11.5 million people are imprisoned worldwide (Fair and Walmsley, 2024; UNODC, 2024). Prisons are the cornerstone of enforcement, deterrence, and rehabilitation. Besides their intrinsic value – emphasized in justice-as-fairness accounts (Rawls, 1971) and codified in the UN Nelson Mandela Rules – sociologists and criminologists have argued for the instrumental value of better in-prison conditions for post-release adjustment and rehabilitation (Haney, 2001; Liebling, 2011). Yet the economics, sociology, and criminology literatures have typically treated prisons as a black box (Western, 2021), rarely examining differences in conditions across establishments and, in particular, the role of governance and management in shaping them. The lack – and importance – of such analysis is well recognized (Wooldredge, 2020): empirical progress has been limited largely because conditions and managerial practices are difficult to observe consistently across prisons.

Using England and Wales as a laboratory, this paper brings the organizational-economics framework on persistent performance differences (Gibbons and Henderson, 2012; Syverson, 2011) and management practices (Bloom et al., 2012) to the study of prison conditions.[1] Incarceration rates are at a historic high in England and Wales, and violence has attracted significant media attention.[2] Besides its intrinsic relevance, this context also allows us to overcome the measurement challenges noted above through several previously unused administrative records. In particular, we focus on different types of violence – against staff, against prisoners, and self-harm – as the main dependent variables. These outcomes not only matter in their own right but also provide a consistently measured proxy for distinct facets of prison conditions. To measure management practices, we draw on systematic and detailed reports from the national prison inspectorate (HMIP). These inspections share features common to management interventions (Bloom et al., 2013) and we use their asynchronous schedule to assess their effectiveness in improving prison conditions.[3]

*These same words were adopted by the United Nations in 2015 when revising the *Standard Minimum Rules for the Treatment of Prisoners*, now named the *Nelson Mandela Rules*. The Mandela Rules are the universally acknowledged benchmark for good prison management and the humane treatment of prisoners.

[1]In the public sector, the framework has been applied to the civil service (Rasul and Rogger, 2018; Fenizia, 2022), procurement (Best et al., 2019), hospitals (Chandra et al., 2013), and schools (Bloom et al., 2015), but not to prisons.

[2]See, e.g., *Financial Times*, 25 Aug 2025, *The Guardian*, 31 Jul 2025, and *BBC*, 16 Oct 2024.

[3]While several other European countries have a similar prison inspection system, the UK is rather unique

We document large differences in violence across prisons that remain unexplained even after conditioning for detailed prisons' characteristics, such as inmates' composition, average cost per prisoner, and staffing level and composition, as well as year fixed effects. Consistent with a framework in which the government responds to shocks in violence, the unexplained variation in violence across establishments *increases*, and its persistence *decreases*, as endogenous responses are controlled for. Still, variation in management practices recorded in inspection reports accounts for a significant share of the unexplained variation in violence. Exploiting the asynchronous rollout for identification, we also show that inspections are associated with a 0.1-0.2 SD reduction in all types of violence that last for at least fifteen months post-inspection. The associated reduction in violence appears to be particularly pronounced when the prison management focuses on implementing fewer, tailored (as opposed to generic) recommendations issued by the inspectors. This suggests that a key challenge in improving conditions might lie in the identification and implementation of appropriate context-specific practices rather than in the lack of information about, or incentives to adopt, boilerplate "best practices". Additional results paint a richer picture of the scope and mechanisms behind the effectiveness of inspections in reducing violence. Taken together, the evidence suggests that alongside traditional levers that are fiscally costly (e.g., capacity expansion) or politically difficult (e.g., amnesties), strengthening management frameworks that promote the diffusion of adequate practices across establishments might offer a scalable, high-return path to improving conditions.

We start by assembling a comprehensive, establishment-level panel for prisons in England and Wales that combines (i) publicly available Ministry of Justice series (e.g., Safety in Custody, population/capacity, costs) and HM Prison & Probation Service operational data; (ii) outputs from His Majestys Inspectorate of Prisons (HMIP)inspection reports, ratings, recommendations, and governor action planstogether with the HMIP Prisoner Survey microdata; and (iii) targeted disclosures obtained under the Freedom of Information Act. The resulting dataset covers inmate demographics, staffing, capacity and crowding, expenditure, security category, program provision, and a rich set of behavioral indicators, providing an exhaustive portrait of all prisons, over a long period of time.

As mentioned, we begin to open the black box by focusing on violence as our key in-prison outcome. Violence is important in its own right and has been widely studied in criminology and sociology as a core indicator of life inside prisons (Wooldredge, 2020). It is also associated with post-release outcomes: e.g., in-prison misconduct and violent infractions (Nagin and Loeffler, 2021; E. F. J. C. v. Ginneken and Palmen, 2023) and low prison morale

in its scope, standardization, and data transparency. On the other hand, the United States has no national prison inspectorate with a comparable prison-by-prison reporting regime.

(Auty and Liebling, 2020) correlate with a higher risk of reoffending, with harsh prison conditions potentially increasing post-release criminal activity (Drago et al., 2011). Finally – and crucial for our purpose – violence is consistently recorded at the establishment level, allowing for comparable measurement across prisons and over time.

Importantly, not all violence is the same. Drawing from established work in psychology and criminology on the emotional underpinnings of offending, victimization, and coping in custody, we interpret violence against staff, violence among inmates, and self-harm as distinct externalizations of *anger*, *fear*, and *despair*, respectively.[4] Each type of violence reflects a specific psychological strain, offering direct insight into the diverse and layered stresses experienced within prison environments. We validate this perspective using individual-level HMIP Prisoner Survey data: establishments with a higher share of prisoners reporting respectful treatment by staff exhibit lower rates of violence against staff; establishments where more prisoners report feeling safe exhibit lower prisoner-on-prisoner violence; and establishments where more prisoners report mental-health problems exhibit higher self-harm.[5]

Applying the persistent-performance-differences (PPDs) lens, we treat violence as the output of a prison production function with inputs such as inmate composition and physical and human capital. After conditioning on such (and other) observables, we uncover large and persistent unexplained cross-prison differences. In a typical year, a prison at the 75th percentile has 113 (184, 709) more incidents per 1,000 prisoners than one at the 25th percentile for violence against staff (against prisoners, self-violence). The interquartile range of residual log-violence – a measure of dispersion adapted, for ease of comparison, from the productivity literature – is 0.348, 0.283, and 0.353 for staff, prisoner-on-prisoner, and self-violence. The extent of dispersion is thus comparable to the within-industry dispersion across manufacturing establishments (Syverson, 2011), and larger than across hospitals (Chandra et al., 2013) (0.290 and 0.230, respectively).

Structural features account for part of the cross-prison differences in violence. For example, higher-security establishments and those with more space per inmate record lower rates, while womens prisons show a distinctive profile with significantly higher incidence of self-harm. Yet, even after conditioning on these factors, a large residual variation remains, suggesting that governance and management practices might play a first-order role – an interpretation in line with evidence from numerous other settings that management quality matters for performance (Bloom and Van Reenen, 2010).[6]

[4]See.e.g, (Haney, 2001; Useem and Reisig, 1999; O'Donnell and Edgar, 1999; Favril et al., 2020).

[5]Throughout the analysis, we show that our main results are largely robust (although, at times, less precise) when restricting attention to "hard" outcomes (serious assaults on staff/prisoners, self-harm requiring hospitalization, homicides, and suicides) which are less subject to mis-reporting. Furthermore, throughout, we focus on individual-level incidents and exclude group-level disorders (e.g., riots).

[6]As further suggestive evidence, governor fixed effects account for a meaningful share of the unexplained

Inspections conducted by His Majesty's Inspectorate of Prisons (HMIP) act as a direct managerial mechanism, systematically diagnosing institutional strengths and weaknesses and areas for improvement. Through on-site observation, interviews with staff and inmates, evaluations of policy implementation and standards compliance, and leadership assessments, inspectors measure prison performance across core dimensions, like safety, dignity, purposeful activity, resettlement, and leadership. Crucially, inspections do not end at diagnosis: prisons are required to respond with formal action plans addressing the identified concerns. This process not only triggers targeted internal reforms but also embeds managerial attention and accountability into prison operations. Moreover, inspection reports and the resulting action plans constitute an exceptionally rich and structured source of both quantitative and qualitative information on prison conditions, operations, and governance. Such granular insight into the practices and functioning of public institutions is rare even in private-sector settings, making this an unusually powerful context for studying how oversight and management interact to shape institutional outcomes.

Econometrically, to identify the causal effect of inspections, we exploit the staggered and repeated timing of inspections using a stacked difference-in-differences (DiD) design with appropriate weights, following the framework of Wing et al. (2024). This approach mitigates bias from heterogeneous treatment effects and ensures credible causal inference. Our design restricts control units to prisons not inspected for at least one year prior, ensuring clean comparisons. While inspection timing is not entirely random - driven in part by dynamic risk assessments based on prison type, prior performance, and inspection history - we show that the time elapsed since the last inspection is a strong predictor of inspection occurrence, whereas lagged violence is not. Our specifications thus control for the time since the last inspection.

We find that inspections lead to substantial and sustained reductions in violence. On average, inspections reduce overall violence by 15 incidents per 1,000 prisoners: 1 incident against staff, 6 among inmates, and 8 of self-harm. These effects grow over time and peak roughly fifteen months after the inspection - suggesting that inspections may initiate long-term institutional change rather than yielding only short-term compliance -. Effects are especially pronounced for self-harm among female inmates and violence towards prisoners in medium- and high-security prisons. These heterogeneous effects highlight how inmate vulnerabilities and organizational structure influence responses to reform. Moreover, we find evidence supporting the role of governors in explaining inspections effectiveness. The

cross-prison variation in violenceincluding when we restrict attention to hard outcomesechoing findings in firms where manager fixed effects explain systematic differences in policies and performance (Bertrand and Schoar, 2003).

heterogeneity of results across violence types supports our interpretation of violence as a multidimensional outcome and opens up numerous an agenda on the underlying mechanisms.

By linking inspections to in-prison outcomes, this paper offers insights into internal prison dynamics with important policy implications. In recent years, there has been an increase in prison spending and a concomitant worsening of performance. Our findings suggest that improving management - not just increasing resources - may be key to enhancing prison outcomes. By diagnosing the internal drivers of prison violence, this paper contributes to a better understanding of how public institutions function and how they might be reformed to serve both societal goals and human dignity more effectively.

Related Literature Our paper contributes to three strands of research on prisons. First, it contributes to the growing literature on prisoner well-being - both within economics (Campaniello et al., 2017; Cunningham et al., 2024; d'Este, 2023; Batistich et al., 2024) and in related disciplines (Useem and Reisig, 1999; Haney, 2001; Fazel et al., 2016; Favril et al., 2020; Wooldredge, 2020) - by systematically documenting the variation and persistence of different forms of violence within prisons. Second, it adds to the literature on prison management, extending beyond the traditional focus on public versus private ownership (Bayer and Pozen, 2005; Mukherjee, 2021; Bales et al., 2005; Spivak and Sharp, 2008). Third, this work connects to research on the determinants of post-release outcomes, including both the consequences of incarceration (Hjalmarsson, 2009; Barbarino and Mastrobuoni, 2014; Bhuller et al., 2020; Galbiati et al., 2021; Humphries et al., 2023) and the role of prison conditions (Katz et al., 2003; Chen and Shapiro, 2007; Drago et al., 2011; Hjalmarsson and Lindquist, 2022; Lotti, 2022; Mastrobuoni and Terlizzese, 2022; Tobón, 2022; Arbour et al., 2024; Alsan et al., 2024). By focusing on inspections as a managerial intervention with in-prison effects, this study helps uncover the mechanisms through which institutional design influences the lived experiences of incarceration.

Our paper also builds on the literature on Persistent Performance Differences (PPDs). Significant variation in performance in public sector agencies has been documented by Rasul and Rogger (2018) and Fenizia (2022) in civil service, by Best et al. (2019) in procurement, by Chandra et al. (2013) in hospitals, and by Bloom, Lemos, et al. (2015) in schools. We enrich this strand of literature by introducing a novel application in the public sector context, the prison system.

Finally, we contribute to the analysis of the role of managers and managerial practices in shaping organizational outcomes. Managers have been shown to be relevant determinants of corporate policies, organizational strategy (Bertrand and Schoar, 2003), and performance (Fenizia, 2022). Recent studies have made progress in measuring managerial practices

(Bloom and Van Reenen, 2010) and have established a positive correlation between management practices and public sector delivery: educational outcomes (Bloom, Lemos, et al., 2015), civil service project completion (Rasul and Rogger, 2018) and health in jails (Alsan et al., 2024). Aligning to the latter, we examine the effects of a management intervention on one measurable outcome of the smooth functioning of prisons, violence.

The remainder of the paper is organized as follows. Section 2 provides details about the institutional background of UK prisons, motivates and validates our choice of violence as the key outcome, and describes the data. Section 3 documents the extent of (unexplained) dispersion in violence across prisons. Section 4 introduces prison inspections, our econometric strategy, and the main results. Section 5 concludes.

2 Institutional Background & Data

2.1 Prisons in the UK

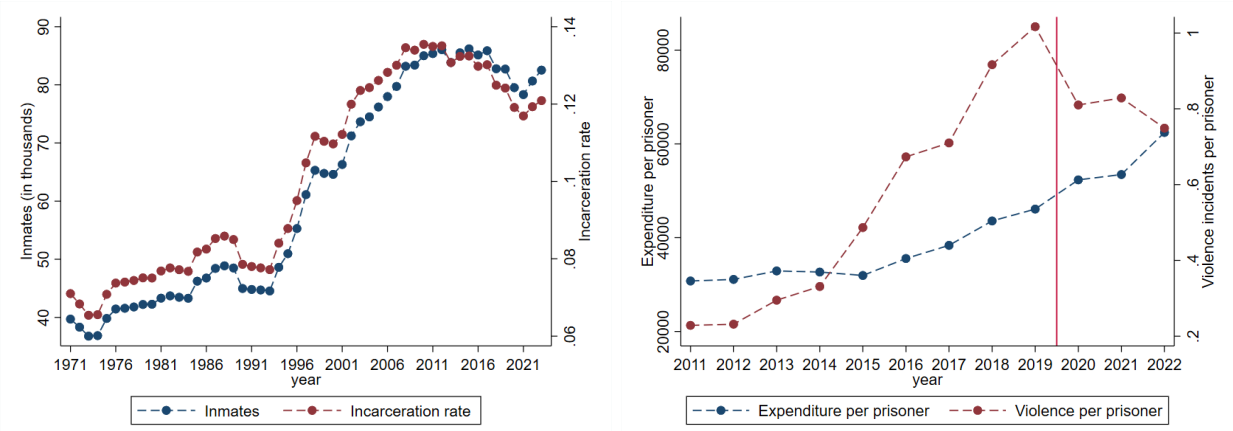
The first form of prison in the United Kingdom we are aware of is the Tower of London. During the late 18^{th} century, imprisonment in American colonies or sailing vessels (hulks) was common. In 1779, the Penitentiary Act introduced a policy of state prisons for the first time, with solitary confinement, separate sections for men and women, and religious instructions. The 19^{th} century was characterized by a pressure to the uniformity and nationalization of the local prisons, followed by major adjustments in the early 20^{th} century: shorter and less likely imprisonment for debt offence, prohibition of imprisonment under age 14, improvements in prison life conditions and rehabilitation.

As of 2024, England and Wales have the highest imprisonment rate in Western Europe.[7] The prison population has almost doubled since 1971, with a sharp increase starting in 1994, as shown in Figure 1. From 2012 to the beginning of the pandemic, we also observed growing levels of within-prison violence. Moreover, as revealed by Figure A1, the prison spending per capita has almost doubled between 2011 to 2022, while its proportion to the Ministry of Justice overall spending has risen by one third.

Nowadays most of the prisons in England and Wales are managed by His Majesty's Prison Service (HMPS) which is part of HM Prison and Probation Service (formerly the National Offender Management Service). As of 2024, there are 122 prisons in England and Wales, 14 of which are operated by private companies, Serco, G4S and Sodexo. Prisons are categorised based on different security levels based on the risk of escape, harm to the

[7]Source: World Prison Brief World Prison Brief Data

Figure 1: **Prison population, violence, and expenditure**



Notes: The left panel plots the trend in the number of inmates and the incarceration rate in prisons in England and Wales from 1971 to 2021. The right panel plots the trend in the prison expenditure per prisoner (left axis) and violence per prisoner (right axis). *Prison spending* refers to the Overall Resource Expenditure, i.e. net expenditure managed and recorded at each prison plus the net expenditure met at regional or national level (source: HMPPS). *Violence* includes self-harm, self-inflicted deaths, escapes, absconds, barricades, hostages, incidents at height, concerted indiscipline, assaults. The red line in the right panel marks the beginning of the Covid-19 pandemic.

public in case of escape, and threat to the control and stability of the prison. Adult male prisoners are allocated to closed (Category A, B, C) or open (Category D) prisons, in order of risk. Women are detained in separate open or closed prisons, and young offenders in Young Offender Institutions (YOI), Secure Training Centres (STCs) or Secure Children's Homes (SCHs), depending on their age.

The dynamics of resource allocation and constraints, workforce experience and coordination, the relationships among stakeholders (*inmates*), and between workforce and stakeholders, suggest examining prisons as *organizations* through the framework of the economics and management literature. Within this context, *Violence in Prisons* emerges as the performance measure of interest, resulting from the interplay of these underlying factors. Notably, prisons are subject to frequent inspections by the HM Inspectorate of Prisons (HMIP), which in purpose and structure are management interventions.[8] These inspections provide us the unique opportunity to analyze the effects of management interventions in public institutions, in a manner similar to consulting in private-sector firms.

[8]Refer to Section 4.1 for details.

2.2 Violence in Prison

Safety from violence is crucial to maintaining an effective penal system that respects humane conditions. The Prison Service committed to the protection of detainees: "Her Majesty's Prison Service serves the public by keeping in custody those committed by the courts. Our duty is to look after them with humanity and help them lead law-abiding and useful lives in custody and after release" (HM Prison Service's statement of purpose). Violence, as the externalization of detainees' psychological discomfort, provides an indicator of the health of the prison environment and of the soundness of the prison as a public organization. Moreover, within-prison violence advises on the evaluation of the welfare effects of incarceration. Conducing to harsh prison conditions, it has been found to increase recidivism (Drago et al., 2011) while diminishing deterrence (Katz et al., 2003)[9]. As such, violence is an essential concern for humanitarian reasons but also from a policy perspective.

From a methodological standpoint, violence[10] is of great convenience for our analysis as it is an objective and measurable outcome. From a substantive point of view, violence unveils the emotional and psychological state of detainees: it offers an unfiltered lens into inmates' perceptions of fair treatment, the prevailing prison culture, their interpersonal relationships with cellmates and staff, and the psychological burden of imprisonment, including feelings of isolation, anxiety, and concern about the future.

Although violence can be seen as the ultimate outcome of the interplay of different feelings, the specific types of violence can shed light on the nature of the discomfort expressed by inmates through violence. The criminology and psychology literature provides an extensive examination of the emotions underlying offending and victimization in prisons, exploring their risk factors and the association with pre-incarceration traits and in-prison experiences (Favril et al., 2020; Useem and Reisig, 1999; Haney, 2001; E. F. v. Ginneken and Wooldredge, 2024; O'Donnell and Edgar, 1999). This body of research suggests that we categorize the manifold facets of inmates' distress into three broad categories:

1. *Anger*, often linked to the dependence on the institutional structure, deprivation, and enforced powerlessness;

2. *Fear*, stemming from interpersonal distrust, suspicion, dangerousness;

[9] Both (Drago et al., 2011) and (Katz et al., 2003) exploit death rate from all causes to proxy for prison harshness.

[10] In the literature, the term prison violence refers to violence at the aggregate level, e.g., protests, and at the individual level, e.g., violence on inmates and staff (Wooldredge, 2020). Throughout our analysis, we will address only prison violence at the individual level to insulate our analysis from the social dynamics of group violence.

3. *Despair*, arising from alienation, psychological distancing, isolation, diminished sense of self-worth, post-traumatic stress.

While they are not directly observable in a consistent manner, violence, in its various forms, is. This motivates our focus on three different observable indicators of violence: *violence on staff* (assaults on staff), *violence on prisoners* (assaults on prisoners and homicides), and *self-violence* (self-harm and suicides). These different manifestations of distress serve as markers of the underlying emotional states of anger, fear, and despair, respectively.

To validate the associations between the three measures of violence and the three forms of distress, we draw on data from the HMIP Prisoner's Survey, administered at the outset of the prison inspections. The survey, offered to a random sample of detainees, covers the prisoner experience from reception to release, including the relationship with the staff, the contacts with family and friends, the time out of cell, healthcare, education, behavior management, together with demographic and background questions (refer to Section A.2.1 for more details). We focus on three questions – "Do most staff here treat you with respect?", "Do you feel safe here?", and "Do you have any mental health problems?" – which, respectively, seem good proxies for the sentiments generating anger, fear, and despair, as described above. We inspect the association between the three measures of violence and the three questions. The top-left panel of Figure 2 shows that establishments with higher rates of self-reported respectful treatment are associated with lower levels of violence against staff. This finding aligns with our hypothesis that such violence is driven by feelings of anger stemming from a sense of powerlessness. The top-right panel reveals a negative relationship between perceived safety and prisoner-on-prisoner violence, consistent with our view that this type of violence reflects feelings of interpersonal distrust and suspicion among inmates. Finally, our third hypothesis posits that alienation, isolation, and post-traumatic stress, conceptualized as forms of despair, are associated with higher self-violence. Our validation exercise in the bottom panel confirms that self-violence is indeed linked to mental health problems.

While using violence as an objective and measurable metric is convenient, concerns arise regarding the accuracy of violence reporting by prison establishments. Assaults and self-harm incidents must first be recorded by officials in paper records (wing observation books), before being entered into the National Offender Management Information System (NOMIS)[11] HMPPS Incident Reporting System (IRS). The IRS is then used to produce internal reports and official statistics, which we rely on for our analysis. It is, however, possible that an incident recorded in the wing observation books is never recorded in the IRS. To monitor data integrity, each establishment is audited once a year, assessing compliance with IRS data

[11]National Offender Management Information System (NOMIS) began its rollout to all prisons in 2010, achieving comprehensive use by 2012. It replaced the Local Inmate Database System (LIDS).

recording quality procedures. According to the July 2018 issue of the Safety in Custody Bulletin, the quality audit across the prison estate in the 2016/17 and 2017/18 performance years revealed around 10% and 11% of under-reporting for assaults and self-harm incidents, respectively. Moreover, practitioners suggest that reporting practices may differ significantly between private and public prisons. Private prisons, which require government authorization to operate, tend to be more meticulous in reporting all incidents, as failing to do so could jeopardize their license renewals.

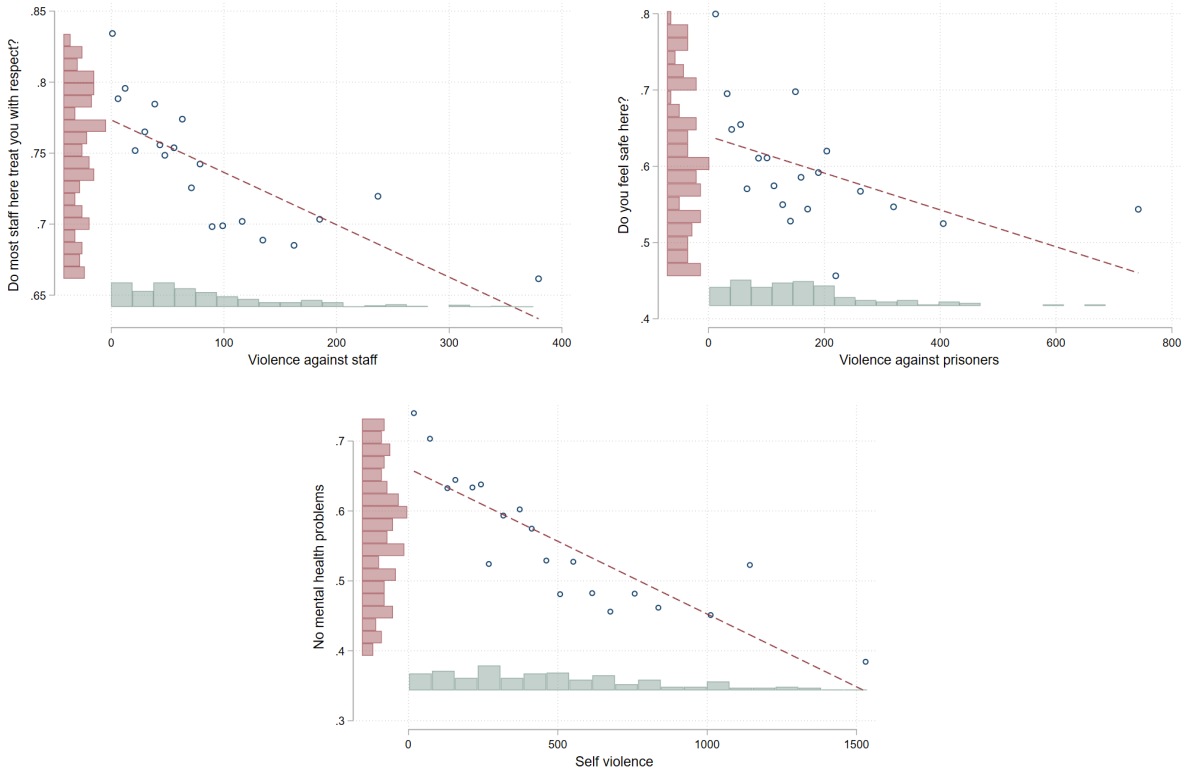
Unfortunately, data on prison-specific IRS quality audits are not available. Given that the ultimate goal of our analysis is to understand the relationship between management practices and violence, under-reporting might complicate the interpretation of our results if a correlation exists between under-reporting and the relationship between violence and management quality. To reassure against this concern, we validate our exercises using *hard* equivalents for each type of violence, i.e. more severe types of violence, which are subject to different reporting procedures. For instance, Prisoner Discipline Procedures Prison Service Instruction (PSI) 05/2018 mandates that serious assaults should be referred for police investigation[12], while deaths are notified to, among others, the police, the next-of-kin, and the coroner, as stipulated by the Deaths in Custody Framework. Accordingly, we define *Serious assaults on staff*, *Serious assaults on prisoners*, *Self-harm requiring hospitalization*, *homicides*, and *suicides* as our *hard* measures of violence. Throughout our work, we will employ these measures to corroborate our findings against under-reporting concerns.

Within-prison violence is a key indicator of the prison system's overall well-being, institutional integrity, and the psychological state of detainees. As such, it is the focus of our analysis. Crucially, it is the distinction between different forms of violence towards staff, between inmates, and self-directed that allows us to understand how prison features and management interventions impact prison conditions by revealing distinct patterns of inmate distress.

[12] According to Prisoner Discipline Procedures Prison Service Instruction (PSI) 05/2018 serious assaults are "those where the injuries include bones, broken skin, or serious bruising, and those where the assault was pre-planned rather than spontaneous, those where the alleged offender has a previous history of violence during the current period in custody, the victim's role within the establishment (e.g. staff), their vulnerability, and the location of the incident, will also be factors". According to PSI, (inter alia) serious assaults should always be referred to the police.

[13] As detailed in Section 2.3, HMIP Prisoners' Survey are conducted at the beginning of each prison full inspection. Therefore, survey responses are available only for the inspection-year pairs in which a full inspection occurred. Out of the 586 inspections, a total of 83,540 prisoners completed the survey. 91%, 64%, and 76% of them completed the above-used questions "Do most staff here treat you with respect?", "Do you feel safe here?", and "Do you have any mental health problems?", respectively.

Figure 2: **Inmates' distress and violence**



Notes: The figure shows the binscatterplots of violence measures, *Violence on staff*, *Violence on prisoners*, and *Self-violence* with, respectively, elicited perceptions of respectful treatment, safety, and mental health, from HMIP Prisoners' Survey from 2003 to 2022[13]. The top-left panel plots *Violence on staff* against survey question "Do most staff here treat you with respect?", the top-right *Violence on prisoners* "Do you feel safe here?", and *Self-violence* against "Do you have any mental health problems?". Violence measures are reported in rates per thousand prisoners and are standardized. Survey responses are expressed in the share of affirmative answers for the respondents.

2.3 Data

This Section outlines the data used in our analysis. We constructed a balanced panel of 114 prisons based in England and Wales, spanning 2004 to 2019. The dataset captures cross-sectional and longitudinal information, mostly from the Ministry of Justice and the HM Prison and Probation Service, on various aspects of the prisons, including inmate demographics, staffing levels, capacity, security classifications, programs offered, and inmates' behavior. We complement these data with reports and surveys from HM Inspectorate of prisons' inspections. We offer a brief description here and refer the reader to Sections 4.2 and A.2 for more details.

Prison Digest Data is primarily sourced from the Ministry of Justice and the HM Prison

and Probation Service. Specifically, we collected information on the prison population, including sex, age, custody status, ethnicity, offense type, and nationality, the cases of deaths, assaults, protests, escapes, absconds, failures to return from temporary release, and releases in error. We also gathered data on the percentage of prisoners held in crowded accommodations, the findings of illicit items, the Community Accommodation Services[14] referrals, and the completion of accredited programs. Workforce data includes the composition by role, the number of working days lost, and the proportion of ethnic minority staff. Other relevant data includes the percentage of positive random mandatory drug tests and the cost per place and prisoner. The mentioned data are publicly available for all public and private prisons except for the workforce and costs data, which are available only for public prisons.

Through a Freedom of Information request, the Ministry of Justice provided the names of the acting governors, which were recorded annually from 2013 to 2022.

From the Centre for Crime and Justice Studies, we collected information on the type of prison, its operator, opening year, land area, and perimeter size. The postal code for each prison is sourced from Google Maps. We also gathered data on the crime rate at the constituency level.

The variables mentioned above are available at different frequencies for different periods and are disaggregated by categories. For detailed information on the periodicity and time coverage, refer to Table A1 in the Appendix.

Table 1 provides descriptive statistics. Panel A reveals that most prisons are male facilities (90.2%), and the majority are Category B (39.4%) or Category C (9.8%), with only a small percentage being private (9.9%). Prison capacity varies significantly, ranging from 14 to 457 thousand square meters and population from 5 to 2080 inmates. Panel B provides the summary statistics for the *within-prison violence measures* in rates per thousand prisoners for each type of violence: Assaults on staff, serious assaults on staff, assaults on prisoners, serious assaults on prisoners, homicides, self-harm incidents, suicides, and self-harm incidents requiring hospitalization. As anticipated in Section 2.2, throughout the analysis, we categorize our measures of violence as follows: *Violence on staff*, which includes all assaults on staff; *Violence on prisoners*, comprising all assaults on prisoners and homicides; and *Self-violence*, which consists of all self-harm incidents and suicides. To validate our results and address potential under-reporting concerns, we also define a stricter subset of these indicators: *Hard violence on staff* (serious assaults on staff), *Hard violence on prisoners* (serious assaults on prisoners and homicides), and *Hard self-violence* (self-harm incidents requiring

[14]The Community Accommodation Service, Tier 2 (CAS-2), is a contracted service that provides short-term accommodation for those who have no suitable accommodation after detention and may otherwise be held in custody.

hospitalization and suicides). The data reveal high levels of violence in prison across all types: On average, prisons experience 81 assaults on staff, 237 assaults on prisoners, 691 cases of self-attacks, and 1 suicide each year. The data also uncovers substantial variation in violence across prisons and years, especially for self-attacks, suggesting huge disparities between facilities and years.

Table A2 in the Appendix reports the summary statistics for the inmates' composition by sex, age, custody status, ethnicity, and offence type, cost per place and prisoner, and the staff composition by role. Table A3 in the Appendix presents the crime rate at the constituency level and a market index of crime at the prison level, for violent crimes and drug-related crimes.

Prison Inspections We complement the Prison Digest data with information from the 818 inspections conducted by the HM Inspectorate of Prisons from August 2004 to December 2019. As detailed in Section 4.1, inspections produce an inspection report, an action plan, and a prisoners' survey. Specifically, for each inspection from 2013 to 2019, we gathered details from the corresponding publicly available inspection report, including the inspection date, the publication date, the inspection team, the number and content of the recommendations issued by the inspection team, and the number of recommendations implemented since the previous inspection. The inspection Action Plan supplements those data with the governor's response (acceptance or refusal of the recommendation, and in case of acceptance, the action plan and target date) to the inspection recommendations. Additionally, we accessed the HMIP Prisoner Survey, detained at UK Data Service, which collects the anonymized inmates' responses to the survey conducted during each inspection. The HM Inspectorate of Prisons Research team provided us with the inspection scores.

Summary statistics about the inspections per se, the inspection report content, and the inspection survey are provided in Section 4.2. More details about the inspection survey are presented in Section A.2.1.

3 Persistent Differences in Violence Across UK Prisons

Descriptive statistics in the previous section revealed substantial variation in violence across prisons. Borrowing the methods from the Persistent Performance Differences (PPDs) literature, we ask whether these differences persist when we control for the inmates' demographics, prison capital, and prison staff levels. While our economic background suggested framing

Table 1: **Prison and within-prison violence descriptives**

	N	Mean	St. Dev.	Min	Max
Panel A: *Prison information*					
Private prison	114	0.099	0.299	0	1
Male prison	114	0.902	0.297	0	1
Category A	114	0.089	0.285	0	1
Category B	114	0.394	0.489	0	1
Category C	114	0.098	0.297	0	1
Open	114	0.053	0.225	0	1
Reception	114	0.089	0.285	0	1
Young Offenders Institutions	114	0.249	0.433	0	1
Victorian prison	114	0.278	0.448	0	1
Prison Area in sq metres	114	119408	82458	14307	457456
Number of years since foundation	114	87.779	69.781	12	430
Number of inmates	114	697.628	346.706	5	2080
Panel B: *Within-prison violence measures*					
Assaults on staff	570	81.283	112.804	0	1340
Serious assaults on staff	570	8.149	11.215	0	100
Assaults on prisoners	570	236.665	371.897	0	5280
Serious assaults on prisoners	570	27.021	33.346	0	660
Homicides	570	0.044	0.246	0	2
Self-harm incidents	570	690.767	4194.661	0	115600
Suicides	570	1.118	2.273	0	40
Self-harm incidents w hospitalization	570	34.111	70.606	0	1800

Notes: The table shows summary statistics for the sample of HMPPS prisons for 2013-2019. Statistics in Panel A are computed on the cross-sectional data. Statistics in Panel B are computed on the longitudinal data, at the prison-year level. Violence measures are reported as the rate per thousand prisoners.

violence as the outcome of the production function of inmates' demographics, prison capital, and prison staff level, it is also consistent with the criminology theories. Specifically, the *Importation theory* posits that pre-existing traits, values, ideologies, and experiences, shaped by inmates' demographics, influence in-prison offending, while the *Deprivation theory* emphasizes the role of prison conditions, regime characteristics and individual confinement experience, proxied by prison capital and prison staff, in shaping in-prison offending (E. F. v. Ginneken and Wooldredge, 2024).

Section 3.1 provides details about the measurement of PPDs. Section 3.2 examines the estimated dispersion and persistence in violence across UK prisons.

3.1 Measuring PPDs

Thanks to the availability of detailed longitudinal data on production activities at the establishment or firm level, researchers have extensively analyzed productivity dispersion. Significant and persistent differences in productivity are observed even within narrowly defined

industries. For example, Syverson (2004) documents a logged total factor productivity (TFP) ratio of 1.92 between the 90^{th} and 10^{th} percentiles in the U.S. manufacturing sector. These productivity disparities exhibit persistence over time, with autoregressive coefficients typically ranging from 0.6 to 0.8 (see Ábrahám and White, 2007; Foster et al., 2008).

We analyze Persistent Differences in *Violence* following the Persistent Productivity Differences (PPDs) literature. Essentially, the counterpart of productivity in our analysis is the unexplained violence within prisons once we account for all the inputs. Similarly to PPDs research, we quantify the dispersion in violence by the interquartile productivity difference divided by the industry median productivity level.

We propose two exercises. First, we estimate the prison fixed effects $\hat{\zeta}_i$ from the regression of the rate of violence per prisoner v_{it} on the inputs Q_{it}, K_{it}, and L_{it}, and year and prison fixed effects, η_t, ζ_i:

$$v_{it} = \alpha Q_{it} + \beta K_{it} + \gamma L_{it} + \eta_t + \zeta_i + \epsilon_{it} \tag{1}$$

where the inmates quality controls Q_{it} include the share of inmates per type of custody, ethnicity, age, offense, and the total number of inmates; prison capital K_{it} controls for the average cost of holding one prisoner for the year from the Direct Resource Expenditure; human capital controls L_{it} controls for the number of personnel, and for the average number of staff per band per prisoner. Second, aligning with Syverson (2004)[15], we estimate the residuals $\hat{\epsilon}_{it}$ from the yearly regression of the log of violence within prisons V_{it} on the logged inputs Q_{it}, K_{it}, and L_{it}, and prison fixed effects, η_t, ζ_i:

$$ln(V_{it}) = \alpha ln(Q_{it}) + \beta ln(K_{it}) + \gamma ln(L_{it}) + \zeta_i + \epsilon_{it} \tag{2}$$

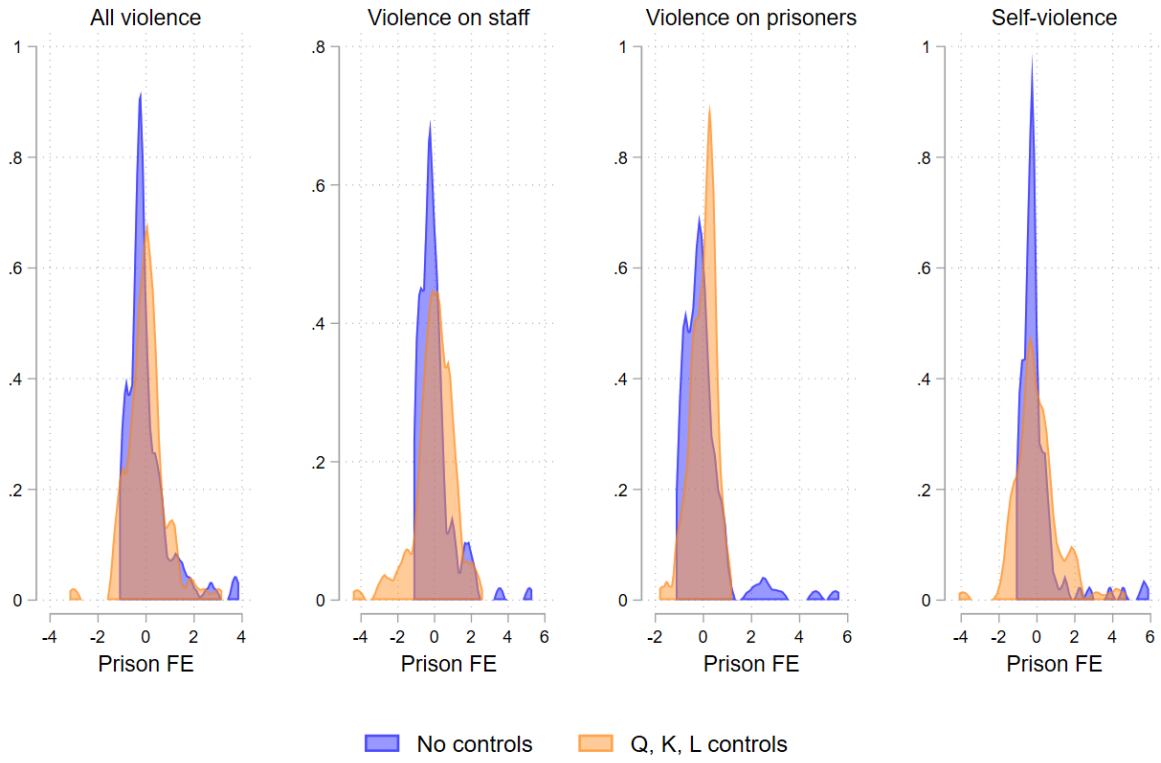
3.2 Dispersion

In this Section, we document the dispersion in violence within prisons across UK prisons. Figure 3 plots the density of prison fixed effects estimated from equation 1, both with and without controls for inmate quality, physical and human capital, across three types of violence: violence on staff, violence on prisoners, and self-violence. Table 4 provides the corresponding estimates for the mean and interquartile range of the prison fixed effects, with and without controls, for each violence measure.

A prison at the 75^{th} percentile of the prison FE distribution experiences 80 (226, 389) more incidents of violence on staff (violence on prisoners, self-violence), per thousand prison-

[15] In Syverson (2004) the plant TFP is computed as the log of the physical output minus a weighted sum of the logged labor, capital, materials, and energy inputs. The weights are the input elasticities, proxied by the industry-level input cost shares. Note that in our analysis the weights are the estimated coefficients at the prison level.

Figure 3: **Density of Prison Fixed Effects**



Notes: The figure plots the density of the prison fixed effects, estimated from equation 1, with and without controls for inmates' quality (Q), physical capital (K), and human capital (L), across three types of violence: violence on staff (left panel), violence on prisoners (center panel), and self-violence (right panel). The outcomes of interest are measured in number of incidents per thousands of prisoners. Only public prisons are included.[16]

ers, than a prison at the 25^{th} percentile, when only year-specific unobservables are controlled for. When controls for inmate composition and physical and human capital are included, the cross-prison dispersion generally increases. Specifically, a prison at the 75^{th} percentile of the prison FE distribution experiences 113 (184, 709) more incidents of violence on staff (violence on prisoners, self-violence), per thousand prisoners, than a prison at the 25^{th} percentile. The observed widening in the interquartile range with the inclusion of controls suggests that differences in inmate composition and input allocation may obscure the underlying heterogeneity in prison violence.

We investigate the margins contributing the most to the observed variation in time-invariant prison violence through three exercises. First, we plot the estimated prison fixed

[16]Data on prison staff levels and composition are available only for public prisons. Therefore, whenever we include these variables in the estimation exercise, we restrict the sample to public prisons.

Figure 4: **IQ of Prison Fixed Effects**

	Prison Fixed Effects			
	All violence	Violence against staff	Violence against prisoners	Self-violence
Mean	986.30	97.99	260.31	627.99
IQ range - No controls	0.70	0.83	0.87	0.62
IQ range - Q, K, L controls	0.83	1.16	0.71	1.13

Notes: The table shows the Interquartile Range (IQ) of the prison fixed effects, estimated with and without controls, for each violence measure. The outcomes of interest are measured in the number of incidents per thousands of prisoners. The IQ is divided by the mean of the outcome measure. Only public prisons are included.

effects from equation 1[17] across key characteristics: ownership (public or private prison), gender (female or male prison), prison function (Category A, Category B, Category C, Category D, Reception, Youth Justice Board), and Victorian prison[18], shown respectively in Figures B1, B2, B3, and B4 in the Appendix. Second, we regress the prison fixed effects for each violence outcome on dummies for the prison function, private ownership, male prison, Victorian design, and prison area, both absolute and per capita (Table B1). Third, we conduct a variance decomposition exercise to assess the importance of different factors by evaluating the loss of fit when excluding time-invariant prison observables from the saturated specification.

The results in Figures B1, B2, B3, and B4, and in Table B1 reveal several key patterns. High-Security prisons (Category A) are associated with fewer cases of violent incidents, across all types, likely reflecting the effectiveness of their stringent security measures. Interestingly, female prisons exhibit significantly higher incidents of self-harm and suicides compared to male prisons. This finding aligns with the anecdotal evidence suggesting that incarcerated women face unique challenges, including heightened gender discrimination, enhanced by the stigma associated with criminality, stress and depression from pregnancies and separation

[17]Note that in order to investigate the ownership as a prison characteristic on violence levels, when we estimate equation 1 we exclude the controls for the staff composition, not available for private prisons. Thus, for the sake of this first exercise, we only include public prisons.

[18]Between 1842 and 1877, a significant prison-building initiative led to the construction or substantial expansion of 90 prisons. These facilities commonly featured a hub-and-spoke design, with cell wings radiating from a central hub. Cells were small, intended for single occupancy, and equipped with features like candle alcoves for Bible reading and larger windows in some cells to provide fresh air for prisoners with tuberculosis. From 1840, the Separate System was implemented, emphasizing solitary confinement and silent contemplation to encourage reform. However, this isolation often led to increased rates of mental illness among inmates. Following the 1865 Prisons Act, the Silent System became prevalent, focusing on silence and hard labor, with strict penalties for communication between prisoners. Currently, 32 Victorian-era prisons remain operational in England and Wales, housing approximately 22,000 prisoners, which accounts for about a quarter of the prison population. The project "The Persistence of the Victorian Prison: Alteration, Inhabitation, Obsolescence and Affirmative Design", directed by the University of Birmingham and the University of Bath, provides the list of the Victorian prisons.

Table 2: **Comparison to PPDs estimates**

Residuals	Interquantile range (75^{th}p-25^{th}p)	Interdecile range (90^{th}p-10^{th}p)
All violence	0.305	0.607
Violence on staff	0.348	0.669
Violence on prisoners	0.283	0.558
Self-violence	0.353	0.719
Syverson, 2004 - Concrete market	0.290	0.651
Chandra et al., 2013 - Hospitals	0.230	0.440

Notes: The table shows the estimated Interquantile range and Interdecile range of the residuals from equation 2 for each violence measure. The numbers shown are means of the annual values from 2013 to 2019 weighting each year equally. The Syverson (2004) statistics refer to the across-industry mean of the within-industry productivity moment.

from newborns, and higher rates of family abandonment. These women also experience physical discomfort in facilities designed for men and marginalization due to being a minority within the prison population. Private, with respect to public prisons, and Victorian, with respect to non-Victorian prisons, show slightly higher violence incidence. Prison area significantly predicts lower violence. Adequate space for cells and communal facilities is presumed to improve mental health, hygiene, and dignity among inmates, while reducing competition for resources and opportunities for violent interactions. As shown in Table B2, prison function and prison area explain 47% (7%, 30%) and 66% (20%, 15%)[19] of the variation in prison fixed effects for violence on staff (violence on prisoners, self-violence), respectively. In comparison, private ownership and Victorian design account for only 1% (13%, 2%) and 11% (1%, 0%) of the explained variability.

The unexplained variation in violence across prisons we observe is comparable in magnitude to the productivity variation across manufacturing plants observed by previous research. One might infer that the dispersion in unexplained violence within centrally governed prisons is notably lower than that across manufacturing plants within the same industry, where prices and competition shape entry and exit. To benchmark the dispersion in prison violence against the private-sector productivity dispersion reported by the literature, we estimate the residuals of log violence from equation 2 year by year. Table 2 reports the average interquartile range of these residuals, with equal weighting for each year. The estimated dispersion for the three types of violence exceeds but remains broadly comparable to the estimates in the private sector by Syverson (2004) and in the public sector by Chandra et al. (2013).

The cross-sectional variations in violence on staff, violence on prisoners, and self-violence

[19]Note that the reported shares reflect the loss of fit resulting from removing one by one the controls from a linear projection of each outcome on the set of fixed effects in the specification controlling for dummies for prison function, private prison, male prison, Victorian prison, and prison area. As the shares are calculated from R^2 when removing the controls one by one, the percentages should not be thought as additive.

exhibit substantial correlations, reflecting underlying behavioral patterns. This is consistent with the expectation that sentiments of anger, fear, and despair - key drivers of such actions - are closely interconnected. Figure B5 in the Appendix presents bilateral scatterplots for the three violent incidents. The prison fixed effects for violence on staff and violence on prisoners show a strong positive correlation, suggesting that anger and fear often coexist and may reinforce each other. A lack of control over one's environment or perceived threats can intensify both emotions: anger may drive confrontational behaviors, while fear triggers defensive responses. Moreover, anger can escalate situations that provoke fear in others, creating a cyclical dynamic. Interestingly, the prison fixed effects for self-violence exhibit a mild positive correlation with violence on staff but a negative correlation with violence on prisoners. This pattern is further supported by the negative loading of self-violence in the Principal Component Analysis (PCA) on prison fixed effects (Figure B6 in Appendix). Despair, often associated with feelings of hopelessness and resignation, appears less compatible with proactive or defensive responses driven by fear, which may explain its weaker connection to violence on prisoners.

One might be concerned about endogenous violence-related factors confounding our dispersion estimates. For example, drugs in prison are both a cause and consequence of violence. Drugs are often used as a coping mechanism for stress or mental illness, but addiction and trafficking exacerbates erratic behavior and conflict, further fuelling violence. To address this concern, in Appendix B.4 we replicate 3 and 4 including a control for the number of drug finds. Similarly, the crime rate in the surroundings of the prison could also be correlated with the error term. For instance, prisons with high levels of violence could contribute to fostering criminal networks locally, or unobserved factors such as cultural attitudes toward crime may influence both crime rates and violence within the prison. To examine this, we construct a "market access" index of crime in the catchment area of the prison, taking an average of the crime rate with weights inversely proportional to the distance constituency-prison. We replicate our analysis in Appendix B.4, controlling for our index of crime. Our estimates remain robust to these additional controls: The distribution of prison fixed effects, with and without drug finds and crime rates, remain similar (see Figures B8 and B9, and Tables B8 and B9).

An additional concern pertains to the potential under-reporting of violence in prisons, as anticipated in Section 2.2. To address this, we re-estimate the prison fixed effects and the interquartile range using the measures least susceptible to misreporting for each type of violence in our analysis. Specifically, we employ *serious staff assaults* for *violence against staff*, *serious prisoner assaults* and *homicides* for *violence against other prisoners*, and *suicides* and *self-harm incidents requiring hospitalization* for *self-violence*. Figure B10 and Tables B10,

B11, and B12 indicate that under-reporting is unlikely to affect our findings substantially.

The analysis above suggests that firms at the 75^{th} percentile experience significantly more cases of violence than prisons at the 25^{th} percentile. To investigate the unexplained dispersion, we first examine its persistence over time. In Section B.2 we estimate the annualized autocorrelation and tri-annual transition matrices. In Section B.3, we inspect the accountability of prison governors in explaining the violence dispersion documented above. Overall, our analysis reveals very weak persistence in in-prison violence, particularly for prisons at the extremes of violence levels. The comparison between the persistence estimates with and without controls for prisoners' and prisons' characteristics and inputs suggests stable prison characteristics and a prison system resource allocation policy that targets high-violence establishments. The latter is also consistent with the widening in the interquartile range with the inclusion of controls, documented earlier in this Section, and the explanatory power of the appointed prison governors, noted in Section B.3.

4 Management and Violence Across UK Prisons

The observed productivity dispersion (PPD) within and across manufacturing industries has spurred extensive research into its underlying drivers. Studies have explored factors such as human and physical capital quality (Ilmakunnas et al., 2004; Sakellaris and Wilson, 2004), information technology (Brynjolfsson et al., 2007), product innovation (Bernard et al., 2010), and firm structure decision (Hortacsu and Syverson, 2009). Additionally, a growing body of literature highlights the positive relationship between managerial practices and managers and organizational performance in both the private sector (Bertrand and Schoar, 2003; Bloom and Van Reenen, 2007; Giorcelli, 2019; Bandiera et al., 2020) and the public sector - schools (Bloom, Lemos, et al., 2015), hospitals (Chandra et al., 2013), public procurement (Best et al., 2019), administrative public agencies (Fenizia, 2022), civil service (Rasul and Rogger, 2018). Fenizia (2022) finds that a one-standard-deviation improvement in managerial quality increases office productivity by 10%. However, these findings may not be generalized across all public agencies, as they differ in the agency mission, output, size, level of bureaucracy, and managers' hiring criteria and incentives.

This study is the first to examine the impact of managers and management practices on organizational performance within the prison system. In Annex B.3, we explore the role of prison governors the senior operational managers in the Prison Service of England and Wales. Our analysis of prison governors is currently constrained by the infrequency of available data (collected annually) and limited demographic and career information.

In the following section, we shift focus to a detailed investigation of prison inspections,

which represent a particularly salient form of management intervention. Inspectors assess the quality of policies and procedures in practice within the prison related to areas like safety, respect, healthcare, education, and resettlement. Findings highlight areas needing improvement, resulting in specific recommendations for the establishment to address. Following the inspection, prison governors and managers create action plans detailing specific steps to address the key concerns and recommendations highlighted by the inspectors. The independent inspection body, HMIP, provides an external perspective on prison operations, holding management accountable for prison status.

Inspections are especially valuable for analysis because they constitute a direct managerial mechanism, objectively diagnosing key concerns and identifying a strategic action plan. Moreover, reports and action plans constitute an incredible source of rich and structured information on prison practices. Such comprehensive insights into managerial practices are exceptionally rare in both the public and private sectors, making this an unusually informative setting for studying management interventions.

4.1 Background on Inspections

The HM Inspectorate of Prisons (HMIP) carries out inspections in the prisons in England and Wales. Chief Inspectors have identified their mandate as "ensure independent inspection of places of detention, report on conditions and treatment and promote positive outcomes for those detained and the public." (*Inspection framework* 2022).

Inspections arise from the obligations of the UK status as a party to the Optional Protocol to the United Nations Convention Against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment (OPCAT). As of 2024, 90 countries are parties to the OPCAT, which provides for a system of international and national visits to all places of detention. At a minimum, OPCAT requires that visiting bodies "are functionally independent with independent personnel, have sufficient expertise, a gender balance, and adequate representation of ethnic and other minorities, are provided with the necessary resources", and "have the powers to regularly examine the treatment of persons deprived of their liberty in places of detention, make recommendations to the relevant authorities with the aim of improving the treatment and conditions of detainees [...] submit proposals and observations concerning existing or draft legislation, access all information concerning the number, location and treatment of all persons deprived of their liberty, access all places of detention, have private interviews with all persons deprived of their liberty as well as any other person who may supply relevant information, choose the places they want to visit and the persons they want

to interview [...]".[20]

Figure 5 presents the timeline of a typical inspection[21]. When an inspection is unannounced, prisons receive no more than 30 minutes' notice. A team of 15-20 inspectors and researchers conducts the assessment over two weeks. The first week of the inspection[22] lasts two days and involves an initial view of the environment and staff-prisoner relationship, along with a preliminary examination of areas of concern, as identified by intelligence, third parties, or previous inspections. During the first two days, meetings with the governor are carried out to ensure a clear understanding of the inspection process and to request documentation. Moreover, detainees and staff surveys are distributed. The second week lasts five days and includes: (1) detailed briefings with governors; (2) the collection of additional documentary evidence, such as observation books, P-Nomis (Prison Service IT system), performance data from the HMPPS performance hub, daily wing entries, care plans and detention and training order/sentence plans; (3) the collection of detainees and staff surveys, and private interviews with detainees and staff; (4) further site visits during different times of the day and at different locations.

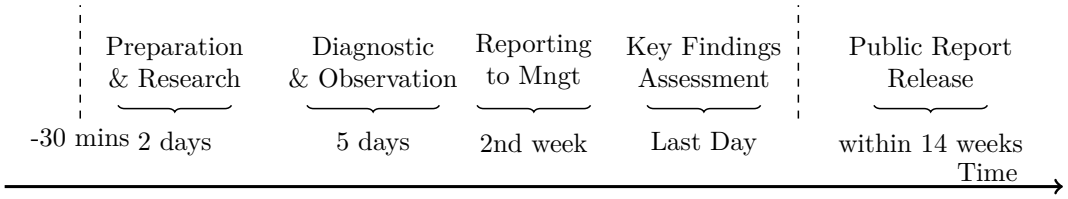
Managers receive daily updates on emerging findings. On the penultimate day of the inspection, the Chief Inspector of Prisons attends to contribute to the assessment process. On the final day of the inspection, the key findings and assessment against the four applicable Healthy establishment tests set out by the HMIP Inspection Framework (*Safety,*

[20]The UKs His Majestys Inspectorate of Prisons (HMIP) runs a routine, largely unannounced national inspection programme with *standardised judgements* against four healthy prison tests and publishes prison-level reports; it also fields a prisoner survey whose microdata are archived for research: HMIP How we inspect; example report pages with scores: HMP Peterborough (Men); survey data via the UK Data Service: HMIP Prisoner Survey. Comparable OPCAT-mandated bodies operate elsewhere in Europe, but outputs are typically less standardised and prison coverage less frequent relative to estate size. In **France**, the Contrôleur général des lieux de privation de liberté (CGLPL) publishes narrative visit reports and recommendations (e.g., annual and visit reports here: CGLPL publications). In **Germany**, the National Agency for the Prevention of Torture reported *66* visits across all places of detention in 2022, with limited prison-specific coverage and no national scoring framework: Annual Report 2022. Nordic NPMs led by Parliamentary Ombudsmen also conduct regular visits (e.g., Swedens JO OPCAT inspections: JO OPCAT reports), but do not provide a directly comparable, standardised prison-level ratings series with public microdata akin to HMIP.

[21]In case of considerably poor healthy establishment tests, ability to follow issued recommendations, or high risk for detainees, established during an inspection, the Chief Inspector informs the Secretary of State for Justice within seven calendar days of the end of the inspection, after which the same communication is published within 24 hours. The Secretary of State has 28 calendar days to respond to the concerns raised in the publicly published letter, called Urgent Notification. The HMIP will then re-inspect the prison. In addition to the inspection program, the HMIP also carries out independent reviews of progress (IRPs), which are meant to examine the prisons' progress with respect to the areas of concern and recommendations identified during the last inspection. By *typical* inspection, and throughout our empirical analysis, we exclude inspections that follow an Urgent Notification and the Independent Reviews of Progress (IRPs).

[22]For announced inspections, the first week of the inspection is conducted on a pre-arranged date, four to six weeks in advance of the inspection.

Figure 5: **Inspection timeline**



Notes: The Figure plots the timeline of the HMIP inspections. When unannounced, 30 minutes' notice is given to the establishment. During two days of the first week, a preliminary view of the establishment and the staff-prisoner relationship, meetings with the governor and detainees, and staff surveys are carried out. During up to five days in the second week of inspection, diagnosis of issues identified during the first week is executed, collecting additional evidence, interviewing detainees and staff, and conducting further site visits. During the two weeks of inspection, governors and managers are daily updated on the inspection team's findings. On the final day of the inspection, the assessment is presented to the establishment management, with the relevant report being published within 14 weeks of completion.

Respect, Purposeful activity, Rehabilitation, and *Release Planning/Resettlement*) are presented to the establishment. Assessment methodology relies on triangulation, meaning that the assessments are always corroborated on at least three different sources, balanced across qualitative and quantitative (e.g., data from the P-Nomis (Prison Service IT system) or the HMPPS performance hub, and conversations with detainees and staff or with third parties).

HMIP aims to publish inspection reports within 14 weeks of completion. The reports highlight the main areas of concern and the detailed recommendations that need to be addressed by prison leaders (e.g., director/governor, Head of Healthcare, Head of Offender Management Unit, Prison Estate) and assesses the progress made in implementing concerns or recommendations from previous inspections, for each healthy establishment test. Prisons are required to respond within three months with an action plan outlining steps taken or planned, including specific timelines and clear ownership for each action. Table A4 in the Appendix provides examples of recommendations, along with the corresponding actions planned or taken in response. Follow-up inspections assess compliance and improvements based on these recommendations.

4.2 Inspections Data

From August 2004 to December 2019, 818 inspections were conducted by the HM Inspectorate of Prisons. As anticipated in Section A1, we gathered data from the inspection report, the action plan, and the prisoners' survey, outputs of the inspection process.

Table 3 provides descriptive statistics about inspections per se, inspection reports' content, and inspection surveys, respectively, in Panels A, B, and C.

On average, between 2004 and 2019, prisons were inspected five times every 28 months.

Table 3: **Inspection descriptive statistics**

	N	Mean	St. Dev.	Min	Max
Panel A: *Inspections*					
Number of inspections per prison	818	7.43	3.13	2	18
Months since last inspection	818	27.94	10.98	1	58
Full inspection	818	0.76	0.43	0	1
Unannounced inspection	818	0.66	0.48	0	1
Size inspection team	818	19.50	3.18	7	36
Panel B: *Inspection reports*					
Months between inspection and report publication	307	4.62	1.06	0	9
Score on Safety	307	2.69	0.87	1	4
Score on Respect	307	2.72	0.72	1	4
Score on Purposeful Activity	307	2.49	0.87	1	4
Score on Resettlement	307	2.70	0.74	1	4
Number of recommendations issued	307	51.56	19.24	10	138
Share of recommendations accepted	307	0.87	0.11	0.5	1
Share of recommendations achieved	307	0.39	0.16	0	0.81
Panel C: *Inspection surveys*					
Number of recipients	307	204.73	24.80	92.92	251.95
Share of recipients	307	0.44	0.28	0.11	1
Number of respondents	307	171.67	22.21	83	228
Share of respondents	307	0.84	0.06	0.58	0.99

Notes: The table reports summary statistics at the inspection level. Statistics *Panel A*[23]is based on the 818 HMIP inspections from 2004 to 2019. Statistics in *Panel B* and *Panel C* are based on 307 inspections from 2013 to 2019.

However, the time between inspections varies significantly, ranging from as short as one month - typically for follow-up inspections - to nearly five years. Section 4.3 will demonstrate that the time between inspections is a predictor of future inspections. Of all inspections conducted, 76% are full inspections, and 66% are unannounced (33% are full unannounced inspections). The inspection is conducted by a team leader and a team of, on average, 19 inspectors, including specialists and partner inspectorates.

The inspection reports, produced by HMIP researchers and inspectors, present the areas of concern and detail the relative recommendations to the establishment, in response to which the action plans are produced, as anticipated in Section 4.1. On average, prisons receive 52 recommendations, they accept almost 90% of them[24], and achieve 40% of the accepted ones. Table A4 provides examples of the recommendations provided by the inspectors and

[23]The size of inspection team is actually deduced from the sample of prison inspections from 2013 to 2019, but included in *Panel A* for relevance. More statistics about inspection report recommendations are provided in Table A5 in the Appendix.

[24]Prison governors and directors can reject recommendations if they deem themselves not the appropriate recipients or if implementation is constrained by resource limitation.

the response of the governor, and Table A5 summarizes statistics about the number of recommendations issued, agreed upon, and achieved, for each Healthy establishment test.

Section A.2.1 provides details about the detainees' survey carried out at the outset of each inspection.

4.3 Inspections Timing

The HMIP operates an unannounced inspection program. In exceptional circumstances, the Chief Inspector may require an inspection to be announced. The timing of inspections is "deliberately unpredictable", with a minimum frequency of one inspection every five years. The timing is based on a dynamic risk assessment that takes into consideration the time elapsed since the last inspection, the functional type, the size and age of the establishment, significant changes to the establishment or changes in leadership, prison performance, and detainee outcomes, consultations with partners, Independent Review of Progress (IRP), and received intelligence.

Through an empirical exercise, we assess the predictability of inspection occurrences based on the prison function, previous levels of violence, and the time elapsed since the last inspection. Specifically, we estimate a logit regression of the inspection occurrence on six lags of *Violence on staff*, *Violence on prisoners*, and *Self-violence*, on the quarter of year since the previous inspection, interacted with prison functions, and the year-month fixed effects:

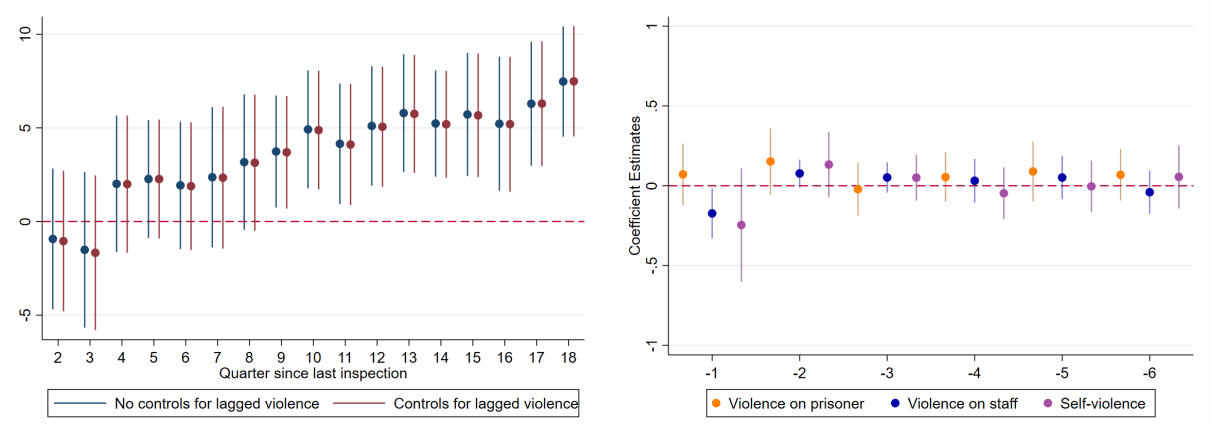
$$\begin{aligned}\text{Inspection}_{it} = &\sum_{l=1}^{6} \beta_l \text{Violence on staff}_{t-l} + \sum_{l=1}^{6} \gamma_l \text{Violence on prisoners}_{t-l} + \sum_{l=1}^{6} \delta_l \text{Self-violence}_{t-l} \\ &+ \sum_{f} \text{function}_i \times \sum_{q=1}^{q=18} \text{quarters since last inspection}_{it} + \text{year-month}_t \end{aligned} \tag{3}$$

The left panel of Figure 6 plots the coefficients on the quarters since the last inspection, estimated from 3, both with and without controls for lagged violence. As expected, inspections become more likely as more time passes since the previous inspection. Importantly, controlling for lagged violence has a negligible impact on the predictive power of the time that has elapsed since the last inspection.

The right panel of Figure 6 shows the estimated coefficients on the lagged violence from 3. The coefficients are individually not statistically different from zero.

Taken together, these findings suggest that, although the timing of inspections could, in theory, be endogenous to past levels of violence, in practice, the time elapsed since the last inspection strongly predicts inspection occurrence, nullifying the predictive power of

Figure 6: **Predicting inspection occurrence**



Notes: The *left* panel plots the estimated coefficients from the logit regression of the inspection dummy on the quarters since the last inspection FE interacted with the prison function FE, for Category C prisons, with and without controls for lagged levels of violence. The *right* panel plots the estimated coefficients from the logit regression of the inspection dummy on the six lags of prisoner assaults, staff assaults, and self-attacks, controlling for the interaction of quarters since the last inspection and prison function FE. The variables of lagged violence are standardized. Year-month FEs are included. SE clustered at the prison level.

violence. Thus, the elapsed time appears to be the dominant driver of inspection timing, mitigating concerns about endogeneity from violence levels.

4.4 Empirical Strategy

As anticipated in the previous section, prisons are inspected multiple times, with inspection timing varying across establishments. To estimate the effect of inspections on within-prison violence under this staggered and repeated adoption setting, we adopt a stacked difference-in-difference (DiD) design.[25] The stacked DiD estimates the change in violence around the inspection time, by aggregating the average treatment effects on the treated (ATT) for inspection cohorts defined by month of inspection occurrence, compared to units that have not been treated in the corresponding window. This stacked design avoids coefficients to be identified by the within-variation coming from late and early treatment adopters, threats to validity in the TWFE approach (Goodman-Bacon, 2021; Chaisemartin and D'Haultfuille, 2020). The following paragraphs describe the construction of our stacked panel, and the econometric specification.

We construct the stacked panel by identifying each month a in which an inspection

[25]Several methods have emerged to support causal inference in staggered adoption designs (Chaisemartin and D'Haultfuille, 2020; Callaway and SantAnna, 2021; Wooldridge, 2021; Gardner, 2022). Cengiz et al. (2019) employs a stacked diff-in-diff which Wing et al. (2024) shows provides a consistent estimate under correctly specified weights.

occurred, and building a corresponding sub-panel (or sub-experiment) centered around that inspection. Specifically, each sub-experiment a includes observations from periods t such that $a - \kappa_{pre} - 1 \leq t \leq a + \kappa_{post}$ for both treated prisons (inspected at time a) and clean controls.[26] Following Wing et al. (2024), clean controls for sub-experiment a are prison-periods not exposed to any inspection within the κ event window surrounding a. For our main specifications, we set $\kappa = 30$ months[27]. Our setting differs from the staggered adoption design in Wing et al. (2024) because units are treated repeatedly. To mitigate related bias, we restrict the *clean* controls to prisons-periods that have not been exposed to inspections within the extended window $[a - \kappa_{pre-extended} - 1, a + \kappa_{post}]$, with $\kappa_{pre-extended} = 30$ and $\kappa_{post} = 14$. This ensures that a control's pre-inspection period is not contaminated by proximity prior inspections.

Vertically concatenating all the sub-experiments yields a dataset at the prison x sub-experiment x year-month level. Table 4 presents summary statistics for the inspections and prisons in our stacked panel. Of the 818 inspections conducted by HMIP between 2004 and 2019, our stacked panel includes 560 inspections for which we could construct sub-experiments as defined above. The majority (71%) are full inspections, while 33% are unannounced (31% are full and unannounced), with an average of 27 months between consecutive inspections. During the 15 months preceding each inspection, prisons experienced on average 53 incidents of violence per month: 4 on staff, 17 between prisoners, and 32 of self-violence. High-severity incidents, such as serious assaults, homicides, suicides, and selharm requiring hospitalization, comprise 90-95% of total incidents across all violence categories. Regarding drug-related outcomes, in the pre-inspection period, 5 drugs every thousand prisoners were found monthly, 63 random mandatory drug test every thousand prisoners were administered, with 7% testing positive.

Within the stacked panel we estimate the following difference-in-difference model:

$$Y_{sat} = \beta_0 + \beta_1 D_{sa} + \beta_2 \text{Post}_t + \beta_3 D_{sa} \cdot \text{Post}_t + \eta_{at} + \epsilon_{sat} \tag{4}$$

where Y_{sat} denotes our violence outcomes - *Violence on staff*, *Violence on prisoners*, and *Self-violence* - for prison s at time t within sub-experiment a. The variable D_{sa} indicates treatment status ($D_{sa} = 1$ if prison s was inspected in period a), and Post_t indicates post-inspection periods. We control for sub-experiment year-month fixed effects η_{sat}. Moreover,

[26] We exclude inspections for which the full κ window cannot be observed.

[27] We set $\kappa = 30$ months to balance capturing longer-term dynamics against retaining a sufficient number of inspections in the sample. Increasing κ would extend the pre- and post-inspection observation window but would substantially reduce the number of eligible prisons for each sub-experiment, since only prisons without inspections over a longer span could be included. In Figures C2 C3, and C4 we replicate our main estimation exercise for different time windows.

Table 4: **Stacked panel descriptive statistics**

	N	Mean	St. Dev.	Min	Max
Panel A: *Inspections information*					
Full inspection	560	0.71	0.45	0.00	1.00
Unannounced inspection	560	0.33	0.47	0.00	1.00
Months since last inspection	560	27.48	10.38	1.00	57.00
Panel B: *Prison information*					
Private	140	0.11	0.32	0.00	1.00
Male	140	0.91	0.28	0.00	1.00
Victorian	140	0.24	0.43	0.00	1.00
Male governor/director	77	0.61	0.49	0.00	1.00
Panel C: *Pre-inspection violence*					
Total violence	10785	52.85	84.82	0.00	1522.65
High-severity total violence	10785	3.18	4.16	0.00	54.26
Low-severity total violence	10785	49.66	83.17	0.00	1508.71
Violence on staff	10785	4.08	7.60	0.00	120.25
High-severity violence on staff	10785	0.39	1.17	0.00	20.00
Low-severity violence on staff	10785	3.69	7.06	0.00	113.92
Violence on prisoner	10785	16.70	31.46	0.00	292.04
High-severity violence on prisoner	10785	1.29	2.40	0.00	38.76
Low-severity violence on prisoner	10785	15.40	30.30	0.00	286.82
Self-violence	10785	32.07	73.64	0.00	1505.23
High-severity self-violence	10785	1.50	2.67	0.00	35.46
Low-severity self-violence	10785	30.57	72.80	0.00	1491.29
Panel D: *Additional pre-inspection outcomes*					
Drug finds	10048	5.12	8.03	0.00	80.88
Number of drug tests	10071	62.93	24.52	0.00	179.58
Number of positive drug tests	10071	4.24	4.90	0.00	67.09

Notes: The table reports summary statistics of the stacked panel with window $\kappa = 30$. *Panel A* includes statistics at the inspection level: whether the inspection was full, unannounced, and the months elapsed since the last inspection. Statistics in *Panel B* provide information at the prison level about whether the prison is private, male, Victorian, and whether the governor/director is male. *Panel C* and *Panel D* collect the statistics of the violence and drug outcomes in the *pre-inspection* period, i.e., 30 months before the inspection, at the prison-month level. The violence and drug figures are measured in the number of incidents per thousand prisoners. *High-severity violence on staff* refers to serious staff assaults, *High-severity violence on prisoner* to serious prisoner assaults and homicides, and *High-severity self-violence* to self-attacks requiring hospitalization and suicides. *Low-severity violence on staff* refers to non-serious staff assaults, *Low-severity violence on prisoner* to non-serious prisoner assaults, and *Low-severity self-violence* to self-attacks not requiring hospitalization. *High-severity total violence* and *Low-severity total violence* sum the respective measures within violence severity across violence types. *Total violence*, *Violence on staff*, *Violence on prisoners*, and *Self-violence* sum the respective measures within violence type across violence severity. *Drug tests* refer to random mandatory drug testing. *Drug finds*, *Number of drug tests*, and *Number of positive drug tests* are only available for the sample period April 2007 to March 2018. *Male governor/director is only available for a sub-sample, mostly after 2010.*

following our discussion about the timing of the inspection in 4.3, we control for the time elapsed from the previous inspection.

We define the sample weights Q_{sa}, which Wing et al. (2024) shows are required for β_3 to identify the parameter of interest:

$$Q_{sa} = \begin{cases} 1 & if D_{sa} = 1 \\ \frac{N_a^D}{N_{\Omega_k}^D} & if D_{sa} = 0 \\ \frac{N_a^C}{N_{\Omega_k}^C} & \end{cases} \tag{5}$$

where $\frac{N_a^D}{N_{\Omega_k}^D}$ refers to the share of treated observations that belong to sub-experiment a, N_a^D, out of all the treated observations, $N_{\Omega_k}^D$; and $\frac{N_a^C}{N_{\Omega_k}^C}$ refers to the share of treated observations that belong to sub-experiment a, N_a^C, out of all the treated observations, $N_{\Omega_k}^C$.

As we will later show, our results are robust to different windows κ on which the sub-experiments are built, to the exclusion of sample weights Q_{sa}, and to linear and non-linear controls for the time elapsed since the previous inspection.

4.5 Evidence

We now turn to the main estimation results from our stacked difference-in-differences approach, examining the effects of inspections on prison violence. Figure 7 plots the estimated coefficients on the treatment for five lead and lag quarters for *Total violence*, *Violence on staff*, *Violence on prisoners*, and *Self-violence*. The plots display a negligible reduction in violent incidents towards the staff, but a considerable decrease in violence among prisoners and self-violence. These reductions are not immediate following the inspection but rather intensify gradually over time, with 12 fewer cases of violence among prisoners and 17 fewer cases of self-violence (37% and 23% of one standard deviation, respectively) fifteen months after the inspection. Given inspections identify key areas of concern and prompt prison governors to implement responsive measures, these inspections could initiate a sustained improvement process in prison governance, conditions, and overall safety, consistent with long-term effects. The delayed effects onset, with null effects on the three months following the inspection, the absence of pre-trends and of a mean-reversion suggest that the reduction we observe is not driven by strategic prison under-reporting prior to the inspectors visit nor by temporary incapacitation due to inspectors physical presence preventing violence.

Table 5 presents the static estimates from our stacked diff-in-diff specification. The results demonstrate that the average prison violence in the 15 months after the inspections is lowered by 15 incidents every thousand prisoners: -1 incident of violence directed at staff,

Table 5: **Inspections on violence**

	Total violence	Violence on staff	Violence on prisoners	Self-violence
Inspected	0.452***	0.445***	0.372***	0.176***
	(0.090)	(0.089)	(0.090)	(0.066)
Post x Inspected	-0.184**	-0.095	-0.203**	-0.107
	(0.074)	(0.079)	(0.083)	(0.079)
Observations	21390	21390	21390	21390
R-squared	0.282	0.252	0.246	0.290
Sub-experiment x Period FE	Yes	Yes	Yes	Yes
Time since last inspection	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1
SE clustered at the prison level.

Notes: The table reports the estimate of the stacked difference-in-difference specification on violence outcomes, *Total violence* in column (1), *Violence on staff* in column (2), *Violence on prisoners* in column (3), and *Self-violence* in column (4). The violence outcomes are measured in number of incidents per thousands of prisoners and standardized. All the specifications include the sub-experiment x year-month fixed effects and the control for the months elapsed since the previous inspection. SE clustered at the prison level. Weights of the stacked panel are computed following Wing et al. (2024).

-6 of prisoner-on-prisoner violence, and -8 in self-inflicted violence per thousand prisoners (10%, 20%, and 11% of one standard deviation, respectively). We find no differences in the inspection treatment effect across inspection characteristics. Table C1 shows that *short* (2-3 days) versus *long* (2-3 weeks) inspections generate similar effects on violence. This similarity likely reflects that both inspection types involve the same core activities – staff interviews, prisoner surveys, facility walkthroughs, document reviews –, with short inspections typically as follow-ups to assess progress on specific issues identified in previous full inspections. If inspections operated primarily through temporary incapacitation – with inspector presence directly preventing violence during the inspection period – we would expect longer inspections to show larger effects. The absence of differential effects reinforces that the violence reduction we observe operates through channels beyond temporary incapacitation. Similarly, Table C2 demonstrates no differential effects between announced versus unannounced inspections. This finding aligns with anecdotal evidence suggesting that the inspection notification period is too brief for meaningful operational changes or adjustments in violence patterns and inmates' interactions.

The violence reduction effects we document prove robust to a comprehensive set of specification checks. These include varying approaches to controlling for time elapsed since previous inspections, alternative fixed effects structures, and different values of the window parameter used in constructing the stacked panel dataset. Figures C1, C2, C3, and C4 in the Appendix plot the estimated diff-in-diff coefficients across more than one hundred specifications. Moreover, given the count data nature of violence incidents, we estimate the inspection effects

Table 6: **Inspections on violence**

	Total violence			Violence on staff			Violence on prisoners			Self-violence		
Severity	All	High	Low	All	High	Low	All	High	Low	All	High	Low
Inspected	0.452***	0.462***	0.436***	0.445***	0.280***	0.432***	0.372***	0.407***	0.354***	0.176***	0.220*	0.170***
	(0.090)	(0.111)	(0.088)	(0.089)	(0.063)	(0.087)	(0.090)	(0.090)	(0.089)	(0.066)	(0.119)	(0.064)
Post x Inspected	-0.184**	-0.070	-0.183**	-0.095	-0.121	-0.082	-0.203**	0.019	-0.212**	-0.107	-0.035	-0.107
	(0.074)	(0.093)	(0.072)	(0.079)	(0.081)	(0.074)	(0.083)	(0.070)	(0.083)	(0.079)	(0.093)	(0.079)
Observations	21390	21390	21390	21390	21390	21390	21390	21390	21390	21390	21390	21390
R-squared	0.282	0.253	0.280	0.252	0.209	0.251	0.246	0.244	0.243	0.290	0.212	0.290
Sub-experiment x Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time since last inspection	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1
SE clustered at the prison level.

Notes: The table reports the estimate of the stacked difference-in-difference specification on violence outcomes, disaggregated by the severity of the incidents. We denote as *Low-severity violence on staff* the non-serious assaults on staff (column 6), *High-severity violence on staff* the serious assaults on staff (column 5), *Low-severity violence on prisoners* the non-serious assaults on prisoners (column 9), *High-severity violence on prisoners* the serious assaults on prisoners and the homicides (column 8), *Low-severity self-violence* the self-attacks not requiring hospitalization (column 12), *High-severity self-violence* the self-attacks requiring hospitalization and suicides (column 11). The violence outcomes are measured in number of incidents per thousands of prisoners and standardized. All the specifications include the sub-experiment x year-month fixed effects and the control for the months elapsed since the previous inspection. SE clustered at the prison level. Weights of the stacked panel are computed following Wing et al. (2024).

in a poisson regression. Count data, which is discrete and often right-skewed, particularly with many zero observations, might violate the linear regression assumptions of normally distributed error term and of continuous response variable. Table C9 confirms that our linear regression results are robust to count data regression.

In Table 6 we disaggregate violence incidents into *High-severity* and *Low-severity* subcategories. The disaggregation reveals which types of violence and so which forms and levels of distress inspections address. Regarding manifestations of fear and despair – culminating in violence on prisoners and self-violence – inspections appear to primarily reduce less severe expressions. Conversely, for anger-related incidents, the strongest effects emerge in more serious forms of violence. The heterogeneity across distress sources and incident severity suggests that inspections, as interventions designed to diagnose institutional weaknesses and recommend improvements for a well-functioning environment, do affect anger-driven violence, and mild forms of fear and despair. However, they do not appear to be effective against severe forms of fear and despair, which might be deeply rooted in individual circumstances that might require more targeted interventions beyond institutional-level ones.

We now turn to heterogeneity tests across prison characteristics to identify which establishment features amplify responsiveness to inspections, thereby informing our subsequent analysis of the mechanisms underlying violence reduction following inspections.

As anticipated in Section 3.2, the literature suggests that incarcerated women face distinct challenges compared to their male counterparts, including heightened gender discrimination compounded by criminal stigma, psychological distress related to pregnancy and separation

from children, higher rates of family abandonment, and discomfort in facilities primarily designed for men. The effectiveness of inspections in improving conditions may vary by prison gender based on inmates' responses to implemented recommendations but also on their psychological reaction to having their needs acknowledged: Inspectors engage directly with inmates through conversations and surveys. When analyzing heterogeneity by prison gender, we find that female prisons exhibit lower levels of violence against staff and other prisoners, while similar rates of self-directed violence, as demonstrated in Table C3. Furthermore, as shown also in Figure 8, inspections appear to significantly reduce self-harm incidents in female facilities (40% of one standard deviation, equivalent to 72 fewer incidents per thousand prisoners[28]).

The structural characteristics and institutional culture of prisons significantly influence how responsive violence rates are to external interventions. In Figure 9 and Table C4 we first explore heterogeneity in violence reduction across establishment category levels. Medium- and high- security prisons exhibit significant higher levels of violence, compared to low-security prisons[29]. According to our estimates, low-security prisons violence levels are unaffected by the inspection occurrence. Conversely, the reduction in violence we earlier observed was coming from medium- and high-security prisons: high-security prisons display a statistically significant reduction of 2 incidents of violence on other prisoners (p-value < 0.01), and medium-security prisons a statistically significant reduction of 6 incidents of violence on other prisoners (p-value < 0.01) and of 16 incidents of self-directed violence (p-value < 0.05). The different responsiveness might be associated with the baseline levels of violence, whereby more room for improvement is found in more violent environments. These differences may alternatively, or additionally, reflect the heterogeneous nature of prisoner populations across security levels, which vary in offences and sentences, and possibly the prisoners and staff behavior and their response to management interventions.

Additionally, prison management may significantly influence the inspection efficacy. The managers' alignment with the inspectors' recommendations, their receptiveness to reform, their oversight capabilities, as well as their professional experience, likely affect the implementation of new practices following the inspection recommendations, thereby generating differential violence reduction effects. To investigate whether some relation exists between management interventions and managers, we compare the inspection estimates with and without governor fixed effects. Figure 10 plots the estimated coefficients. The inclusion of controls for the governor fixed effects reduces the inspection coefficient by approximately half

[28]The pre-inspection self-violence incidents in female prisons are on average 132 every thousand prisoners, with a standard deviation of 180.

[29]We classify as high-security prisons Category A and Category B prisons, as medium-security prisons Category C and reception, and as low-security prisons Category D and YJB establishments.

across all violence types, suggesting that the specific management and leadership practices do influence how violence respond to inspections. The availability of new data on governors[30] will enable to determine whether governor experience is responsible the role of governors in inspection effectiveness.

On top of management and institutional features, we investigate whether other detainees' behaviours, in themselves spy of an healthy environment, are affected by inspections. Drug misuse has increase by 50% between 2012/13 and 2017/18 and it is one of the biggest challenges facing the criminal justice system today (HM Prison and Probation Service, 2019)[31]. As reported in Table 4, in the pre-inspection period, the prisons in our sample were finding drugs 5 times, testing 63 inmates, 4 of them testing positive, every thousand prisoners every month. Inspectors often suggest, among the report recommendations, to identify strategies to reduce drug demand and supply, to store medicines safely, to increase the drug testing program resources, and to improve treatments to drug-dependent detainees. Figures 11 plot the dynamic diff-in-diff estimates for drug finds, Random Mandatory drug tests, and the share of positive tests out of the Random Mandatory drug tests, around the inspection time. The graphs demonstrate that drug finds and drug testing increase gradually after the inspection, with an average post-inspection increase of, respectively, 1.5 and 3.5 cases every thousand prisoners each month (19% and 15% of one standard deviation, see Table C6). However, the share of positive drug testing, while stable around zero after the inspection, shows a slight but stable increase in the fifteen months prior to the inspection, which likely was one of the factors that predicted the inspection occurrence. This evidence suggests that inspections increased drug detection but not deterrence: it may be that the combination of unobserved prevention and treatment measures and the modest increase in detection we observe were insufficient to disrupt the established drug contraband networks and the inmates' drug dependencies.

The misuse of drugs also contributes to a cycle of disruption and violence (HM Prison and Probation Service, 2019). Drugs misuse and its traffic are identified as causes of anger and intimidation towards prisoners and staff, and self-harm. To address the possibility that the violence reductions we observe are driven by the increased post-inspection drug detection we

[30] Governors' tenure is provided by the Ministry of Justice through a Freedom of Information (FOI) request, for almost the entire sample of prisons (130 out of 140 prisons), for the sample period 2010-2019. Therefore the current data prevents us from inferring the out-of-sample governor tenure. Extending coverage to 1990 would instead allow to measure the governors tenure in the prison system.

[31] HM Prisons use a variety of drug testing and drug finding approaches. Prisoners can be required to undertake monthly random Mandatory Drug Testing (rMDT) or Suspicion-Based Drug Testing, or can volunteer for Compact-Based Drug Testing. Moreover, staff dedicate time to search and prevent contraband, with the help of specialist teams and intelligence. Prisoners who are found positive are subject to sanctions, and are supported to access treatment.

find, we augment specification 4 with controls for drug finds and drug testing. As shown in Tables C7 and C8, estimates remain virtually unchanged when accounting for this potential mechanisms.

Overall, our empirical analysis of prison inspections indicates that they significantly reduce violence against inmates and self-directed violence. Notably, inspections appear particularly effective in reducing self-harm among female inmates and violence on prisoners in medium- and high-security prisons. Moreover, governors' identity seem to play a key role in the responsiveness of prisons to inspections. These findings suggest that individual prisoner challenges, along with institutional design and organizational culture, management, and leadership, potentially influence how prisons respond to management interventions.

4.6 Mechanisms

While the previous Section identified establishment characteristics that may enhance responsiveness to interventions, this section investigates the mechanisms behind inspections effectiveness. We argue that the documented reductions in violence are primarily attributable to recommendations implementation rather than by short-term strategic suppression of violence. A temporary suppression of violence during inspections cannot account for the increasing effects observed in Figure 7, where virtually no effect is observable in the first three months. The timing of the effects instead supports the hypothesis of sustained improvements triggered by inspection. We propose, and provide supportive evidence, that the "consulting" dimension of inspections – wherein inspectors formulate recommendations tailored to each establishment – has been conducive to prison conditions improvement and, consequently, to violence reduction. As anticipated in Section 4.1, HMIP inspectors on the final day of the inspection present to the establishment an assessment and a list of recommendations, based on qualitative and quantitative sources (e.g., data from P-Nomis, HMPPS performance, conversations with detainees and staff). The assessment and the recommendations are categorized into four categories from the four Healthy establishment test as set out by the HMIP Inspection Framework, *Safety, Respect, Purposeful Activity*, and *Rehabilitation and release planning*. Progress on these recommendations is reviewed at the following inspection.

On average, prisons receive 52 recommendations, they accept almost 90% of them[32], and achieve 40% of the accepted ones. Table A4 provides examples of the recommendations provided by the inspectors and the response of the governor, and Table A5 summary statistics about the number of recommendations issued, agreed upon, and achieved.

[32]Prison governors and directors can reject recommendations if they deem themselves not the appropriate recipients or if implementation is constrained by resource limitation.

Recommendations constitute detailed and tailored guidelines on issues that require intervention on the part of the prison. A high achievement rate indicates that most of the concerns raised have been resolved. Heterogeneity in treatment effect by achievement rate would point to recommendations accounting for violence reductions. However, the sign of the relationship is ambiguous. Prisons with relatively favorable baseline conditions, requiring only minor adjustments, may achieve high achievement rates, with negligible effects on violence. Conversely, establishments undertaking substantial changes to address systemic issues may gain a smaller share of recommendations but with significant improvements in violence reduction. An alternative hypothesis is that a positive relationship exists between achievement rate and violence reduction, independently of the baseline prison conditions or complexity of the recommendation implementation.

To investigate this relationship, we estimate the specification in 4[33], allowing the treatment effect to vary across sub-experiments. We then examine whether the sub-experiment-specific effects correlate with the share of recommendations achieved by the prisons treated in that sub-experiment. We focus on recommendations achieved – out of those agreed to – across all categories, and directed to the prison governor or director.

Figure 12 presents the average treatment effect for the three violence outcomes, comparing prisons with *Above-* and *Below-* median achievement rates. Prisons with below-median achievement rates exhibit a slightly larger reduction in violence following the inspection, although differences are not statistically significant. This pattern suggest that achievement rates alone are not conclusive about whether recommendations drive violence reduction. If anything, the previous figure points to greater violence reductions in low-achiever prisons, consistent with three possible explanations: (i) the inspector's consulting service is not the main driver of violence reduction, (ii) prisons disregard recommendations, or (iii) prisons prioritize a small set of high-impact recommendations.

To investigate these options, we construct a measure of *recommendation prevalence* reflecting how idiosyncratic each recommendation is. In particular, we assign a set of semantic similarity scores to each prison recommendation based on its textual proximity to other recommendations issued in the same year and within the same category, but in different prisons[34].

[33] Note that the specification in 4 is estimated on our full stacked sample, but the correlation analysis of the heterogeneous treatment effects on the reduced sample for which we have information on the recommendations issued, accepted and achieved, i.e., on 226 inspections taking place between 2013 and 2019.

[34] We employ Doc2Vec, which is an unsupervised machine learning algorithm that generates fixed-length vector representations of variable-length text documents. It embeds these into a continuous vector space, where proximity between vectors reflects semantic similarity between recommendations. We then use cosine similarity to compare a recommendation r against all other recommendations r' from the same year and category but from different prisons, computing a distribution of similarity scores for each recommendation.

Figure 13 presents the average inspection effects for the three violence outcomes, comparing prisons with *Low-* and *High achievement rate* (below and above median) and *Low-* and *High relative prevalence*, i.e., below and above median relative prevalence of recommendations achieved to prevalence of recommendations not achieved. Among the high-achievers, prevalence does not appear to matter. In contrast, the stronger negative effect observed for low-achievers in Figure 12 is driven by prisons that addressed the most idiosyncratic (low-prevalence) recommendations, with almost no effect in prisons focusing on common (high-prevalence) recommendations.

Taken together, these results suggest that the consulting dimension of inspections – through the identification and notification of areas of concern and the provision of possible solutions – contributes, at least in part, to the post-inspection reduction in violence on prisoners and self-violence. Both a high achievement rate and a low achievement rate, with a focus on less common issues, are associated with a greater reduction in violence.

5 Conclusion

Rising expenditures on the prison system, alongside deteriorating conditions and incarceration experiences, underscore the need to understand what drives variation in in-prison outcomes. Violence within prisons is not merely a security concern - it is a behavioral expression of detainees' distress and a diagnostic of institutional health -. Yet, standard observable inputs such as inmate composition, staffing levels, and physical infrastructure fail to fully account for the persistent differences in violence across prisons.

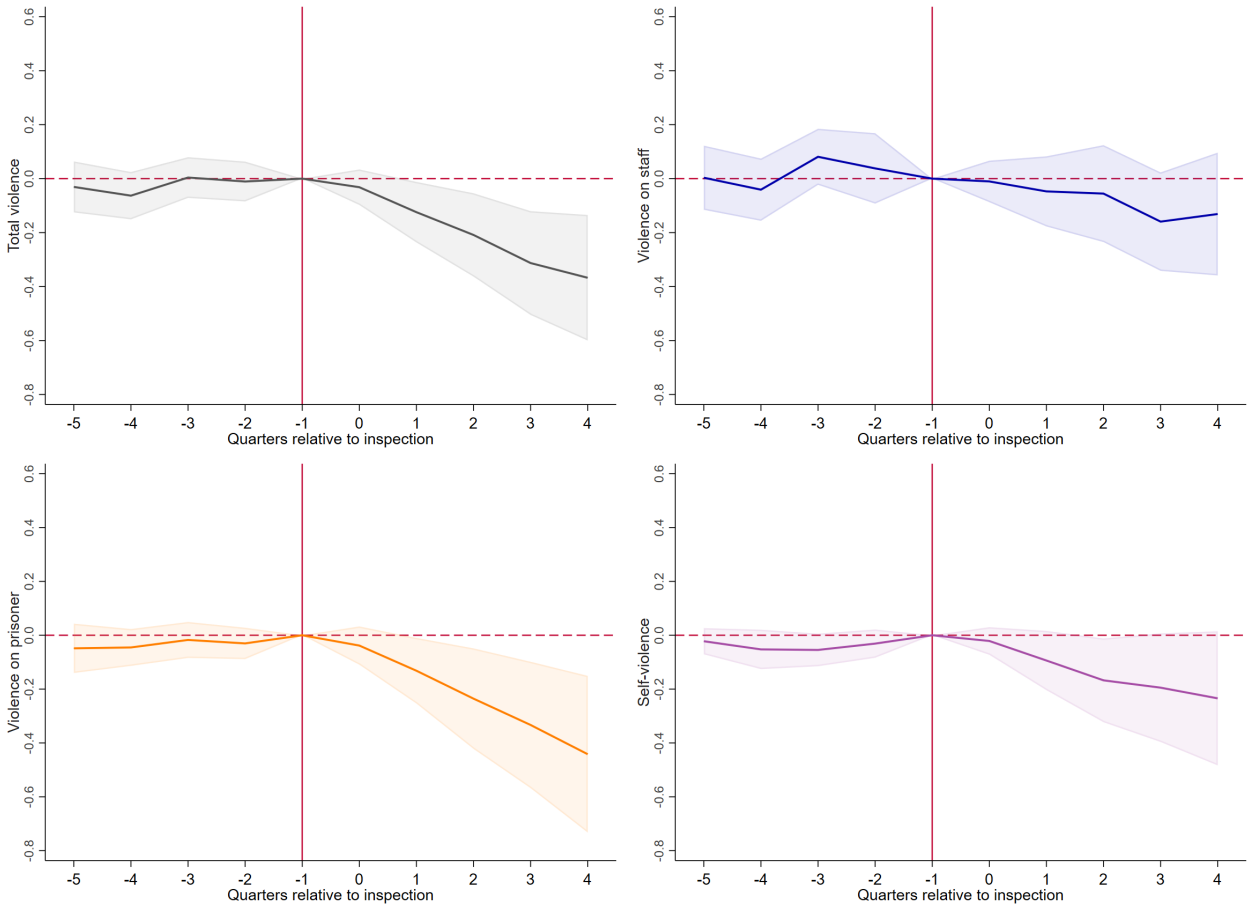
Structural explanations alone risk overlooking the internal dynamics that shape prison life. Inspections by His Majesty's Inspectorate of Prisons (HMIP) offer a rare window into these dynamics. We draw on comprehensive inspection reports, which include qualitative assessments of institutional performance, documentation of good practices, areas identified for improvement, and required follow-up action plans. Among public-sector interventions, prison inspections are perhaps the closest analog to private-sector consulting. The reductions in violence we observe may reflect a broader (re)organizational process initiated by inspection feedback and subsequent institutional change.

Future work will leverage the rich textual and procedural detail in these reports to identify the specific management practices associated with violence reduction. This will help address a key question in public management: are performance gaps driven by a lack of knowledge about effective practices or by constraints that prevent their implementation? Understanding

We then aggregate these scores across r' using the mean, median, percentiles (90th, 75th, 25th, 10th), and Gini coefficient to produce *prevalence scores*.

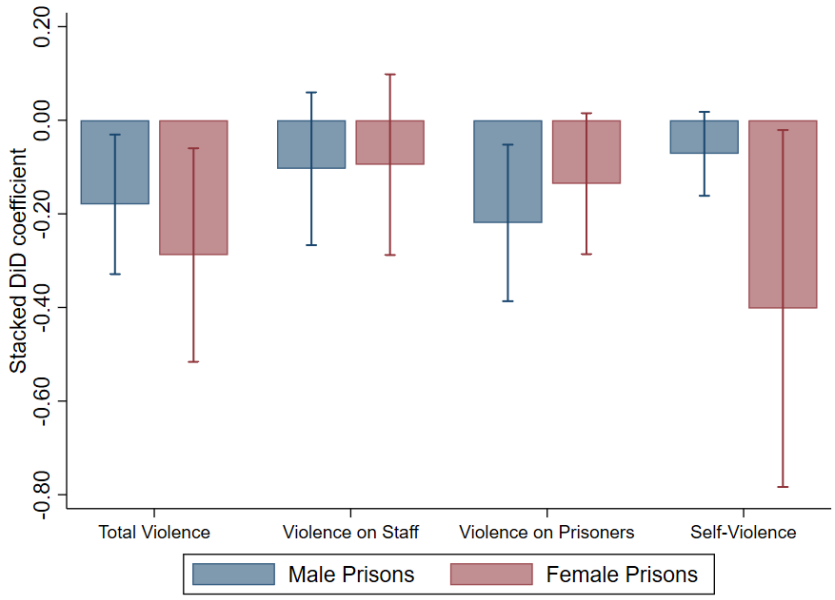
this distinction is crucial for designing reforms that enhance the functioning and humanity of prison institutions.

Figure 7: **Inspections on violence**



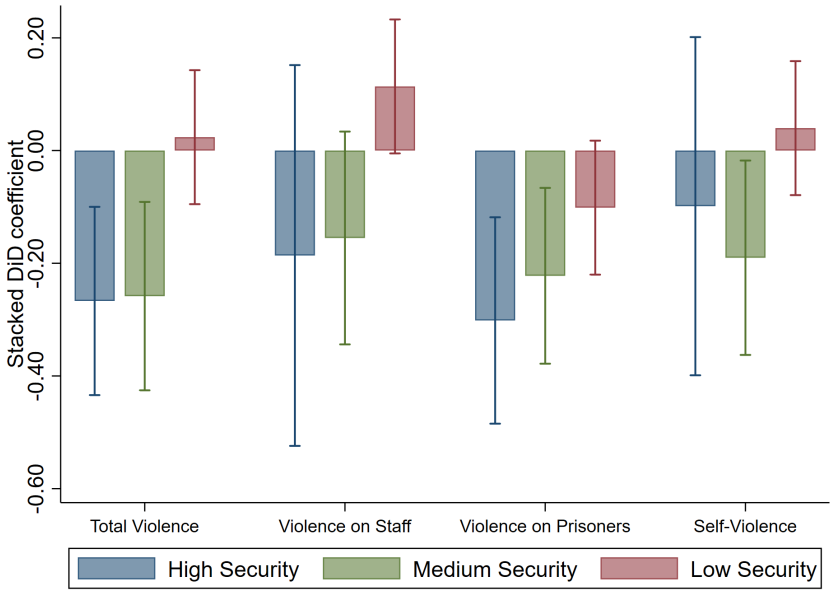
Notes: The figure plots the dynamic estimated of the stacked difference-in-difference specification on violence outcomes, *Total violence* in the top left panel, *Violence on staff* in the top right panel, *Violence on prisoners* in the bottom left panel, and *Self-violence* in the bottom right panel. The labels on the x-axis refer to quarters relative to the inspection month, i.e., $t = 0$ refers to the three months following the inspection. The violence outcomes are measured in number of incidents per thousands of prisoners and standardized. All the specifications include the sub-experiment x year-month fixed effects and the control for the months elapsed since the previous inspection. SE are clustered at the prison level. Weights of the stacked panel are computed following Wing et al. (2024).

Figure 8: **Inspections on violence by prison gender**



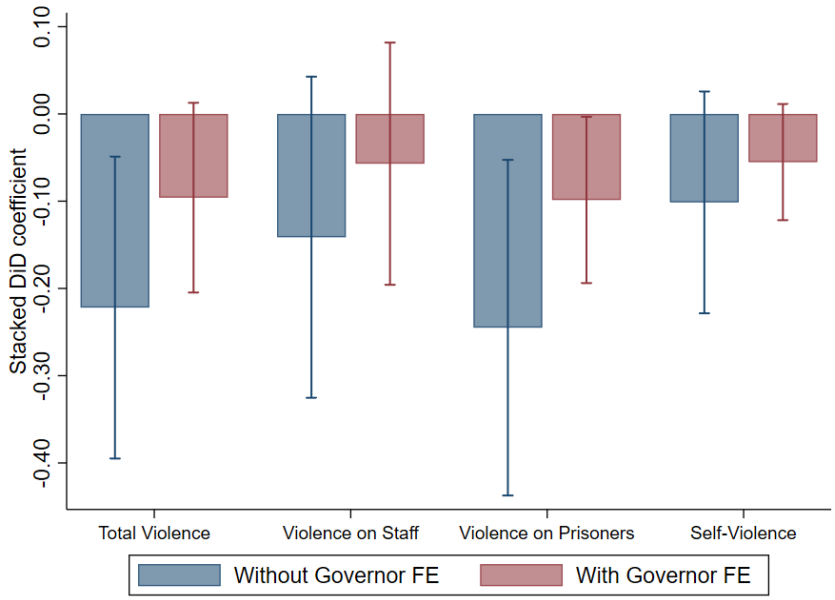
Notes: The Figure plots the estimates of the stacked triple difference-in-difference specification for heterogeneity in the prison gender. The violence outcomes are measured in number of incidents per thousands of prisoners and standardized. All the specifications include the sub-experiment x year-month fixed effects and the control for the months elapsed since the previous inspection. SE clustered at the prison level. Weights of the stacked panel are computed following Wing et al. (2024). Please refer to the relative Table C3.

Figure 9: **Inspections on violence by establishment security level**



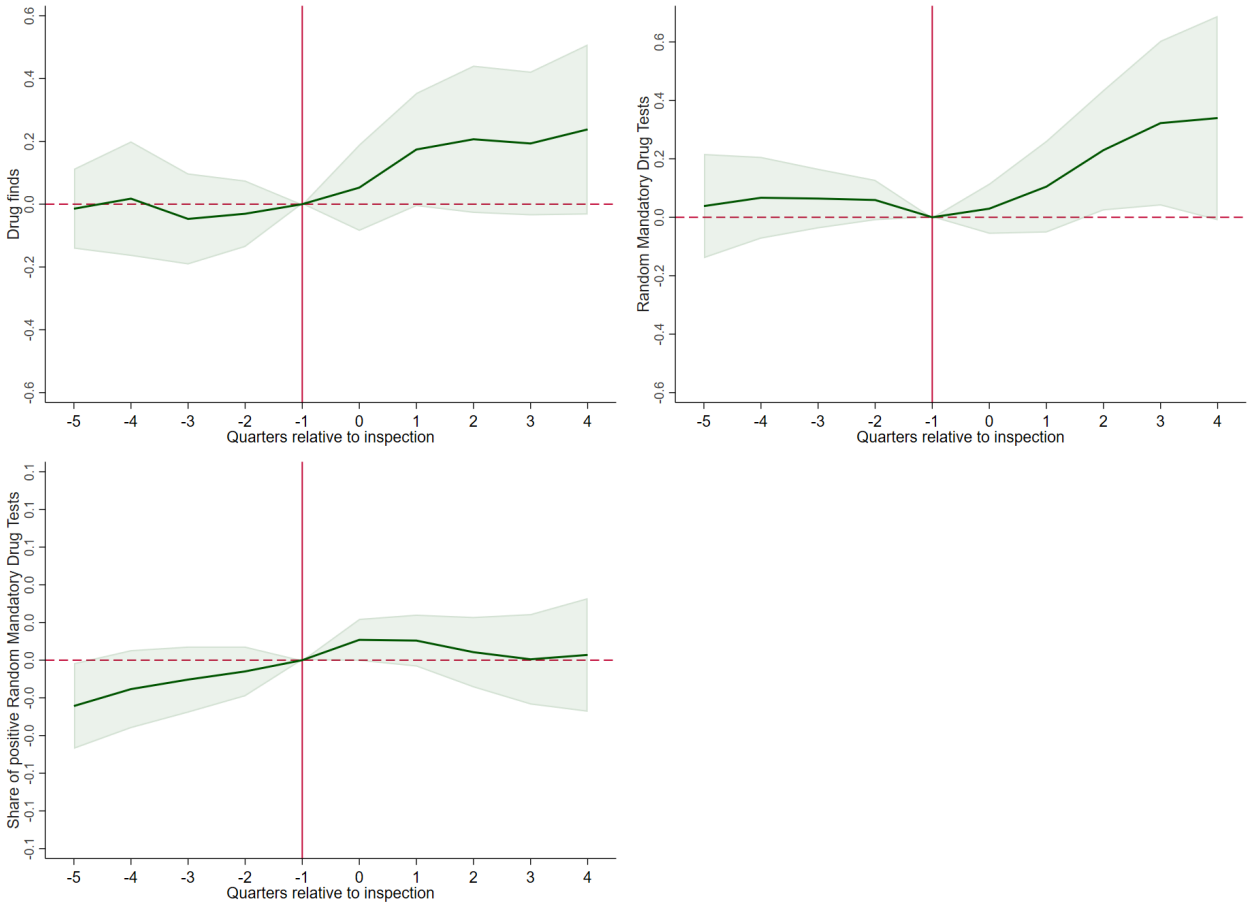
Notes: The Figure plots the estimates of the stacked triple difference-in-difference specification for heterogeneity by the establishment security level. We classify as *High security* prisons Category A and Category B prisons, as *Medium security* prisons Category C and reception, and as *Low security* prisons Category D and YJB establishments. The number of prisons in our stacked panel are 25, 97 and 24, respectively for *High security*, *Medium security*, and *Low security* prisons. The violence outcomes are measured in number of incidents per thousands of prisoners and standardized. All the specifications include the sub-experiment x year-month fixed effects and the control for the months elapsed since the previous inspection. SE clustered at the prison level. Weights of the stacked panel are computed following Wing et al. (2024). Please refer to Table C4.

Figure 10: **Inspections on violence & Prison leadership**



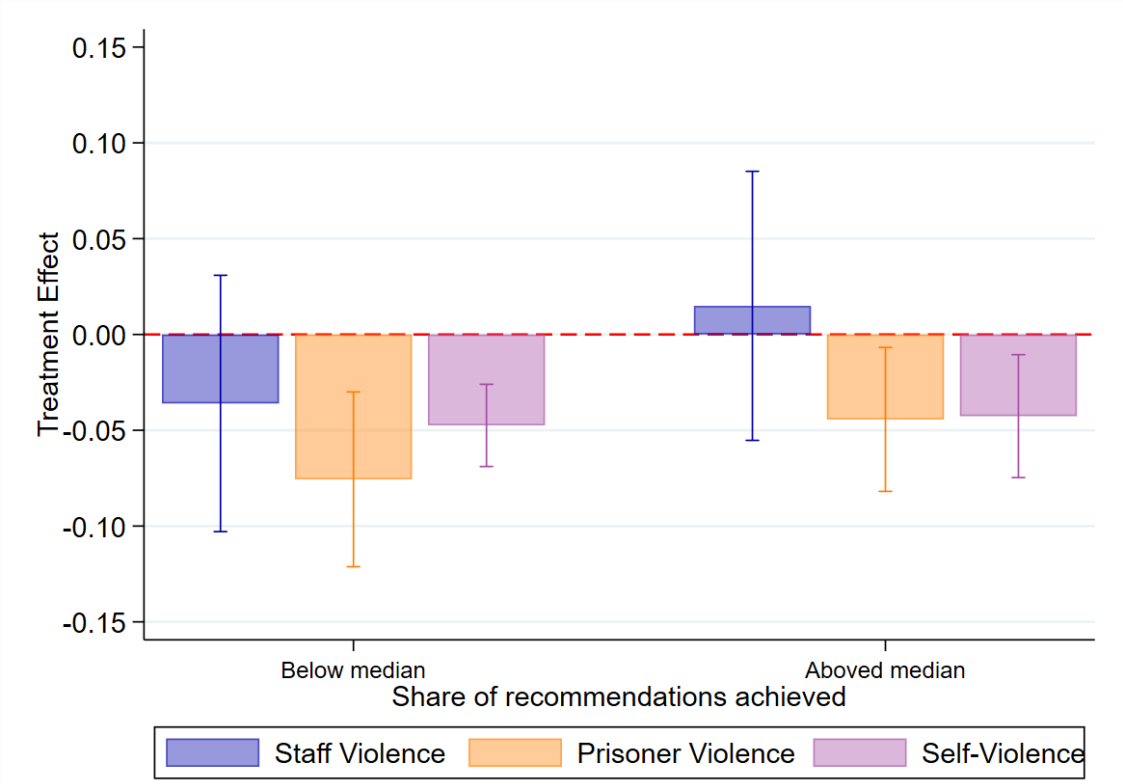
Notes: The figure plots the estimates of the stacked difference-in-difference specification on violence outcomes with and without controls for governor fixed effects. In the stacked sample we observe 236 governors in 130 prisons. Note that the baseline *Without Governor FE* is estimated on the subsample for which governors' names are available. The violence outcomes are measured in number of incidents per thousands of prisoners and standardized. All the specifications include the sub-experiment x year-month fixed effects and the control for the months elapsed since the previous inspection. SE clustered at the prison level. Weights of the stacked panel are computed following Wing et al. (2024). Please refer to Table C5.

Figure 11: **Inspections on drug use**



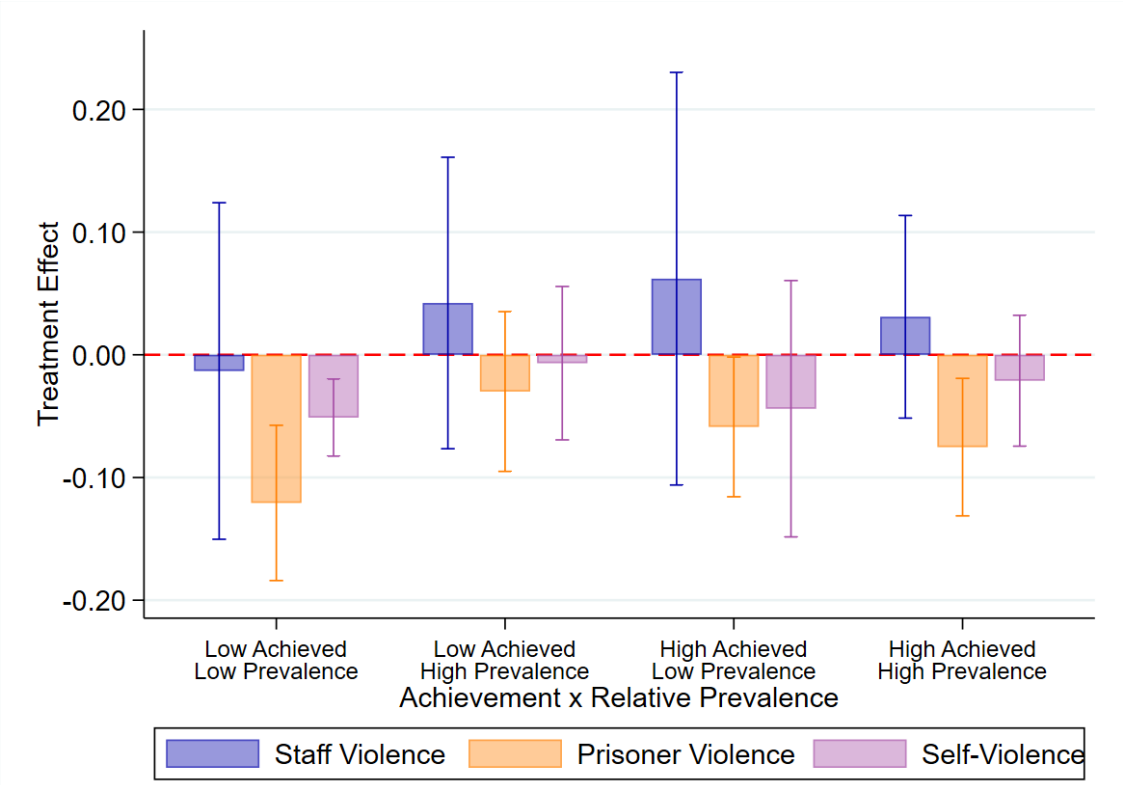
Notes: The figure plots the dynamic estimated of the stacked difference-in-difference specification on drug outcomes, *Drug finds* in the top left panel, *Random Mandatory drug testing* in the top right panel, and *Share of positive Random Mandatory drug testin* in the bottom left panel. The labels on the x-axis refer to quarters relative to the inspection month, i.e., $t = 0$ refers to the three months following the inspection. The drug finds and the drug tests are measured in number of incidents per thousands of prisoners and standardized. All the specifications include the sub-experiment x year-month fixed effects and the control for the months elapsed since the previous inspection. SE are clustered at the prison level. Weights of the stacked panel are computed following Wing et al. (2024). Refer to Table C6.

Figure 12: **Heterogeneity of inspection effect by share of recommendations achieved**



Notes: The figure plots the heterogeneous treatment effect estimated through the stacked difference-in-difference specification on violence outcomes, *Violence on staff* blue, *Violence on prisoners* in orange, *Self-violence* in purple, by levels of share of recommendations achieved from one inspection to the following. We augment the specification in 4, allowing the treatment effect to vary at the sub-experiment level. Prisons are assumed to be subject to the treatment effect of the sub-experiment in which they were inspected. As the data on recommendations are available only for 226 inspections, the figure statistics are generated on a sample of $N = 226$. The figure plots the average treatment effect by levels of share of recommendations achieved (below and above median). The violence outcomes are measured in number of incidents per thousands of prisoners and standardized. We consider the share of recommendations achieved out of the agreed ones and only those directed to the governor. P-values for the difference in estimates between *below* and *above* median are 0.13, 0.18, and 0.70, for violence on staff, violence on prisoners, and self-violence, respectively. All the specifications include the sub-experiment x year-month fixed effects and the control for the months elapsed since the previous inspection. SE are clustered at the prison sub-experiment level. Weights of the stacked panel are computed following Wing et al. (2024).

Figure 13: **Heterogeneity of inspection effect by share of recommendations achieved and recommendation prevalence**



Notes: The figure plots the heterogeneous treatment effect estimated through the stacked difference-in-difference specification on violence outcomes, *Violence on staff* blue, *Violence on prisoners* in orange, *Self-violence* in purple, by levels of share of recommendations achieved from one inspection to the following. We augment the specification in 4, allowing the treatment effect to vary at the sub-experiment level. Prisons are assumed to be subject to the treatment effect of the sub-experiment in which they were inspected. As the data on recommendations are available only for 226 inspections, the figure statistics are generated on a sample of $N = 226$. The figure plots the average treatment effect by levels of share of recommendations achieved (below and above median) and by level of the prevalence of achieved recommendations relative to the prevalence of not achieved ones (below and above median). The violence outcomes are measured in the number of incidents per thousand prisoners and standardized. We consider the share of recommendations achieved out of the agreed ones and only those directed to the governor. Prevalence at the recommendation level is measured as the median across the prevalence of the recommendation with all other recommendations in the same category and year, but in a different prison. P-values for the difference in estimates between *Low Prevalence* and *High Prevalence* are 0.55, 0.06, and 0.22 within *Low achieved*, and 0.75, 0.69, 0.38 within *High achieved*, for violence on staff, violence on prisoners, and self-violence, respectively. All the specifications include the sub-experiment x year-month fixed effects and the control for the months elapsed since the previous inspection. SEs are clustered at the prison sub-experiment level. Weights of the stacked panel are computed following Wing et al. (2024).

References

Ábrahám, A. and T. Kirk White (2007). "The Dynamics of Plant-level Productivity in U.S. Manufacturing". In: *Working paper, University of Rochester*.

Alsan, Marcella, Antony M. Barnett, Peter Hull, and Crystal Yang (2024). *Something Works in US Jails: Misconduct and Recidivism Effects of the IGNITE Program*. Working Paper w32282. National Bureau of Economic Research.

Arbour, William, Guy Lacroix, and Sylvie Marchand (2024). "Prison Rehabilitation Programs and Recidivism: Evidence from Variations in Availability". In: *Journal of Human Resources*.

Auty, Katherine M and Alison Liebling (2020). "Exploring the relationship between prison social climate and reoffending". In: *Justice Quarterly* 37.2, pp. 358–381.

Bales, William D., Laureen E. Bedard, Scott T. Quinn, David T. Ensley, and Gina P. Holley (2005). "Recidivism of Public and Private State Prison Inmates in Florida". In: *Criminology Public Policy* 4.1, pp. 57–82.

Bandiera, Oriana, Andrea Prat, Stephan Hansen, and Raffaella Sadun (2020). "CEO Behavior and Firm Performance". In: *Journal of Political Economy* 128.4, pp. 1325–1369.

Barbarino, Alessandro and Giovanni Mastrobuoni (2014). "The Incapacitation Effect of Incarceration: Evidence from Several Italian Collective Pardons". In: *American Economic Journal: Economic Policy* 6.1, pp. 1–37.

Batistich, Mary Kate, William N. Evans, Tyler Giles, and Rebecca Margolit-Chan (Nov. 2024). *Therapy to Reduce Violence and Improve Institutional Safety During Incarceration*. NBER Working Papers 33147. National Bureau of Economic Research, Inc.

Bayer, Patrick and David E. Pozen (2005). "The Effectiveness of Juvenile Correctional Facilities: Public versus Private Management". In: *The Journal of Law and Economics* 48.2, pp. 549–589.

Bernard, Andrew B., Stephen J. Redding, and Peter K. Schott (Mar. 2010). "Multiple-Product Firms and Product Switching". In: *American Economic Review* 100.1, pp. 70–97.

Bertrand, M. and A. Schoar (2003). "Managing with Style: The Effect of Managers on Firm Policies". In: *Quarterly Journal of Economics* 118.4, pp. 1169–1208.

Best, M. C., J. Hjort, and D. Szakonyi (2019). *Individuals and Organizations as Sources of State Effectiveness*. Tech. rep. Working Paper 23. National Bureau of Economic Research.

Bhuller, Manudeep, Gordon B. Dahl, Katrine V. Løken, and Magne Mogstad (2020). "Incarceration, Recidivism, and Employment". In: *Journal of Political Economy* 128.4, pp. 1269–1324.

Bloom, N., R. Lemos, R. Sadun, and J. Van Reenen (2015). "Does Management Matter in Schools?" In: *The Economic Journal* 125.584, pp. 647–674.

Bloom, N. and J. Van Reenen (2007). "Measuring and Explaining Management Practices Across Firms and Countries". In: *The Quarterly Journal of Economics* 122.4, pp. 1351–1408.

— (2010). "Why Do Management Practices Differ across Firms and Countries?" In: *Journal of Economic Perspectives* 24.1.

Brynjolfsson, Erik, A. McAfee, M. Sorell, and F. Zhu (2007). "Scale without Mass". In: *Harvard Business School Working Paper* 7.16.

Callaway, Brantly and Pedro H.C. SantAnna (2021). "Difference-in-Differences with multiple time periods". In: *Journal of Econometrics* 225.2, pp. 200–230.

Campaniello, Nadia, Theodoros M. Diasakos, and Giovanni Mastrobuoni (2017). "Rationalizable Suicides: Evidence from Changes in Inmates Expected Length of Sentence". In: *Journal of the European Economic Association* 15.2, pp. 388–428.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). "The Effect of Minimum Wages on Low-Wage Jobs". In: *The Quarterly Journal of Economics* 134.3, pp. 1405–1454.

Chaisemartin, Clément de and Xavier D'Haultfuille (2020). "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects". In: *American Economic Review* 110.9, pp. 2964–96.

Chandra, A., A. Finkelstein, A. Sacarny, and C. Syverson (2013). *Healthcare Exceptionalism? Productivity and Allocation in the U.S. Healthcare Sector*. Tech. rep. Working Paper 19200. National Bureau of Economic Research.

Chen, M. Keith and Jesse M. Shapiro (2007). "Do Harsher Prison Conditions Reduce Recidivism? A Discontinuity-Based Approach". In: *American Law and Economics Review* 9.1, pp. 1–29.

Cunningham, Scott, Jennifer A. Seward, Karen Clay, and Victoria S. Vigliotti (2024). "Adverse Impacts of Mental Health Needs Assessment on Jail Outcomes: Evidence from Transition Age Youth and Adults". In: *Journal of Human Resources* 59.S, S282–S316.

d'Este, Rocco (2023). *Inside Prisons: The Effects of a Blanket Ban on Legal Highs*. Available at SSRN 3992273.

Drago, Francesco, Roberto Galbiati, and Pietro Vertova (2011). "Prison Conditions and Recidivism". In: *American Law and Economics Review* 13.1, pp. 103–130.

Fair, Helen and Roy Walmsley (2024). *World Prison Population List (14th edition)*. Reports ~11.5 million people held in penal institutions worldwide. Institute for Crime & Justice Policy Research (ICPR), Birkbeck, University of London. URL: https://www.prisonstudies.org/.

Favril, L., R. Yu, K. Hawton, and S. Fazel (2020). "Risk factors for self-harm in prison: a systematic review and meta-analysis". In: *Lancet Psychiatry* 7.8, pp. 682–691.

Fazel, Seena, Alison J. Hayes, Katherine Bartellas, Massimo Clerici, and Robert Trestman (2016). "Mental Health of Prisoners: Prevalence, Adverse Outcomes, and Interventions". In: *The Lancet Psychiatry* 3.9, pp. 871–881.

Fenizia, A. (2022). "Managers and productivity in the public sector". In: *Econometrica* 90.3, pp. 1063–1084.

Foster, L., J. Haltiwanger, and C. Syverson (2008). "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" In: *American Economic Review* 98.1, pp. 394–425.

Galbiati, Roberto, Aurelie Ouss, and Adrien Philippe (2021). "Jobs, News and Reoffending after Incarceration". In: *The Economic Journal* 131.633, pp. 247–270.

Gardner, J. (2022). "Two-stage differences in differences." In.

Gibbons, Robert and Rebecca Henderson (2012). *What do managers do?: Exploring persistent performance differences among seemingly similar enterprises*. Harvard Business School.

Ginneken, Esther F. J. C. van and Hanneke Palmen (2023). "Is There a Relationship Between Prison Conditions and Recidivism?" In: *Justice Quarterly* 40.1, pp. 106–128. DOI: 10.1080/07418825.2022.2040576.

Ginneken, Esther F.J.C. van and John Wooldredge (2024). "Offending and victimization in prisons: New theoretical and empirical approaches". In: *International Journal of Law, Crime and Justice* 77.

Giorcelli, M. (2019). "The Long-Term Effects of Management and Technology Transfers". In: *American Economic Review* 109.1, pp. 121–152.

Goodman-Bacon, A. (2021). "Difference-in-differences with variation in treatment timing". In: *Journal of Econometrics* 225.2, pp. 254–277.

Haney, C. (2001). *The Psychological Impact of Incarceration: Implications for Post-Prison Adjustment*. Tech. rep. Office of the Assistant Secretary for Planning and Evaluation.

Hjalmarsson, Randi (2009). "Juvenile Jails: A Path to the Straight and Narrow or to Hardened Criminality?" In: *The Journal of Law and Economics* 52.4, pp. 779–809.

Hjalmarsson, Randi and Matthew J. Lindquist (2022). "The Health Effects of Prison". In: *American Economic Journal: Applied Economics* 14.4, pp. 234–270.

HM Prison and Probation Service (2019). *Prison Drugs Strategy*. Tech. rep.

Hortacsu, Ali and Chad Syverson (2009). "Why do Firms Own Production Chains?" In: *US Census Bureau Center for Economic Studies Paper* 09-31.

Humphries, John Eric, Aurelie Ouss, Kamelia Stavreva, Megan Stevenson, and Winnie van Dijk (2023). *Conviction, Incarceration, and Recidivism: Understanding the Revolving Door*. Available at SSRN.

Ilmakunnas, P., M. Maliranta, and J. Vainiomäki (2004). "The Roles of Employer and Employee Characteristics for Plant Productivity". In: *Journal of Productivity Analysis* 21.3, pp. 249–276.

Inspection framework (Mar. 2022). HM Inspectorate of Prisons.

Katz, Lawrence, Steven D. Levitt, and Ellen Shustorovich (2003). "Prison Conditions, Capital Punishment, and Deterrence". In: *American Law and Economics Review* 5.2, pp. 318–343.

Liebling, Alison (2011). "Moral performance, inhuman and degrading treatment and prison pain". In: *Punishment & Society* 13.5, pp. 530–550. DOI: 10.1177/1462474511422159.

Lotti, Giovanni (2022). "Tough on Young Offenders: Harmful or Helpful?" In: *Journal of Human Resources* 57.4, pp. 1276–1310.

Mastrobuoni, Giovanni and Daniele Terlizzese (2022). "Leave the Door Open? Prison Conditions and Recidivism". In: *American Economic Journal: Applied Economics* 14.4, pp. 200–233.

Maue, Casey C, Marshall Burke, and Kyle J Emerick (2020). *Productivity Dispersion and Persistence Among the World's Most Numerous Firms*. Working Paper 26924. National Bureau of Economic Research.

Mukherjee, Anita (2021). "Impacts of Private Prison Contracting on Inmate Time Served and Recidivism". In: *American Economic Journal: Economic Policy* 13.2, pp. 408–438.

Nagin, Daniel S. and Charles E. Loeffler (2021). "The Impact of Incarceration on Recidivism". In: *Annual Review of Criminology* 4, pp. 133–152. DOI: 10.1146/annurev-criminol-030920-112506.

O'Donnell, Ian and Kimmett Edgar (1999). "Fear in Prison". In: *The Prison Journal* 1.

Prison Governance (Oct. 2019). House of Commons Justice Committee.

Rasul, I. and D. Rogger (2018). "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service". In: *Economic Journal* 128, pp. 413–446.

Rawls, John (1971). *A Theory of Justice*. Cambridge, MA: Belknap Press of Harvard University Press.

Sakellaris, P. and D. J. Wilson (2004). "Quantifying Embodied Technological Change". In: *Review of Economic Dynamics* 7.1, pp. 1–26.

Spivak, Andrew L. and Susan F. Sharp (2008). "Inmate Recidivism as a Measure of Private Prison Performance". In: *Crime Delinquency* 54.3, pp. 482–508.

Syverson, Chad (2004). "Product Substitutability and Productivity Dispersion". In: *The Review of Economics and Statistics* 86.2, pp. 534–550.

— (2011). "What Determines Productivity?" In: *Journal of Economic Literature* 49.2, pp. 326–365.

Tobón, Santiago (2022). "Do Better Prisons Reduce Recidivism? Evidence from a Prison Construction Program". In: *Review of Economics and Statistics* 104.6, pp. 1256–1272.

UNODC (2024). *Global prison population and trends*. https://www.unodc.org/. UNODC Research Brief: estimates 11.5 million people imprisoned worldwide in 2022.

Useem, B. and M. D. Reisig (1999). "Collective action in prisons: protests, disturbances, and riots". In: *Criminology* 37, pp. 735–760.

Western, Bruce (2021). "Inside the Box: Safety, Health, and Isolation in Prison". In: *Journal of Economic Perspectives* 35.4, pp. 97–122.

Wing, Coady, Seth Freedman, and Alex Hollingsworth (2024). "Stacked Difference-in-Differences". In: *NBER Working Paper*, w32054.

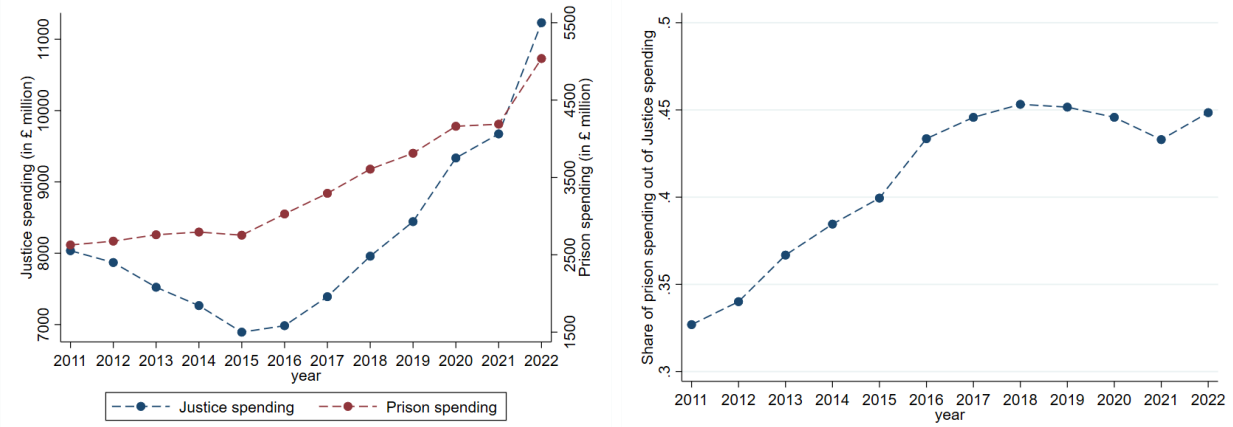
Wooldredge, John (2020). "Prison Culture, Management, and In-Prison Violence". In: *Annual Review of Criminology* 3.1, pp. 165–188.

Wooldridge, Jeffrey (2021). "Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators". In.

Appendix A: Institutional Background & Data

A.1 Background

Figure A1: **Prison spending**



Notes: The left panel plots the trend in Justice and Prison spending in real terms in England and Wales from 2011 to 2022. The right panel plots the trend in Prison spending as a share of Justice spending in England and Wales from 2011 to 2022. *Justice spending* refers to the Resource DEL (excluding depreciation) by the departmental group of Justice (source: HM Treasury). *Prison spending* refers to the Overall Resource Expenditure, i.e. net expenditure managed and recorded at each prison plus the net expenditure met at regional or national level (source: HMPPS).

A.2 Data

A.2.1 Inspection Survey

At the beginning of every HMIP inspection, a detainee survey is carried out. We accessed the HMIP Prisoner Survey, detained at UK Data Service, which collects the anonymized inmates' responses.

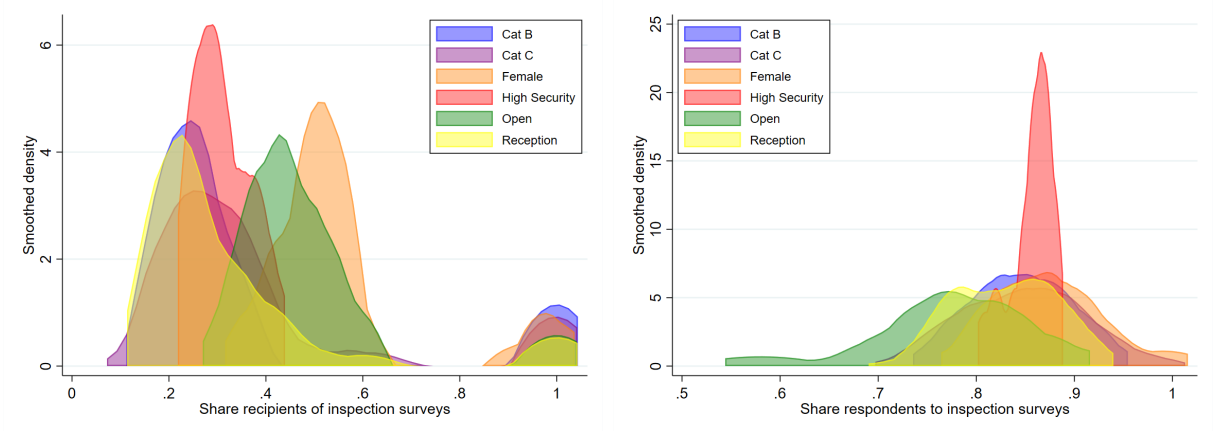
The questionnaire is offered to a random sample of detainees, whose size is calculated to ensure that the findings can be generalised to the entire population of the establishment. Smaller establishments may offer the questionnaire to the entire population. On average, it was distributed to 205 prisoners (43% of the establishment population). Participation is then voluntary. On average, 172 inmates, 28% of the prison population and 84% of the addressed sample, completed the survey.

Figure A2 reveals that the recipient rate in the population is heterogeneous across prison functions, while the response rate in the survey recipients sample is uniform across establishment functions. Figure A3 suggests that while the sample size addressed by HMIP

researchers does vary across establishments to comply with population representativeness criteria, detainees' response rate is fairly homogenous across establishments.

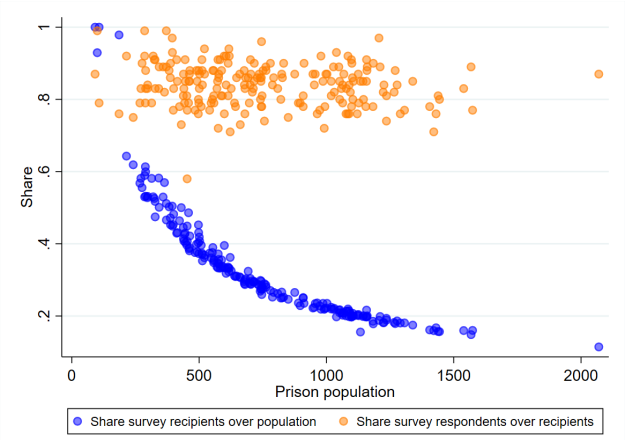
The survey consists of questions covering the prisoner experience from reception to release, such as the relationship with the staff, the contacts with family and friends, the time out of cell, healthcare, education, behavior management, as well as demographic and background questions.

Figure A2: **Prisoner Survey Recipients and Respondents by prison function**



Notes: The figure plots the density of the share of recipients (*left panel*) and respondents to the Inspection Prisoners' Survey (*right panel*), by prison function. The share of recipients is calculated as the fraction of inmates to whom the survey was distributed over the total establishment population. The share of respondents is calculated as the fraction of survey respondents over the survey recipients.

Figure A3: **Prisoner Survey Recipients and Respondents by prison size**



Notes: The figure plots the share of recipients (*left panel*) and respondents to the Inspection Prisoners' Survey (*right panel*) against the prison size. The share of recipients is calculated as the fraction of inmates to whom the survey was distributed over the total establishment population. The share of respondents is calculated as the fraction of survey respondents over the survey recipients.

A.2.2 Data sources & statistics

[1] *Market index of crime, violent crimes and drug-related crimes*, refers to the average constituency-level crime rate, per capita, for violence crime and drugs for each prison, with weights inversely proportional to the distance constituency-prison.

Table A1: **Data sources**

Variable	Periodicity	Categories	Time Coverage
Prison Population Data			
Population	Quarterly	Age (9 groups) Custody type (3 classes) Offence type (11 classes) Ethnicity (4 classes)	2015 – 2023
	Yearly	Incentive status (4 classes)	2016 – 2023
Safety in Custody			
Death	Monthly	Type of death (5 classes) Age (11 groups) Ethnicity (4 classes)	2000 – 2023
Assaults	Monthly	Type of assault (5 categories)	2003 – 2023
Self-harm	Monthly	Type of self-harm (2 categories)	2003 – 2023
Escapes	Yearly		1996 – 2023
Absconds	Yearly		1996 – 2023
Temporary Release Failures	Yearly		2012 – 2023
Releases in Error	Yearly		2007 – 2020
Percentage of prisoners held in crowded accommodation	Yearly	Single/double crowded accommodation	1996 – 2023
Protest	Yearly	Type of protest (5 classes)	2007 – 2023
Finds	Monthly	Type of find (3 categories) Type of find (13 categories)	2015 – 2023 2007 – 2023
Drug Finds	Monthly	Drug type (18 categories)	2000 – 2023
Community Accommodation Service (CAS)	Yearly		2016 – 2023
Reform			
Accredited Programme	Yearly	Programme type (5 categories)	2016–2021
Percentage of positive random mandatory drug tests for traditional drug	Yearly		1999–2020
Random mandatory drug tests administered	Monthly		2007-2018
Positive drug samples	Yearly	Drug type (18 categories)	2020
Random mandatory drug testing outcomes	Monthly	Outcome type (6 categories)	2007-2018
Staff			
Staff	Quarterly	Role (20 classes)	2013–2023
Staff ethnic minority	Yearly		2020–2023
Staff sickness	Yearly		2013–2022
Cost			
Resource Expenditure	Yearly	Direct/Overall	2013–2022
Cost per Place	Yearly	Direct/Overall	2013–2022
Cost per Prison	Yearly	Direct/Overall	2013–2022
Governors			
Governors' name	Yearly		2013–2022
Crime			
Crime at the constituency level	Yearly	Type of crime (14 categories)	2013–2023

Notes: The table outlines various categories related to prison data, including population statistics, safety, reform efforts, staff details, costs, governors' information, and crime data. The data spans multiple years and frequencies of collection.

Table A2: **Inmates' composition, Physical capital and Human capital**

	Mean	St. Dev.	Min	Max
Inmates' composition				
Total population	697.628	346.706	5	2080
Share of Population in Remand	0.109	0.145	0	1
Share of Population Sentenced	0.881	0.166	0	1
Share of Population Non Criminal	0.009	0.080	0	1
Share of Asian/Asian British	0.076	0.049	0	0
Share of Black/African/Caribbean/Black British	0.129	0.103	0	0
Share of Mixed/Multiple ethnic groups	0.044	0.027	0	0
Share of Other ethnic group	0.010	0.015	0	0
Share of White	0.732	0.165	0	1
Share of Not recorded ethnicity	0.009	0.031	0	1
Share of Population in age < 21	0.096	0.236	0	1
Share of Population in age 21 - 29	0.289	0.111	0	1
Share of Population in age 30 - 39	0.287	0.101	0	0
Share of Population in age 40 - 49	0.177	0.064	0	0
Share of Population in age 50 - 59	0.099	0.062	0	0
Share of Population in age 60 - 69	0.035	0.038	0	0
Share of Population in age 70 +	0.017	0.023	0	0
Share of Violence against the person	0.267	0.153	0	1
Share of Sexual offences	0.147	0.235	0	1
Share of Robbery	0.105	0.061	0	0
Share of Theft and Bulgarly	0.138	0.079	0	0
Share of Drug Offences	0.153	0.100	0	0
Share of Civil offences	0.014	0.012	0	0
Share of Fraud Offences	0.022	0.032	0	0
Share of Other offences	0.154	0.106	0	1
Share of population in sentence < 6m	0.034	0.049	0	0
Share of population in sentence 6m < x < 1y	0.029	0.032	0	0
Share of Population in sentence 1y < x < 4y	0.211	0.162	0	1
Share of Population in sentence 4y < x < 10y	0.283	0.181	0	1
Share of Population in sentence > 10y	0.150	0.163	0	1
Share of Population in life sentence	0.104	0.156	0	1
Share of Population in IPP sentence	0.054	0.060	0	1
Share of Population unsentenced	0.135	0.191	0	1
Physical capital				
Cost per Prisoner, from Direct Resource Expenditure	27277.057	13738.850	0	112669
Human capital				
Total staff	334.885	137.169	56	858
Bands 9-11 / Senior Managers (Operational) per prisoner	0.008	0.038	0	1
Bands 6-8 / Managers per prisoner	0.018	0.076	0	2
Band 5 / Custodial Managers per prisoner	0.027	0.077	0	2
Band 4 / Supervising Officer per prisoner	0.038	0.152	0	4
Bands 3-4 / Prison Officer (incl. specialists) per prisoner	0.299	1.091	0	29
Band 2 / Operational Support per prisoner	0.093	0.404	0	11
Senior Civil Servant per prisoner	0.006	0.039	0	1
Bands 9-11 / Senior Managers (Non Operational) per prisoner	0.007	0.039	0	1
Bands 5-8 / Managers per prisoner	0.021	0.076	0	2
Band 4 / Executive Officer per prisoner	0.011	0.051	0	1
Band 3 / Administrative Officer per prisoner	0.059	0.227	0	6
Band 2 / Administrative Assistant per prisoner	0.010	0.038	0	1
Other Admin per prisoner	0.007	0.039	0	1
Instructional Officer per prisoner	0.008	0.012	0	0
Psychology per prisoner	0.013	0.040	0	1
Chaplaincy per prisoner	0.009	0.038	0	1
Healthcare - Nursing Grades per prisoner	0.007	0.039	0	1
Healthcare - Other healthcare per prisoner	0.006	0.039	0	1
Industrial / Facilities per prisoner	0.026	0.152	0	4
Other Grades (including Probation Service grades) per prisoner	0.018	0.077	0	2

Notes: The table shows the summary statistics of the inmates' composition, the prison costs, and the workforce composition.

Table A3: **Crime rate**

	Mean	St. Dev.	Min	Max
Number of violent crimes per capita - constituency level	0.020	0.011	0.002	0.096
Number of drugs-related crimes per capita - constituency level	0.002	0.002	0.000	0.030
Market index of violent crimes	0.008	0.015	0.000	0.132
Market index of drugs-related crimes	0.001	0.003	0.000	0.020

Notes: The table shows the summary statistics of the crime rates at the constituency level and a market index of crime at the prison level[1] from 2013 to 2019 for violent crimes and drug-related crimes.

Table A4: **Prison Inspection Report Recommendations and Action Plan**

Recommendation	Agreed	Response	Responsible owner	Target date
Safety:				
Courts, escorts and transfers				
Women judged suitable for open conditions should not have to travel in cellular vehicles.	Not Agreed	This recommendation cannot be agreed because... Limited non-cellular vehicles are available; however, these are used for prisoners with special needs such as mobility issues... does not have the resources to move prisoners into the open estate in these non-cellular vehicles...	Prison Escort Contract Service (PECS)	
Early days in custody				
First night interviews should be carried out in a confidential setting.	Agreed	A system is now in place to ensure interviews are held on a one-to-one basis using the interview room and classroom... ensuring confidentially.	Governor	Completed
Security				
All requested suspicion drug tests should be completed within set guidelines.	Accepted subject to resources	The Suspicion Mandatory Drug Testing has been difficult to achieve this year due to resourcing pressures... The committed Resourcing Plan from HMPPS.. will provide the resourcing resilience to deliver this security work with consistency.	Head of Security	3 months
Respect:				
Legal rights				
Prisoners should have access to legal textbooks.	Accepted	Greater prisoner access to the library will be addressed by the introduction of new staffing profiles... Interim arrangements have been made in the form of a mobile library provision...	Head of Reducing Reoffending	5 months
Faith and religious activity				
The negative perceptions of Muslim prisoners about access to chaplains should be explored and addressed.	Accepted	There will be Chapel/Muslim representatives on each residential unit... A World Faith event will be organised... A Muslim chaplain will visit all units at least once a week	Head of Chaplaincy	10 days
Health services				
Newly initiated medicines should be ordered and supplied in a timely manner, to ensure that treatment begins promptly.	Accepted	The current process is for prescribers to mark such prescriptions as urgent which are then faxed to the supplier the same day and delivered same or next day. Prescribers have been reminded of this process and a quarterly audit has been initiated.	Head of Healthcare	Completed
Purposeful activity:				
Time out of cell				
A full regime which starts and finishes on time should be in place for all prisoners.	Partly agreed	... Modifications to the regime have been introduced which will enable less deviation from the core day... This will increase hours spent in purposeful activity...	Governor	Completed
Education, skills and work activities				
Leaders and managers should integrate English and maths into sessions effectively.	Accepted	... has rescheduled Maths and English sessions to maximise use of the current resource and education spaces using the model of a college day whereby prisoners in workshops are required to attend Maths and English once a week.	Governor	Completed
Outcomes and achievements				
A greater range of accreditation opportunities and qualifications should be available for work activities.	Agreed	The Industries Manager will review where accreditation is available and where opportunities exist for prisoners to work towards qualifications (e.g., the bike workshop)	Governor	3 months
Rehabilitation and release planning:				
Children and families and contact with the outside world				
Prisoners should have access to parenting and relationship courses.	Agreed	The Head of Reducing Reoffending will introduce a parenting and relationship course with a view to commencing during...	Governor	11 months

Notes: The table collects recommendations identified by the inspectors, in the context of full HMIP inspections, and the establishment response, including whether the recommendation was accepted, partially accepted or not accepted, the action taken (from the moment of the inspection to the moment of the submission of the action plan), the identified responsible owner for each action, and the target date. Note that: (1) To each recommendations multiple actions taken or planned can be listed, and each can be assigned to a different responsible owners; (2) Given that the action plan should be available three months after

Table A5: **Inspection report recommendations descriptive statistics**

	Mean	St. Dev.	Min	Max
Panel A: *Across all categories*				
Number of recommendations issued	51.56	19.24	10	138
Share of recommendations accepted	0.87	0.11	0.5	1
Share of recommendations achieved	0.39	0.16	0	0.81
Panel B: *Safety*				
Number of recommendations issued	12.34	6.19	1	38
Share of recommendations accepted	0.90	0.11	0.4	1
Share of recommendations achieved	0.42	0.22	0	1
Panel C: *Respect*				
Number of recommendations issued	19.53	8.34	2	54
Share of recommendations accepted	0.85	0.15	0.38	1
Share of recommendations achieved	0.38	0.23	0	1
Panel D: *Purposeful Activity*				
Number of recommendations issued	9.90	4.26	1	29
Share of recommendations accepted	0.89	0.13	0.3	1
Share of recommendations achieved	0.39	0.25	0	1
Panel E: *Rehabilitation*				
Number of recommendations issued	9.22	4.20	1	23
Share of recommendations accepted	0.86	0.17	0.2	1
Share of recommendations achieved	0.40	0.19	0	0.89

Notes: The table shows summary statistics for the number of recommendations issued by the inspection team at the end of the inspection, together with the share of recommendations accepted by the prison governor and the share of achieved. The statistics are reported at the inspection level, for the 307 inspections from 2013 to 2019. In *Panel A* we consider all recommendations, while in the following four panels we disaggregate by the Healthy establishment test as set out by the HMIP Inspection Framework, *Safety, Respect, Purposeful Activity*, and *Rehabilitation and release planning*.

Table A6: **Prevalence Score Descriptive Statistics by Category**

	Mean	Sd	Min	Max
Panel A: *Across all categories*				
Mean	0.32	0.04	0.22	0.43
Median	0.31	0.03	0.22	0.42
p90	0.49	0.04	0.34	0.69
p75	0.40	0.04	0.29	0.52
p25	0.23	0.03	0.13	0.34
p10	0.17	0.04	0.05	0.28
Gini	0.78	0.03	0.67	0.86
Panel B: *Safety*				
Mean	0.32	0.03	0.22	0.42
Median	0.31	0.03	0.22	0.40
p90	0.48	0.04	0.34	0.60
p75	0.40	0.03	0.29	0.49
p25	0.23	0.03	0.13	0.33
p10	0.16	0.04	0.06	0.28
Gini	0.77	0.03	0.68	0.85
Panel C: *Respect*				
Mean	0.32	0.04	0.22	0.43
Median	0.32	0.03	0.22	0.42
p90	0.48	0.04	0.35	0.62
p75	0.40	0.04	0.29	0.51
p25	0.24	0.04	0.13	0.34
p10	0.17	0.04	0.05	0.28
Gini	0.78	0.03	0.67	0.86
Panel C: *Purposeful Activity*				
Mean	0.33	0.03	0.23	0.42
Median	0.32	0.03	0.22	0.41
p90	0.50	0.05	0.37	0.69
p75	0.41	0.04	0.30	0.52
p25	0.23	0.03	0.14	0.33
p10	0.17	0.03	0.06	0.28
Gini	0.77	0.03	0.68	0.86
Panel D: *Rehabilitation*				
Mean	0.32	0.03	0.23	0.42
Median	0.31	0.03	0.22	0.41
p90	0.49	0.04	0.37	0.63
p75	0.40	0.03	0.30	0.51
p25	0.23	0.03	0.14	0.33
p10	0.16	0.03	0.07	0.27
Gini	0.77	0.02	0.68	0.85

Statistics calculated for each inspection category

Notes: The table displays summary statistics at the recommendation level for the our measure of prevalence score by category. To compute the *recommendation prevalence score* we assign a set of semantic similarity scores to each prison recommendation based on its textual proximity to other recommendations issued in the same year and within the same category but in different prisons. In particular, we employ Doc2Vec, which is an unsupervised machine learning algorithm that generates fixed-length vector representations of variable-length text documents. It embeds these into a continuous vector space, where proximity between vectors reflects semantic similarity between recommendations. We then use cosine similarity to compare a recommendation r against all other recommendations r' from the same year and category but from different prisons, computing a distribution of similarity scores for each recommendation. We then aggregate these scores across r' using the mean, median, percentiles (90th, 75th, 25th, 10th), standard deviation, and Gini coefficient to produce *prevalence scores*.

Table A7: **Correlation Matrix of Prevalence Scores**

	Recommendation prevalence score						
	Mean	Median	p90	p75	p25	p10	Gini
Mean	1						
Median	0.980***	1					
p90	0.888***	0.828***	1				
p75	0.954***	0.934***	0.932***	1			
p25	0.954***	0.954***	0.735***	0.845***	1		
p10	0.919***	0.903***	0.681***	0.784***	0.969***	1	
Gini	0.683***	0.718***	0.318***	0.497***	0.846***	0.882***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table displays the correlation between the different measures of the *recommendation prevalence score*, where the score at the recommendation level is computed as the *mean, median,* 90^{th}, 75^{th}, 25^{th}, 10^{th} *percentile*, or *Gini coefficient* of the prevalence score between the recommendation and each other recommendation in other prisons within same category and year.

Appendix B: Persistent Differences in Violence

B.1 Further results in Dispersion

Table B1: **Prison function, ownership, Victorian structure, and prison area**

	Prison Fixed Effects		
	Violence on staff	Violence on prisoners	Self-violence
Category A	-68.433*	-55.167	-656.447**
	(37.473)	(40.498)	(266.254)
Category B	2.462	-32.522	-403.200
	(34.875)	(37.690)	(247.797)
Category D	143.478***	-84.435**	-366.004
	(38.489)	(41.596)	(273.473)
Reception	-19.590	42.122	-383.736*
	(28.934)	(31.270)	(205.585)
YJB	119.877	38.711	61.884
	(104.816)	(113.276)	(744.738)
Private	19.451	148.404***	212.563
	(33.645)	(36.360)	(239.053)
Female prison	56.228	3.371	1091.822***
	(38.412)	(41.512)	(272.924)
Victorian prison	63.768**	32.608	59.824
	(29.106)	(31.455)	(206.804)
Prison area in sq metres	0.001***	0.001***	0.003**
	(0.000)	(0.000)	(0.001)
Prison area per prisoner	-0.586***	-0.607***	-1.296
	(0.112)	(0.121)	(0.793)
Observations	108	108	108
R-squared	0.318	0.575	0.258

*** p<0.01, ** p<0.05, * p<0.1

Notes: The table shows the estimates of the regression of prison fixed effects for *Violence on staff*, *Violence on prisoners*, and *Self-violence*, estimated on inmates' quality Q, capital K controls, and year FE, on *Prison functions* dummies, *Private prison* dummy, *Victorian prison* dummy, and *Prison area* in squared meters in absolute terms and relative to the inmates' population. The excluded category is Category C.

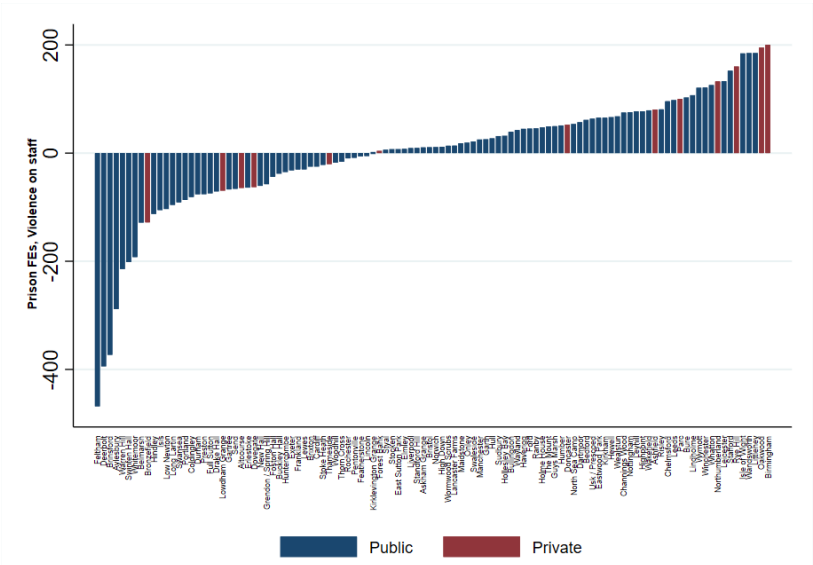
Table B2: **Decomposition of violence based on loss of fit**

Decomposition based on loss of fit (% of R^2)			
Outcome:	**Violence on staff**	**Violence on prisoners**	**Self-violence**
Prison function	47.2	6.7	30.5
Private dummy	0.7	12.7	2.3
Male dummy	67.1	23.1	48.18
Victorian dummy	10.6	0.8	0.2
Area	66.0	20.0	14.7
R^2	0.32	0.57	0.26
Adj R^2	0.25	0.53	0.18
Obs	108	108	108

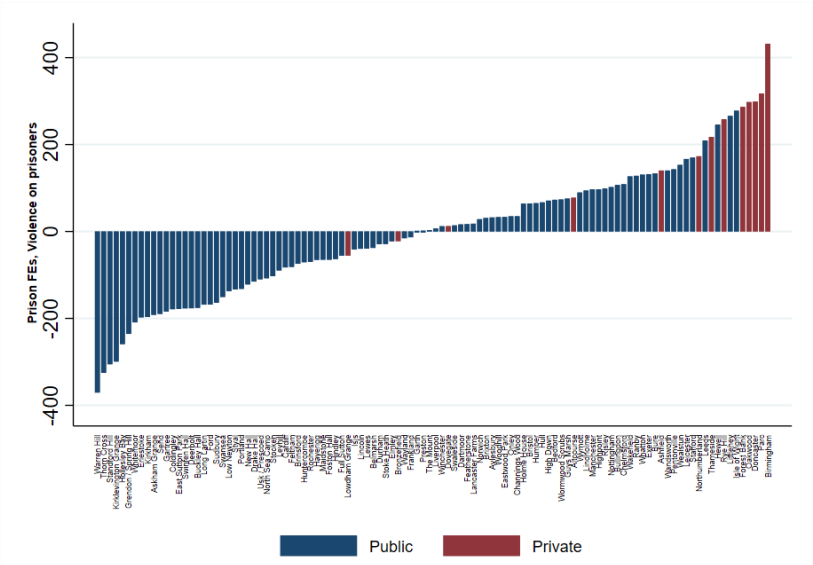
Notes: Each entry reflects the loss of fit resulting from removing the controls in the rows from a linear projection of each outcome on the set of fixed effects in the specification controlling for {Prison Function, Private prison dummy, Male prison dummy, Victorian prison dummy, Prison area in squared meters}. The loss of fit is computed as the share over the fit in the full model: $\frac{(R^2_I - R^2_{I-1})}{R^2_{I-1}}$. The outcomes are measured in number of incidents per thousand inmates.

Figure B1: **Prison Fixed Effects for Private VS Public prisons, ranked by level**

(a) Violence on staff



(b) Violence on prisoners



(c) Self-violence

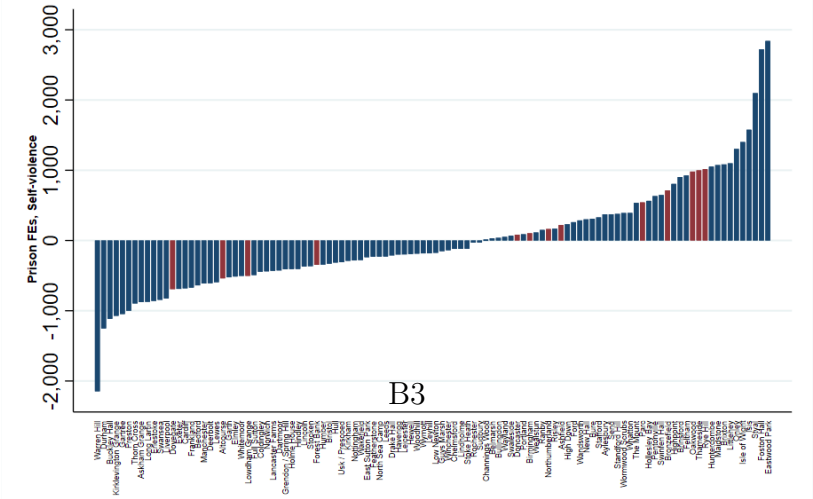
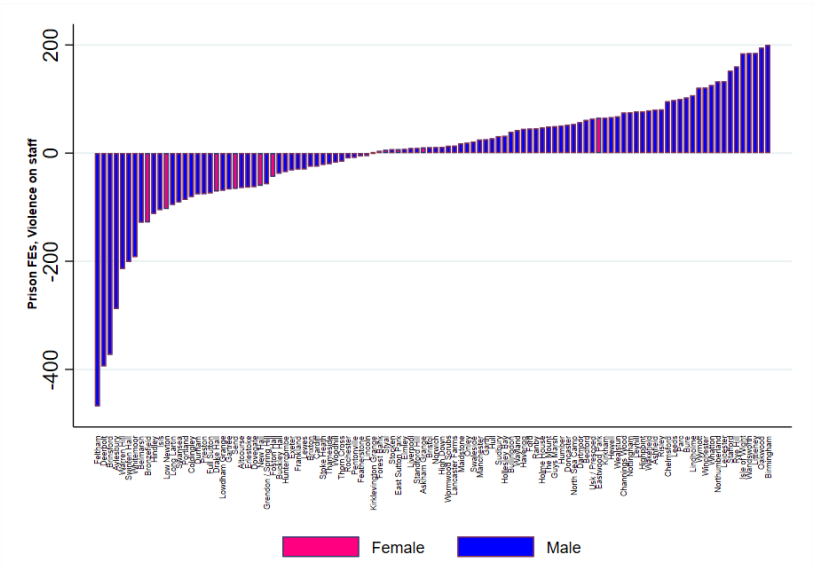
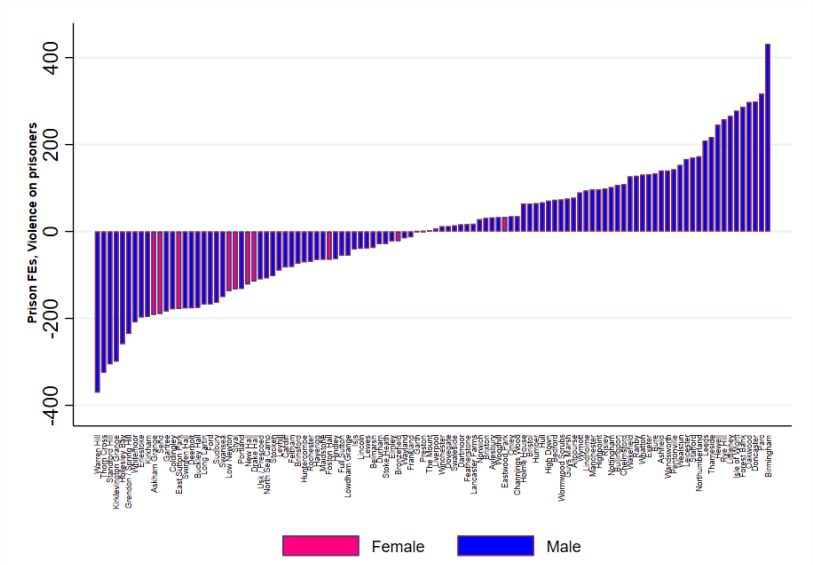


Figure B2: **Prison Fixed Effects for Female VS Male prisons, ranked by level**

(a) Violence on staff



(b) Violence on prisoners



(c) Self-violence

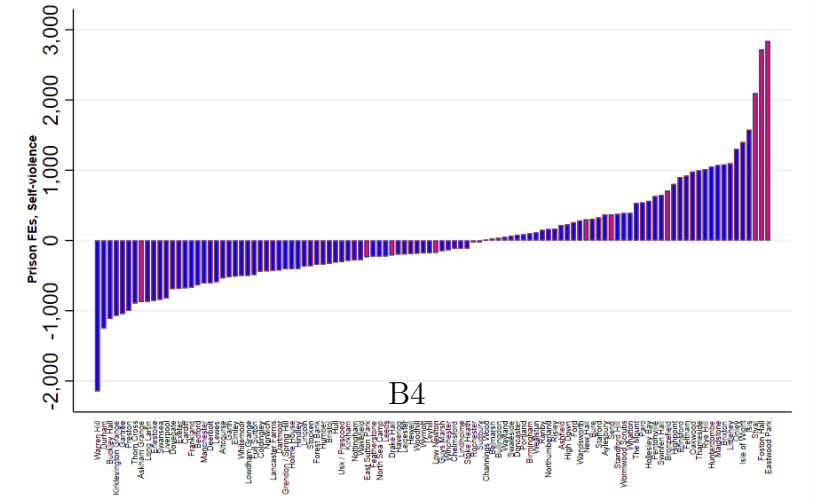
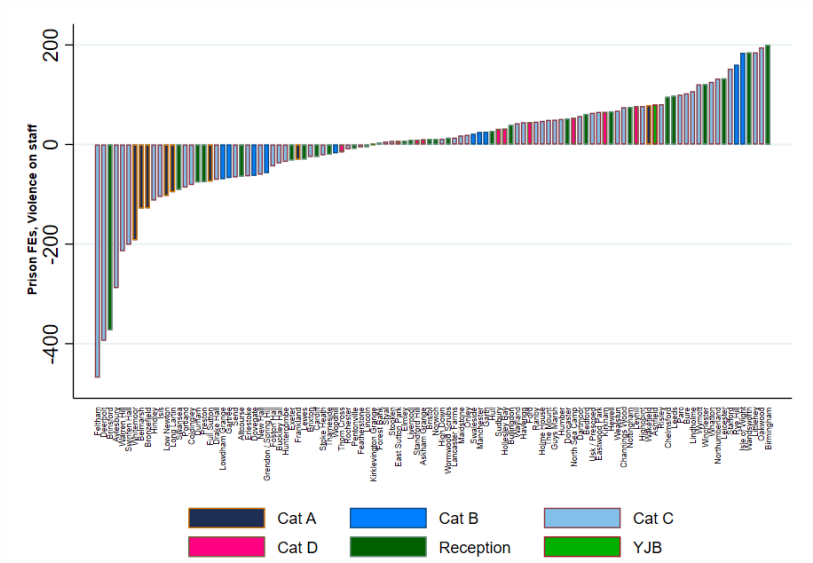
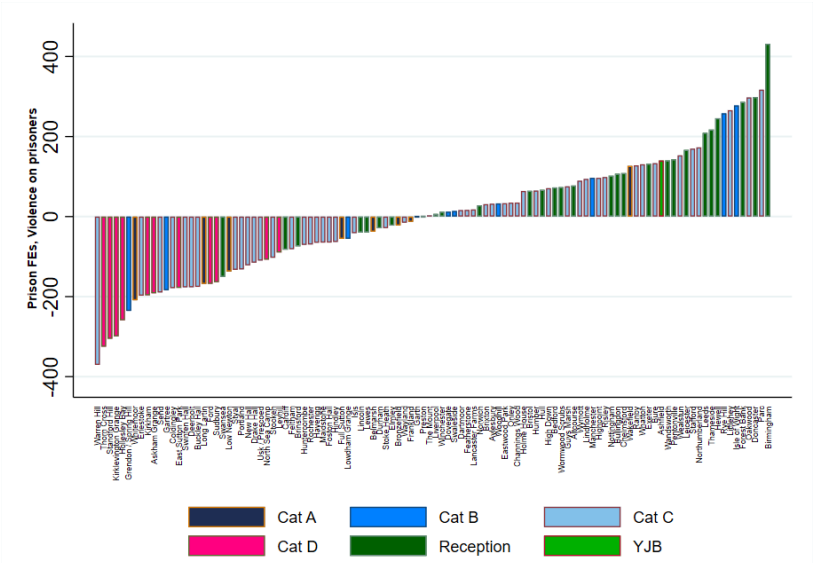


Figure B3: **Prison Fixed Effects by prison function, ranked by level**

(a) Violence on staff



(b) Violence on prisoners



(c) Self-violence

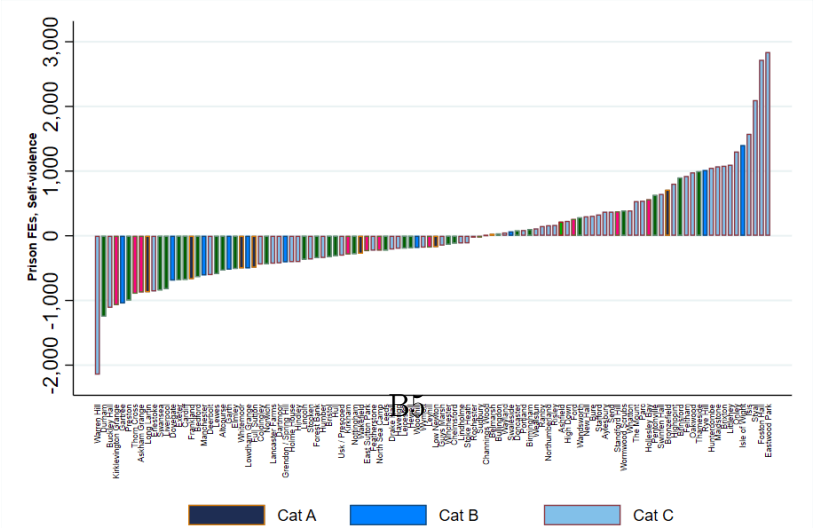
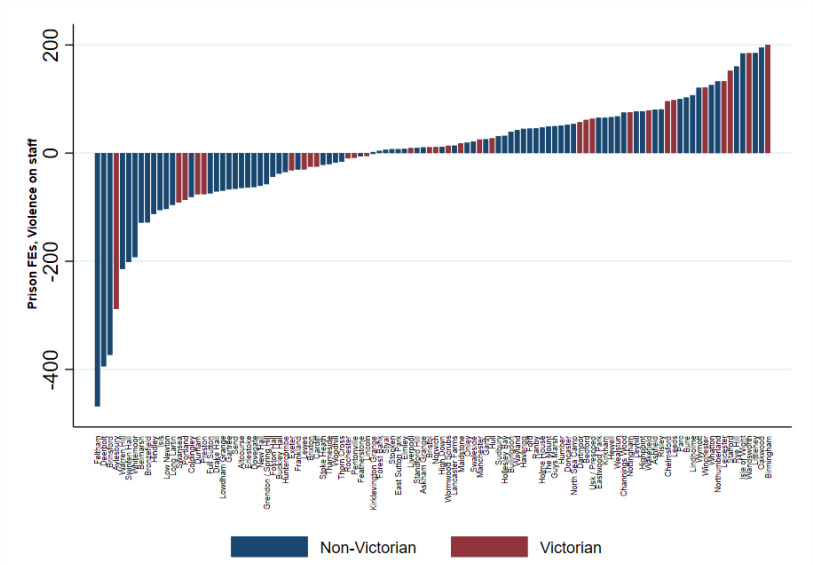
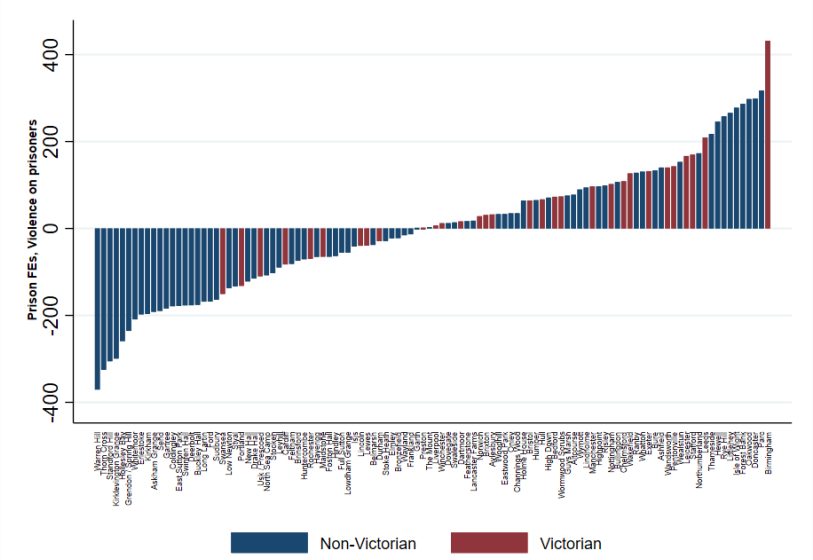


Figure B4: **Prison Fixed Effects for Victorian VS non-Victorian prisons, ranked by level**

(a) Violence on staff



(b) Violence on prisoners



(c) Self-violence

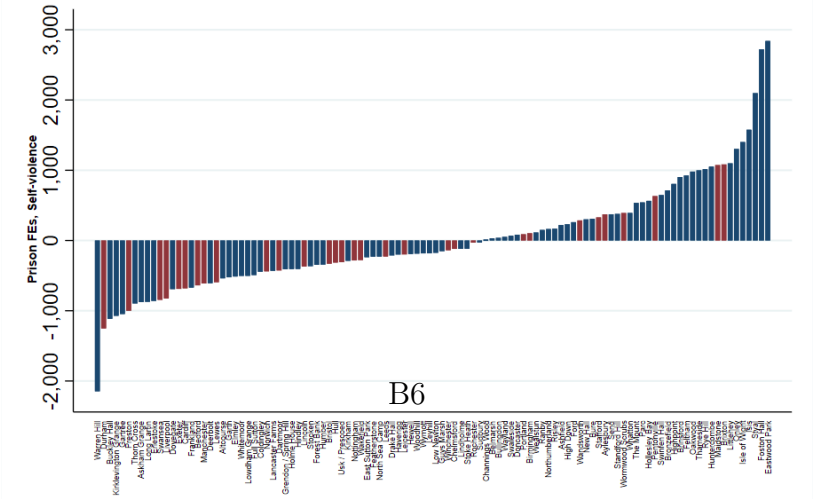
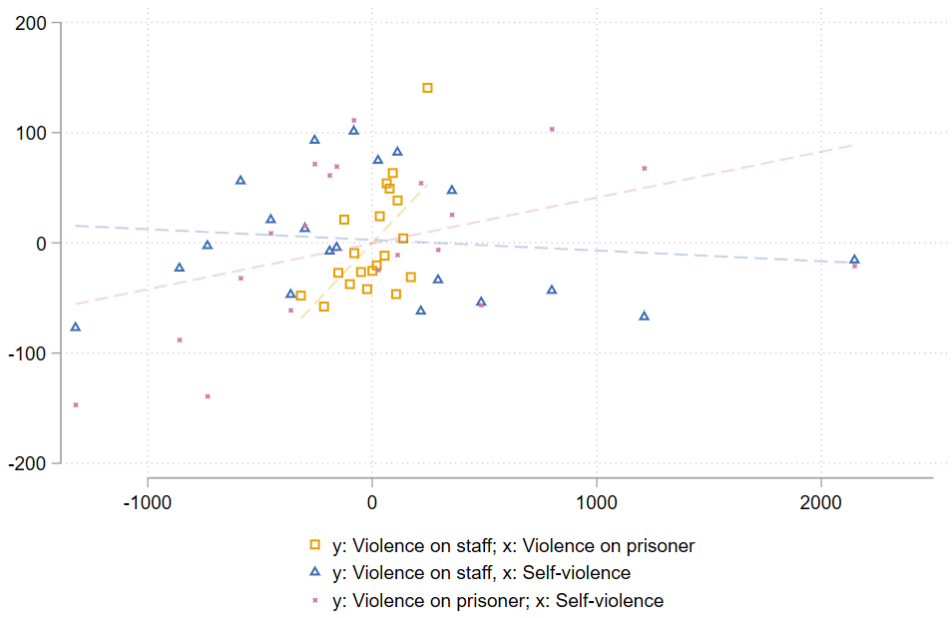


Figure B5: **Bilateral relationship between types of violence**

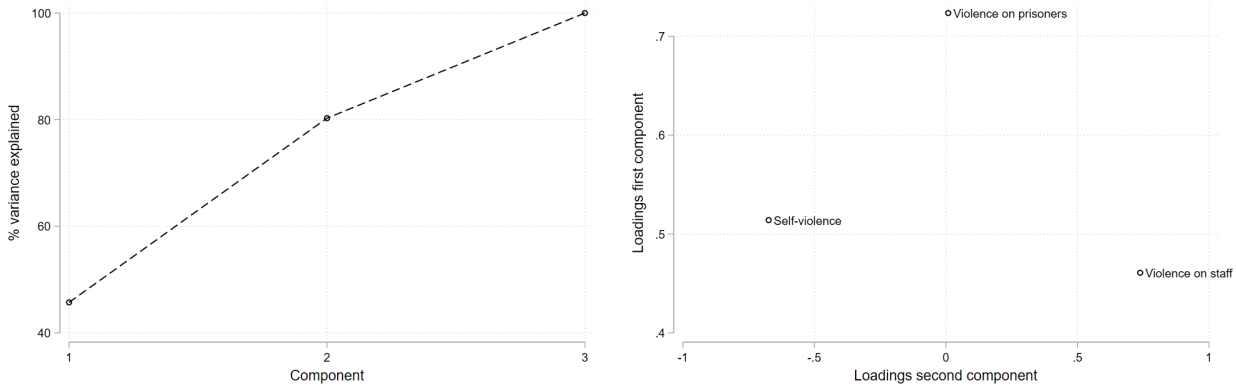


Notes: The figure plots the bilateral scatterplots of the the estimated Prison Fixed Effects for *Violence on staff*, *Violence on prisoners*, *Self-violence*. The outcomes of interest are measured in number of incidents per thousands of prisoners.

Figure B6: **PCA on Prison Fixed Effects**

(a) Variance explained

(b) Loadings



Notes: The Principal Component Analysis is performed on the Prison Fixed Effects resulting from the regressions of *Violence on staff*, *Violence on prisoners*, and *Self-violence* on the set of controls. Panel (a) plots the variance explained by the three components, and panel (b) the relative loadings.

B.2 Persistence

The analysis above suggests that firms at the 75^{th} percentile experience significantly more cases of violence than prisons at the 25^{th} percentile. To investigate the unexplained dispersion, we first examine its persistence over time.

To quantify the extent to which large differences in unexplained violence persist year-over-year, we estimate an OLS regression of contemporaneous violence residuals (from equation 1) on their one-year-lagged counterparts. The model uses NeweyWest standard errors to account for heteroskedasticity and first-order autocorrelation:

$$v_{it} = \delta + \theta v_{it-1} + \kappa_{it} \tag{B1}$$

Table B3 presents the estimated AR(1) coefficients for each type of violence, separately for the *conditional* and *unconditional* residuals[1]. Comparing these within each violence type reveals distinct patterns. The conditional specification exhibits substantial persistence year to year across all violence types. The conditional specification exhibits significantly lower persistence than the unconditional one (dropping from 0.56-0.33 to 0.21-0.07)[2]. If the *Q, K, L* factors we control for were exogenous to the dynamic process of violence, the coefficient should not differ across specifications. The substantial difference instead suggests that some factors stable over time and determinants of violence create spurious persistence in violence and/or that resources, such as prison workforce, are allocated based on the in-prison conditions. The results are consistent with a prison system resource allocation policy that responds to violence levels.

The AR(1) model above estimates the year-by-year persistence, capturing short-term relationships between consecutive years. However, it does not provide insights into how prisons transition between different levels of violence over longer time horizons. To address this, we complement the AR(1) analysis by computing transition matrices from the first three years of our sample to the last three years. Transition matrices also allow us to uncover differences in the persistence dynamics of prisons across different parts of the distribution that a single AR(1) coefficient may obscure.

Tables B4, B5, and B6 present the transition matrix for the distribution of residuals from equation 1, for each violence type. Each entry (q, r) reflects the probability that a prison

[1] *Unconditional* refers to the residuals being obtained from the regression of violence measures on the set of fixed effects only, while *Conditional* to the residuals from the full specification, as in equation 1.

[2] Overall, the observed persistence is between one-half and one-fourth of the estimates from the existing firm-productivity literature. For example, Maue et al. (2020) and Foster et al. (2008) find autocorrelation coefficients of 0.37-0.44 among African agricultural firms and 0.75-0.8 among U.S. manufacturing firms, respectively.

Table B3: **AR(1) of violence residuals**

	Residuals							
	Unconditionally				Conditionally			
All violence	0.425***				0.178			
	(0.103)				(0.110)			
Violence on staff		0.561***				0.066		
		(0.125)				(0.070)		
Violence on prisoners			0.499***				-0.060	
			(0.181)				(0.061)	
Self-violence				0.327***				0.212*
				(0.125)				(0.114)
Observations	452	452	452	452	380	380	380	380

*** p<0.01, ** p<0.05, * p<0.1

Notes: The table shows the estimates from the OLS regression of contemporaneous violence residuals from equation 1 on the one-year-lagged residuals with NeweyWest standard errors robust to heteroskedasticity and first-order autocorrelation for *violence on staff*, *violence on prisoners*, and *self-violence*. *Unconditionally* refers to the residuals being obtained from the regression of violence measures on the set of fixed effects only, while *Conditionally* to the residuals from the full specification, as in equation 1.

at the q^{th} quantile of the 2014-2015 prisons distribution remains in, or transition to, the r^{th} quantile in the 2018-2019 distribution. On average, more than half of the prisons in a given quantile either stay in the same quantile or move up or down by just one quantile, indicating a moderate level of mobility. However, the transition matrices reveal that prisons in the worst-performing quantile (highest levels of violence) and the best-performing quantile (lowest levels) are more "stuck" in their positions. This suggests that extreme levels of violence, whether high or low, are more persistent over time compared to middle-range quantiles. Combined with the results from the AR(1), the results are consistent with resource allocation successfully improving conditions in middle-range violence prisons, whereas extreme cases are resistant to interventions.

As for the dispersion analysis in Section 3.2, in Appendix we replicate the persistence analysis (1) including a control for the number of drugs finds and a "market access" index of crime (see Tables B13 and B14), and (2) on violence measures that are less susceptible to misreporting, i.e. *serious staff assaults*, *serious prisoner assaults* and *homicides*, and *suicides* and *self-harm requiring hospitalization* (see Table B15).

Overall, our analysis reveals very weak persistence in in-prison violence, particularly for prisons at the extremes of violence levels. The difference between the conditional and unconditional specifications reflects a combination of stable prison characteristics and a resource allocation policy that targets high-violence establishments.

Table B4: **Transition matrix of violence on staff**

	Violence on staff				
	1	2	3	4	5
1	31.58	15.79	21.05	31.58	0.00
2	15.79	21.05	5.26	36.84	21.05
3	26.32	26.32	21.05	10.53	15.79
4	15.79	21.05	10.53	10.53	42.11
5	10.53	15.79	42.11	10.53	21.05

Notes: Transition matrices are computed as the probability for a prison at the q^{th} quintile in 2014-2015 to be in the r^{th} quintile in 2018-2019 of the distribution of residuals, where the residuals are computed from the regression of the level of *Violence on staff* per thousand prisoners on Q, K, L controls and year FE.

Table B5: **Transition matrix of violence on prisoners**

	Violence on prisoner				
	1	2	3	4	5
1	36.84	26.32	10.53	21.05	5.26
2	36.84	21.05	15.79	15.79	10.53
3	10.53	21.05	21.05	21.05	26.32
4	15.79	10.53	36.84	26.32	10.53
5	5.26	21.05	15.79	10.53	47.37

Notes: Transition matrices are computed as the probability for a prison at the q^{th} quintile in 2014-2015 to be in the r^{th} quintile in 2018-2019 of the distribution of residuals, where the residuals are computed from the regression of the level of *Violence on prisoners* per thousand prisoners on Q, K, L controls and year FE.

Table B6: **Transition matrix of self-violence**

	Self-violence				
	1	2	3	4	5
1	36.84	26.32	15.79	10.53	10.53
2	21.05	21.05	21.05	26.32	10.53
3	15.79	10.53	21.05	26.32	26.32
4	31.58	21.05	15.79	10.53	21.05
5	0.00	21.05	21.05	26.32	31.58

Notes: Transition matrices are computed as the probability for a prison at the q^{th} quintile in 2014-2015 to be in the r^{th} quintile in 2018-2019 of the distribution of residuals, where the residuals are computed from the regression of the level of *Self-violence* per thousand prisoners on Q, K, L controls and year FE.

B.3 Prison Managers

Governors play a vital role in the effective running of a prison, working in highly complex and unpredictable environments while managing vulnerable people. According to the National Careers Service, the governors' day-to-day tasks include "managing prison security, standards, and budgets, supervising prisoners and making inspections, supporting vulnerable prisoners and those at risk of self-harm, carrying out disciplinary procedures, managing and motivating prison staff, updating records and write reports, developing the prison to meet government targets, working with other professionals, such as medical staff, social workers and probation officers". However, according to the Justice Committee, there is a lack of clarity over the extent to which governors have full control over their prisons: their control on capital expenditure is limited and facilities management contracts are national arrangements.

Prison governors work typically from 37 to 41 hours per week, under an annual salary ranging from £47,500 to £75,500. There are two hiring pathways: (1) Start as a prison officer, gain several years of experience, including line management, and progressively move into senior management positions, eventually reaching the governor role; (2) Apply through a management training scheme, such as the Senior Leadership Scheme or the Leadership Development Programme.

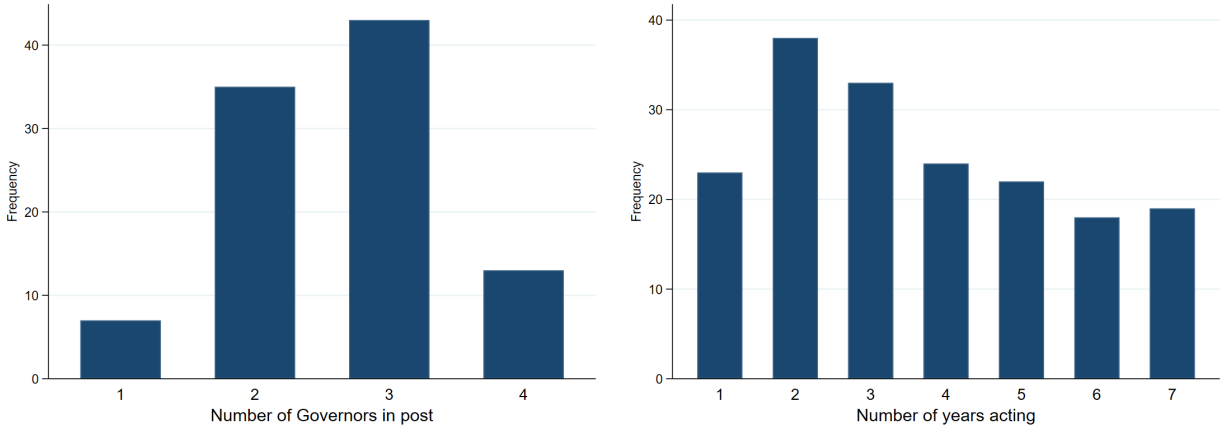
Over the period 2013-2019, among the 98 British prisons we observe, 177 distinct governors acted. 97% of them acted at more than one prison over the sample period, and 36% at exactly two prisons. The governors are predominantly males (69%) and work in male prisons - 95% of the male governors work in male prisons, while 83% of the female governors work in male prisons.

Our analysis reveals substantial turnover in prison leadership. Figure B7 panel (a) shows the number of prisons with one, two, three, or four governors in post between 2013 and 2019. 43 prisons out of 98 changed three governors in 7 years. The average length of service for governors in the given prison was 2.6 years. Figure B7 panel (b) plots the number of governors by years in service as governor.

The impact of such high turnover on prison performance is complex. Leadership transitions can serve strategic purposes, such as deploying effective governors to struggling facilities. However, when governors depart, their initiatives and successful practices may not be sustained by their successors. Furthermore, a governor's effectiveness in one prison does not guarantee success in another. Recognizing these concerns, the House of Commons Justice Committee recommended that "HMPPS should work on the principle that where possible governors remain at one prison for at least five years before being moved to other parts of the Service" and "review incentive structures to see what more can be done to incentivize

Figure B7: **Governors and prisons**

(a) Number of prisons by number of governors in post (b) Number of governors by number of years acting



Notes: The left panel plots the histogram of the number of prisons by number of governors in post between 2013 and 2019. The right panel shows the histogram of the number of governors by number of years acting between 2013 and 2019.

individuals to stay in the Service" (*Prison Governance* 2019).

We conduct a loss-of-fit exercise to quantify the relative importance of governors' fixed effects in driving variation in within-prison violence. Specifically, we regress *violence on staff*, *violence on other prisoners*, and *self-violence* on Q, K, L controls, as well as year, prison, and governor fixed effects, then measure the reduction in the R^2 when each set of controls (Q, K, L, and governors' fixed effects) is removed individually. Table B7 presents the results. We find that governors' fixed effects explain over 8% (3%, 0.2%) of the variation in *violence on staff* (*violence on other prisoners*, *self-violence*), versus 0.23% (2.9%, 0.3%) explained by measures of inmates' composition, physical and human capital.

These findings suggest that prison governors play a crucial role in shaping prisoner-on-staff and prisoner-on-prisoner violence. The relatively limited explanatory power of governors' fixed effects for self-violence may reflect the deeper psychological roots of that behaviour, likely less susceptible to managerial practices of top managers, who operate at a distance from the day-to-day life of detainees.

In the Table B16 we replicate the loss-of-fit exercise for the *hard* measures of violence, i.e. *serious staff assaults*, *serious prisoner assaults* and *homicides*, and *suicides* and *self-harm requiring hospitalization*, and show that the results are robust to the use of measures less likely to be under-reported.

However, our data does not allow us to disentangle the mechanisms through which governors influence the violence levels. Further research could explore whether this effect arises

Table B7: **Decomposition of violence based on loss of fit**

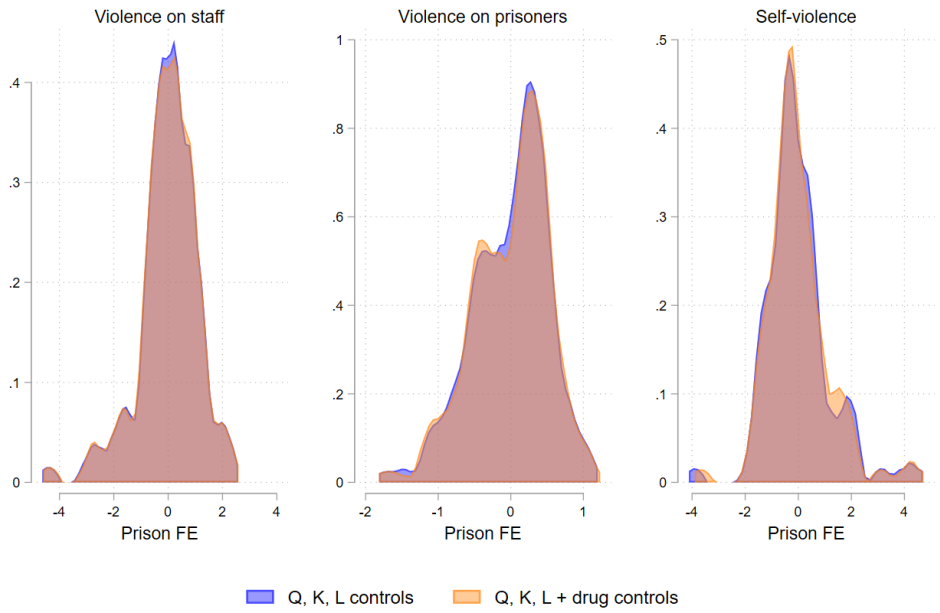
Decomposition based on loss of fit (% of R^2)			
Outcome:	Violence on staff	Violence on prisoners	Self-violence
Inmates' composition Q	0.6	2.3	0.0
Physical Capital K	0.8	0.1	0.0
Human Capital L	0.9	0.5	0.3
Governor	8.2	3.7	0.2
R^2	0.93	0.96	0.99
Adj R^2	0.93	0.96	0.99
Obs	592	592	592

Notes: Each entry reflects the loss of fit resulting from removing the controls in the rows from a linear projection of each outcome on the set of controls for inamtes' composition Q, physical capital K, human capital L, and governors' fixed effects, while controlling for year fixed effects and prison fixed effects, for each violence measure. The loss of fit is computed as the share over the fit in the full model: $\frac{(R^2_I - R^2_{I-1})}{R^2_{I-1}}$. The outcomes are measured in number of incidents per thousand inmates.

from the policies they implement, their oversight of officers, or the institutional culture they foster.

B.4 Robustness

Figure B8: **Density of Prison Fixed Effects controlling for drugs**



Notes: The figure plots the density of the prison fixed effects, estimated from equation 1, with and without the *Drug control*, with controls for inmates' quality (Q), physical capital (K), and human capital (L), across the three types of violence: violence on staff (left panel), violence on prisoners (center panel), and self-violence (right panel). The outcomes of interest are measured in number of incidents per thousands of prisoners. *Drug control* refers to the number of drug finds per thousand prisoners. The estimated Prison Fixed Effects for the *Q, K, L controls* specification are the same as the conditional estimates in Figure 3.

Table B8: **IQ of Prison Fixed Effects controlling for drugs**

	Violence on staff	Violence on prisoners	Self-violence
Mean	97.99	260.31	627.99
IQ range - Q, K, L controls	1.17	0.74	1.12
IQ range - Q, K, L + drug controls	1.18	0.73	1.22

Notes: The table shows the Interquartile Range (IQ) of the prison fixed effects, estimated with and without *Drug control*, with controls for inmates' quality (Q), physical capital (K), and human capital (L), for each violence measure. The outcomes of interest are measured in the number of incidents per thousand of prisoners. The Interquantile Range is divided by the mean of the outcome measure. *Drug control* refers to the number of drug finds per thousand prisoners. The estimates reported here for *IQ range - Q, K, L controls* are computed as in Table 4.

Table B9: **IQ of Prison Fixed Effects controlling for crime**

	Violence on staff	Violence on prisoners	Self-violence
Mean	97.99	260.31	627.99
IQ range - Q, K, L controls	1.16	0.71	1.13
IQ range - Q, K, L + market index of crime controls	1.18	0.66	1.17

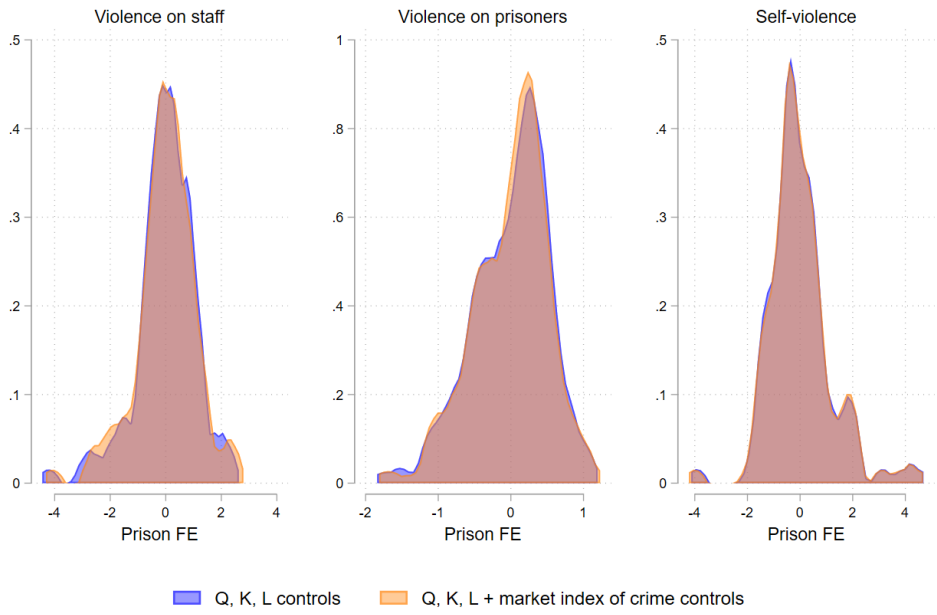
Notes: The table shows the Interquartile Range (IQ) of the prison fixed effects, estimated with and without *market index of crime controls*, with controls for inmates' quality (Q), physical capital (K), and human capital (L), for each violence measure. The outcomes of interest are measured in the number of incidents per thousand of prisoners. The Interquantile Range (IQ) is divided by the mean of the outcome measure. *Market index of crime* refers to the average constituency-level crime rate for violent crime and drugs for each prison, with weights inversely proportional to the distance constituency-prison. The estimates reported here for *IQ range - Q, K, L controls* are the same as in Table 4.

Table B10: **IQ of Prison Fixed Effects**

	Prison Fixed Effects					
	Violence on staff		Violence on prisoners		Self-violence	
	All	Hard	All	Hard	All	Hard
Mean	97.99	9.62	260.31	30.65	627.99	36.61
IQ range - No controls	0.83	1.06	0.87	0.92	0.62	0.85
IQ range - Q, K, L controls	1.16	1.56	0.71	1.02	1.13	1.41

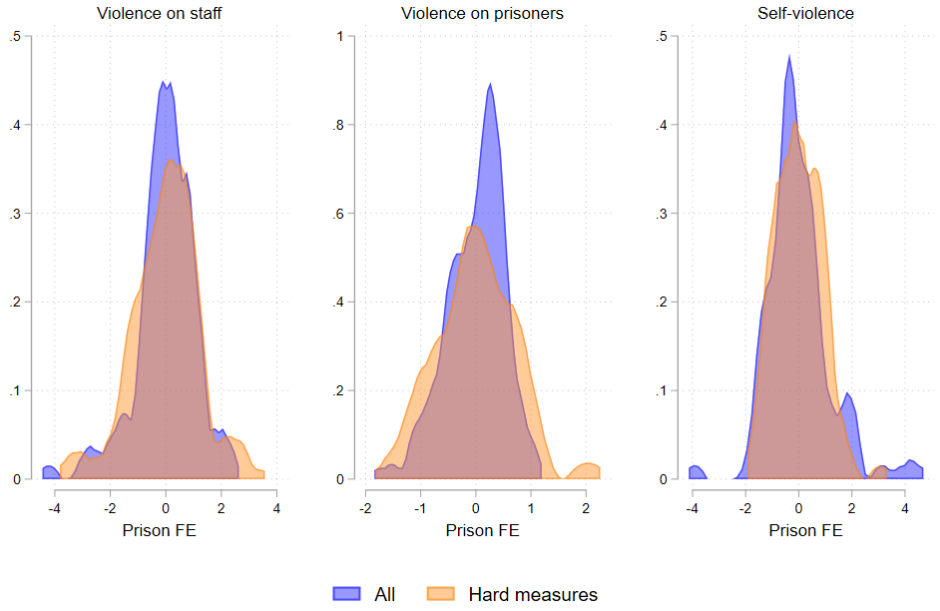
Notes: The table shows the Interquartile Range (IQ) of the prison fixed effects, estimated with and without controls for inmates' quality (Q), physical capital (K), and human capital (L), on *All* violence versus *Hard measures* of violence across the thee types of violence. The outcomes of interest are measured in number of incidents per thousands of prisoners. The Interquartile Range (IQ) is divided by the mean of the outcome measure. *Hard measures* include serious assaults on staff within the *Violence on staff*, serious assaults on prisoners and homicides within the *Violence on prisoners*, and suicides and self-harm incidents that required hospitalization within the *Self-violence*. The estimates reported here for *IQ range - No controls* and *IQ range - Q, K, L controls* for *All* violence are the same as in Table 4.

Figure B9: **Density of Prison Fixed Effects controlling for crime**



Notes: The figure plots the density of the prison fixed effects, estimated from equation 1, with and without the *Market index of crime controls*, with controls for inmates' quality (Q), physical capital (K), and human capital (L), across the three types of violence: violence on staff (left panel), violence on prisoners (center panel), and self-violence (right panel). The outcomes of interest are measured in number of incidents per thousands of prisoners. *Market index of crime* refers to the average constituency-level crime rate for violent crime and drugs for each prison, with weights inversely proportional to the distance constituency-prison. The estimated Prison Fixed Effects for the *Q, K, L controls* specification are the same as the conditional estimates in Figure 3.

Figure B10: **Density of Prison Fixed Effects**



Notes: The figure plots the density of the prison fixed effects, estimated from equation 1, with controls for inmates' quality (Q), physical capital (K), and human capital (L), on *All* violence versus *Hard measures* of violence across the three types of violence: violence on staff (left panel), violence on prisoners (center panel), and self-violence (right panel). The outcomes of interest are measured in number of incidents per thousands of prisoners. *Hard measures* include serious assaults on staff within the *Violence on staff*, serious assaults on prisoners and homicides within the *Violence on prisoners*, and suicides and self-harm incidents that required hospitalization within the *Self-violence*. Prison Fixed Effects are estimated in the conditional regression. The estimated Prison Fixed Effects for *All* violence are the same as the conditional estimates in Figure 3.

Table B11: **Decomposition of Hard violence based on loss of fit**

Decomposition based on loss of fit (% of R^2)			
Outcome:	Hard violence on staff	Hard violence on prisoners	Hard self-violence
Prison function	27.1	29.0	54.5
Private dummy	3.4	12.9	0.0
Male dummy	51.2	12.2	16.8
Victorian dummy	5.9	7.3	0.0
Area	47.1	0.6	27.5
R^2	0.40	0.65	0.48
Adj R^2	0.33	0.61	0.43
Obs	108	108	108

Notes: Each entry reflects the loss of fit resulting from removing the controls in the rows from a linear projection of each outcome on the set of fixed effects in the specification controlling for {Prison Function, Private prison dummy, Victorian prison dummy, Prison area in squared meters}. The loss of fit is computed as the share over the fit in the full model: $\frac{(R^2_I - R^2_{I-1})}{R^2_{I-1}}$. The outcomes are measured in number of incidents per thousand inmates. *Hard violence on staff* includes serious assaults on staff, *Hard violence on prisoners* serious assaults on prisoners and homicides, and *Self-violence* includes self-harm requiring hospitalization and suicides.

Table B12: **Comparison to PPDs estimates**

Residuals	Interquantile range (75^{th}p-25^{th}p)	Interdecile range (90^{th}p-10^{th}p)
Hard violence on staff	0.357	0.703
Hard violence on prisoners	0.350	0.693
Hard self-violence	0.424	0.820
Syverson, 2004 - Concrete market	0.290	0.651
Chandra et al., 2013 - Hospitals	0.230	0.440

Notes: The table shows the estimated Interquartile range and Interdecile range of the residuals from equation 2 for each *Hard* measure of violence. The numbers shown are means of the annual values from 2013 to 2019 weighting each year equally. The Syverson (2004) statistics refer to the across-industry mean of the within-industry productivity moment. *Hard measures* include serious assaults on staff within the *Violence on staff*, serious assaults on prisoners and homicides within the *Violence on prisoners*, and suicides and self-harm incidents that required hospitalization within the *Self-violence*.

Table B13: **AR(1) of violence residuals**

	Residuals					
	No drug finds control			Drug finds control		
Violence on staff	0.063			0.063		
	(0.070)			(0.070)		
Violence on prisoners		-0.065			-0.064	
		(0.061)			(0.062)	
Self-violence			0.213*			0.196*
			(0.114)			(0.117)
Observations	376	376	376	376	376	376

*** p<0.01, ** p<0.05, * p<0.1

Notes: The table shows the estimates from the OLS regression of contemporaneous violence residuals from equation 1, controlling for *drug finds*, on the one-year-lagged residuals with NeweyWest standard errors robust to heteroskedasticity and first-order autocorrelation for *violence on staff*, *violence on prisoners*, and *self-violence*. *No drug finds control* refers to the residuals being obtained from the regression of violence measures on the set of fixed effects and Q, K, L controls, as in equation 1, while *Drug finds control* to the residuals from the same specification while controlling also for the number of drug finds per thousand prisoners.

Table B14: **AR(1) of violence residuals**

	Residuals					
	No market index of crime control			Market index of crime control		
Violence on staff	0.066			0.044		
	(0.070)			(0.071)		
Violence on prisoners		-0.060			-0.060	
		(0.061)			(0.061)	
Self-violence			0.212*			0.212*
			(0.114)			(0.114)
Observations	380	380	380	380	380	380

*** p<0.01, ** p<0.05, * p<0.1

Notes: The table shows the estimates from the OLS regression of contemporaneous violence residuals from equation 1, controlling for *drug finds*, on the one-year-lagged residuals with NeweyWest standard errors robust to heteroskedasticity and first-order autocorrelation for *violence on staff*, *violence on prisoners*, and *self-violence*. *No market index of crime controls* refers to the residuals being obtained from the regression of violence measures on the set of fixed effects and Q, K, L controls, as in equation 1, while *Market index of crime controls* to the residuals from the same specification while controlling also for the *Market index of crime*.

Table B15: **AR(1) of violence residuals**

	Residuals					
	Unconditionally			Conditionally		
Hard violence on staff	0.179*			-0.148**		
	(0.092)			(0.073)		
Hard violence on prisoners		0.153*			-0.168***	
		(0.080)			(0.057)	
Hard self-violence			0.093			-0.017
			(0.078)			(0.066)
Observations	452	452	452	380	380	380

*** p<0.01, ** p<0.05, * p<0.1

Notes: The table shows the estimates from the OLS regression of contemporaneous violence residuals from equation 1 on the one-year-lagged residuals with NeweyWest standard errors robust to heteroskedasticity and first-order autocorrelation for *Hard* measures of *violence on staff*, *violence on prisoners*, and *self-violence*. *Unconditionally* refers to the residuals being obtained from the regression of violence measures on the set of fixed effects only, while *Conditionally* to the residuals from the full specification, as in equation 1. *Hard measures* include serious assaults on staff within the *Violence on staff*, serious assaults on prisoners and homicides within the *Violence on prisoners*, and suicides and self-harm incidents that required hospitalization within the *Self-violence*.

Table B16: **Decomposition of hard violence based on loss of fit**

Decomposition based on loss of fit (% of R^2)			
Outcome:	Hard violence on staff	Hard violence on prisoners	Hard self-violence
Inmates' composition Q	2.2	2.9	0.5
Physical Capital K	0.6	0.1	0.0
Human Capital L	1.4	1.0	0.6
Governor	7.9	7.1	2.6
R^2	0.84	0.90	0.7
Adj R^2	0.71	0.81	0.95
Obs	592	592	592

Notes: Each entry reflects the loss of fit resulting from removing the controls in the rows from a linear projection of each outcome on the set of controls for inmates' composition Q, physical capital K, human capital L, and governors' fixed effects, while controlling for year fixed effects and prison fixed effects, for each violence measure. The loss of fit is computed as the share over the fit in the full model: $\frac{(R^2_I - R^2_{I-1})}{R^2_{I-1}}$. The outcomes are measured in number of incidents per thousand inmates. *Hard measures* include serious assaults on staff within the *Violence on staff*, serious assaults on prisoners and homicides within the *Violence on prisoners*, and suicides and self-harm incidents that required hospitalization within the *Self-violence*.

Appendix C: Inspections as Management Interventions

C.1 Further Results

Table C1: **Inspection on violence by inspection length**

	Total violence	Violence on staff	Violence on prisoners	Self-violence
Inspected	0.508***	0.496***	0.449***	0.171*
	(0.127)	(0.119)	(0.136)	(0.102)
Inspected × Long inspection	-0.062	-0.057	-0.086	0.006
	(0.075)	(0.066)	(0.084)	(0.086)
Post × Inspected	-0.259**	-0.162	-0.269**	-0.138
	(0.105)	(0.114)	(0.110)	(0.111)
Post × Inspected × Long inspection	0.081	0.072	0.071	0.035
	(0.061)	(0.069)	(0.056)	(0.051)
Observations	21390	21390	21390	21390
R-squared	0.282	0.252	0.247	0.290
Sub-experiment x Period FE	Yes	Yes	Yes	Yes
Time since last inspection	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1
SE clustered at the prison level.

Notes: The table reports the estimate of the static stacked triple difference-in-difference specification for heterogeneity in the inspection type, *short* versus *long* inspections. The violence outcomes are measured in number of incidents per thousands of prisoners and standardized. All the specifications include the sub-experiment x year-month fixed effects and the control for the months elapsed since the previous inspection. SE clustered at the prison level. Weights of the stacked panel are computed following Wing et al. (2024).

Table C2: **Inspection on violence by inspection announcement**

	Total violence	Violence on staff	Violence on prisoners	Self-violence
Inspected	0.506***	0.485***	0.428***	0.198***
	(0.095)	(0.094)	(0.095)	(0.072)
Inspected x Unannounced inspection	-0.305***	-0.229***	-0.316***	-0.124
	(0.086)	(0.081)	(0.098)	(0.086)
Post x Inspected	-0.213***	-0.114	-0.233***	-0.121
	(0.078)	(0.083)	(0.088)	(0.087)
Post x Inspected x Unannounced inspection	0.038	0.014	0.041	0.028
	(0.051)	(0.061)	(0.041)	(0.045)
Observations	21390	21390	21390	21390
R-squared	0.292	0.258	0.256	0.291
Sub-experiment x Period FE	Yes	Yes	Yes	Yes
Time since last inspection	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1
SE clustered at the prison level.

Notes: The table reports the estimate of the static stacked triple difference-in-difference specification for heterogeneity in the inspection type, *unannounced* versus *announced* inspections. The violence outcomes are measured in number of incidents per thousands of prisoners and standardized. All the specifications include the sub-experiment x year-month fixed effects and the control for the months elapsed since the previous inspection. SE clustered at the prison level. Weights of the stacked panel are computed following Wing et al. (2024).

C.2 Robustness

Table C3: **Inspections on violence by prison gender**

	Total violence	Violence on staff	Violence on prisoners	Self-violence
Post x Inspected	-0.180**	-0.104	-0.219**	-0.072
	(0.076)	(0.083)	(0.085)	(0.046)
Female prison	-0.207*	-0.285*	-0.157	-0.012
	(0.122)	(0.146)	(0.096)	(0.159)
Post x Inspected x Female prison	-0.108	0.009	0.084	-0.330*
	(0.111)	(0.095)	(0.055)	(0.184)
Observations	21390	21390	21390	21390
R-squared	0.312	0.255	0.258	0.575
Sub-experiment x Period FE	Yes	Yes	Yes	Yes
Time since last inspection	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1
SE clustered at the prison level.

Notes: The table reports the estimate of the stacked triple difference-in-difference specification for heterogeneity in the prison gender. The violence outcomes are measured in number of incidents per thousands of prisoners and standardized. All the specifications include the sub-experiment x year-month fixed effects and the control for the months elapsed since the previous inspection. SE clustered at the prison level. Weights of the stacked panel are computed following Wing et al. (2024).

Table C4: **Inspections on violence by establishment security level**

	Total violence	Violence on staff	Violence on prisoners	Self-violence
Inspected	1.216***	1.160***	1.391***	0.121
	(0.285)	(0.259)	(0.328)	(0.164)
Post x Inspected	0.024	0.114	-0.101	0.040
	(0.084)	(0.114)	(0.085)	(0.061)
High security	0.762***	0.739***	0.389***	0.547***
	(0.172)	(0.226)	(0.119)	(0.107)
Medium security	0.380***	0.252**	0.213**	0.370***
	(0.091)	(0.097)	(0.089)	(0.095)
Post x Inspected x High security	-0.291***	-0.300	-0.200**	-0.139
	(0.090)	(0.194)	(0.083)	(0.143)
Post x Inspected x Medium security	-0.282***	-0.269**	-0.121	-0.230***
	(0.087)	(0.117)	(0.079)	(0.061)
Observations	21390	21390	21390	21390
R-squared	0.331	0.297	0.355	0.307
Sub-experiment x Period FE	Yes	Yes	Yes	Yes
Time since last inspection	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1
SE clustered at the prison level.

Notes: The table reports the estimate of the stacked triple difference-in-difference specification for heterogeneity by the establishment security level. We classify as *High security* prisons Category A and Category B prisons, as *Medium security* prisons Category C and reception, and as *Low security* prisons Category D and YJB establishments. The number of prisons in our stacked panel are 25, 97 and 24, respectively for *High security*, *Medium security*, and *Low security* prisons. The violence outcomes are measured in number of incidents per thousands of prisoners and standardized. All the specifications include the sub-experiment x year-month fixed effects and the control for the months elapsed since the previous inspection. SE clustered at the prison level. Weights of the stacked panel are computed following Wing et al. (2024).

Table C5: **Inspections on violence & Prison leadership**

	Total violence		Violence on staff		Violence on prisoners		Self-violence	
Inspected	0.530***	0.322***	0.521***	0.296***	0.427***	0.280***	0.217***	0.131***
	(0.095)	(0.076)	(0.093)	(0.086)	(0.093)	(0.076)	(0.073)	(0.041)
Post x Inspected	-0.222**	-0.096*	-0.141	-0.057	-0.245**	-0.099**	-0.101	-0.055
	(0.088)	(0.055)	(0.094)	(0.071)	(0.098)	(0.049)	(0.065)	(0.034)
Observations	17450	17450	17450	17450	17450	17450	17450	17450
R-squared	0.315	0.747	0.284	0.628	0.276	0.765	0.233	0.737
Sub-experiment x Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Governor FE	No	Yes	No	Yes	No	Yes	No	Yes
Time since last inspection	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1
SE clustered at the prison level.

Notes: The table reports the estimate of the stacked difference-in-difference specification on violence outcomes with and without controls for governor fixed effects. In the stacked sample we observe 236 governors in 140 prisons. Note that the baseline in columns (1)-(3)-(5)-(7) is estimated on the subsample for which governors' names are available. The violence outcomes are measured in number of incidents per thousands of prisoners and standardized. All the specifications include the sub-experiment x year-month fixed effects and the control for the months elapsed since the previous inspection. SE clustered at the prison level. Weights of the stacked panel are computed following Wing et al. (2024).

Table C6: **Inspections on drug detection and deterrence**

	Drug finds	Drug tests	Share positive drug tests
Post x Inspected	0.185**	0.147*	0.182**
	(0.074)	(0.088)	(0.081)
Observations	20485	20176	20176
R-squared	0.268	0.171	0.261
Sub-experiment x Period FE	Yes	Yes	Yes
Time since last inspection	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1
SE clustered at the prison level.

Notes: The figure plots the dynamic estimated of the stacked difference-in-difference specification on drug outcomes, *Drug finds* in the top left panel, *Random Mandatory drug testing* in the top right panel, and *Share of positive Random Mandatory drug testin* in the bottom left panel. The labels on the x-axis refer to quarters relative to the inspection month, i.e., $t = 0$ refers to the three months following the inspection. The drug finds and the drug tests are measured in number of incidents per thousands of prisoners and standardized. All the specifications include the sub-experiment x year-month fixed effects and the control for the months elapsed since the previous inspection. SE are clustered at the prison level. Weights of the stacked panel are computed following Wing et al. (2024).

Table C7: **Inspections on violence & drug finds**

	Total violence		Violence on staff		Violence on prisoners		Self-violence	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post x Inspected	-0.204***	-0.198***	-0.117	-0.115	-0.223**	-0.216***	-0.108	-0.104
	(0.075)	(0.073)	(0.080)	(0.078)	(0.086)	(0.082)	(0.078)	(0.074)
Drug finds		-0.033		-0.013		-0.037		-0.022
		(0.042)		(0.037)		(0.037)		(0.052)
Observations	20485	20485	20485	20485	20485	20485	20485	20485
R-squared	0.287	0.287	0.263	0.263	0.264	0.265	0.209	0.210
Sub-experiment x Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time since last inspection	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

Notes: The figure plots the dynamic estimated of the stacked difference-in-difference specification on violence outcomes, controlling for *Drug finds* in columns (2)-(4)-(6)-(8). The baseline in columns (1)-(3)-(5)-(7), without controls for drug finds, are estimated on the sample for which drug finds data are available. The violence outcomes and the drug finds are measured in number of incidents per thousands of prisoners and standardized. Drug finds refer to month-level number of incidents in which drugs were found. All the specifications include the sub-experiment x year-month fixed effects and the control for the months elapsed since the previous inspection. SE clustered at the prison level. Weights of the stacked panel are computed following Wing et al. (2024).

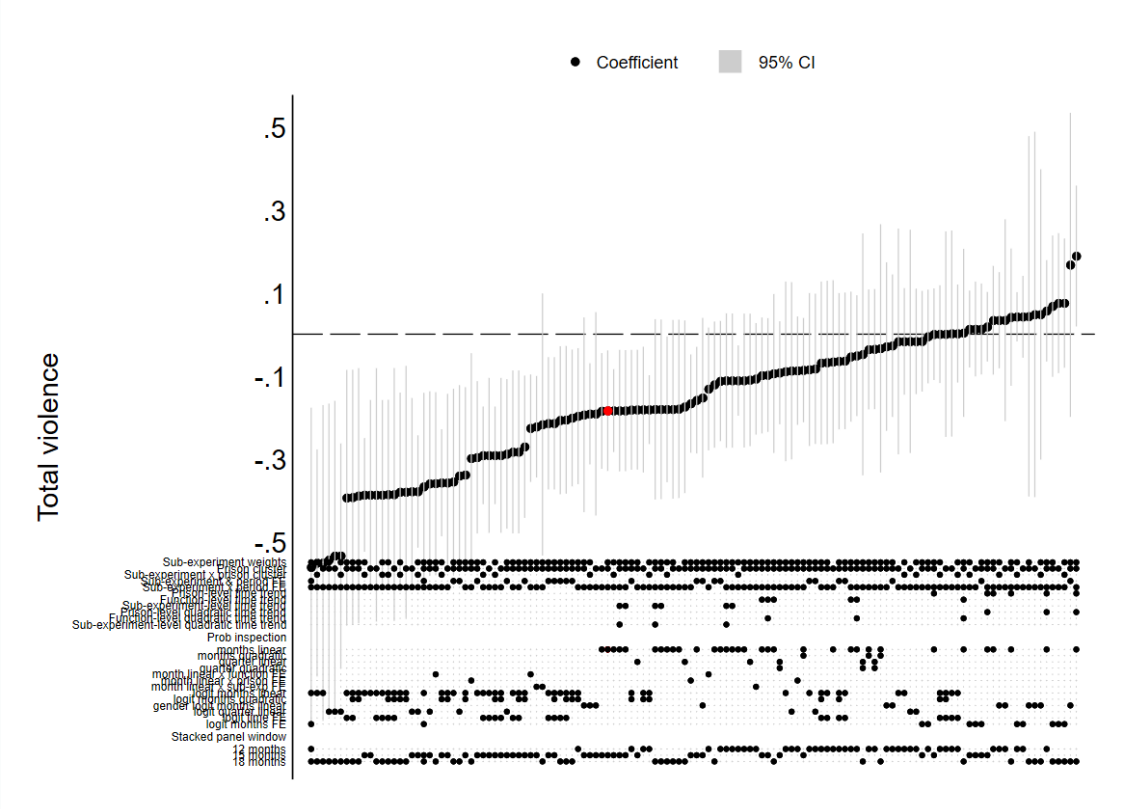
Table C8: **Inspections on violence & drug tests**

	Total violence		Violence on staff		Violence on prisoners		Self-violence	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post x Inspected	-0.231***	-0.251***	-0.141*	-0.147*	-0.227***	-0.243***	-0.108	-0.164**
	(0.076)	(0.076)	(0.083)	(0.083)	(0.082)	(0.086)	(0.078)	(0.083)
Number of tests		0.137		0.037		0.110		0.152
		(0.101)		(0.093)		(0.121)		(0.098)
Observations	20176	20176	20176	20176	20176	20176	20485	19772
R-squared	0.265	0.280	0.248	0.249	0.244	0.253	0.209	0.224
Sub-experiment x Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time since last inspection	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

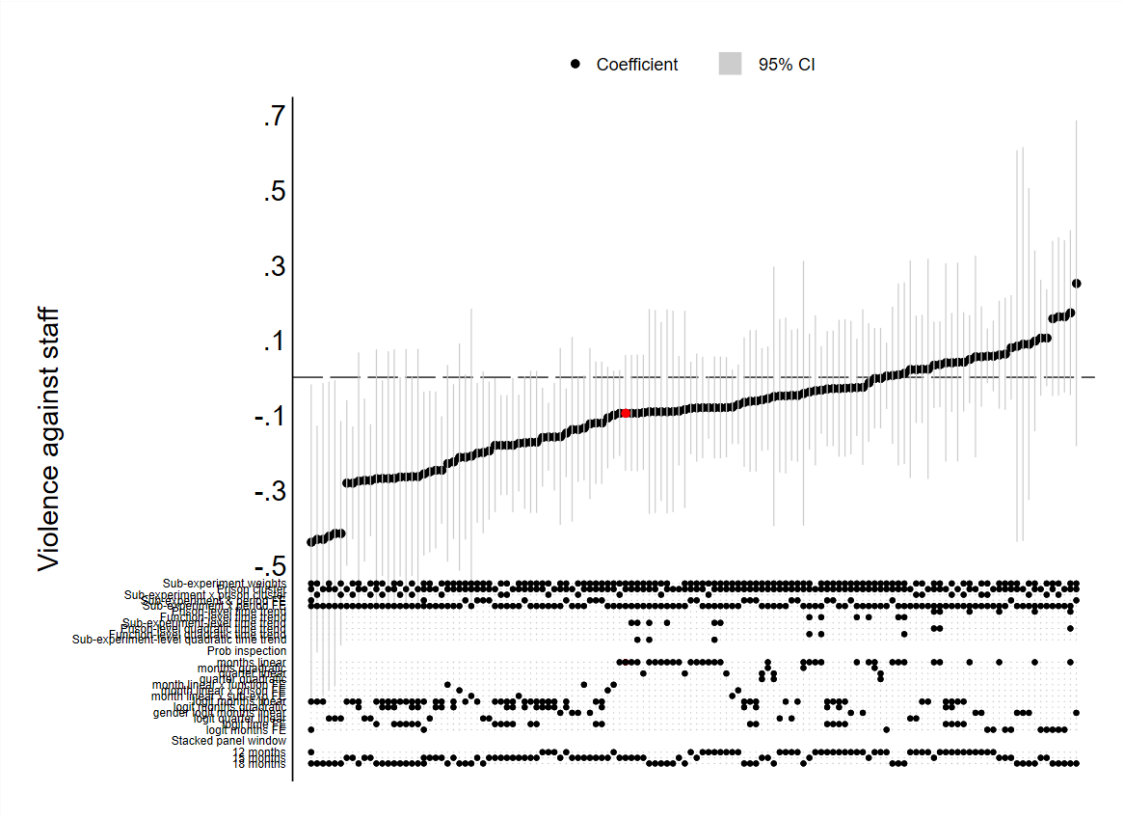
Notes: The figure plots the dynamic estimated of the stacked difference-in-difference specification on violence outcomes, controlling for *Number of tests* in columns (2)-(4)-(6)-(8). The baseline in columns (1)-(3)-(5)-(7), without controls for drug tests, are estimated on the sample for which drug finds data are available. *Drug tests* refer to the number of Random Mandatory drug tests administered each month. The violence outcomes and the drug tests are measured in number of incidents per thousands of prisoners and standardized. All the specifications include the sub-experiment x year-month fixed effects and the control for the months elapsed since the previous inspection. SE clustered at the prison level. Weights of the stacked panel are computed following Wing et al. (2024).

Figure C1: **Alternative specifications**



Notes: The figure plots the difference-in-difference coefficient estimated from 158 different specifications. The coefficient in red is from our baseline specification. Each specification includes different fixed effects, controls for the time elapsed since the last inspection, or is estimated in the stacked panel assembled with a different window. *Months linear* refers to a linear control in the months since the last inspection; *Months quadratic* to a quadratic control in the months since the last inspection; *Sub-experiment & period FE* and *Sub-experiment x period FE* include fixed effects for the sub-experiment and period, separately or interacted. *Prob inspection* denotes all the specifications in which we control for the probability of inspection estimated in a logit regression of the inspection occurrence on different forms of controls for the time since last inspection: linear control for the month (*logit month linear*), linear and quadratic control for the month (*logit month quadratic*), linear control for the month interacted with the prison gender FE(*gender logit month linear*), linear control for the quarter (*logit quarter linear*), and months fixed effects (*logit months FE*). Moreover, we estimate all the specifications mentioned above in three stacked panels, built with $\kappa_{pre} = 12$, $\kappa_{pre} = 15$, and $\kappa_{pre} = 18$.

Figure C2: **Alternative specifications**



Notes: Footnotes from Table C1 also apply here.

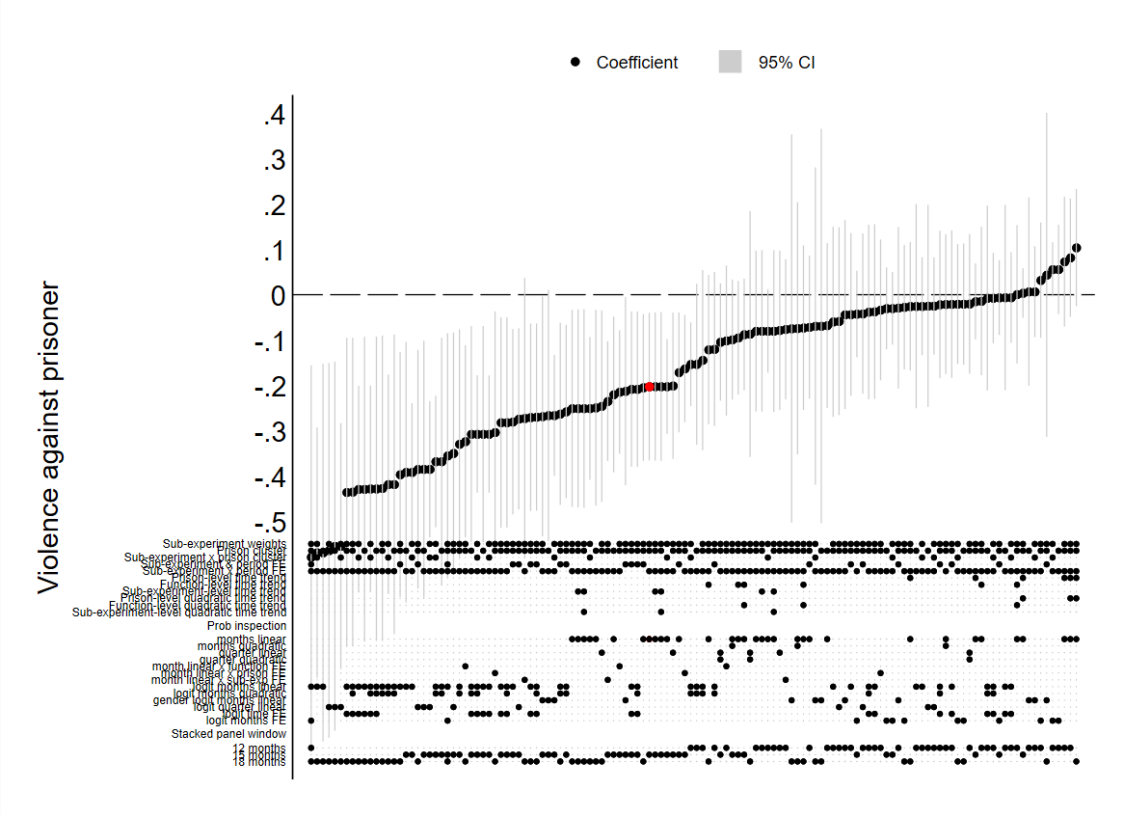
Table C9: **Count data regression**

	Total violence	Violence on staff	Violence on prisoners	Self-violence
Inspected	0.783***	1.144***	1.114***	0.597**
	(0.235)	(0.266)	(0.227)	(0.257)
Post x Inspected	-0.313**	-0.333	-0.345**	-0.283*
	(0.142)	(0.230)	(0.174)	(0.162)
Observations	21390	21311	21374	21390
R-squared				
Outcome mean	28.71	2.33	8.64	17.75
Sub-experiment x year-month FE	Yes	Yes	Yes	Yes
Month since last	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1
SE clustered at the prison-level.

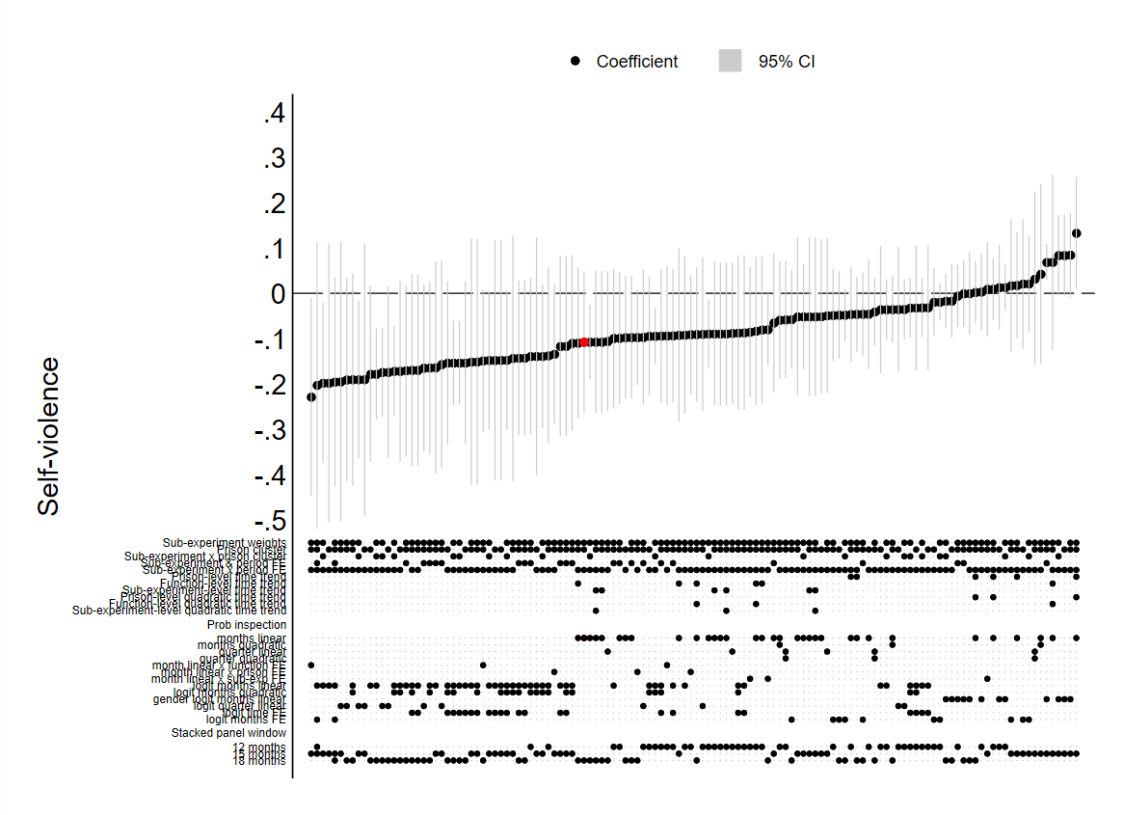
Notes: The table reports the estimate of the static stacked poisson difference-in-difference specification on violence outcomes, *Total violence* in column (1), *Violence on staff* in column (2), *Violence on prisoners* in column (3), and *Self-violence* in column (4). The coefficients included in the Table are 1 − Incidence Rate Ratio (IRR). The violence outcomes are measured in number of incidents per thousands of prisoners and standardized. All the specifications include the sub-experiment x year-month fixed effects and the control for the months elapsed since the previous inspection. SE clustered at the prison level. Weights of the stacked panel are computed following Wing et al. (2024).

Figure C3: **Alternative specifications**



Notes: Footnotes from Table C1 also apply here.

Figure C4: **Alternative specifications**



Notes: Footnotes from Table C1 also apply here.