# Standing in Prisoners' Shoes: A Randomized Trial on How Incarceration Shapes Criminal Justice Preferences

Arto Arman, Andreas Beerli, Aljosha Henkel, Michel André Maréchal*

January 31, 2026

**Abstract**

We study how incarceration experience shapes preferences for criminal justice policies. In collaboration with a newly opened prison, we conducted a randomized field experiment that offered citizens the opportunity to experience up to two days of incarceration, closely replicating the real-life journey of inmates. Providing citizens with a chance to gain firsthand incarceration leads to a significant shift in punitive attitudes, with participants becoming less supportive of harsh criminal justice policies and donating more money to organizations advocating more moderate justice policies. Although individuals overestimated the wellbeing of actual prisoners, the intervention did not alter these beliefs. This suggests that the observed changes in policy preferences are driven more by personal experience than by revised beliefs about the burden of confinement. By randomizing institutional exposure outside the laboratory, our study highlights the causal role of personal experience in the formation of policy preferences.

**Key Words**: Incarceration, Field Experiment, Personal Experience, Criminal Justice Policy, Punitive Attitudes, Prison

**JEL Codes**: C93, D83, K14, P37

*Arman: KOF Swiss Economic Institute, ETH Zurich (arman@kof.ethz.ch). Beerli: KOF Swiss Economic Institute, ETH Zurich (beerli@kof.ethz.ch). Henkel: KOF Swiss Economic Institute, ETH Zurich (henkel@kof.ethz.ch). Maréchal: Department of Economics, University of Zurich (michel.marechal@econ.uzh.ch). The study was pre-registered at the AEA RCT Registry (AEARCTR-0009150) and was reviewed and approved by the Human Subjects Committee of the Faculty of Economics, Business Administration, and Information Technology at the University of Zurich (OEC IRB #2022-013). We thank the prison management of the Prison Zurich West and the Department of Justice and Home Affairs of the Canton of Zurich for the cooperation. We greatly appreciate research assistance provided by Ermir Binakaj, Gisela Fonseca, Benjamin Huber, Finan Marugg, Morris Niklaus, and Domenico Sorrentino. We thank Alain Cohn, Jennifer Doleac, Uri Gneezy, Dominik Hangartner, Claire Hill, Matt Lowe, Wladislaw Mill, Joshua B. Miller, Theo Offerman, Arnaud Philippe, Christopher Roth, Sally Sadoff, Julia Seither, Alex Sutherland, Santiago Tobon, Jeroen van de Ven, David Yanagizawa-Drott and seminar and workshop participants at ETH Zurich, University of Zurich, UCSD Rady School of Management, SNoPE ETH Zurich, IMEBESS Riga, Universidad del Rosario Bogotá, ESA Bogotá, AFE London, WFEEB Heilbronn, NCBEE Copenhagen, WECB Freiburg, ViCE, WEC Jr. UPenn, Law and Rationality Workshop, CSEE Bonn, and Matterhorn Symposium Brig for their helpful comments.

# 1 Introduction

During the past few decades, many regions around the world have experienced a notable shift toward more punitive criminal justice policies. This trend is most evident in the widespread increase in mass incarceration, which has resulted in nearly 11 million confined individuals worldwide (Fair & Walmsley, 2021). While public opinion is considered a key factor in shaping criminal justice policies and judicial outcomes (Brace & Boyea, 2008; Pratt, 2007; Enns, 2016; Jennings et al., 2017), most individuals have very limited knowledge about the everyday realities of life behind bars (Bryant & Morris, 1998; Doble, 1995; Farkas, 1993; Hough & Roberts, 2005; Roberts & Hough, 2005). Prisons, by their very nature, remain largely inaccessible to citizens, leaving public perceptions confined to media portrayals and fiction (Cecil & Leitner, 2009; Wilson & O'Sullivan, 2004).This raises the question of whether public attitudes toward incarceration are based on an accurate understanding of prison life, and whether society's stance on criminal justice policies would remain unchanged if people experienced incarceration firsthand.

In this paper, we shed light on these questions by taking advantage of a unique randomized controlled trial that provided regular citizens with the opportunity to gain firsthand incarceration experience. The Department of Justice and Home Affairs of the Canton of Zurich in Switzerland organized a four-day test run in March 2022 for a newly built prison to test its operational and security procedures under simulated real-life conditions. Volunteering participants could experience up to 48 hours of incarceration, following standard procedures closely mirroring a real inmate's journey from intake to release, including strip searches, confiscation of all belongings, and a strictly regulated daily routine. Participants were locked in their prison cells for most of the day and were only able to leave those for a 10-minute shower and a 90-minute window reserved for a courtyard walk. In sum, the unique setup provided participants with an experience as close as possible to real incarceration.[1]

As approximately 750 interested volunteers exceeded the prison's capacity, we randomly assigned applicants into a treatment and a control group. The prison authorities were given the discretion to select any applicant from the treatment group and schedule a test run slot based on cell occupancy, the applicants' availability, gender, and smoking status. In contrast, we instructed prison authorities not to invite applicants from the control group to

[1] According to reports comparing prison conditions across developed countries, incarceration experience in Switzerland generally ranks toward the less severe to moderate end of the spectrum, but exhibits significant heterogeneity across various dimensions (e.g., World Health Organization: Regional Office for Europe, 2023; Aebi & Cocco, 2024). For further details, see Section 2.

participate in the test run. This design enables us to estimate the intention-to-treat effect of providing the opportunity to gain firsthand incarceration experience. We conducted surveys with applicants before and after the test run to gauge their *punitive attitudes* and *beliefs about wellbeing in prison*, using both questionnaire items and behavioral measures involving real financial consequences, designed to mitigate social desirability and experimenter demand effects (Bursztyn et al., 2025).

We document several findings. First, compared to the control group, subjects who received the chance to participate in the test run became substantially less supportive of punitive policies. Moreover, they donated more money to organizations advocating for more moderate, rather than a harsher, criminal justice policy. We observe a significant reduction in punitive justice attitudes, exceeding 0.3 standard deviations. This treatment effect corresponds to about 30% of the difference in punitive attitudes between the United States and Norway — two countries often regarded as representing opposite ends on the spectrum in terms of the punitiveness of their justice systems.[2] Since participation was voluntary, our subjects are not representative of the general population. They are more educated, express greater institutional trust, and include a greater share of people working in justice-related professions. Because these characteristics are common among groups that are politically engaged or in positions to influence the criminal justice system, they are a relevant population for examining responses to incarceration. Moreover, we examine the generalizability of our results by using data on demographics, preferences, and attitudes from an additional survey conducted with a sample representative of the Swiss voting population with respect to age, gender, and income, and applying entropy balancing to reweight the experimental sample. Our main findings remain robust after reweighting the sample to better reflect the characteristics of the general population.

Second, we observe that the public significantly overestimates the wellbeing of actual prisoners. However, when asked to predict their own wellbeing in prison, their estimates align more closely with the self-reported experiences of actual prisoners. This finding suggests that the public has an intuitive sense of how they might feel if incarcerated, despite lacking firsthand experience. The discrepancy between perceptions of actual prisoners' wellbeing and their own anticipated wellbeing in prison indicates that citizens may view prisoners as better equipped to cope with the challenges of incarceration than they themselves would be.

[2]The comparison of punitiveness is based on the percentage of the public favoring imprisonment for a recidivist burglar, as reported by Van Dijk et al. (2007) who used data from the 2005 International Crime and Victim Survey (ICVS) and the 2005 European Survey of Crime and Safety (EU ICS).

Third, the opportunity to participate in the test run did not significantly affect beliefs about wellbeing in prison, which is consistent with the nuanced nature of the public's misperceptions. Our analysis of secondary outcomes further reveals that we find no evidence that the intervention affected broader beliefs about the criminal justice system, including trust in legal institutions and perceptions of procedural fairness. Together with the null effects on beliefs about prisoner wellbeing, these findings suggest that the reduction in punitive attitudes is more plausibly due to personal experience with incarceration than to revised beliefs about prison conditions or institutional functioning. One way to interpret these results is through the lens of salience theory (Bordalo et al., 2020, 2022). When forming punitive justice attitudes, individuals evaluate multiple attributes of incarceration (i.e., different dimensions that people care about when judging punishment), such as its deterrent and rehabilitative functions, the wellbeing of inmates, and the fairness of justice institutions. According to salience theory, not all attributes receive equal attention when individuals form policy preferences. Consequently, punitive attitudes depend not only on individuals' beliefs about these attributes but also on the attentional weights they assign to each one. In our setting, firsthand incarceration made certain attributes more salient and increased their weights by directing attention to the lived experience of imprisonment. In a follow-up survey conducted about eleven months after the test run with a nonrandomized group of selected politicians, judges, and journalists (see Online Appendix C), respondents most frequently recalled experiences related to the loss of control and autonomy and to long waiting times and boredom, followed by the poor physical environment of the prison and the emotional burden of loneliness. These memories point to a shift in attentional weights: by making the emotional burden of imprisonment more prominent, the incarceration experience plausibly increased the weight placed on inmate wellbeing relative to other considerations such as deterrence or retribution. This shift could in turn have led to greater compassion and increased support for more moderate punishment, even without changing beliefs about prison wellbeing. This is consistent with our finding that the intervention affected punitive attitudes but not beliefs about prisoner wellbeing, procedural fairness, or trust in the justice system. At the same time, we cannot completely rule out more specific belief changes. For example, our belief measures may not fully capture changes in beliefs about the rehabilitative or retributive effects of imprisonment.

Our paper connects to several strands of the literature. First, a growing body of research in economics and psychology highlights the importance of personal experience (e.g., Foster & Rosenzweig, 1995; Malmendier, 2021; Hertwig & Erev, 2009). This literature has relied

almost exclusively either on non-experimental observational data or on laboratory settings.[3] Observational studies suggest that personal experiences can have long-lasting and profound effects on beliefs, preferences, and economic decisions. These effects span a wide range of outcomes such as inflation expectations, consumption behavior, attitudes toward redistribution, job preferences, and broader policy preferences (Giuliano & Spilimbergo, 2025; Malmendier & Wachter, 2024; Roth & Wohlfart, 2018; Cotofan et al., 2023; Washington, 2008; Alesina & Fuchs-Schündeln, 2007). In the context of crime, Philippe (2024) examines how criminals learn about legal changes, underscoring the significance of direct sentencing experience compared to announcement and information from peers. Similarly, Haselhuhn et al. (2012) find that experience with fines increases compliance in returning video rentals.[4] A complementary laboratory literature compares the roles of experience relative to information in judgment and decision making (Weber et al., 2004; Hertwig et al., 2004; Barron & Erev, 2003; Simonsohn et al., 2008; Miller & Maniadis, 2012; Herz & Taubinsky, 2018). For example, Conlon et al. (2022), demonstrate that individuals value information they discover themselves far more than information provided by others. We contribute to this literature by using a controlled field experiment to investigate how immersive, real-world experiences shape preferences for criminal justice policy. Our design provides causal evidence that firsthand institutional exposure can influence preferences and public opinion, fundamental ingredients of political decision-making.

Second, previous studies have examined the potential of prison visits and encounters with incarcerated individuals as interventions to deter at-risk youth and former offenders from future crime (e.g., "Scared Straight"). Petrosino et al. (2013)'s meta-analysis of randomized controlled trials indicates that these programs do not only fail to deter crime but often increase crime among at-risk youth. In contrast to Scared Straight and similar initiatives, our intervention targets the general population and focusses on firsthand incarceration experiences rather than confrontations with incarcerated individuals and their narratives of crime and prison life. Furthermore, we do not aim to evaluate deterrence effects but seek to understand the broader attitudinal impacts of incarceration experience.[5]

[3] One notable exception is Andries et al. (2024), who use an immersive virtual reality intervention to experimentally simulate the experiences and struggles of unauthorized migrants.

[4] See also Dušek & Traxler (2022)for related evidence on the impact of punishment experience on future compliance, as well as Lochner (2007) for evidence on how arrest experiences shape beliefs about the probability of future arrest.

[5] Prior research has also examined prison tours as an educational tool, primarily for criminology students (e.g., Calaway et al., 2016; Long & Utley, 2018; Murdoch, 2020). These studies are often constrained by small sample sizes and lack of randomization, limiting the generalizability of their findings.

Third, our study relates to a vast literature highlighting that citizens are often poorly informed about crime and the criminal justice system (Apel, 2022). For example, many people hold profound misconceptions about sentencing practices and their severity (Roberts et al., 2022; Hough & Roberts, 2012). Additionally, citizens often misunderstand the demographic composition of prisoners and the realities of life in prison (see, e.g., Bryant & Morris, 1998; Roberts & Hough, 2005). Doble (1995) and Farkas (1993), for example, document that many people mistakenly believe prisoners spend their time idly or engaged primarily in leisure activities. However, little attention has been paid to the public perceptions of inmate wellbeing. We provide novel evidence that citizens systematically overestimate the wellbeing of prisoners. Interestingly, when asked to predict their own wellbeing in prison, citizens report levels more similar to those reported by actual prisoners. This finding suggests that citizens think prisoners are better able to cope with incarceration.

Finally, our study also connects to research on the effects of information on attitudes toward the criminal justice system. These studies have experimentally varied the provision of information on sentencing practices, sentencing guidelines, the availability of alternative sanctions, and narrative accounts of incarceration (Mitchell & Roberts, 2011; Doble, 2002; Harney, 2023). Rather than relying on information-based interventions, which typically produce moderate effects on preferences (Haaland et al., 2023), we offer participants a deeply immersive experience of incarceration. This approach complements the existing literature on information provision by leveraging firsthand exposure to foster changes in attitudes, akin to Andries et al. (2024) who combined virtual reality with information provision to shape attitudes toward immigration.

# 2 Background

## 2.1 The Swiss incarceration experience in international comparison

Switzerland incarcerates 73 individuals per 100,000 citizens, a figure comparable to Norway or Germany, which have incarceration rates of 54 and 67 individuals per 100,000 citizens, respectively. In stark contrast, the United States has an incarceration rate of 614 inmates per 100,000 citizens (Institute for Crime and Justice Policy Research, 2024). Reports comparing prison conditions across developed countries suggest that the Swiss incarceration experience varies substantially across different dimensions. However, it is generally considered less harsh

than in other developed countries, particularly the United States. Swiss prisoners — similar to other European inmates — typically serve shorter sentences for equivalent offenses compared to their U.S. counterparts. For instance, only 5.8% of Swiss prisoners serve sentences exceeding ten years, in contrast to 52.4% of U.S. inmates serving such sentences in 2019 (Kazemian & Galleguillos, 2025; Aebi & Triago, 2021). By comparison, the proportions of inmates serving sentences over ten years are 8.9% in Norway and 18% France in 2019 (Aebi & Triago, 2021). Inmates in Swiss correctional facilities benefit from access to education, vocational training programs, and opportunities for paid work, which are not guaranteed in all European countries (World Health Organization: Regional Office for Europe, 2023).

However, the Swiss National Commission for the Prevention of Torture has repeatedly criticized the stringent conditions in pretrial detention (National Commission for the Prevention of Torture (NCPT), 2017, 2022). Until recent years, pretrial detainees in most Cantons were confined to their cells for up to 23 hours a day, with minimal or no contact with the outside world. The NCPT has also highlighted issues such as overcrowding in some Swiss prisons and insufficient access to therapeutic measures. Consistent with these findings, Aebi & Cocco (2024) report that the suicide rate among Swiss prison and pretrial detention inmates is one of the highest in Europe, at 202 per 100,000 inmates, compared to the European median of 53 per 100,000 inmates.

## 2.2 Test run for a newly built prison

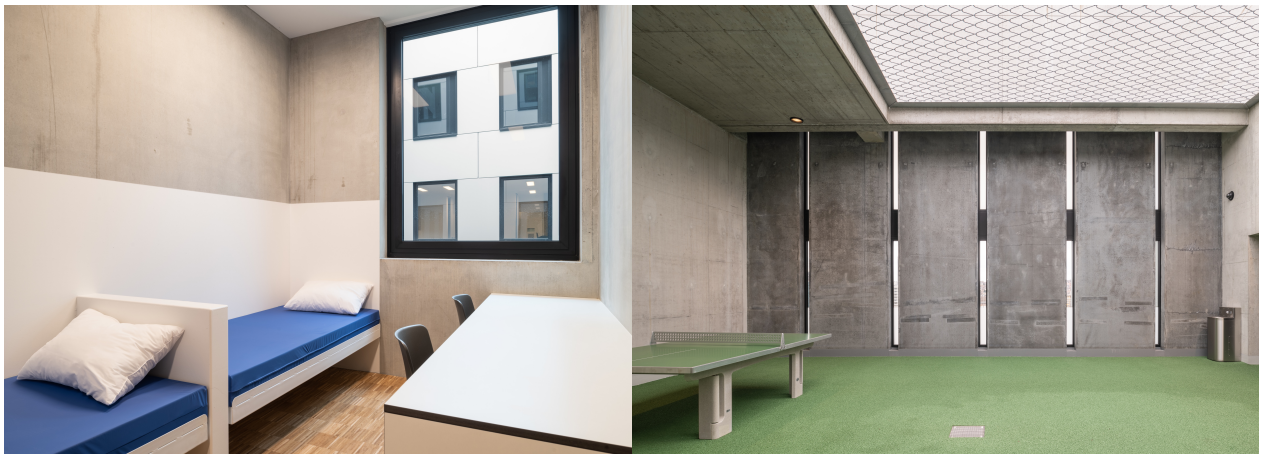
The Department of Justice and Home Affairs of the Canton of Zurich in Switzerland organized a four-day test run for a newly built prison to test its operational and security procedures under simulated real-life conditions. Volunteering participants could experience up to 48 hours of incarceration, following standard procedures closely mirroring a real inmate's journey from intake to release. After participants arrived at the prison (individually or via police transport)[6], the prison staff confiscated all belongings. Subsequently, participants underwent a strip search[7] and received their prison uniform. Next, they were put into a waiting cell for an undetermined amount of time, which lasted more than 3 hours for some participants before being transferred into their assigned prison cell. While the prison features both single and double cells, the majority of cells used for the test run (53 out of 57) were double cells

[6]The police transported a small subset of participants to the prison to simulate the organizational challenge of a large group of inmates entering at the same time.

[7]The strip search was the only voluntary element of the test run. Despite this, in the online application process more than 80% of all applicants opted into having the strip search.

Figure 1: Images of the prison

(a) Double cell (b) Courtyard



*Note* This figure shows a double cell in panel (a) and the courtyard in panel (b) of the prison in which participants were incarcerated during the test run. Source: © Baudirektion Kanton Zürich / Till Forrer. Used with permission.

measuring 13 square meters.[8] See Figure 1 for images of a double cell and the courtyard. Although applicants were asked for their availability on each of the four days of the test run during the online application process (see Section 3.1 for details), participants were only informed of their admission time. They were not told their exact time of release, mimicking, though to a lesser extent, the uncertainty about the duration of incarceration that actual inmates would face.

The test run participants followed the same strictly regulated daily routine as regular inmates. They were locked up in their prison cells for the vast majority of the day and could only leave the cell for a 10-minute shower in the morning and during a 90-minute window in the afternoon reserved for the courtyard walk. For an overview of the daily schedule, see Table 1.

The key feature of the test run is that participants' incarceration experience was as close as possible to a real inmate's journey from intake to release. Two aspects, however, were different: (i) participants had the possibility of terminating their incarceration with an personal "code word" that they defined with the guards during the entry process, (which, according to the prison staff, only two participants made use of), and (2) participants would not share their cells with suspected criminals, but regular citizens with clean criminal records and no ongoing proceedings. Together with the fact that they were held in a modern uncrowded

[8] The prison features four additional single cells and five additional double cells that were not used for the test run.

Table 1: Daily schedule

| Time | Schedule |
|---|---|
| 6:45 AM | Wake-up call |
| 7:15 AM | Breakfast in cell |
| 8:30 AM | Kiosk at cell door (to borrow books or buy snacks / cigarettes) |
| 10:00 AM to 11:00 AM | 10-minute slot for daily shower |
| 12:00 PM | Lunch in cell |
| 1:15 PM to 2:45 PM | 90-minute window for courtyard walk |
| 5:00 PM | Dinner in cell |
| 10:00 PM | Lights off |

*Note*: This table shows the daily routine participants followed during the prison test run mirroring the routines of regular inmates.

prison[9], our estimates will likely represent a lower bound of the effects of firsthand incarceration experience.

Anecdotal accounts reveal that participants often struggled with the lack of autonomy and the psychological burden of incarceration. Participants' narratives repeatedly pointed to the loss of control over their decisions, as particularly salient aspect of the test run, illustrated by statements such as: "The experience of not being able to decide for oneself about one's own actions (when to eat, go for a walk, etc.) is very profound." and "The feeling of simply not knowing what is happening, of being powerless, and of no longer being able to go outside." Other accounts highlighted the emotional toll of incarceration, emphasizing the monotony: "In no other place does time pass so quickly as in the courtyard. Back in the cell, I am overcome by a great wave of boredom. I feel like I'm a dog and only look forward to eating and going for a walk."[10]

# 3 Experimental design

## 3.1 Online application process and randomization

After the announcement of the prison test run by the Department of Justice and Home Affairs of the Canton of Zurich through various media outlets, the authorities opened an online application portal in early February 2022. Applicants were only eligible to participate in the

[9] At peak times, 77 test run participants were incarcerated simultaneously. Given the prison's capacity for 124 inmates, this corresponds to a maximum occupancy share of 62%.

[10] These quotes stem from an additional qualitative survey conducted with participants guaranteed a slot in the test run (politicians, judges, and media professionals; see Online Appendix C), as well as from newspaper articles written by participating media professionals.

test run if they were adults living in the Canton of Zurich and had a clean criminal record and no ongoing proceedings. The authorities were overwhelmed with applications and ultimately left with a sample of 747 eligible applicants from the general population.[11] This number exceeded the available capacities for the test run, providing an opportunity to randomize access. We randomly assigned eligible applicants into a treatment and a control group, stratified by employment in the public sector. We implemented the stratified randomization because we expected a high number of applications from public sector employees working for the justice department who might have visited a prison before.

We provided the prison authorities with the list of individuals from the treatment group, and they had the discretion to select any applicant from this group and schedule a slot in the test run based on cell occupancy as well as applicants' availability, gender, and smoking status. To provide prison authorities flexibility in scheduling the cell allocation plan, we oversampled the treatment group relative to the number of available slots. Consequently, we provided them with a list of 200 people for the 163 initially planned test run slots, consisting of the 182 subjects who were randomized into the treatment group and the 18 VIPs who were guaranteed a slot.[12] On the contrary, we instructed prison authorities not to invite applicants from the control group for the test run.

## 3.2 Baseline survey and invitation for the test run

After the randomization, we invited all applicants to complete the baseline survey.[13] We included the questionnaire measures for each of the primary outcomes to increase the precision of our regression analyses.

After the baseline survey and eight days prior to the start of the test run, prison authorities

[11] In addition to the applicants from the general population, 65 further people were guaranteed a slot in the test run. These included 18 VIPs (politicians and employees in managing positions at the Justice Department), 28 judges, and 19 media professionals. While we invited the VIPs and judges to our surveys, the Department of Justice and Home Affair did not want members of the media to participate in the surveys. All of these additional 65 people were not randomized and are therefore excluded from all analyses in this paper.

[12] After the randomization, we were informed that the 163 slots did not only have to accommodate the 18 VIPs, but in addition 28 judges and 19 media people, which further decreased the number of available slots for applicants randomized into the treatment group (see Section 4.2 for further details).

[13] Due to the relatively short window for the prison authorities to create the test run allocation plan, we had to conduct the randomization before we administered the baseline survey. However, the outcome of the randomization was communicated to most of the applicants only after they completed the baseline survey. Of the 418 subjects included in the regression analyses, just three completed the baseline survey after they were informed about the outcome of the randomization. We show in Table B.1 that our results are robust to excluding these three subjects.

communicated whether the applicants were invited for the test run or not, explaining that the decision was based on a lottery draw due to excess demand. Subjects from the treatment group invited to participate in the test run received additional information about their dates and times of entry and release. The test run took place between March 24th and 27th, 2022.

## 3.3 Endline survey

After the test run, we invited all applicants to our endline survey. To incentivize participation, we offered a raffle in which every fourth participant received 10 Swiss francs, conditional on survey completion. In the endline survey we elicited our two primary outcomes: (i) punitive attitudes and (ii) beliefs about wellbeing in prison. For each outcome, we included both a behavioral and a questionnaire measure. As pre-registered, we created an index for each outcome out of the two individual measures using principal component analysis, but also report results for each of the two individual measures separately.

To create the *punitive justice index*, we combine the following two measures. First, using a donation task, we measured subjects' support for, or opposition to, policies targeting a tougher criminal justice system, dubbed *punitive justice donations*. For this task, we gave them the option to direct a monetary donation of up to 5 Swiss francs (corresponding to roughly 5 US dollars and 40 cents) to a civic organization that either actively supports or actively opposes such policies. We asked the following question:
*"Are you more inclined to direct us to donate to an organization that supports a tougher penal system, or to an organization that supports a more moderate penal system?"*[14]
Conditional on their answer, they had to indicate how much of 5 Swiss francs we should donate on their behalf.[15] Second, we measured subjects' *punitive attitudes* by asking them how strongly they support strict criminal prosecution and harsh sentencing on a Likert-scale from 0 (do not support at all) to 6 (completely support).[16].

To construct the index of *beliefs about wellbeing in prison*, we proceed as follows. First, we elicited *beliefs about the wellbeing of actual prisoners*, using an incentivized guessing task.

[14]We randomized the sequence of the two options in both the question text and in the corresponding response options, to mitigate potential ordering effects.

[15]We deliberately withheld the organizations' identities to prevent subjects from making donation decisions based on their subjective perceptions.

[16]Using data from an additional survey conducted with a representative sample of the Swiss voting population, we find that a lower score on the punitive attitudes index correlates with having stronger policy preferences for (i) less strict prison conditions (e.g., open prisons), (ii) a greater focus on reintegration into society (e.g., rehabilitative offers for continuing vocational training), and (iii) an increased use of alternative sanctions as opposed to incarceration (e.g., electronic ankle bracelets, community service). For graphical visualizations of these correlations, see Online Appendix Figure A.1.

We surveyed actual inmates in a Swiss prison (the Prison Pfäffikon in the Canton of Zurich) about their personal wellbeing on a scale from 0 to 10, where 0 means they are extremely bad and 10 means they are extremely good. In the endline survey, we asked our study participants to guess the average value of the actual prison inmates' answers. We encouraged them to be as precise as possible, with the three most accurate answers each winning 50 Swiss francs. Second, we elicited subjects' own expected wellbeing in prison (dubbed *citizens' anticipated wellbeing in prison*) by asking them to imagine they would be incarcerated for six months and to indicate what they think their personal wellbeing during incarceration would be, again using an 11-point Likert-scale.

As secondary outcomes, we also elicited participants beliefs about trust in institutions and their perceptions of procedural fairness.[17] Online Appendix Figure A.2, summarizes the study timeline.

## 3.4 Sample size

678 of the 747 applicants (90.8%) participated in the baseline survey, with no significant difference in response rates between the experimental groups (90.4% in control and 91.8% in treatment group, $p = 0.594$, Wilcoxon rank-sum test). 435 of these 747 invited applicants participated in the endline survey. The final sample of 435 subjects consists of 126 subjects in the treatment group (from the 182 subjects originally randomized into it), and of 309 subjects in the control group (from originally 565 subjects).[18] We discuss attrition in more detail in Section 4.5. For the regression analysis, the sample consists of 418 subjects for punitive justice outcomes, and 416 subjects for wellbeing outcomes respectively, due to the inclusion of baseline variables and two subjects dropping out of the endline survey.

[17] To elicit trust in legal institution, we asked participants to indicate how much trust they have in the Swiss legal system on a Likert-scale from 0 (no trust at all) to 10 (complete trust). To measure perceptions of procedural fairness, we asked them to indicate how fairly affected persons are treated by the Swiss justice system from 0 (not fair at all) to 6 (very fair).

[18] We initially pre-registered sample sizes of 185 and 560 subjects for the treatment and control groups, respectively. However, these figures were incorrectly recorded in the pre-registration due to a document versioning error.

# 4 Results

## 4.1 Sample description and randomization check

The average age in our sample is 39 years and nearly 51% of subjects are male. Almost every fifth subject is a public sector employee at the canton of Zurich (18%), while 23% work in a justice-related profession. 27% of subjects have previously visited a prison (see Table 2). This relatively high proportion of subjects who have previously visited a prison can be attributed to the sizable share of individuals employed in a justice-related profession.[19] While these descriptive statistics illustrate that our experimental sample does not represent the general population, we find similar results if we re-weight observations using additional data from a representative sample (see Table 6).

Table 2 contains the balance check for our sample. Subjects in the treatment and control groups are comparable on all characteristics elicited during the application process and the baseline survey. Additionally, when all observable characteristics are regressed on an indicator for treatment assignment, the F-test ($p = 0.925$) reveals that we cannot reject the null hypothesis that all coefficients are jointly zero.

## 4.2 First stage results

We intentionally oversampled the treatment group relative to the number of available slots in the test run to provide more flexibility to the prison authorities when creating the cell allocation plan. Additionally, given the voluntary nature of the test run, some invited subjects from the treatment group canceled their participation shortly before the test run or simply did not show up. In a first step, we therefore analyze how treatment assignment affected incarceration status and duration. We estimate the following regression model using Ordinary Least Squares (OLS):

$$Y_i = \alpha + \beta_1 T_i + \beta_2 Z_i + \epsilon_i \tag{1}$$

The dependent variable $Y_i$ specifies whether subject $i$ participated in the test run, the total number of hours subject $i$ was incarcerated during the test run, and whether subject $i$ spent at least one night (i.e., between 11:00 PM and 6:00 AM) in prison during the test run. The dummy variable $T_i$ indicates the treatment assignment, $Z_i$ is a dummy for the strata

[19] The Spearman's correlation between working in a justice-related profession and previous prison visit is 0.58 ($p = 0.000$).

Table 2: Balance Check

| | Full Sample | | Control | | Treatment | | |
|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | *p-val* |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Punitive Attitudes (baseline) | 4.376 | (1.37) | 4.377 | (1.40) | 4.372 | (1.29) | 0.656 |
| Anticipated Wellbeing in Prison (baseline) | 4.170 | (2.45) | 4.141 | (2.41) | 4.240 | (2.54) | 0.856 |
| Procedural Fairness (baseline) | 4.732 | (0.90) | 4.721 | (0.90) | 4.760 | (0.90) | 0.445 |
| Trust in Institutions (baseline) | 7.971 | (1.41) | 7.939 | (1.37) | 8.050 | (1.52) | 0.139 |
| Criminal Attitudes | 1.368 | (1.23) | 1.384 | (1.25) | 1.331 | (1.19) | 0.764 |
| Willingness to take Risk | 4.885 | (2.12) | 4.852 | (2.11) | 4.967 | (2.14) | 0.684 |
| Age | 38.885 | (13.14) | 39.152 | (13.65) | 38.230 | (11.81) | 0.844 |
| Male | 0.510 | (0.50) | 0.511 | (0.50) | 0.508 | (0.50) | 0.949 |
| Available on Thursday | 0.563 | (0.50) | 0.573 | (0.50) | 0.540 | (0.50) | 0.527 |
| Available on Friday | 0.703 | (0.46) | 0.725 | (0.45) | 0.651 | (0.48) | 0.125 |
| Available on Saturday | 0.770 | (0.42) | 0.744 | (0.44) | 0.833 | (0.37) | 0.045 |
| Available on Sunday | 0.651 | (0.48) | 0.628 | (0.48) | 0.706 | (0.46) | 0.119 |
| Interested in Overnight Stay | 0.924 | (0.27) | 0.926 | (0.26) | 0.921 | (0.27) | 0.860 |
| Consented Strip Search | 0.766 | (0.42) | 0.790 | (0.41) | 0.706 | (0.46) | 0.063 |
| Smoker | 0.145 | (0.35) | 0.142 | (0.35) | 0.151 | (0.36) | 0.821 |
| Works in Public Sector | 0.184 | (0.39) | 0.181 | (0.39) | 0.190 | (0.39) | 0.821 |
| Justice-Related Profession | 0.230 | (0.42) | 0.226 | (0.42) | 0.240 | (0.43) | 0.756 |
| Previous Prison Visit | 0.270 | (0.44) | 0.276 | (0.45) | 0.256 | (0.44) | 0.678 |
| Tertiary Education | 0.615 | (0.49) | 0.589 | (0.49) | 0.678 | (0.47) | 0.092 |
| F-test: | | | | | | | p = 0.925 |

*Note*: The table reports descriptive statistics for the full sample and separately for the treatment and control groups. It includes baseline measures of punitive attitudes, citizens' anticipated wellbeing in prison, perceptions of procedural fairness, trust in institutions, criminal attitudes, and willingness to take risks. Additionally, it reports subjects' age, the proportion of subjects who identify as male, indicate availability on each of the four test run days (Thursday to Sunday), are interested in staying overnight, consent to a strip search, are smokers, work in the public sector, work in a justice-related profession, have previously visited a prison, and completed tertiary education. Column (7) contains $p$-values from $\chi^2$-tests for binary variables and Wilcoxon rank-sum tests for continuous variables, respectively. The $p$-value at the bottom is based on an F-test that evaluates whether all covariates are jointly equal to zero in a linear regression predicting treatment assignment. The number of observations for variables elicited at baseline is 418, while the number of observations for variables from the application data is 435.

fixed effect (i.e., working in public sector), and$\epsilon_i$ is the idiosyncratic error term. Table 3 shows that the treatment had a significant positive effect on all three measures of test run participation: Subjects in the treatment group had a 57.7 percentage points higher likelihood of attending the test run ($p = 0.000$, column 1). Individuals in the treatment group were incarcerated for an additional 11.6 hours on average ($p = 0.000$, column 2). Participants who actually attended the test run spent on average 21 hours in prison, with a maximum of 48 hours. Finally, column (3) shows that the treatment increased the likelihood of staying at least one night in prison by 48 percentage points ($p = 0.000$).[20]

[20] 13 subjects from the control group nevertheless participated in the test run. Discussions with the prison authorities revealed that three of those subjects were treated as VIPs (i.e., prison authorities guaranteed them a slot from the beginning without informing us beforehand), and the other ten subjects were accidentally invited by the prison authorities because of organizational mistakes. All of these 13 subjects participated in the baseline and endline survey and are thus included in the final sample, resulting in the positive numbers

Table 3: First Stage Regression

| | Attended Test Run | Hours in Prison | Overnight in Prison |
|---|---|---|---|
| | (1) | (2) | (3) |
| Treatment Group | 0.577*** | 11.583*** | 0.480*** |
| | (0.045) | (1.105) | (0.046) |
| Strata FE | ✓ | ✓ | ✓ |
| Control group mean | 0.042 | 0.954 | 0.036 |
| Observations | 435 | 435 | 435 |

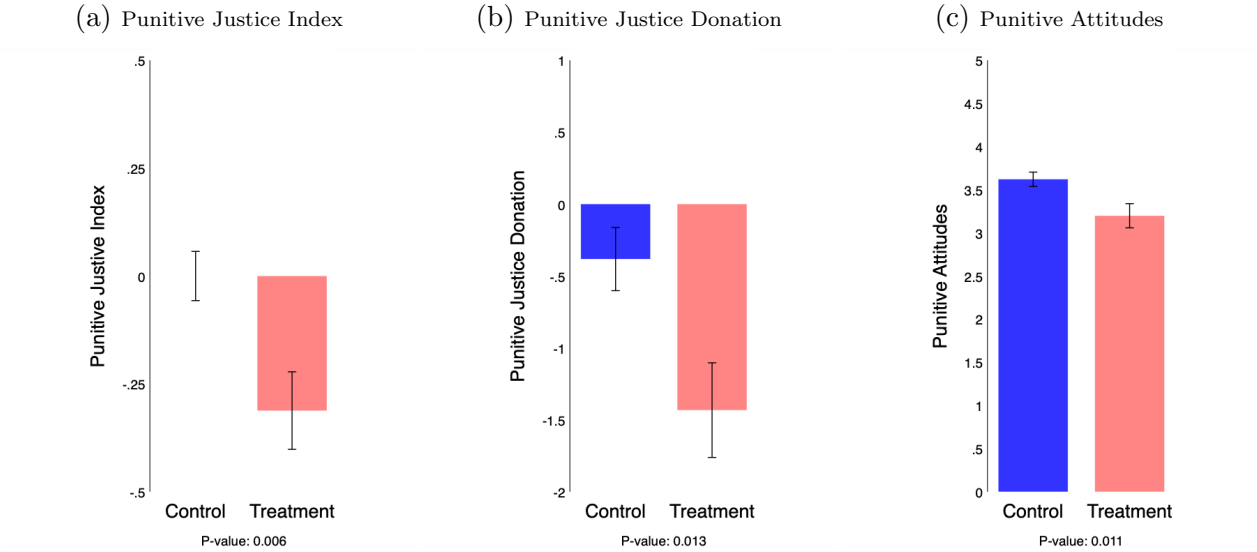
*Note*: This table shows the effect of treatment assignment on subjects' participation in the prison test run (i.e., treatment take-up) using OLS regressions. Robust standard errors are reported in parentheses. The dependent variables are a dummy indicating attendance at the test run (column 1), the hours subjects spent in prison (column 2), and a dummy for whether they spent a night in prison (column 3; defined as being in the cell between 11pm and 6am), respectively. All regressions include strata fixed effects, i.e., a dummy for subjects working in the public sector. Levels of significance: $^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$.

## 4.3 Effects on punitive attitudes

Panel (a) in Figure 2 depicts the average score from the punitive justice index in the endline survey (normalized to a control group mean of 0 and standard deviation of 1) by treatment status. A higher score indicates stronger support for punitive justice policies. Subjects in the treatment group scored on average -0.3 standard deviations lower than those in the control group ($p = 0.006$, Wilcoxon rank-sum test). Panels (b) and (c) in Figure 2 show the results separately for the two individual components of the punitive justice index. We coded all donations in favor of a more moderate criminal justice system with negative values and donations in favor of a tougher criminal justice system with positive values. Our variable *punitive justice donations* thus ranges from -5 Swiss francs (if respondents made the largest possible donation for a more moderate criminal justice system) to +5 Swiss francs (if the respondents made the largest possible donation for a tougher criminal justice system). Panel (b) highlights that punitive justice donations were on average roughly 1 Swiss franc lower in the treatment group than in the control group, highlighting greater support for a more moderate criminal justice system in the treatment group ($p = 0.013$, Wilcoxon rank-sum test). We find similar results for the questionnaire measure of punitive attitudes (see panel c). While the control group reported an average score of 3.6 on the punitive attitude scale, the punitive attitudes are significantly lower in the treatment group with an average score of 3.2 ($p = 0.011$, Wilcoxon rank-sum test).

for the control group means.

Figure 2: Effect on Punitive Attitudes




*Note*: This figure shows averages for different measures of punitive attitudes in the endline survey (N = 435) by treatment status. Error bars represent standard errors of the mean. Panel (a) reports the average score of the punitive justice index (normalized to a control group mean of 0 and standard deviation of 1). A higher score indicates stronger support for punitive justice policies. Panel (b) depicts the average punitive justice donation in the donation task (ranging from -5 to +5 Swiss francs). Donations in favor of a more moderate (a tougher) criminal justice system are coded with negative (positive) values, respectively. Panel (c) shows the average response to the punitive attitudes survey question: how strongly they support strict criminal prosecution and harsh sentencing on a scale from 0 (do not support at all) to 6 (completely support).

The regression analysis corroborates the previous non-parametric results. We estimate the following linear regression model using OLS:

$$Y_i = \alpha + \beta_1 T_i + \beta_2 X_i + \beta_3 Z_i + \epsilon_{ic} \tag{2}$$

The dependent variable $Y_i$ specifies subject $i$'s punitive justice index, punitive justice donation, or punitive attitudes in the endline survey (normalized to a control group mean of 0 and standard deviation of 1). It is regressed on a dummy variable $T_i$ indicating the treatment assignment. The vector $X_i$ represents a subset of controls elicited during the application process or during the baseline survey. We used pre-registered control variables elicited at baseline, namely: punitive attitudes, anticipated wellbeing in prison, trust in institutions and procedural fairness (see Section D in the Online Appendix for the corresponding survey questions). Additionally, we control for criminal attitudes, willingness to take risk, age, gender, previous prison visit, working in a justice-related profession, and completed ter-

tiary education.[21] $Z_i$ are the strata fixed effects. We correct standard errors to account for dependencies in the error term $\epsilon_{ic}$ for individuals who spent time in the same cell.

Column 1 in Table 4 contains the result for the punitive justice index, controlling for a basic pre-registered set of baseline outcomes. We find a significant decrease in the punitive justice index by 0.33 standard deviations ($p = 0.000$, t-test). In Column 2, we present the result for the punitive justice index, including the full set of controls. We find a similar significant decrease in the punitive justice index of 0.33 standard deviations ($p = 0.000$, t-test). Columns 3 and 4 show the results for the punitive justice donation and punitive attitude question separately, including the same set of controls. Both regressions report significant treatment effects of -0.3 standard deviations ($p = 0.000$ and $p = 0.001$, t-tests).

To assess the impact of attending the test run on punitive attitudes, we also estimated the local average treatment effect (LATE) through a two-stage least squares approach, with treatment assignment as an instrument for test run participation. The results, presented in column (1) of Online Appendix Table B.2 indicate a LATE of -0.59 standard deviations on the punitive justice index ($p = 0.000$, t-test).[22]

To summarize, the opportunity to gain firsthand incarceration experience significantly increased support for a more moderate justice system.[23] To gauge the magnitude of the effect, we use survey data from the 2005 International Crime and Victim Survey (ICVS) and the 2005 European Survey of Crime and Safety (EU ICS) compiled by Van Dijk et al. (2007). In particular, we use the percentage of survey respondents favoring imprisonment for a recidivist burglar[24] to compare the United States and Norway, two countries commonly seen to represent two opposing ends of the spectrum in terms of the punitiveness of their justice systems (Bhuller et al., 2020). On average, the difference in punitive attitudes between these two countries is 18 percentage points (47% in the U.S. and 29% in Norway, respectively) or 1.1 standard deviations.[25] Thus, our intention-to-treat effect accounts for about 30% of the gap in punitive attitudes between the United States and Norway.

[21] We present results throughout the paper for both specifications, using only pre-registered controls and the full set of controls. Both specifications yield similar findings.

[22] See Online Appendix Table B.2 for estimates of the LATE on the two individual components of the punitive justice index. We find a LATE of -0.53 for punitive justice donations ($p = 0.001$, t-test) and -0.53 for punitive attitudes ($p = 0.000$, t-test).

[23] Interestingly we find a consistent correlational pattern: Online Appendix Table B.3 shows that previous prison visits are significantly negatively associated with punitive attitudes in the baseline survey and the punitive justice index at endline for subjects in the control group.

[24] The recidivist burglar was specified as a man aged 21 who is found guilty of burglary for the second time, having stolen a color television

[25] The average percent favoring imprisonment is 32.8% for the 32 included countries, with a standard deviation of 16.2%.

Table 4: Regression Analysis: Punitive Attitudes

| | Punitive Justice Index | | Punitive Justice Donation | Punitive Attitudes |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Treatment Group | -0.326*** | -0.331*** | -0.300*** | -0.297*** |
| | (0.081) | (0.081) | (0.092) | (0.080) |
| Strata FE | ✓ | ✓ | ✓ | ✓ |
| Baseline Outcomes | ✓ | ✓ | ✓ | ✓ |
| Criminal and Risk Attitudes | × | ✓ | ✓ | ✓ |
| Previous Prison Visit | × | ✓ | ✓ | ✓ |
| Socio-Demographics | × | ✓ | ✓ | ✓ |
| Adjusted R-squared | 0.443 | 0.447 | 0.258 | 0.489 |
| Observations | 418 | 418 | 418 | 418 |

*Note*: This table shows the effect of treatment assignment on punitive attitudes in the endline survey, using OLS regressions. Standard errors (in parentheses) are corrected for clustering of the error term for individuals who spent time together in the same cell. In columns 1 and 2, the dependent variable is the punitive justice index. In column 3 and 4, the dependent variables are punitive justice donations and punitive attitudes, respectively. All three dependent variables are normalized to a control group mean of 0 and standard deviation of 1. All regressions include baseline outcomes and strata fixed effects, i.e., a dummy for subjects working in the public sector. Baseline outcomes include measures of anticipated wellbeing in prison, punitive attitudes, trust in institutions, and perceptions of procedural fairness. Columns 2 to 4 include further controls for baseline criminal attitudes and risk attitudes, whether subjects previously visited a prison and socio-demographics, including subjects' age, and dummies for males, tertiary education, and employment in a justice-related profession. Levels of significance: $^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$.

We also explore whether treatment effects are heterogenous across different subpopulations. We examine the two pre-registered moderators: previous prison visit and individuals working in the public sector. Additionally, we perform the same exploratory analysis for female and male subjects, individuals working in a justice-related profession, and subjects with low and high baseline punitive attitudes (based on a median split). Online Appendix Figure A.4 indicates that, the treatment effect remains stable within the range of -0.2 and -0.4 standard deviations for all subpopulations, suggesting the treatment had a broadly uniform effects across those subgroups. We summarize these findings in the following result:

**Result 1:** *The opportunity to gain firsthand incarceration experience significantly reduces punitive attitudes.*

## 4.4 Effects on beliefs about wellbeing in prison

We first assess the accuracy of citizens' perception of wellbeing in prison, by comparing the subjective wellbeing of actual inmates from a Swiss prison (the Prison Pfäffikon) with the control group's beliefs regarding these prisoners' wellbeing. Panel (a) in Figure 3 reveals that the control group's beliefs about the wellbeing of actual prisoners are far from reality. While the prisoners indicated an average wellbeing of 3.2, control group subjects estimated the wellbeing to be around 6.2 on average ($p = 0.000$, Wilcoxon rank-sum test).[26] Although we cannot rule out that some inmates strategically underreported their wellbeing, such behavior is unlikely to entirely explain the observed difference between wellbeing of actual prisoners and the control's beliefs. First, the inmate survey was completely anonymous, reducing incentives for strategic misreporting. Second, closing a gap of roughly three points on a 0 to 10 scale would require an implausibly large downward bias in responses.

Panel (b) in Figure 3 compares the control group's beliefs about actual prisoners' wellbeing with their own anticipated wellbeing in prison. While the control group subjects estimated an average wellbeing of 6.2 for prisoners in Pfäffikon, they indicated an average anticipated wellbeing in prison of 4.3 for themselves ($p = 0.000$, Wilcoxon rank-sum test). This suggests that they expect prisoners to cope better with imprisonment than they themselves would.[27] In addition, they might perceive prisoners' outside options to be worse than life inside prison. These findings are summarized in the following result:

**Result 2:** *Subjects in the control group overestimate the wellbeing of actual prisoners, but anticipate their own wellbeing in prison to be significantly lower than the wellbeing of prisoners.*

Turning to the effects of treatment assignment, we find no differences in beliefs about wellbeing in prison between the treatment and control group. Panel (a) in Online Appendix Figure A.3 illustrates that the average wellbeing beliefs index (normalized to a control group mean of 0 and standard deviation of 1) is almost identical between the treatment and the

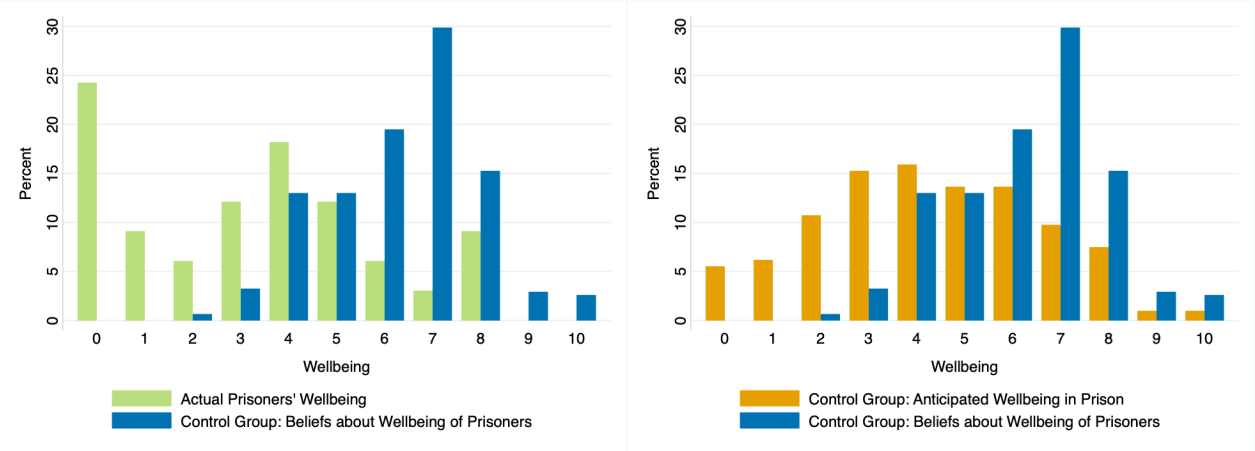
[26] Although the level of wellbeing estimated by the control group is much higher than the wellbeing indicated by actual prisoners, it is still substantially lower ($p = 0.000$, t-test) than the wellbeing in the general population which is 8.02 in 2022 according to Swiss Household Panel (SHP) data (Tillmann et al., 2022). Note that the questions are not identical across data sources as the SHP measures life satisfaction in contrast to wellbeing. However, the SHP also uses a similar 11-point Likert scale (*"In general, how satisfied are you with your life if 0 means "not at all satisfied" and 10 means "completely satisfied"?"*).

[27] This finding is consistent with psychological literature finding biases towards out-group individuals on the dimension of pain (Hoffman et al., 2016; Staton et al., 2007; Todd et al., 2000).

Figure 3: Beliefs about Wellbeing in Prison

(a) Actual Prisoners' Wellbeing compared with Control Group's Beliefs about the Wellbeing of Actual Prisoners

(b) Control Group's Anticipated Wellbeing in Prison compared with Control Group's Beliefs about the Wellbeing of Actual Prisoners




*Note*: Panel (a) shows distributions of i) the wellbeing of actual prisoners from the prison of Pfäffikon (N = 33) on a 11-point scale from 0 (extremely bad) to 10 (extremely well) and ii) the control group's incentivized beliefs about the wellbeing of actual prisoners from the prison of Pfäffikon (N = 308). Panel (b) shows distributions of i) the control group's anticipated wellbeing in prison using an 11-point Likert-scale (N = 308) and ii) the control group's incentivized beliefs about the wellbeing of actual prisoners from the prison of Pfäffikon (N = 308).

control group ($p = 0.443$, Wilcoxon rank-sum test). Similarly, Panels (b) and (c) in Online Appendix Figure A.3 show no significant differences between experimental conditions for either citizens' beliefs about the wellbeing of actual prisoners or their anticipated wellbeing in prison. Subjects in the control and treatment group estimated the wellbeing of actual prisoners to be 6.2 and 6.25, respectively ($p = 0.893$, Wilcoxon rank-sum test). Likewise, the average citizens' anticipated wellbeing in prison is 4.3 in the control group and 4.4 in the treatment group ($p = 0.735$, Wilcoxon rank-sum test).

The regression results, presented in Table 5, support the nonparametric analysis and reveal no significant differences for any of the beliefs about wellbeing ($p > 0.87$, t-tests).

We also explored whether heterogenous treatment effects could obscure this null result. Following the same approach used for punitive attitudes, we regress treatment assignment on the wellbeing beliefs index across different subsamples. The analysis reveals no significant treatment effects in any subsample (see Online Appendix Figure A.5). These findings are summarized in the following result:

**Result 3:** *The opportunity for firsthand incarceration experience does not affect beliefs about wellbeing in prison.*

Table 5: Regression Analysis: Beliefs about Wellbeing in Prison

| | Wellbeing Beliefs Index | | Beliefs About Prisoners' Wellbeing | Anticipated Wellbeing in Prison |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Treatment Group | 0.005 | 0.001 | -0.011 | 0.013 |
| | (0.091) | (0.091) | (0.101) | (0.083) |
| Strata FE | ✓ | ✓ | ✓ | ✓ |
| Baseline Outcomes | ✓ | ✓ | ✓ | ✓ |
| Criminal and Risk Attitudes | × | ✓ | ✓ | ✓ |
| Previous Prison Visit | × | ✓ | ✓ | ✓ |
| Socio-Demographics | × | ✓ | ✓ | ✓ |
| Adjusted R-squared | 0.251 | 0.253 | 0.098 | 0.308 |
| Observations | 416 | 416 | 416 | 416 |

*Note*: This table shows the effect of treatment assignment on beliefs about wellbeing in prison, using OLS regressions. Standard errors (in parentheses) are corrected for clustering of the error term for individuals who spent time together in the same cell. In columns 1 and 2, the dependent variable is the wellbeing beliefs index. In columns 3 and 4, the dependent variables are beliefs about the wellbeing of actual prisoners and the anticipated wellbeing in prison, respectively. All three dependent variables are normalized to a control group mean of 0 and standard deviation of 1. All regressions include baseline outcomes and strata fixed effects, i.e., a dummy for subjects working in the public sector. Baseline outcomes include measures of anticipated wellbeing in prison, punitive attitudes, trust in institutions, and perceptions of procedural fairness. Columns 2 to 4 include further controls for baseline criminal attitudes and risk attitudes, whether subjects previously visited a prison and socio-demographics, including subjects' age, and dummies for males, tertiary education, and employment in a justice-related profession. Levels of significance: $^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$.

### 4.4.1 Additional Beliefs About the Justice System

The intervention could also shift other beliefs about the criminal justice system, such as perceptions of deterrence, rehabilitation, or retribution. While we do not measure these beliefs directly, we collected two related measures as secondary outcomes: trust in legal institutions and perceived procedural fairness. These measures capture participants' perceptions of whether the justice system generally functions effectively and justly. If participants concluded from their incarceration experience that prison does not deter or rehabilitate, this would likely reduce their trust in the system. Similarly, if they thought that incarceration does not result in deserved punishment, we would expect lower assessments of procedural fairness. However, as shown in Online Appendix Table B.4, the treatment had no effect on either measure ($p = 0.974$ and $p = 0.795$, respectively). While these broad measures do not completely rule out changes in more specific beliefs about deterrence, rehabilitation, or

retribution, they do highlight that the treatment did not affect confidence in the justice system. Taken together, the null effects on beliefs suggest that lived incarceration experience, rather than belief updating about prison conditions or the functioning of the justice system, is driving the shift in punitive justice attitudes.

## 4.5 Robustness: External validity, attrition, and demand effects

Next, we explore the external validity of our findings and assess their robustness by addressing attrition and experimenter demand effects.

**External validity** Individuals who volunteered for the prison test run are likely to differ from the general population. To evaluate the generalizability of our findings, we conducted an additional survey in collaboration with CINT, using an online sample (N=497) broadly representative of the Swiss voting population with respect to age, gender, and income.[28] While we cannot account for all possible dimensions, we cover key differences in attitudes and personality (e.g., punitive attitudes, beliefs about wellbeing in prison, willingness to take risk, trust in institutions) as well as socio-economic characteristics (e.g., education, working in justice-related profession). Online Appendix Table B.5 compares the experimental sample with the representative sample from the Swiss voting population in terms of socio-economic characteristics and all baseline attitudes measured in our experiment.[29] It is reassuring that our sample does not differ significantly from the representative sample in the two main outcome dimensions: punitive attitudes and anticipated wellbeing in prison.

However, our experimental sample differs markedly from the general population in several characteristics. At the same time, some of these characteristics arguably make our participants a particularly relevant group: individuals who are highly educated, have greater trust in institutions, or work in justice-related professions are also more likely to be politically active or to be in positions to influence the criminal justice system. We nevertheless assess the robustness of our results by reweighting our sample to better reflect the Swiss voting population. Specifically, we use entropy balancing, as described in Hainmueller (2012), to

[28] These data come from a separate online experiment conducted as part of a bachelor's thesis. In the study, participants in the treatment group were shown statements from prisoners about the most stressful aspects of prison life. For the purposes of the present analysis, we rely solely on data from participants in the control group.

[29] The wording of the survey questions on punitive attitudes differs slightly between the experimental sample at baseline and the representative sample. We asked subjects in our experiment "How strongly do you support strict and harsh prosecution of criminals?", whereas in the representative sample, the question was phrased "How strongly do you support strict criminal prosecution and harsh sentencing?".

Table 6: Regression Analysis; Effects on Punitive Attitudes; Using Weights to Reflect the Swiss Voting Population

| | Punitive Justice Index | | Punitive Justice Donation | Punitive Attitudes |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Treatment Group | -0.650*** | -0.664*** | -0.694*** | -0.561*** |
| | (0.171) | (0.157) | (0.138) | (0.188) |
| Strata FE | ✓ | ✓ | ✓ | ✓ |
| Baseline Outcomes | ✓ | ✓ | ✓ | ✓ |
| Criminal and Risk Attitudes | × | ✓ | ✓ | ✓ |
| Previous Prison Visit | × | ✓ | ✓ | ✓ |
| Socio-Demographics | × | ✓ | ✓ | ✓ |
| Adjusted R-squared | 0.564 | 0.577 | 0.470 | 0.593 |
| Observations | 418 | 418 | 418 | 418 |

*Note*: This table shows the effect of treatment assignment on punitive attitudes in the endline survey, using weighted OLS regressions (i.e., the sample is re-weighted based on the means and covariance of the 11 baseline variables shown in Table B.5). Standard errors (in parentheses) are corrected for clustering of the error term for individuals who spent time together in the same cell. In columns 1 and 2, the dependent variable is the punitive justice index. In column 3 and 4, the dependent variables are punitive justice donations and punitive attitudes, respectively. All three dependent variables are normalized to a control group mean of 0 and standard deviation of 1. All regressions include baseline outcomes and strata fixed effects, i.e., a dummy for subjects working in the public sector. Baseline outcomes include measures of anticipated wellbeing in prison, punitive attitudes, trust in institutions, and perceptions of procedural fairness. Columns 2 to 4 include further controls for baseline criminal attitudes and risk attitudes, whether subjects previously visited a prison and socio-demographics, including subjects' age, and dummies for males, tertiary education, and employment in a justice-related profession. Levels of significance:*$p < .1$, **$p < .05$, ***$p < .01$.

reweight the experimental sample according to the mean and variance of all the eleven variables measured in the representative sample. Table 6 and Online Appendix Table B.6 present the results for the effects on punitive attitudes and beliefs about wellbeing in prison, respectively, using the reweighted sample. The findings indicate that when we reweight the sample to more accurately reflect the characteristics of the general population, the effect becomes even larger and remains statistically significant ($p \leq 0.003$, t-tests). By contrast, the treatment effect on beliefs about the wellbeing in prison remains small and insignificant even after re-weighting ($p \geq 0.884$, t-tests).

**Attrition** The survey response rate is lower in the control group (54.7%) than in the treatment group (69.2%, $p < 0.001$, Wilcoxon rank-sum test). However, as emphasized by Ghanem et al. (in press) differential attrition rate tests may overreject internal validity, as equal attrition rates are not a necessary condition for internal validity. Because we have access to baseline outcome measures, we implement their recommended outcome-specific

tests for attrition bias. Reassuringly, we fail to reject internal validity for the respondent subpopulation ($p = 0.416$ for punitive attitudes and $p = 0.803$ for wellbeing in prison) as well as for the full study population ($p = 0.440$ for punitive attitudes and $p = 0.406$ for wellbeing in prison). We nevertheless compute bounds for the treatment effect as a more conservative robustness check in Online Appendix Table B.7 and Online Appendix Table B.8. Using Lee (2005)'s approach, we find that the bounds are relatively wide, making them less informative for our conclusions. However, we leverage our sequential data collection efforts by applying the approach proposed by Behaghel et al. (2015). In particular, assuming monotonicity in the relationship between subjects' reluctance to respond and the number of attempts required, we use the number of reminders sent during the endline survey to estimate tighter bounds. We find that the treatment effect bounds range from -0.42 (lower bounds) to -0.20 (upper bounds), depending on the measure of punitive attitudes. All lower bounds are statistically significant on the 1% level, but the upper bounds reach marginal significance only for punitive attitudes ($p = 0.063$, t-tests).

Second, to further address potential concerns about attrition, we make use of the fact that both outcome dimensions are measured before and after the intervention. This enables us to conduct a difference-in-differences analysis, which remains valid as both treatment and control groups follow parallel trends, even if they are not perfectly balanced.[30] The results from this within-subjects approach fully align with our main analysis. As shown in Table 7, the treatment reduces punitive attitudes by approximately 0.3 standard deviations ($p = 0.000$, t-tests), while there is no significant effect on anticipated wellbeing in prison ($p \geq 0.669$).[31] Taken together, these results suggest that our findings are unlikely to be driven by selective attrition.

**Demand effects** Participants might have inferred the study's hypothesis or perceived cues about socially desirable responses and adjusted their responses accordingly (Zizzo, 2010; De Quidt et al., 2018). While we cannot completely rule out such demand effects, two pieces of evidence suggest they are unlikely to explain our results. First, although the treatment reduced punitive attitudes sizably, we find no effect on beliefs about wellbeing in prison. If participants were simply trying to provide responses they believed the experimenters wanted,

[30] Note that, the wording of the punitive attitudes question slightly differs between the baseline and endline survey. At baseline we asked subjects "How strongly do you support strict and harsh prosecution of criminals?" and in the endline survey participants responded to the question "How strongly do you support strict criminal prosecution and harsh sentencing?" The response options and scale was identical across survey waves.

[31] The minimal detectable effect for a within-subjects approach with our sample is 0.13 standard deviations.

Table 7: Within-Subject Analysis: Punitive Attitudes and Citizens' Anticipated Wellbeing in Prison

| | Punitive Attitudes | | Wellbeing in Prison | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Treatment Group | -0.005 | -0.002 | 0.046 | 0.031 |
| | (0.099) | (0.093) | (0.110) | (0.108) |
| After Test Run | -0.058 | -0.040 | -0.072 | -0.045 |
| | (0.045) | (0.047) | (0.053) | (0.054) |
| Treatment Group x After Test Run | -0.296*** | -0.298*** | -0.000 | 0.039 |
| | (0.083) | (0.081) | (0.103) | (0.102) |
| Strata FE | ✓ | ✓ | ✓ | ✓ |
| Baseline Outcomes | × | ✓ | × | ✓ |
| Criminal and Risk Attitudes | × | ✓ | × | ✓ |
| Previous Prison Visit | × | ✓ | × | ✓ |
| Socio-Demographics | × | ✓ | × | ✓ |
| Adjusted R-squared | 0.010 | 0.213 | -0.003 | 0.067 |
| Observations | 836 | 832 | 832 | 832 |

*Note*: This table presents difference-in-difference estimates for the main outcomes, using OLS regressions. 'Treatment Group' is a dummy for treatment assignment. 'After Test Run' is a dummy for answers in the endline survey compared to the baseline survey, and 'Treatment Group × After Test Run' is the corresponding interaction term. Standard errors (in parentheses) account for dependencies across multiple observations for the same individual and for clustering among individuals who spent time together in the same cell. The dependent variables are punitive attitudes (columns 1 and 2) and citizens' anticipated wellbeing in prison (columns 3 and 4). All dependent variables are normalized to a control group mean of 0 and standard deviation of 1. All regressions include strata fixed effects, i.e., a dummy for subjects working in the public sector. Columns 2 and 4 further include otucomes measured at baseline (i.e., citizens' anticipated wellbeing in prison in columns 1 and 2, respectively punitive attitudes in columns 3 and 4), as well as baseline measures of trust in institutions, and perceptions of procedural fairness, criminal attitudes. risk attitudes, whether subjects previously visited a prison, and socio-demographics (sbjects' age, and dummies for males, tertiary education, and work in justice-related professions). Levels of significance: $^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$.

we would expect both outcomes to shift in the same direction. Second, our primary outcome consists of a self-reported measure and a donation task with real financial consequences. Because the behavioral measure should be less susceptible to social desirability bias and demand effects (Bursztyn et al., 2025), the fact that treatment effects are very similar across both outcomes further reduces the plausibility of this explanation.

A different concern is that participants may have engaged in expressive responding (Malka & Adelman, 2023): They may have exaggerated their responses to signal their prior policy stance to the experimenter. In this case, the treatment would not change punitive attitudes but would instead induce individuals to express stronger versions of their pre-existing attitudes. However, when we split the sample at the median of baseline punitive attitudes, both initially "lenient" and "tough" subjects become less punitive by a similar amount (see low

and high punitive attitudes subgroups in Online Appendix Figure A.4). If the treatment primarily triggered expressive responding, we would expect treated lenient subjects to become less punitive and treated tough subjects to become more punitive relative to their controls (i.e., polarization rather than the parallel reductions we observe).

# 5 Discussion and conclusion

How does incarceration experience affect societal attitudes toward criminal justice policies? In a randomized controlled trial, we offered regular citizens the opportunity to experience up to two days of incarceration in a newly constructed prison, designed to simulate real-life conditions and replicate the procedures experienced by actual inmates. Subjects who received the chance to experience incarceration firsthand became substantially less supportive of harsh criminal justice policies and donated more money to an organization advocating for a more moderate criminal justice system. Because participants were incarcerated alongside fellow citizens rather than convicted criminals and could terminate their incarceration upon request, our estimates likely represent a lower bound of the effects of firsthand incarceration experience. These findings remained robust after reweighting our sample to better reflect the demographics and attitudes of the general voting population. While the outcomes we examine are not direct measures of policy change, they constitute meaningful, policy-relevant indicators of public sentiment. Public attitudes are considered a key driver of criminal justice policy (Enns, 2016; Brace & Boyea, 2008; Pratt, 2007). For example, using representative survey data and administrative voting records from Switzerland, we find that attitudes toward law and order significantly predict the outcome of a popular initiative to amend the law by introducing lifelong custody for certain offenders.[32] Beyond public attitudes, donations to advocacy organizations represent active political engagement and are arguably less susceptible to social desirability bias than survey responses (Haaland et al., 2023).

What mechanisms drive these changes in attitudes toward criminal justice policies? One possibility is that citizens underestimate the harshness of life behind bars and update their

[32] We use data from the nationally representative VOX surveys covering the years from 1996 to 2016 (GfS-Forschungsinstitut, 2017), which include the following item measuring attitudes toward law and order *"Would you prefer a Switzerland where peace and order are given little emphasis, or a Switzerland where peace and order are strongly emphasized? [1 = little emphasis, ..., 6 = strong emphasis]."* We averaged responses at the community level (including only communities with at least 30 survey responses) and at the cantonal level, and link these averages to administrative voting records on support for a 2004 popular initiative proposing lifelong custody for untreatable and dangerous sex offenders and violent criminals. Regardless of the level of aggregation, public attitude significantly predicts voting outcomes. Results are similar when using self-reported voting outcomes instead of administrative data (see Online Appendix Table B.9).

beliefs following their firsthand experience. Our results, however, do not support this hypothesis: we find no significant treatment effects on beliefs about wellbeing in prison. We also find no treatment effect on beliefs capturing confidence in and fairness of the criminal justice system more generally. We further explore whether interactions with fellow participants sharing the same cell influenced punitive attitudes. To test for peer effects, we examine whether cellmates' baseline punitive attitudes correlate with the endline attitudes of participants who shared a cell. We find no significant correlation ($p \geq 0.139$, t-tests), and the sign of the correlation is opposite to what we would expect (Online Appendix Table B.10). While the coefficient estimate is noisy, it does not indicate that peer effects are responsible for the observed changes in attitudes toward criminal justice policies. Although our data does not allow us to determine the precise mechanism, salience theory (Bordalo et al., 2020, 2022) offers a plausible framework for interpreting our findings. In particular, it is consistent with the observation that attitudes changed even though the intervention did not affect beliefs. Firsthand incarceration experience made certain attributes of imprisonment, such as loss of control and autonomy, long waiting times, and boredom, particularly salient (see Online Appendix C), thereby plausibly increasing the attentional weight these attributes receive in the decision process relative to other aspects of incarceration, such as deterrence. By rendering the emotional burden of incarceration more prominent, the intervention may have increased the relative weight placed on inmate wellbeing. This shift could in turn have triggered greater compassion and increased support for moderate punishment, even in the absence of changes in beliefs about wellbeing in prison.

Taken together, our findings underscore the crucial role of personal experience in shaping support for criminal justice reform. Our field experiment thus contributes to the growing literature on the impact of personal experiences on beliefs and preferences, which has largely relied on observational data or laboratory settings. Moreover, by extending this literature to the domain of criminal justice, we provide novel evidence of how real-world experiences causally influence public opinion. Our findings raise a broader question about how closely public attitudes toward criminal justice are actually connected to the realities of incarceration. We show that even a brief experience can meaningfully shift policy preferences. Although we do not claim that post-experience attitudes are more valid, the shift we observe suggests that such preferences are often formed without close engagement with prison life.

From a policy perspective, we do not see widespread incarceration simulations as a practical or scalable policy intervention. However, our results suggest that experience-based

educational tools (such as guided prison tours[33] or virtual simulations) could play a role in broadening the perspectives of key stakeholders in the justice system, including judges, lawyers, and policymakers. Future work could explore how best to design and implement such interventions.

[33]For example, the Frederick Douglass Project for Justice (https://www.douglassproject.org/) offers guided prison visits in the United States. See also Murdoch (2020) and Long & Utley (2018) on the use of prison tours for educational purposes.

# References


Aebi, M. F., & Cocco, E. (2024). *Prisons and prisoners in europe 2023: Key findings of the SPACE 1 report.* Retrieved 2025-08-18, from https://wp.unil.ch/space/files/2025/04/space_i_2023_key_findings.pdf

Aebi, M. F., & Triago, M. (2021). *Space i - 2019 – council of europe annual penal statistics: Prison populations.* Retrieved 2026-01-05, from https://api.unil.ch/iris/server/api/core/bitstreams/93f9e680-dcc7-4e70-8b32-86e7604e3a84/content

Alesina, A., & Fuchs-Schündeln, N. (2007). Good-bye lenin (or not?): The effect of communism on people's preferences. *American Economic Review*, *97*(4), 1507–1528.

Andries, M., Bursztyn, L., Chaney, T., Djourelova, M., & Imas, A. (2024). *In their shoes: Empathy through information* (Working Paper No. 32569). National Bureau of Economic Research (NBER).

Apel, R. (2022). Sanctions, perceptions, and crime. *Annual Review of Criminology*, *5*(1), 205–227.

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of behavioral decision making*, *16*(3), 215–233.

Behaghel, L., Crépon, B., Gurgand, M., & Le Barbanchon, T. (2015). Please call again: Correcting nonresponse bias in treatment effect models. *Review of Economics and Statistics*, *97*(5), 1070–1080.

Bhuller, M., Dahl, G. B., Løken, K. V., & Mogstad, M. (2020). Incarceration, recidivism, and employment. *Journal of political economy*, *128*(4), 1269–1324.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2020). Memory, attention, and choice. *Quarterly Journal of Economics*, *135*(3), 1399–1442.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2022). Salience. *Annual Review of Economics*, *14*(1), 521–544.

Brace, P., & Boyea, B. D. (2008). State public opinion, the death penalty, and the practice of electing judges. *American Journal of Political Science*, *52*(2), 360–372.

Bryant, P. T., & Morris, E. (1998). What does the public really think? *Corrections Today*, *59*(1), 26–29.

Bursztyn, L., Haaland, I. K., Röver, N., & Roth, C. (2025). *The social desirability atlas* (Working Paper No. 33920). National Bureau of Economic Research (NBER).

Calaway, W. R., Callais, T., & Lightner, R. (2016). Going to prison: The effect of a prison tour on students' attitudes toward punitiveness. *Journal of Criminal Justice Education*, *27*(3), 432–448.

Cecil, D. K., & Leitner, J. L. (2009). Unlocking the gates: An examination of MSNBC investigates–lockup. *The Howard Journal of Criminal Justice*, *48*(2), 184–199.

Conlon, J. J., Mani, M., Rao, G., Ridley, M., & Schilbach, F. (2022). *Not learning from others* (Working Paper No. 30378). National Bureau of Economic Research (NBER).

Cotofan, M., Cassar, L., Dur, R., & Meier, S. (2023). Macroeconomic conditions when young shape job preferences for life. *Review of Economics and Statistics*, *105*(2), 467–473.

De Quidt, J., Haushofer, J., & Roth, C. (2018). Measuring and bounding experimenter demand. *American Economic Review*, *108*(11), 3266–3302.

Doble, J. (1995). Crime and corrections: The views of the people of oklahoma. *Journal of the Oklahoma Criminal Justice Research Consortium*, *2*, 4–29.

Doble, J. (2002). Attitudes to punishment in the u.s.: Punitive and liberal opinions. In J. V. Roberts & M. Hough (Eds.), *Changing attitudes to punishment: Public opinion, crime and justice* (pp. 148–162). Willan Publishing.

Dušek, L., & Traxler, C. (2022). Learning from law enforcement. *Journal of the European Economic Association*, *20*(2), 739–777.

Enns, P. K. (2016). *Incarceration nation*. Cambridge University Press.

Fair, H., & Walmsley, R. (2021). *World prison population list.* Retrieved 2025-08-18, from https://www.prisonstudies.org/sites/default/files/resources/downloads/world_prison_population_list_13th_edition.pdf

Farkas, S. (1993). Pennsylvanians prefer alternatives to prison. *Overcrowded Times*, *4*(2), 1–13.

Foster, A. D., & Rosenzweig, M. R. (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of political Economy*, *103*(6), 1176–1209.

GfS-Forschungsinstitut. (2017). *Vox-analysen eidg. urnengänge.* (Zusammenarbeit mit den politikwissenschaftlichen Instituten der Universitäten Bern, Genf und Zürich. Dataset)

Ghanem, D., Hirshleifer, S., & Ortiz-Becerra, K. (in press). Testing attrition bias in field experiments. *Journal of Human Resources*.

Giuliano, P., & Spilimbergo, A. (2025). Aggregate shocks and the formation of preferences and beliefs. *Journal of Economic Literature*, *63*(2), 542–597.

Haaland, I., Roth, C., Stantcheva, S., & Wohlfart, J. (2025, December). Understanding economic behavior using open-ended survey data. *Journal of Economic Literature*, *63*(4), 1244–80.

Haaland, I., Roth, C., & Wohlfart, J. (2023). Designing information provision experiments. *Journal of Economic Literature*, *61*(1), 3–40.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political analysis*, *20*(1), 25–46.

Harney, J. (2023). The power of empathy: Experimental evidence of the impact of perspective-focused interventions on support for prison reform. *Criminal Justice Policy Review*, *34*(1), 20–42.

Haselhuhn, M. P., Pope, D. G., Schweitzer, M. E., & Fishman, P. (2012). The impact of personal experience on behavior: Evidence from video-rental fines. *Management Science*, *58*(1), 52–61.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological science*, *15*(8), 534–539.

Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in cognitive sciences*, *13*(12), 517–523.

Herz, H., & Taubinsky, D. (2018). What makes a price fair? an experimental study of transaction experience and endogenous fairness views. *Journal of the European Economic Association*, *16*(2), 316–352.

Hoffman, K. M., Trawalter, S., Axt, J. R., & Oliver, M. N. (2016). Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. *Proceedings of the National Academy of Sciences*, *113*(16), 4296–4301.

Hough, M., & Roberts, J. (2005). *Understanding public attitudes to criminal justice*. McGraw-Hill Education (UK).

Hough, M., & Roberts, J. V. (2012). Public opinion, crime and criminal justice. In M. Maguire, R. Morgan, & R. Reiner (Eds.), *The oxford handbook of criminology* (5th ed., pp. 279–300). Oxford University Press.

Institute for Crime and Justice Policy Research. (2024). *World prison brief.* Retrieved 2025-08-18, from https://www.prisonstudies.org/country/switzerland

Jennings, W., Farrall, S., Gray, E., & Hay, C. (2017). Penal populism and the public thermostat: Crime, public punitiveness, and public policy. *Governance*, *30*(3), 463–481.

Kazemian, L., & Galleguillos, S. (2025). A global comparison of long prison sentences. *Journal of Criminal Justice*, *96*, 102341.

Lee, D. S. (2005). *Training, wages, and sample selection: Estimating sharp bounds on treatment effects* (Working Paper No. 11721). National Bureau of Economic Research (NBER).

Lochner, L. (2007). Individual perceptions of the criminal justice system. *American Economic Review*, *97*(1), 444–460.

Long, J. J., & Utley, M. E. (2018). An afternoon spent behind bars: The impact of touring a correctional facility on student learning. *Journal of Correctional Education (1974-)*, *69*(3), 32–48.

Malka, A., & Adelman, M. (2023). Expressive survey responding: A closer look at the evidence and its implications for american democracy. *Perspectives on Politics*, *21*(4), 1198–1209.

Malmendier, U. (2021). Exposure, experience, and expertise: Why personal histories matter in economics. *Journal of the European Economic Association*, *19*(6), 2857–2894.

Malmendier, U., & Wachter, J. A. (2024). Memory of past experiences and economic decisions. In *The oxford handbook of human memory, two volume pack: Foundations and applications.* Oxford University Press.

Miller, J., & Maniadis, Z. (2012). *The weight of personal experience: an experimental measurement* (Working Paper No. 452). Innocenzo Gasparini Institute for Economic Research (IGIER).

Mitchell, B. J., & Roberts, J. (2011). Public attitudes towards the mandatory life sentence for murder: putting received wisdom to the empirical test. *Criminal Law Review*, *6*, 456–468.

Murdoch, D. J. (2020). Examining student participation in pedagogical correctional tours: "educational and a lesson for the soul and humanity"? *Criminology, Criminal Justice, Law & Society*, *21*(2), 57–75.

National Commission for the Prevention of Torture (NCPT). (2017). *Report of the national commission for the prevention of torture NCPT on the third universal periodic review of switzerland.* Retrieved 2025-08-18, from https://uprdoc.ohchr.org/uprweb/downloadfile.aspx?filename=4245&file=EnglishTranslation

National Commission for the Prevention of Torture (NCPT). (2022). *Report of the national commission for the prevention of torture on the forth universal periodic review of switzerland.* Retrieved 2025-08-18, from https://www.sem.admin.ch/dam/nkvf/it/data/Stellungnahmen/220713-stn-upr-switzerland.pdf.download.pdf/220713-stn-upr-switzerland.pdf

Petrosino, A., Turpin-Petrosino, C., Hollis-Peel, M. E., & Lavenberg, J. G. (2013). 'scared straight' and other juvenile awareness programs for preventing juvenile delinquency. *Cochrane Database of Systematic Reviews*(4).

Philippe, A. (2024). Learning by offending: How do criminals learn about criminal law? *American Economic Journal: Economic Policy*, *16*(3), 27–60.

Pratt, J. (2007). *Penal populism.* Routledge.

Roberts, J. V., Bild, J., Pina-Sánchez, J., & Hough, M. (2022). *Public knowledge of sentencing practice and trends.* Retrieved 2025-08-18, from https://www.researchgate.net/profile/Jose-Pina-Sanchez/publication/360438799_Public_knowledge_of_sentencing_practice_and_trends/links/652f8e136725c324011486a1/Public-knowledge-of-sentencing-practice-and-trends.pdf

Roberts, J. V., & Hough, M. (2005). The state of the prisons: Exploring public knowledge and opinion. *The Howard Journal of Criminal Justice*, *44*(3), 286–306.

Roth, C., & Wohlfart, J. (2018). Experienced inequality and preferences for redistribution. *Journal of Public Economics*, *167*, 251–262.

Simonsohn, U., Karlsson, N., Loewenstein, G., & Ariely, D. (2008). The tree of experience in the forest of information: Overweighing experienced relative to observed information. *Games and Economic Behavior*, *62*(1), 263–286.

Staton, L. J., Panda, M., Chen, I., Genao, I., Kurz, J., Pasanen, M., ... Cykert, S. (2007). When race matters: Disagreement in pain perception between patients and their physicians in primary care. *Journal of the National Medical Association*, *99*(5), 532.

Tillmann, R., Voorpostel, M., Antal, E., Dasoki, N., Klaas, H., Kuhn, U., ... Ryser, V.-A. (2022). The swiss household panel (SHP). *Journal of Economics and Statistics (Jahrbuecher fuer Nationaloekonomie und Statistik)*, *242*(3), 403–420.

Todd, K. H., Deaton, C., D'Adamo, A. P., & Goe, L. (2000). Ethnicity and analgesic practice. *Annals of emergency medicine*, *35*(1), 11–16.

Van Dijk, J., Van Kesteren, J., & Smit, P. (2007). *Criminal victimisation in international perspective.* Retrieved 2025-08-18, from https://repository.wodc.nl/bitstream/handle/20.500.12832/1204/ob257_full_text_tcm28-69406.pdf

Washington, E. L. (2008). Female socialization: How daughters affect their legislator fathers' voting on women's issues. *American Economic Review*, *98*(1), 311–332.

Weber, E. U., Shafir, S., & Blais, A.-R. (2004). Predicting risk sensitivity in humans and lower animals: Risk as variance or coefficient of variation. *Psychological review*, *111*(2), 430.

Wilson, D., & O'Sullivan, S. (2004). *Images of incarceration: Representations of prison in film and television drama*. Waterside Press.

World Health Organization: Regional Office for Europe. (2023). *Status report on prison health in the WHO european region 2022*. Retrieved 2025-08-18, from https://iris.who.int/bitstream/handle/10665/365977/9789289058674-eng.pdf?sequence=1

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, *13*(1), 75–98.