Use of Reentry Support Services and Recidivism: a Field Experiment Varying Dosage

Marco Castillo* Sera Linardi† Ragan Petrie‡§

December 8, 2025

Abstract

Many previously incarcerated individuals are rearrested following release from prison. We investigate whether encouragement to use reentry support services reduces rearrest. Field experiment participants are offered a monetary incentive to complete different dosages of visits, either three or five, to a support service provider. The incentive groups increased visits, and one extra visit reduces rearrests three years after study enrollment by 8.6 percentage points. Black participants are more likely to take up treatment and benefit the most from visits. The study accounts for first-stage heterogeneity and documents sizable heterogeneous treatment effects.

Keywords: recidivism, reentry support services, dosage effects, field experiment

JEL Codes: K42, C93

*Castillo: Department of Economics, Texas A&M University; IZA; CESifo Research Network; marco.castillo@tamu.edu

†Linardi: Graduate School of Public and International Affairs, University of Pittsburgh; linardi@pitt.edu

‡Petrie: Department of Economics, Texas A&M University; CESifo Research Network; rpetrie@tamu.edu

§We thank our partner aftercare service provider for their support and willingness to collaborate with us on the study. The research was funded by the Bernard and Audre Rapoport Foundation, a College of Liberal Arts at Texas A&M University Seed Grant and the University of Pittsburgh Central Research Development Funds. The study has IRB approval from University of Pittsburgh (PRO17020307) and Texas A&M University (IRB2018-0488D). The study is pre-registered at the American Economics Association RCT Registry (AEARCTR-0003375). The registration includes the study description and pre-analysis plan.

1 Introduction

Evidence suggests the period after release from incarceration is critical. Individuals often attempt to re-integrate into society without basic needs secured (Roman and Travis, 2006; Geller and Curtis, 2011), i.e. housing, clothing or a cell phone, and the connection and social support of friends and family (Denney et al., 2014). This makes the transition challenging. Of the 400,000 individuals released from state prisons in 2005, almost half are re-arrested within the first year and one-third within the first six months (Durose et al., 2014). Reentry services in the United States are intended to provide support for this transitional period, however utilization is low. The average inmate completes just under 0.6 rehabilitation programs a year in jail (Kuziemko, 2013), and less than 50% of individuals released to parole with a referral to community treatment attended any session at all (Prendergast et al., 2003).

Increased usage of reentry services might make the transition smoother, more likely to be successful and ultimately reduce recidivism. However, selection and low compliance rates prevent obtaining causal estimates and assessing their external validity. Importantly, due to the high heterogeneity in services, it is unclear what mix of services and quantity, i.e., dosage, should be used (see Manski, 2025, for a discussion of identification of optimal dosages in medical trials). In this paper, we report a field experiment that explores both dosage effects and heterogeneous treatment selection. With a partner aftercare service provider, we conduct a field experiment that encourages previously-incarcerated individuals to use support services by offering a monetary incentive when a goal of a certain number of visits is met.

Crucially, our experimental design does not require participants to use a predetermined set of services but rather those deemed most adequate by the participant. By allowing participants to choose service usage, we reduce the chance of mismatching treatments to participants. We find that the use of reentry services is effective at reducing arrests up to three years after study enrollment. We also find significant heterogeneity in the demand for reentry services and effectiveness: Black participants are more likely to visit the provider

and experience the largest reduction in rearrest rates, whereas Non-Black participants are far less responsive to treatment.

Reentry services are available through federal, state, nonprofit and privately-funded providers.[1] Services aim to help navigate post-incarceration life and can be comprehensive, i.e. providing food, clothing, identification cards, housing and job referrals, training, counseling and peer support. Several randomized controlled trials have evaluated these types of programs (Grommon et al., 2013; Cook et al., 2015; Wiegand and Sussell, 2016; D'Amico and Kim, 2018) by randomizing a treatment group of voluntary participants to receive specialized services. The estimated intent-to-treat effects on reducing recidivism are mixed. Many participants do not fully utilize all services offered or attend all sessions of a service program. Thus, treatment service programs are often not completed as designed, i.e. some participants do not "comply" with treatment.[2] Mixed findings could result from ineffective programs or incomplete treatment. Treatment dosage may be low, not by design but, because participation in the program was incomplete. Support services might be more effective at reducing recidivism if participants would use them more frequently or receive the intended dose. In our study, we fix the reentry services available and vary the dosage of services used by offering different incentivized visit goals. We then examine whether more service usage causally reduces recidivism.

The recent literature on optimal treatment assignment shows that there might be gains

[1] The National Reentry Resource Center provides information for re-entrants and listings of all Second Chance Act grantees in the U.S. to help re-entrants connect with service providers in their communities. Prisoner reentry services in the United States are made up of an informal assortment of government and nonprofit organizations, and the overall usage of reentry services is difficult to estimate (Nhan et al., 2017).

[2] Grommon et al. (2013) report compliance rates decline during the year-long program of a substance abuse treatment program. In the first phase of the program, participants completed an average of 6.5 hours a week of the 10 hour planned treatment. In review papers, Doleac et al. (2019); Doleac (2023) notes several possible reasons for mixed ITT results. Individual programs cannot overcome the large barriers to successful reentry. Wrap-around services may be challenged to deliver multiple programs. If case management is part of these services, it implies a higher frequency of personal contact and potentially more scrutiny and monitoring. D'Amico and Kim (2018) find a slight increase in the total number of rearrests for those in the treatment group, likely due to more intensive case management. Increasing the level of supervision for probationers and parolees either has no impact on the likelihood of committing new offenses (Lane et al., 2005; Barnes et al., 2012; Boyle et al., 2013; Hyatt and Barnes, 2017) or increases recidivism (Lee, 2022). Prendergast et al. (2015) uses monetary incentive and finds no effects on attendance at a five-month long community substance abuse treatment among prisoner and parolees.

by improving the matching of treatments to different populations (see Kitagawa and Tetenov, 2018; Athey and Wager, 2021) and by allowing some participants to choose to be treated or not (Ida et al., 2022). Heterogeneity in selection into treatment and treatment heterogeneity in observables and unobservables might explain disparate effects of reentry policies. As shown by Abadie et al. (2024), heterogeneity in treatment participation might also be of consequence in estimating the average causal response parameters. Our experimental design addresses the issue of heterogeneity and treatment selection by recruiting participants from diverse backgrounds and encouraging reentry services without limiting them to a predetermined set of options. The advantage of this approach is that it uncovers marked differences in demand for services by race. We then explore the effect of different service usage on reducing recidivism.

The field experiment is implemented in partnership with a Pittsburgh, PA support service provider for the previously incarcerated. Individuals are recruited into the study at the provider's site and randomized into three groups that vary the number of required visits to the provider for a fixed monetary incentive. The Control group has no required visits or monetary incentive. Participants in the Easy treatment need to complete three visits to receive a $50 incentive, and those in the Hard treatment must complete five visits to receive the $50 incentive. The number of visits chosen for the Easy and Hard treatments were calibrated to historical data from our partner and were chosen to be attainable and potentially encourage enough service usage to facilitate a more successful transition to non-prison life. The design focuses on visits, rather than use of particular services, to allow participants to choose services that best meet their needs. We vary the dosage, rather than the incentive, because this gives us variation in visits and we can test the causal impact of service usage on recidivism.

The main outcomes we examine are number of visits to the provider within a year of study enrollment and probability of arrest up to three years after enrollment. We link data from our intake survey with the participants, administrative data on participants' visits

from our partner and publicly-available, administrative data on arrests from the state of Pennsylvania. These data allow us to examine frequency and timing of visits, services used and frequency and timing of arrests by treatment and to explore heterogeneity in treatment response.

Our results show that the incentivized goal treatments did increase visits relative to the Control group. Those in the Easy group completed more visits than those in the Hard group, suggesting that the five visit goal was more difficult to achieve. The Easy treatment resulted in 32% of participants completing 3-4 visits and 18% completing 5+ visits. The proportion completing 5+ visits in the Hard treatment is similar to the Easy treatment (19%), and the proportion completing 3-4 visits (14%) is no different than the Control.

We find a significant difference by race in visits completed. The number of visits in the treated conditions is 42% larger for Black participants than Non-Black participants. These differences are apparent in each treatment arm of the study. The number of visits by race in the Control group is not significantly different (p-value = 0.1754), thus we do not find support for the existence of pre-existing differences in visits between populations.[3] In our sample, Black participants are more disadvantaged. They are slightly less educated, have more children and have fewer alternative places to go for support. But, they are no more likely to be unemployed, and thus have more free time, nor do they live closer to the service provider. While disadvantage might explain higher use of support services, we find no clear set of covariates that predict differences in visit completion by race.

The treatment effects on rearrest in the full sample are noisy for both arms of the experiment. To fully exploit the variation on visits generated by our treatments, we estimate the effect of visits on rearrests using 2SLS and treatment assignment as instruments. Because treated Black participants increase their visits far more than treated non-Black participants, we follow advice by Abadie et al. (2024) to address this heterogeneous first stage. In particular, we interact treatment assignment with covariates. In the full sample, we find that one

[3]We confirm this using record data from a random sample of 200 non-participants (p-value = 0.1931).

extra visit reduces rearrests by 8.6 (s.e. 4.3) percentage points. Black participants experience a 12.5 (s.e. 5.4) percentage point reduction in rearrest for one extra visit. Given the lower response to treatment by Non-Blacks, we have little power to assess the effect of visits on their rearrests. We verify our results are robust to biases due to many weak instruments (Kolesár, 2013), and we reproduce our main findings using alternative measures of long-term outcomes, i.e., criminal offense, summary, misdemeanor, or felony charges. Our study shows heterogeneity in treatment take-up translates into heterogeneity in treatment effects.

Participants choose the services they want to use during a visit, and we explore which services are more effective. While completing any visit is negatively associated with rearrest, we find that visits that included counseling, mentoring or peer or job-related support are associated with the largest decline in arrests. Black participants are more likely to use these services than non-Black participants. While we cannot distinguish whether Black participants chose the most effective treatments or these treatments are more effective for Black participants, these findings suggest more work is needed in improving treatment assignment.

A crucial question is whether dosage itself matters. This is, in general, difficult to elucidate in a heterogeneous treatment effect model with unordered treatments (Mogstad and Torgovitsky, 2024). We develop, and test, conditions under which our experimental design identifies dosage effects and confirm that a 3-4 visit dosage is effective in reducing recidivism. Black participants who complete this dosage are 19.6 (s.e. 7.1) percentage points less likely to be rearrested. Completing 5+ visits has no significant effect, nor does either dosage have a significant effect for Non-Blacks. Our finding that treatment effects on recidivism vary by reentry service visit dosage echoes findings in Rose and Shem-Tov (2021) who look at incarceration effects and exploit discontinuities in sentencing guidelines. In their study, incarceration reduces recidivism, but the effect diminishes with the dose of imprisonment. Our focus is on the critical period of reentry, and we use experimental variation and derive conditions to identify causal dosage effects. Our findings provide insights on whether the effect of different dosages is due to nonlinear treatment effects or selection into treatments.

We find evidence consistent with the latter: those completing the low dosage are likely to be of a different response type, and differ on some observable characteristics, from those completing the high dosage.

Our first contribution is to show the importance and impact of heterogeneous response to treatment in the context of reentry services aimed to reduce recidivism. We identify a heterogeneous first-stage demand for reentry services by race. This has not been examined before in previous experimental studies on reentry services (Grommon et al., 2013; Cook et al., 2015; Wiegand and Sussell, 2016; D'Amico and Kim, 2018) and only examined to a limited extent in the broader literature on crime.[4] Differences in demand for treatment by race may be difficult to assess in some studies due to targeting or because compliance is not an issue.[5] By uncovering differential first-stage responses to treatment (i.e. number of visits), and accounting for this in our main analysis of treatment on re-arrest, our study illustrates barriers some populations face to access beneficial services. Identifying the nature of these barriers is an important next task.

The second contribution of our paper is establishing the importance of dosage effects in reentry services on reducing recidivism.[6] We explicitly incorporate dosage into our design, thus allowing a test of the causal effect of visit dosage on recidivism. By taking into account noncompliance and heterogeneous response to treatment, we establish the effect of an additional visit on reducing recidivism. We also establish that a small dosage of visits is more effective at reducing recidivism than a large dosage. This result is consistent with lower dosages attracting those who benefit most from the services provided. Indeed, we find evidence that Black participants positively select into treatment.

[4]Differences in compliance by race are scant in a restorative justice program (Shem-Tov et al., 2024) and non existence in Michigan's IGNITE program (Alsan et al., 2024). Compliance is often addressed in the full sample, i.e. Heller et al. (2017)'s study on the effect of the Becoming a Man program, but not separated by race. Other literatures have examined compliance differences by race, i.e. Kling et al. (2007) find a larger proportion of Black families among compliers in a housing voucher program (see Table A1 in that paper).

[5]For instance, participants in Bhatt et al. (2024) are ninety-seven percent Black. Quasi-experimental studies using judge designs rarely face a compliance problem.

[6]Research on within-prison programs finds that the timing of programs relative to release date can affect recidivism (Arbour, 2022; Papp et al., 2021).

By setting different visit dosage goals, our study design allows us to examine the effect of these goals and speak to the goal-setting literature.[7] Goal-setting theory (Locke and Latham, 1990) is based on the premise that conscious goals affect action (Ryan, 1970). Goals need to be challenging, but attainable, to motivate completion of a task (Zimmerman et al., 1992). If goals are too much of a stretch, they will not be achieved (Sitkin et al., 2017; Markovitz, 2012; Ordóñez et al., 2009), and monetary stakes can also influence goal achievement (Corgnet et al., 2015; Goerg and Kube, 2012). It is not clear-cut how to set challenging, yet attainable, goals, and a certain goal that works well for one individual might not for another. The heterogeneity we observe in the first-stage confirms the differential impact of a certain goal for ex-inmates in a reentry support services setting.

In terms of generalizability of our empirical results (List, 2020), our sample is a subset of clients to a large aftercare service provider in Pittsburgh who visited the facility during an 18-month period. Almost all clients invited to enroll in the study did. Since we have administrative records on visits, we have no attrition on this outcome. While we cannot guarantee complete administrative records of arrests, as we only have records for Pennsylvania, we think that we have most because evidence suggests our sample has limited geographical mobility. In terms of scaling our findings, the effect of visit dosage on rearrest may change as our encouragement design is extended to other populations and settings. This is because, while aftercare service providers likely focus on similar services, aftercare services might differ in content across settings. We view the visit dosage effects as an initial insight (List, 2020). Replications need to be completed to understand if the visit dosage effects apply to other previously-incarcerated populations and well as other service providers in other settings.

The paper proceeds as follows. Section 2 describes the field experiment design, our partner support service setting and field implementation. Section 3 describes the data sets used and linked for the analysis. Section 4 presents summary statistics of our participant

[7]Some studies use light-touch and nudge-type interventions to address recidivism, including mental health outreach (Batistich et al., 2021) and reminders (Fishbane et al., 2020). Our approach requires effort from participants via visits to the service provider and includes monetary incentives.

sample and services used. Section 5 describes our methods. Section 6 reports on intent-to-treat estimates on visits and probability of arrest. We discuss heterogeneous treatment effects, average causal response estimates and specific dosage effects on re-arrest. Section 7 presents robustness checks. Section 8 presents a benefit-cost analysis of the intervention, and Section 9 concludes.

2 Field experiment

The field experiment is designed to examine how encouragement of different dosages of reentry services affects recidivism. The design employs encouragement of service usage via increased visits, rather than random assignment to service access. It also focuses on visits, rather than providing a particular service, to allow participants to use services that meet their needs. Our design fits squarely with the study setting and our partner's requirement that use of their services is not denied to any individual. Plus, it does not withhold potentially beneficial services to anyone.

Our aim is to understand the effect of service usage on the probability of arrest. As such, we fix the monetary incentive upon reaching the goal and vary the number of visits needed to reach the goal. An alternative approach would have been to fix the visit goal and vary the incentive to reach the goal. We do not use this latter design approach for two reasons. First, we do not know what would be the appropriate goal in this setting. Second, we want to understand how the dosage of visits affects rearrests. This requires that treatments exogenously alter the required number of visits. Had we fixed the number of visits and altered the incentive, we would have a binary outcome, i.e. visit goal met or not. To understand dosage effects, we need variation in visits.[8]

[8]By fixing the payment ($50) and varying the visits to receive the payment (i.e. 3 or 5 visits), the per visit effective payment varies across treatments (i.e. $17, $10). An alternative design that varies dosage but keeps the effective per visit payment constant would be to fix the payment per visit (say at $10 a visit) and vary the number of visits that need to be completed to receive any payment (i.e. 3 or 5 visits). While the effective per visit payment would be constant, the final payment would differ across treatments ($30, $50). Our partner thought it important to keep final payments fixed across treatments. We cannot think of a

2.1 Aftercare services

We partnered with an aftercare service provider (ASP) in Pittsburgh, PA and employed research assistants to be on site to implement the field experiment. The ASP is a non-profit that provides comprehensive support services to previously incarcerated individuals.[9] Their reintegration program includes a variety of services, including material assistance (i.e. bus passes, use of computers and phones, clothing), informational resources, referrals, support services (i.e. peer support groups, mentoring) and guidance regarding employment, housing, other social services and obtaining an identification card. The ASP provides these services in-house and via referrals to other service providers, such as housing lists and mental health services, in the area.[10]

Our partner ASP is among the largest providers of comprehensive services to the previously incarcerated in Allegheny County, PA. They provide services on-site at their office and have a support program run within the Allegheny County Jail. Most clients come to know of our partner from referrals, the in-jail program and word of mouth. Based on our partner's records, in 2015 and 2016, prior to the field experiment, they served 811 individuals. The majority of their clients (67%) were most recently arrested in Allegheny County, and almost all in Pennsylvania. The top three services used were computer usage, bus cards, and ID assistance. Most clients (61%) came to the ASP only once, 28% came 2-4 times, and the remaining 11% visited 5 or more times. Clients who came more frequently were more likely to use the computer and obtain work-related services, whereas those who visited less frequently were unlikely to seek employment help. Activities delivered by our partner center around servicing individuals' needs, not advocating for the use of particular services. During client intake, staff ask what the individual wants to work on that day. Staff may suggest

design that would not require changing two elements across treatments.

[9] Our setting is support services for post-incarceration. This differs from studies that explore programs and interventions aimed at at-risk youth and preventing criminal behavior (i.e. Heller, 2014; Blattman et al., 2017).

[10] A full list of services provided by the ASP, and whether the use of that service during a visit would count towards a "valid" visit (for the experiment), is in Table A.1.

that clients take toiletries or browse the clothing closet.

2.2 Design

Individuals arrive at our partner ASP to use services and are invited to be part of a study on use of aftercare services and recidivism. Participants could be first-time or continuing clients.[11] Upon agreement, a research assistant orally completes the intake survey with the participant and records the responses. The survey includes questions on contact information, date of birth, most recent incarceration date and location, demographics and education.[12]

Upon completion of the intake survey, participants are randomized into one of three groups: a control group and two treatment groups.[13] All groups were presented with a business-size card, the content of which varied depending on treatment assignment. On the front of the card, there was the provider's logo, address and phone number. On the back of the card, a research assistant filled in the participant's name, an identification code and the survey date. This procedure ensures the card is unique to the participant and could not be shared.

For the two treatment groups, the front of the card also included 5 blank boxes. Each time a participant in the treatment groups visited the service provider and used at least one "valid" service, a research assistant or staff member would put their initials in the box.[14] Boxes would be initialized for each visit, not each service. The enrollment visit does not count towards the visits goal. Any valid visit after enrollment is initialized on the card. A

[11] Seventy-three percent of our sample are first-time visitors.

[12] Materials used in the field experiment and intake survey questions are in Appendix B. IRB approval is from University of Pittsburgh (PRO17020307) and Texas A&M University (IRB2018-0488D). The study was pre-registered in 2018 at the American Economics Association RCT Registry (AEARCTR-0003375) and includes the study description and analysis plan. The analysis plan is available in Appendix B.

[13] Randomization was done by the research assistant shuffling 12 blank, opaque envelopes and allowing the participant to choose one. Each envelope contained a card for one of the three treatments. There were four envelopes for each treatment group, so each treatment had an equal probability of being assigned. The participant and research assistant were blind to which envelope contained which treatment.

[14] Most services provided were counted as valid (Table A.1). However, if a participant came in to pick up a bus pass, make a personal phone call or use the computer for personal activities (i.e. checking social media, online search for something unrelated to core provider services), those did not count as a valid service. Each visit is recorded in the ASP's database, including name of the client, date of the visit and services used. Clients are required to present identification to use services.

participant who came in to the ASP once and used three valid services would get one box initialed, just as a participant who came in once and used one valid service. Once all 5 boxes were initialed, the card could be traded in for a prepaid Visa debit card loaded with $50. One of the treatment groups already had two of the boxes initialed on the card and thus only had to complete 3 visits to get the Visa card (Easy treatment). The other group had no initials on the card and thus had to complete 5 visits to get the Visa card (Hard treatment). The treatment groups needed to complete the required visits within a year to get the $50 Visa card. Initials and visits were validated using the provider's visit records prior to issuing a participant the Visa card. The research project fully utilized electronic record keeping, and participants were presented with physical cards to increase saliency. Participants were informed that lost cards would be replaced and discrepancies in visits recorded on the card would be resolved in accordance with electronic visit records.

The Control group also received a card, but it did not have the 5 blank boxes. The front of the card included a statement that encouraged the holder to use at least 5 services within the year. Images of the cards used for the Control group, Easy treatment and Hard treatment are in Appendix B. Upon completion of the intake survey and random assignment to treatment, all participants are given a bus pass from the Port Authority of Allegheny County for one week of unlimited rides (valued at $24). A bus pass is one of the most common services that brings clients to the ASP.

The two treatments, Easy and Hard, keep the encouragement of the $50 Visa card constant and changed the cost to get the reward, i.e. by having to complete 3 or 5 visits. By encouraging repeated exposure to aftercare, the participant may develop a relationship with the provider staff and with a positive peer group of other clients using provider services. The card also provides a tangible way to keep track of service usage.

2.3 Implementation

Previous incarceration is a requirement to receive services at the ASP, thus all individuals who came to the office were eligible for invitation to be part of the study. Recruitment began at the beginning of October 2018 and continued until mid-March 2020 when in-person services at the provider were shut down due to lockdowns initiated by the emergence of COVID-19.[15] Research assistants were on site at the provider for partial days, four days a week, totaling 18 hours per week, and invited any individual who came into the office or called on the phone to be part of the study. Most individuals who were invited to be part of the study agreed to do so. There were very few refusals. In total, 490 individuals were recruited to be part of our study: 166 in the Control group, 164 in the Easy treatment and 160 in the Hard treatment. Roughly 23.6% of the ASP's clients during our study period are treated.[16] We test for interference with an augmented sample consisting of our experimental sample and a sample of non-treated clients. We cannot reject the null hypothesis of no interference (Table A.2).

To determine what effect size we are powered to detect, we use the probability of being re-arrested within 20 months of release from incarceration based on a U.S. Department of Justice report on recidivism (Durose et al., 2014). The probability of re-arrest is 56%. In the analysis, we use a 36-month window as our outcome variable as this provides additional time for re-arrest that might have been disrupted due to the COVID-19 pandemic lockdowns. Under these assumptions, we are powered to detect a 15 percentage point reduction in re-arrest given our sample size (power=0.8, alpha=0.05). We note that our sample includes individuals who were released from incarceration within the previous year, as well as those who were released several years prior. Our power calculations use the arrest rate at 20 months after release, but our sample includes those who had been released more than 20

[15]The Governor of Pennsylvania closed all non-essential businesses and issued a stay-at-home order on March 19 and 23 (https://pittsburghpa.gov/mayor/covid-updates).

[16]According to the ASP reports, unduplicated clients were 361, 1442 and 1092 in 2018, 2019 and 2020 respectively.

months prior. Thus, with our sample, the effect size we are powered to detect may be slightly smaller than 15 percentage points.

The lockdowns and business closures during the early phase of the COVID-19 pandemic affected participants' ability to visit the ASP for services for several months, in addition to likely affecting the ability to commit crimes. Between March-June 2020, the provider's office was closed, but staff called existing clients weekly to check on material needs and mental health. From June 2020 through April 2021, the provider returned to offering all of its services to new and existing clients through a combination of phone calls, video-chat and in-person appointments. During all periods of office closures, staff delivered basic necessities such as food, clothing, IDs, toiletries, and cleaning supplies to a large number of clients. Peer support group meetings and the mentor program transitioned from in-person meetings to video conferences. Since April 2021, the ASP has returned to providing all services in person. Given these disruptions, we control for month and year of enrollment in our analysis. All results are robust to the inclusion of these controls.

3 Data

There are three sources of data used in the analysis. The first is the data collected from the intake survey with our recruited sample and includes treatment assignment.

The second is from the ASP's administrative data on background characteristics of the client, i.e. date of birth, sex and race of the client, and detailed information on visits. Each time a client visits the provider, the visit is recorded in their digital records, including the client's name, date of visit and services used. A visit is coded as "valid" if it was to use a provider service, such as housing search, food pantry, clothing, mentoring, support group, family services, employment services or obtaining an identification card (Table A.1 lists services and whether they counted towards a valid visit). Personal use of the phone or computer is not counted as valid. We broadly categorize services for the analysis into

short-run necessities (i.e. food, housing, clothing, transportation, identification card) and longer-run needs (employment, family services, mentoring, peer group support).

The third is public data from the Unified Judicial System (UJS) of Pennsylvania.[17] This data set includes criminal offense and arrest records in Pennsylvania. Currently, there is no single data source that combines criminal offense and arrest records across all states. Finding these data in all states would require a state-by-state search. This is not done because of limited resources and it would likely yield few additional results.[18] Thus, our outcome variable is a lower bound on the total number of criminal offenses and arrest records a participant could have. We obtained records for our sample from January 2011 through July 2023.

For our analysis, a criminal offense is defined as an encounter with law enforcement that resulted in a record entry in the UJS data, and an arrest is when the criminal offense produced an arrest. Not all criminal offenses end up in an arrest (roughly one in two offenses lead to an arrest). For example, violations, such as traffic infractions or failure to pay court fees that do not result in an arrest, do not count as an arrest. We do not distinguish between an arrest where charges were dropped or sustained. In the analysis, we use arrests that appear in the UJS data up to three years after the participant was recruited into our study.

4 Sample description

4.1 Characteristics

Table 1 provides summary statistics for our experimental sample, based on responses to the intake survey. Across all treatments, 28% of participants are female, 46% are classified by

[17] We also searched federal crimes in the Public Access to Court Electronic Records (PACER) system. These data do not have birth dates, so we could not verify that a match on name was a valid match. Thus, we do not use these data in our analysis.

[18] Evidence suggests the previously incarcerated participants in our study have limited mobility. All participants, and all clients at the ASP, resided in western Pennsylvania. Most participants (75%) resided in 28 zip codes in Pittsburgh. The remaining 25% resided in 30 different zip codes throughout southwestern Pennsylvania.

the ASP as Black, <1% as Hispanic/Multiple and 53% as White. The average age is 42.6 years, 7% are married or with a partner and have two children. The average education level is a high school diploma, 25% were employed at the time of intake and 31% knew of the ASP while they were incarcerated. The participants are a mix of recently arrested and those who have not been arrested for a while. Almost half have been arrested in the three years prior to the start of our study in 2018, with the average year of last arrest being 2015. The length of the most recent incarceration was 741 days (2 years). This means that, on average, our participants were within a year of release from incarceration when they enrolled in our study. Over half of our sample (56%) provided an institutional address at study enrollment, i.e. parole office, halfway house, rehabilitation facility, homeless shelter. Our sample is similar to the incarcerated population in Pennsylvania, but with more women, where 48% are Black, 5% are women and the average age is 44 years old.[19]

The table also reports summary statistics separately for the Control, Hard and Easy groups. The groups are balanced on all characteristics. The final table column reports the p-value of an F-test of equality of coefficients across the three groups.

4.2 Services used

Participants used a variety of services at the ASP during visits, and the composition changes over time. Figure 1 illustrates the types of services used during the first through fifth+ visit for those who completed 1, 2, 3, 4 or 5+ valid visits within one year of study enrollment.[20] Services are grouped into short-run necessities (i.e. food, housing, clothing, transportation, identification card) and longer-run needs (i.e. employment, family services, mentoring, peer group support). The figure shows that, on the first visit, 70-80% of services used are for

[19]See Pennsylvania Department of Corrections Dashboard, https://dashboard.cor.pa.gov/us-pa/narratives/prison/2

[20]Participants complete a different number of visits. If we fix the visit and look at services used during that visit, we confound the composition of participant visit types. To address this, Figure 1 fixes the participant visit type by number of visits completed, i.e. 1, 2, 3, 4, or 5+, and shows the services used by that group at the first, second, third, fourth and fifth or more visit. We only use valid visits in the figure. If a participant came to pick up a bus pass or gift card, that is not a valid visit. Services used by Black and non-Black participants are similar (Figure A.3) as they are for first-time visitors (see Figure A.4).

Table 1: Sample description and balance across treatment groups

	All	s.d.	Control	s.d.	5 visits	s.d.	3 visits	s.d.	F-test	p-val.
Female	0.28	0.45	0.29	0.45	0.30	0.46	0.25	0.43	0.56	0.57
Black	0.46	0.50	0.45	0.50	0.49	0.50	0.44	0.50	0.58	0.56
Age	42.60	10.96	42.59	10.57	42.02	11.06	43.19	11.28	0.46	0.63
Married/Partnered	0.07	0.25	0.08	0.27	0.06	0.23	0.07	0.25	0.29	0.75
Education (years)	12.24	1.65	12.25	1.73	12.15	1.64	12.30	1.59	0.34	0.71
Employed	0.25	0.44	0.26	0.44	0.20	0.40	0.30	0.46	2.28	0.10
Knew the aftercare	0.31	0.46	0.28	0.45	0.33	0.47	0.32	0.47	0.43	0.65
Number of children	2.04	2.12	2.13	2.02	2.01	2.09	1.98	2.27	0.26	0.77
Has other support	0.44	0.50	0.48	0.50	0.42	0.49	0.41	0.49	0.78	0.46
Year of last arrest	2,015.41	5.57	2,014.95	6.34	2,015.41	5.77	2,015.88	4.43	1.11	0.33
Duration of last incarceration	741.47	1,328.64	697.21	1,337.48	747.58	1,301.81	777.16	1,352.32	0.13	0.88
Arrested 36m prior to study	0.48	0.50	0.48	0.50	0.49	0.50	0.48	0.50	0.02	0.98
Institutional housing	0.56	0.50	0.55	0.50	0.56	0.50	0.57	0.50	0.10	0.90
Observations	490		166		160		164			

Notes: The last two columns report the F-test statistic and corresponding p-value of a joint test of equality across the three treatment groups (Control, Hard, Easy). Numbers are average, and s.d. is standard deviation. Female is a dummy variable for female. Black is a dummy variable for black participant. Age is in years. Married/Partnered is a dummy variable for being in a relationship. Education is in years. Employed is a dummy variable for being currently employed at the time of enrollment into the study. Knew the aftercare is a dummy variable for having heard of our partner ASP while incarcerated. Number of children is number of children. Has other support is a dummy variable for having access to other support services. Year of last arrest is year. Duration of last incarceration is length of most previous incarceration in days. Arrested 36m prior is a dummy variable for having been arrested at least once during the 36 months prior to enrollment into the study. Institutional housing is a dummy variable for providing an institutional address at study enrollment, i.e. parole office, halfway house, rehabilitation facility, homeless shelter.

short-run necessities. The main pattern is that this proportion tends to drop each subsequent visit and is especially pronounced for those who completed 5+ visits.

For those who go for 5+ visits, it is only from the fifth visit onward does demand for services that address longer-run needs, i.e. employment, become more prominent. These findings suggest that programs that focus exclusively on employment and training of the previously incarcerated need to address short-run necessities first.[21]

5 Methods

Our two primary methods to analyze the data are the following. Methods to analyze specific dosage effects are described in Section 6.3.

First, a concern with multiple treatments is that estimates of a model with multiple

[21]Studies focusing on employment assistance for the previously incarcerated find employment referrals to be ineffective (Farabee et al., 2014), provide benefits in the short run (Cook et al., 2015) and offer no significant effect on labor market outcomes (D'Amico and Kim, 2018).

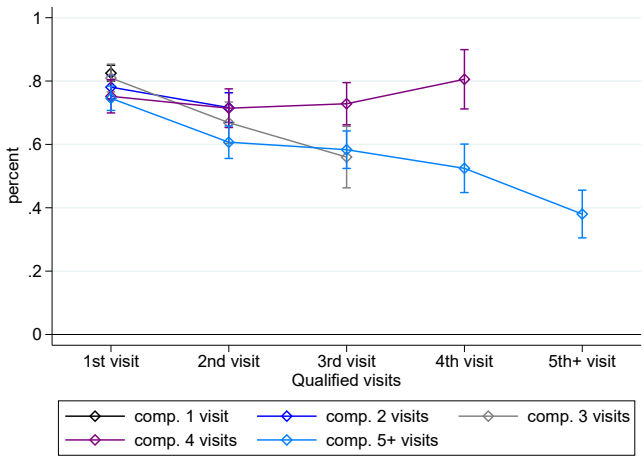


Figure 1: Percent of services used for short-run necessities (i.e. clothing, housing, food, ID, transportation) by visit number

Notes: The figure shows usage of services for short-run necessities during the first, second, third, fourth and fifth+ visit. Each line shows usage for those who completed 1, 2, 3, 4 or 5+ valid visits within one year of study enrollment. Error bars denote standard errors. Sample used for the figure are those who completed 5+ visits (n=77), 4 visits (n=47), 3 visits (n=52), 2 visits (n=98) and 1 visit (n=159).

treatments (i.e. Easy, Hard) and controls for covariates, with the Control treatment as the omitted group, will estimate a convex combination of treatment effects (Goldsmith-Pinkham et al., 2024). Thus, all regression results reported in Section 6 implement the Goldsmith-Pinkham et al. (2024) multiple treatment contamination bias robust estimates, which is a fully saturated model of treatments interacted with demeaned covariates. This approach produces uncontaminated treatment effect estimates for each treatment and works well when there is strong overlap in the covariate distribution in each treatment (which is our case). We use the Stata multe command, which implements the Goldsmith-Pinkham et al. (2024) approach, to estimate intent to treat (ITT) effects. All regressions include a rich set of controls to account for demographics, housing stability and timing of most recent arrest.

Second, we find heterogeneity by race in the response to treatment in the first stage, i.e. the number of visits completed within a year of enrollment (Section 6.1). We account for this first-stage heterogeneity in our average causal response (ACR) estimates by following a

suggestion by Abadie et al. (2024) of interacting treatment assignment and covariates in the first-stage for number of visits completed.[22] The gains in asymptotic mean squared error of instrumental variable estimators are quite large when there is substantial heterogeneity in the first-stage parameters, as is the case in our setting. Thus we account for heterogeneous response to treatment in the first stage in the 2SLS regressions. To assess the possibility of bias due to the presence of many weak instruments, we also conduct estimations using Kolesár (2013) UJIVE procedure to obtain unbiased estimates. Our main findings are not affected.

6 Treatment effects

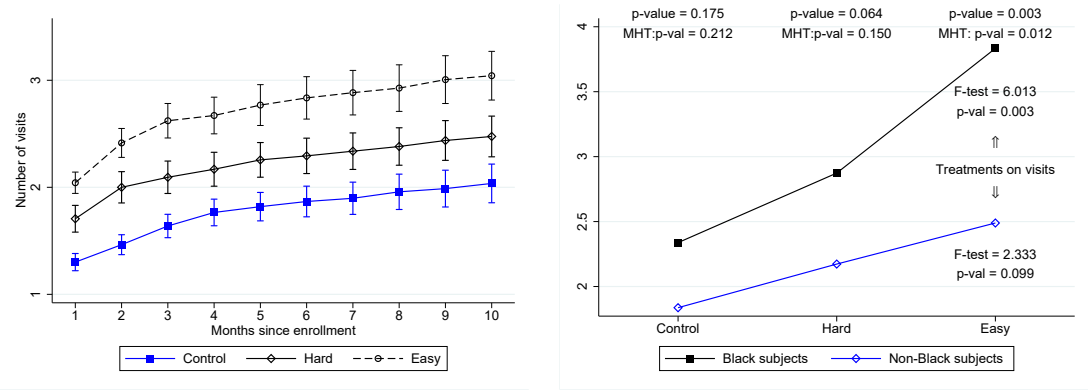
We examine treatment effects on the number of visits within one year of study enrollment and the probability of arrest within three years. To examine treatment effects on rearrest, we present intent to treat (ITT) estimates of the Easy and Hard treatments and average causal response (ACR) estimates, accounting for first-stage heterogeneity as in Abadie et al. (2024), using total number of visits. We also examine nonlinearity in dosage effects and confirm that our results are robust.

6.1 Visits

Participants had one year to complete the required number of visits to the ASP for the $50 incentive in the Easy and Hard treatments. The average number of visits completed after one year in the Easy treatment is 3.07, in the Hard treatment is 2.52 and in the Control is 2.06. These are significantly different from one another.[23] This confirms that our encouragement design worked as intended and increased the number of visits for those offered the $50 incentive compared to those who were not offered an incentive. Panel (a)

[22] Abadie et al. (2024) call this a weighted average causal effect when the second stage is heterogeneous. We present evidence of heterogeneity in the first and second stage.

[23] Hard v. Control t-test = 1.74 (p-value = 0.0828). Easy v. Control t-test = 3.4672 (p-value = 0.0006). Hard v. Easy t-test = 1.88 (p-value = 0.0616).



(a) Average visits

(b) Total visits by race

Figure 2: Treatment effects on number of visits

Notes: Both panels include visits completed within one year of study enrollment. Error bars denote standard errors. Panel (a) shows the average cumulative number of visits per treatment group at each month following study enrollment. Panel (b) shows the total number of visits a year after recruitment by treatment split by Black and Non-Black participants. P-values for difference in means tests between Black and Non-Black participants, both t-tests and adjusted for multiple hypotheses testing (MHT), are reported on the top of the figure. F-tests for differences across treatments, for each group, are reported on the right side of the figure.

of Figure 2 tracks how the average number of visits changed over the 12 months following enrollment. We confirm that visits are significantly the highest in Easy, then Hard and then Control at every month since enrollment.[24]

The treatments also reduced the time taken to complete the required visits for the $50 incentive. Participants in the Easy treatment completed 3-4 visits in half the time or less than those in the Control and Hard treatments. Those in the Hard treatment completed 5 or more visits in two-thirds the time or less than those in the Control and Easy treatments.[25] Forty seven percent of participants completed the required number of visits, or more, within two months of recruitment in the Easy treatment while 19% of participants did so in the Hard treatment. These numbers are 56% and 22% for Black and non-Black participants in the Easy treatment and 40% and 17% in the Hard treatment. We conclude that the goal in

[24]Figure A.1 shows the cumulative distribution functions of visits across treatments. First-order stochastic dominance tests show that Easy FOSD Hard (p-value = 0.625, p-value = 0.0026) FOSD Control (p-value = 0.8912, p-value = 0.0213).

[25]The median number of days to complete 3-4 visits is 41 in the Control, 42 in Hard and 17 in Easy. The median number of days to complete 5 visits is 71 in Control, 33 in Hard, and 52 in Easy.

Table 2: ITT EFFECTS ON NUMBER OF VALID VISITS WITHIN A YEAR OF ENROLLMENT

	All	Non-Black	Black
Hard	0.423	0.120	0.882
	(0.244)	(0.307)	(0.327)
Easy	1.192	0.436	1.970
	(0.319)	(0.253)	(0.628)

Notes: Number of observations is 490 for All, 225 for Black participants and 265 for non-Black participants. Standard errors in parentheses. ITT effects are calculated using Stata `multe` command, which implements Goldsmith-Pinkham et al. (2024) multiple treatment contamination bias robust estimates. We interact treatment and race to expand the number of effective treatments. Heterogeneous treatment effects are estimated jointly but presented in separate columns. Regressions control for sex, race, residence status (temporary housing, rehab center, homeless) and whether the participant had an arrest in the three years before study enrollment.

the Easy treatment is easier to achieve than in the Hard treatment.

Heterogeneous treatment effects on visits by race are shown in Panel (b) in Figure 2.[26] This shows total number of visits completed one year after recruitment by race and treatment condition. The total number of visits for Black and non-Black participants is not significantly different in the Control condition, but they differ significantly in the treatment conditions. We provide within-treatment comparisons and the overall effect of treatments on visits. For the latter, we run separate regressions of the impact of treatments on the number of visits by race. Table 2 presents these estimates.[27] Treatment assignment has a larger and more significant effect on Black participants than non-Black participants (F-stat for joint significance is 2.33 for Non-Blacks and 6.01 for Blacks). When we estimate the effect of visits on arrests in Section 6.2, we follow Abadie et al. (2024) to account for this first-stage heterogeneity.

We examine observable differences between Black and Non-Black participants to understand first-stage heterogeneity by race. Black participants are older, have fewer years of

[26]Comparison by race is included in the study pre-registration and pre-analysis plan. We do not explore heterogeneity by sex, as we originally specified, as there is insufficient variation (see Table 1).

[27]The differential effect of the Hard and Easy condition on visits is not mechanical, i.e., failing to complete five visits can be due to arrests, which is more likely as time elapses. Table A.3 shows that the results are similar if we restrict the sample to those not arrested in the first six months after recruitment.

education, more children and fewer support alternatives (Table A.4). However, they are not more likely to be unemployed (76%) and thus have more free time. It is also not a question of proximity. There is no significant difference in distance from place of residence to the ASP facilities for Black and non-Black participants (p-value = 0.2736). We find no significant difference in visits by Black and Non-Black participants in the Control group, suggesting there are no underlying differences in propensity to visit. We do find positive selection on unobservables in that Black participants are more likely to select into treatment because they have more to gain.[28]

6.2 Arrests

Our main results are presented in Table 3. This table reports intention-to-treat effects of the Easy and Hard treatments and the effect of visits on the probability of arrest within three years of study enrollment.[29]

The first two columns show estimates for all participants, and the last four columns present estimates by race.[30] Columns 1, 3 and 5 report ITT estimates of the Easy and Hard conditions. The Hard condition estimates are positive and small in magnitude. The Easy condition estimates are negative in the full sample and negative and significant in the sample for Blacks. They are positive and noisy for Non-Blacks. Column 2 shows the ACR effect in the full sample of an extra visit is a 8.6 percentage point decrease in the probability of rearrest in the three years since recruitment.[31] In other words, those completing two extra visits (the intended goal in the Easy treatment given the pre-intervention median of one visit

[28] In Appendix C, we present evidence of positive selection into visits to the ASP for Black participants based on estimates of marginal treatment effects (Heckman and Vytlacil, 2007) that use visit variation produced by our experimental design. Black participants who are least resistant to increase visits to the ASP are the ones for whom the reduction in rearrest is the highest. There is no evidence of heterogeneous marginal treatment effects for Non-Black participants.

[29] There were no constraints on when participants in the Control, Easy and Hard treatment groups could visit the ASP for services. We test for interference using the pooled experimental and a synthetic control group and find no evidence (Table A.2).

[30] These are estimated separately, however, results are similar if estimated jointly.

[31] Estimates are similar when controlling for being a first-time visitor (Table A.6). Estimates not accounting for contamination bias and first-stage heterogeneity are in Table A.5.

Table 3: Treatment effects on 3-year re-arrest rate

	All		Non-Black		Black	
	ITT	TOT	ITT	TOT	ITT	TOT
Hard	0.022		0.005		0.039	
	(0.049)		(0.064)		(0.074)	
Easy	-0.011		0.112		-0.162	
	(0.049)		(0.063)		(0.074)	
Visits		-0.086		0.175		-0.125
		(0.043)		(0.124)		(0.054)
MHT p-value				0.163		0.078

Notes: Number of observations is 490 for All, 225 for Black participants and 265 for non-Black participants. Standard errors in parentheses. ITT effects are calculated using Stata `multe` command, which implements Goldsmith-Pinkham et al. (2024) multiple treatment contamination bias robust estimates. We interact treatment and race to expand the number of effective treatments. Heterogeneous treatment effects are estimated separately. ACR estimates are the effect of the number of visits on the probability of being arrested up to 36 after enrollment. The estimates implement Abadie et al. (2024) approach to deal with first-stage heterogeneity. Regressions control for sex, race, residence status (temporary housing, rehab center, homeless) and whether the participant had an arrest in the three years before study enrollment. Missing information is handled by replacing missing items with a zero and adding a dummy for missing data. We use Stata rwolf2 to estimate the MHT corrected p-values of the ACR by race according to (Romano and Wolf, 2005).

per year) experience an average 17 percentage point decrease in the probability of arrest in the three years since enrollment. This is close to a 57% decrease from the baseline probability of arrest of 30%.[32]

Columns 4 and 6 show that the negative treatment effect of visits on arrest rates are present for Black participants only. Non-Black participants do not exhibit this effect. Given that Non-Black participants were less responsive to treatment, and thus the first stage for them is weak (Non-Black participant F-stat = 2.33), we are not powered to detect effects on rearrest. We report the results on Non-Blacks for completeness, but these estimates are likely biased and should be interpreted with more caution.[33] With these caveats, we conclude

[32]The astute reader will note that the ACR estimates are more precise than ITT. Due to multiple treatments in our setting, it is unclear that ITT results would be more precise. The ACR estimates, by construction, use all the variation across treatments in the first stage to obtain more precise estimates of the effect of an extra visit on rearrests.

[33]To assess potential bias in our estimates due to many weak instruments, we re-estimate the ACRs using Kolesár (2013)'s UJIVE method (Table A.7). Our results are very similar.

that the findings in Table 3 suggest that a potential reason for Black participants to be more likely to comply with treatment is that the treatment was beneficial.[34]

6.3 Dosage effects

We investigate the nonlinear effect of visits on rearrest by examining different visit dosages. Because visit dosage can take multiple levels and there is more than one treatment, we derive conditions to estimate dosage effects. We explain how different treatment dosages affect recidivism and the conditions required to measure this causal link, considering participant compliance.

We define compliance as taking values 0, 1, or 2 that correspond to 0-2 visits (0), 3-4 visits (1), and 5 or more visits (2). Treatment assignment is defined by variable $Z \in \{0, 1, 2\}$ for Control, Hard and Easy treatment conditions. The corresponding counterfactual outcomes are $Y_i(0, z)$, $Y_i(1, z)$, and $Y_i(2, z)$. Each participant responds to treatment assignment according to $D_i : \{0, 1, 2\} \to \{0, 1, 2\}$. The observed outcome is determined as follows:

$$Y_i = \sum_j \mathbf{1}\{Z_i = j\} Y_i(D_i(j)) \tag{1}$$

Assumption 1: i. $Y_i(D, Z) = Y_i(D)$ for all D and Z. ii. $(Y_i(D), D_i(Z))$ are independent of Z for all D and Z.

Assumption 1.i is an exclusion restriction, and Assumption 1.ii is an independence assumption. They establish that treatment assignments are valid instruments.[35] In our experiment, there are 27 different ways to respond to the incentives offered.[36] To discipline our analysis, we start by making a monotonicity assumption.

Assumption 2 (Monotonicity): $D_i(j) \geq D_i(0)$ for all i and $j = 1, 2$.

The assumption says that visits are increasing in incentives. The assumption does not claim

[34] This is consistent with positive selection for Black participants, as evidenced in marginal treatment effects (Appendix C), and discussed in footnote 28.

[35] We follow Lee and Salanié (2020)'s presentation of the necessary conditions for valid instruments.

[36] Possible patterns of the three binned visit behaviors across the three treatments are $3^3 = 27$.

that either the Easy or Hard treatment elicits higher compliance. Evidence of monotonicity can be seen in Figures 2 and A.1.

Monotonicity implies that the ITT estimates for the Easy treatment capture the effect on participants who increased visits from 0-2 to 3-4, 0-2 to 5+ or 3-4 to 5+ with respect to the Control group. The ITT estimates for the Hard treatment capture the effect on participants who increased visits from 0-2 to 3-4, 0-2 to 5+ or 3-4 to 5+ with respect to the Control group.[37]

Let β_E^{ITT} and β_H^{ITT} denote the ITT estimate of the Easy and Hard treatments. Let $i_z \rightarrow j_{z'}$ denote that a participant switches from compliance level i when $Z = z$ to j when $Z = z'$. Let $P(i_z \rightarrow j_{z'})$ represent the proportion of participants in the population who follow this pattern of behavior. Under Assumption 1 and Assumption 2, we have that: $\beta_E^{ITT} = P(0_C \rightarrow 1_E)E[Y(1) - Y(0)|0_C \rightarrow 1_E] + P(0_C \rightarrow 2_E)E[Y(2) - Y(0)|0_C \rightarrow 2_E] + P(1_C \rightarrow 2_E)E[Y(2) - Y(1)|1_C \rightarrow 2_E]$ and $\beta_H^{ITT} = P(0_C \rightarrow 1_H)E[Y(1) - Y(0)|0_C \rightarrow 1_H] + P(0_C \rightarrow 2_H)E[Y(2) - Y(0)|0_C \rightarrow 2_H] + P(1_C \rightarrow 2_H)E[Y(2) - Y(1)|1_C \rightarrow 2_H]$.

This exercise illustrates that the ITT estimators combine the effects of different treatment dosages. We cannot rely on ITT estimates by themselves to identify the effect of specific visit dosage on recidivism. Nonetheless, we can use the expressions to determine the conditions needed to identify the causal effect of specific visit dosages. For instance, economic rationality suggests that a participant would not switch from 0-2 visits if assigned to the Control to 3-4 visits if assigned to the Hard condition.[38] Such behavior would not secure rewards. If the participant does not need the reward to complete 3-4 visits, they should complete them

[37]The classification we adopt is not without a potential cost. Angrist and Imbens (1995) show that binarizing a continuous treatment can lead to a violation of the exclusion restriction if the chosen thresholds do not fully capture changes in the intensive margin across treatments. To examine potential intensive margin variation within our defined categories, we test if the number of visits of those completing less than three visits is different between treatments and control and if the number of visits of those completing at least five visits is different between treatments and controls. We find no evidence consistent with changes in the intensive margin of visits. This suggests our discretization does not generate a violation of the exclusion restriction.

[38]Both 0-2 visits and 3-4 visits are available in Control and Hard at the same latent price, i.e. the non-observed incentives are fixed. Switching across treatments would reveal that both options are mutually revealed preferred to each other. Such behavior could be consistent with someone who planned to do more but failed, however, we assume this behavioral pattern is absent.

in the Control as well. We should then not expect a statistically significant increase in 3-4 visits between the Control and Hard treatments. Figure A.2 shows behavior consistent with this condition. The proportion of participants completing 3-4 visits in Control and Hard are the same.[39] The main implication of this observation is that we can identify the effect of 5+ visits on rearrest by estimating the effect of being assigned to the Hard treatment compared to the Control, excluding the Easy condition. We note that the effect of 5+ visits combine the effect on those who switch from 0-2 visits in the Control condition to 5+ visits in the Hard condition and those who switch from 3-4 visits in the Control condition to 5+ visits in the Hard condition. That is, it combines the treatment effect of two different response types.

Having excluded the possibility that 3-4 visits increase in the Hard treatment, we explore conditions to identify the effect of 3-4 visits on rearrest. The formulas above suggest we can determine the effect of 3-4 visits by comparing the Easy and Hard treatments, excluding the Control, if two conditions hold. First, the proportion of participants who switch to 5+ visits is similar in the Easy and Hard treatments, and second, those who switch to 5+ visits in the Easy and Hard conditions are the same participants. The first condition is testable since we observe visits, and Figure A.2 validates it. The proportion of 5+ visits in Easy and Hard are similar ($\sim$ 19%).[40] We assess the validity of the second condition by comparing the composition of the first five visits for those who completed at least five visits in the Easy and Hard treatments. We find no statistical difference in the composition of these visits, providing evidence that the maintained identification assumption is not violated.[41]

Dosage effects are presented in Table 4 under the conditions outlined above. The es-

[39] There is no significant difference in the proportion of 3-4 visits between the Control and Hard treatments for the whole sample (p-value = 0.7373) and the subsamples of Black (p-value = 0.8560) and Non-Black (p-value = 0.6998) participants.

[40] There is no significant difference in the proportion of 5+ visits between the Easy and Hard treatments for the full sample (p-value = 0.8040) and the subsamples of Black (p-value = 0.8560) and Non-Black (p-value = 0.5711) participants.

[41] To test the hypothesis, we average the dummy variables indicating a visit was for a particular purpose (e.g., clothing). We compare these averages across treatments. The statistical tests all suggest no differences in behavior: jobs (p-value = 0.851), bus pass (p-value = 0.438), legal advice (p-value = 0.938), clothing (p-value = 0.357), housing (p-value = 0.464), social and mentoring (p-value = 0.783).

Table 4: Treatment effects of different dosages on 3-year re-arrest rate

	All	Non-Black	Black
3-4 visits	-0.042	0.092	-0.196
	(0.050)	(0.067)	(0.071)
5+ visits	0.025	0.012	0.040
	(0.049)	(0.065)	(0.073)

Notes: Number of observations is 490 for All, 225 for Black participants and 265 for non-Black participants. Standard errors in parentheses. Estimates of 3-4 visits are obtained in the subsample of the Easy and Hard treatment. Estimates of the 5+ visits are obtained in the subsample of the Hard treatment and Control. ITT effects are calculated using Stata `multe` command. Regressions control for sex, race, residence status (temporary housing, rehab center, homeless) and whether the participant had an arrest in the three years before study enrollment. Missing information is handled by replacing missing items with a zero and adding a dummy for missing data.

timates of 3-4 visits are obtained in the subsample excluding the Control condition. The estimates of 5+ visits are obtained excluding the Easy treatment. All regressions control for sex, race, residence status (temporary housing, rehab center, homeless) and whether the participant had an arrest in the three years before study enrollment.

The table shows no significant dosage effects for the full sample or Non-Blacks. For Black participants, completing 3-4 aftercare service visits significantly decreases recidivism by 19.6 percentage points. Completing 5+ visits has no significant effect. These results suggest that dosage effects are nonlinear, with a lower dosage of visits, i.e. 3-4 in our sample and setting, reducing recidivism and a higher dosage of visits, i.e. 5+ in our sample, having no significant effect.

The estimates for 3-4 visits pertain to a particular response type: those who completed 3-4 visits in the Easy condition and complete 0-2 visits in the Hard condition (and by economic rationality would also complete 0-2 visits in the Control condition as well). The response type to which the dosage effect of 3-4 visits pertains is different from the response types used to estimate the dosage effect of 5+ visits.[42] Indeed, to determine whether larger dosages

[42]Evidence on complier characteristics suggests that those who complete 3-4 visits are different than those who complete 5+ visits (Table A.8): more likely to be Black, have less education, less likely to have other support, less likely to be in institutional housing.

are less effective than smaller ones, we would need to identify the effect of 0-2, 3-4, and 5+ visits on the same response type, i.e. for full compliers. Full compliers, however, would violate the unordered monotonicity pattern we observe in our data, whereby the number of visits are weakly larger in the Easy condition. As discussed earlier, identification of dosage effects in our setting requires such a monotonicity assumption. In other words, absent such an assumption, dosage effects are not point identified for any response type. In sum, we identify a significant effect of 3-4 visits and a null effect of 5+ visits, but we cannot conclude this is due to diminishing returns without making additional identifying assumptions that would allow us to extrapolate treatment effects across response types.

7 Robustness

Our study shows that one additional valid visit reduces rearrest by 8.6 percentage points in the overall sample and by 12.5 percentage points for Black participants. The effect is large, especially compared to other studies on reentry service programs (D'Amico and Kim, 2018; Grommon et al., 2013; Cook et al., 2015; Wiegand and Sussell, 2016) which find mixed and small effects on recidivism. We explore possible reasons for our findings.

If instead of using arrests, we use the probability of a criminal offense within three years of recruitment, we find a similar, but noisier, estimate (see Table A.9). Similar patterns obtain if we concentrate on arrests involving a summary, misdemeanor or felony charge. The test of the joint hypothesis that arrests, criminal offenses, felonies, misdemeanors, and summary are jointly insignificant is $\chi^2(5) = 11.20$, p-value $= 0.0476$. While results using alternative outcomes are noisy, they point in the same direction as arrests. We conclude that arrests provide a good summary statistic for treatment effects.

Other reasons for our large effect are sample selection and treatment endogeneity. Our sample is a subset of the population of previously arrested individuals who visit the ASP. While records from the ASP show that the modal number of visits per person is one, our

participants might be more sensitive to treatment since they find it useful to get help with basic needs. We cannot discard this channel. Our experiment does provide causal estimates of additional services usage by this population. In other words, a potential benefit of the ASP might be their ability to attract those who have the potential to benefit from their services. In our setting of encouragement of service usage, without constraints on types of services used, we observe positive selection for Black participants (see Appendix C). This suggests our treatment can produce a significant reduction in recidivism, especially for the previously incarcerated who are more likely to benefit from these services. Additional studies can help elucidate whether these gains extend to a broader set of ex-inmates.

Some types of visits may be associated with larger reductions in arrests. Table A.10 presents ACR estimates of rearrest across six categories of visits instead of the total number of visits. Completing any type of visit is negatively associated with rearrest (except for legal services which are in low demand), and visits that included counseling, mentoring and peer or job-related support are associated with the largest declines in arrests. Black participants are more likely to use these types of services than non-Black participants (Figure A.5).

Finally, participants may have learned that additional services were useful, i.e., the final dosage was larger than planned in the experimental design. To assess this hypothesis, we calculate the increase in demand for services above what was required to claim the rewards. We compare the control condition and the Easy condition since this condition made it easier to achieve visit goals. We find that 7% of participants in the control and 14% in the Easy condition completed three valid visits. However, 17% of participants completed strictly more than three visits in the control, while 36% completed more than three visits in the Easy condition. While we cannot discard the hypothesis that participants completed more visits as an insurance policy, this is also consistent with participants finding the services beneficial. This analysis also shows that the marginal participant is more likely to complete four instead of three visits. As discussed in Appendix C, these participants experienced the largest gains.

8 Benefit-Cost Analysis

Using back-of-the-envelope calculations, we assess the dollar value of services used for those who completed 3-4 visits in the Easy treatment and 5+ visits in the Hard treatment. We combine services into five broad categories: bus pass, clothing/food, legal assistance to obtain an ID, housing search assistance, job search assistance and peer support and mentoring, and we calibrate a value for each service.[43] The average value to a participant of services used when completing 3-4 visits is $257.81 and when completing 5+ visits is $500.80.[44] These values are sizeable and exceed the value of the $50 gift card received upon visit goal completion. An alternative measure of value per visit can be obtained using the financial reports from the ASP. Using reports from 2021 and assuming two visits per year and 1092 visitors per year, we find that the average value of services used per visit is $603.

We follow Alsan et al. (2024)'s approach to calculate the costs to society that were avoided due to the reduction in crime induced by our treatment. The UJS dataset provides a gradation of each type of crime and a description. Summary counts are not disaggregated; misdemeanors and felonies are grouped into four categories. To measure cost, we first find the crime description that captures at least fifty percent of the counts in each category. We then use Miller et al. (2021)'s cost calculations for various crimes to provide an average cost for each count, using a simple average of the crimes we find. We then calculate the sum or the max over all of these costs and counts associated with an arrest three years since study enrollment. These costs are calculated only if the counts are associated with an arrest. Table A.9 shows the average causal effect of visits on the cost of crime (last two columns). We find

[43] A bus pass costs $24. We assume a clothing/food visit is worth $20, given that the cost of obtaining used clothing at Goodwill and donated food. A legal assistance visit to obtain an ID is assumed to be $60 for labor (3 hours of assistance * $20/hour) since obtaining an ID for low-income, homeless individuals is free in Pennsylvania. A housing search visit is assumed to be $80 for labor (4 hours of assistance * $20/hour) since assistance includes counseling, referrals, search on in-house computers and help filling out applications. A job search visit is assumed to be $80 for labor (4 hours of assistance * $20/hour) since assistance includes resume development, counseling, identifying employment opportunities and online search. A peer support/mentoring visit is assumed to be $50 to account for the opportunity cost of the peer or mentor's time.

[44] The average number of units/visits for bus pass, clothing/food, legal, housing, job and peer support are: 2.0, 1.65, 0.38, 0.69, 0.90 and 0.52 for 3-4 visits and 3.03, 2.87, 0.53, 1.50, 1.57 and 1.87 for 5+ visits.

that one extra visit is associated with a $2,092.02 reduction in the societal cost of crime. Our estimates, while noisy, suggest the additional services used due to the experimental treatments pass the benefit-cost analysis.

9 Conclusion

We investigate whether increasing the number of visits to use support services for previously-incarcerated individuals can decrease recidivism. Given the heterogeneity in services needed by these individuals, we implement an encouragement design that experimentally varies the dosage of visits, with no restrictions on the types of services used. We find that one extra visit decreases the probability of arrest three years after enrollment by 8.6 percentage points. These findings from the full sample hide a large degree of heterogeneity. Black participants are more likely to increase visits than non-Black participants when encouraged to do so (3.3 v. 2.3 completed visits), and they are less likely to be rearrested after completing visits. The dosage effect of visits on reducing rearrests is nonlinear. Black participants who complete 3-4 visits are 19.6 percentage points less likely to be rearrested, but there is no effect for those completing 5+ visits. Given the low response to treatment by Non-Blacks, we are not powered to detect treatment effects on rearrests for this group. Not all services are equally effective at reducing rearrest, with peer support, mentoring and job-related services having the largest impacts, and Black participants are more likely to choose these more effective services.

Our experiment adds new knowledge to the existing mixed evidence on the impact of reentry services on reducing recidivism. The findings from our study suggest that the amount and composition of reentry services might be miscalibrated. Our experimental design uncovers these features by allowing self-selection into types of services and encouraging different service levels. Our analysis shows that heterogeneity is important in all stages of the intervention: who selects into treatment, what services participants select, and who benefits from

it. The study speaks to the importance of considering first-stage heterogeneity, heterogeneous treatment effects and who might benefit the most from treatment. A fruitful avenue to explore would be deploying research designs to help uncover the most effective aspects of reentry support services to aid those seeking help and embed those in program design.

References

Abadie, A., Gu, J., and Shen, S. (2024). Instrumental variable estimation with first-stage heterogeneity. *Journal of Econometrics*, 240(2):105425.

Alsan, M., Barnett, A. M., Hull, P., and Yang, C. (2024). "something works"' in us jails: Misconduct and recidivism effects of the ignite program. Technical report, National Bureau of Economic Research.

Andresen, M. E. (2018). Exploring marginal treatment effects: Flexible estimation using stata. *The Stata Journal*, 18(1):118–158.

Andresen, M. E. and Huber, M. (2021). Instrument-based estimation with binarised treatments: issues and tests for the exclusion restriction. *The Econometrics Journal*, 24(3):536–558.

Angrist, J. D. and Imbens, G. W. (1995). 2-stage least-squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Association*, 90(430):431–442.

Arbour, W. (2022). Can recidivism be prevented from behind bars? evidence from a behavioral program. Technical report, Working Paper.

Athey, S. and Wager, S. (2021). Policy learning with observational data. *Econometrica*, 89(1):133–161.

Barnes, G. C., Hyatt, J. M., Ahlman, L. C., and Kent, D. T. (2012). The effects of low-intensity supervision for lower-risk probationers: Updated results from a randomized controlled trial. *Journal of Crime and Justice*, 35(2):200–220.

Batistich, M. K., Evans, W., and Phillips, D. (2021). Reducing re-arrests through light touch mental health outreach. Technical report, University of Notre Dame.

Bhatt, M. P., Heller, S. B., Kapustin, M., Bertrand, M., and Blattman, C. (2024). Predicting and preventing gun violence: An experimental evaluation of readi chicago. *The quarterly journal of economics*, 139(1):1–56.

Blattman, C., Jamison, J., and Sheridan, M. (2017). Reducing crime and violence: Experimental evidence from cognitive behavioral therapy in liberia. *American Economic Review*, 107(4):1165–1206.

Boyle, D. J., Ragusa-Salerno, L. M., Lanterman, J. L., and Marcus, A. F. (2013). An evaluation of day reporting centers for parolees: Outcomes of a randomized trial. *Criminology & Public Policy*, 12(1):119–143.

Brinch, C. N., Mogstad, M., and Wiswall, M. (2017). Beyond late with a discrete instrument. *Journal of Political Economy*, 125(4):985–1039.

Cook, P. J., Kang, S., Braga, A. A., Ludwig, J., and O'Brien, M. E. (2015). An experimental evaluation of a comprehensive employment-oriented prisoner re-entry program. *Journal of Quantitative Criminology*, 31(3):355–382.

Corgnet, B., Gómez-Miñambres, J., and Hernán-Gonzalez, R. (2015). Goal setting and monetary incentives: When large stakes are not enough. *Management Science*, 61(12):2926–2944.

Denney, A. S., Tewksbury, R., and Jones, R. S. (2014). Beyond basic needs: Social support and structure for successful offender reentry. *Journal of Qualitative Criminal Justice & Criminology*, 2:39–67.

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

Doleac, J. L. et al. (2019). Wrap-around services don't improve prisoner reentry outcomes. *Journal of Policy Analysis and Management*, 38(2):508–514.

Durose, M. R., Cooper, A. D., and Snyder, H. N. (2014). *Recidivism of prisoners released in 30 states in 2005: Patterns from 2005 to 2010*, volume 28. US Department of Justice, Office of Justice Programs, Bureau of Justice.

D'Amico, R. and Kim, H. (2018). Evaluation of seven second chance act adult demonstration programs: Impact findings at 30 months. Technical report, Social Policy Research Associaties.

Farabee, D., Zhang, S. X., and Wright, B. (2014). An experimental evaluation of a nationally recognized employment-focused offender reentry program. *Journal of Experimental Criminology*, 10(3):309–322.

Fishbane, A., Ouss, A., and Shah, A. (2020). Behavioral nudges reduce failure to appear for court. *Science*, 370(682):909–931.

Geller, A. and Curtis, M. A. (2011). A sort of homecoming: Incarceration and the housing security of urban men. *Social Science Research*, 40:1196–1213.

Goerg, S. J. and Kube, S. (2012). Goals (th)at work: Goals, monetary incentives, and workers' performance. Technical report, Max Planck Institute for Research on Collective Goods.

Goldsmith-Pinkham, P., Hull, P., and Kolesár, M. (2024). Contamination bias in linear regressions. *American Economic Review*, 114(12):4015–51.

Grommon, E., Davidson II, W. S., and Bynum, T. S. (2013). A randomized trial of a multimodal community-based prisoner reentry program emphasizing substance abuse treatment. *Journal of Offender Rehabilitation*, 52(4):287–309.

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

Heller, S. (2014). Summer jobs reduce violence among disadvantaged youth. *Science*, 346(6214):1214–1223.

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

Hyatt, J. M. and Barnes, G. C. (2017). An experimental evaluation of the impact of intensive supervision on the recidivism of high-risk probationers. *Crime & Delinquency*, 63(1):3–38.

Ida, T., Ishihara, T., Ito, K., Kido, D., Kitagawa, T., Sakaguchi, S., and Sasaki, S. (2022). Choosing who chooses: Selection-driven targeting in energy rebate programs. Technical report, National Bureau of Economic Research.

Kitagawa, T. and Tetenov, A. (2018). Who should be treated? empirical welfare maximization methods for treatment choice. *Econometrica*, 86(2):591–616.

Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.

Kolesár, M. (2013). Estimation in an instrumental variables model with treatment effect heterogeneity. Technical report.

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

Lane, J., Turner, S., Fain, T., and Sehgal, A. (2005). Evaluating an experimental intensive juvenile probation program: Supervision and official outcomes. *Crime & Delinquency*, 51(1):26–52.

Lee, L. M. (2022). Halfway home? residential housing and reincarceration. *American Economic Journal: Applied Economics*.

Lee, S. and Salanié, B. (2020). Filtered and unfiltered treatment effects with targeting instruments.

List, J. (2020). Non est disputandum de generalizability? a glimmpse into the external validity trial. Technical report, NBER Working Paper 27535.

Locke, E. and Latham, G. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice Hall.

Manski, C. F. (2025). Using limited trial evidence to credibly choose treatment dosage when efficacy and adverse effects weakly increase with dose. *Epidemiology*, 36(1):60–65.

Markovitz, D. (2012). The folly of stretch goals. *Harvard Business Review*.

Miller, T. R., Cohen, M. A., Swedler, D. I., Ali, B., and Hendrie, D. V. (2021). Incidence and costs of personal and property crimes in the usa, 2017. *Journal of benefit-cost analysis*, 12(1):24–54.

Mogstad, M. and Torgovitsky, A. (2024). Instrumental variables with unobserved heterogeneity in treatment effects. In *Handbook of Labor Economics*, volume 5, pages 1–114. Elsevier.

Nhan, J., Bowen, K., and Polzer, K. (2017). The reentry labyrinth: The anatomy of a reentry services network. *Journal of Offender Rehabilitation*, 56(1):1–19.

Ordóñez, L., Schweitzer, M., Galinsky, A., and Bazerman, M. (2009). Goals gone wild: The systematic side effects of over-prescribing goal setting. *Academy of Management Perspectives*, 21(1).

Papp, J., Wooldredge, J., and Pompoco, A. (2021). Timing of prison programs and the odds of returning to prison. *Corrections*, 6(2):124–149.

Prendergast, M., Anglin, M., Burdon, W., and Messina, N. (2003). Evaluation of the 1,000-bed expansion of therapeutic community treatment programs for prisoners: Final report

(california department of corrections contract 97.355). *Los Angeles: Integrated Substance Abuse Programs, University of California, Los Angeles*.

Prendergast, M. L., Hall, E. A., Grossman, J., Veliz, R., Gregorio, L., Warda, U. S., Van Unen, K., and Knight, C. (2015). Effectiveness of using incentives to improve parolee admission and attendance in community addiction treatment. *Criminal Justice and Behavior*, 42(10):1008–1031.

Roman, C. G. and Travis, J. (2006). Where will i sleep tomorrow? housing, homelessness, and the returning prisoner. *Housing Policy Debate*, 17:389–418.

Romano, J. and Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. *Econometrica*, 73(4):1237–1282.

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

Ryan, T. (1970). *Intentional behavior*. New York: Ronald Press.

Shem-Tov, Y., Raphael, S., and Skog, A. (2024). Can restorative justice conferencing reduce recidivism? evidence from the make-it-right program. *Econometrica*, 92(1):61–78.

Sitkin, S., Miller, C., and See, K. (2017). The stretch goal paradox. *Harvard Business Review*, 95(1).

Wiegand, A. and Sussell, J. (2016). Evaluation of the re-integration of ex-offenders (rexo) program: Final impact report.

Zimmerman, B. J., Bandura, A., and Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal*, 29(3):663–676.

APPENDICES INTENDED FOR ONLINE PUBLICATION

A Extra figures and tables

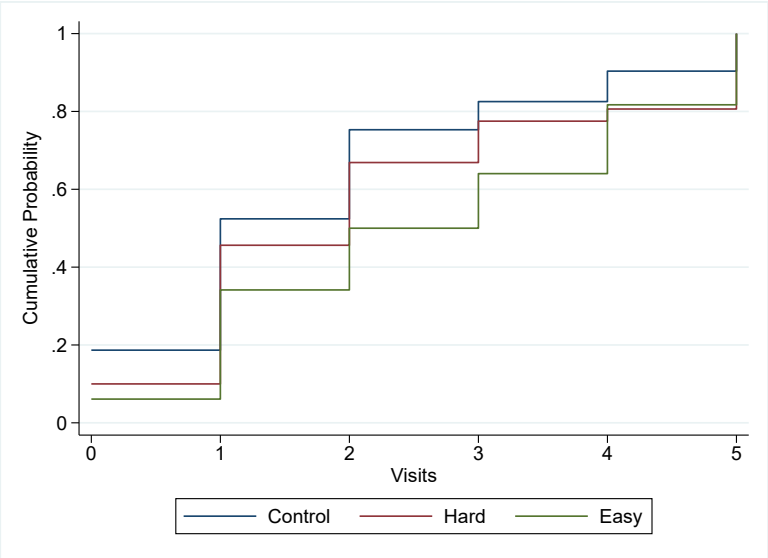
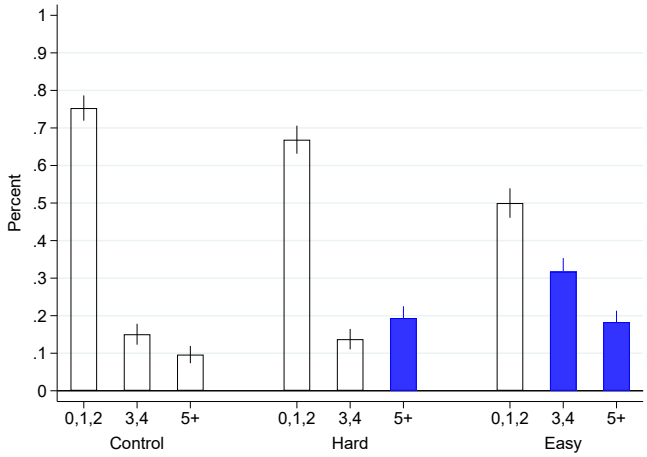


Figure A.1: Distribution of number of visits by treatment

Notes: The figure shows the distribution of number of visits by treatment. First-order stochastic dominance tests show that Easy FOSD Hard (p-value = 0.625, p-value = 0.0026) FOSD Control (p-value = 0.8912, p-value = 0.0213).



Notes: Blue indicates the proportion in Easy and Hard that completed the required number of visits for the $50 incentive. Error bars denote standard errors.

Figure A.2: Total number of visits completed 12 months after enrollment

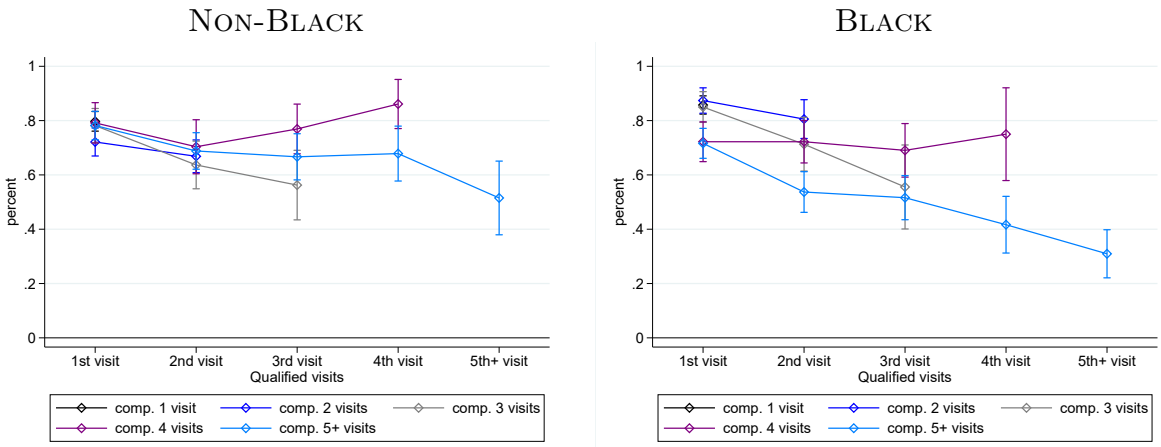


Figure A.3: Percent of services used for short-run necessities (i.e. clothing, housing, food, ID, transportation) by visit number - Non-Black and Black participants

Notes: The figure shows usage of services for short-run necessities during the first, second, third, fourth and fifth+ visit. Each line shows usage for those who completed 1, 2, 3, 4 or 5+ visits within one year of study enrollment. Error bars denote standard errors. Sample used for the figure are those who completed 5+ visits, 4 visits, 3 visits, 2 visits and 1 visit.

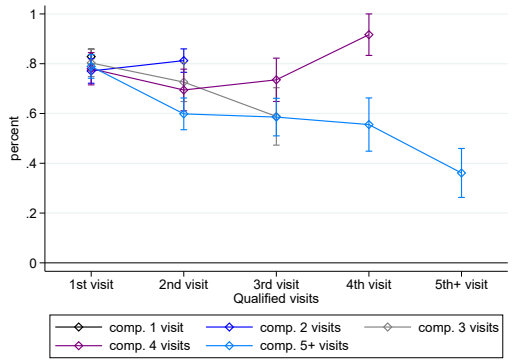


Figure A.4: Percent of services used for short-run necessities (i.e. clothing, housing, food, ID, transportation) by visit number - First-time visitors

Notes: The figure shows usage of services for short-run necessities during the first, second, third, fourth and fifth+ visit. Each line shows usage for those who completed 1, 2, 3, 4 or 5+ valid visits within one year of study enrollment. Error bars denote standard errors. Sample used for the figure are those who completed 5+ visits (n=77), 4 visits (n=47), 3 visits (n=52), 2 visits (n=98) and 1 visit (n=159).

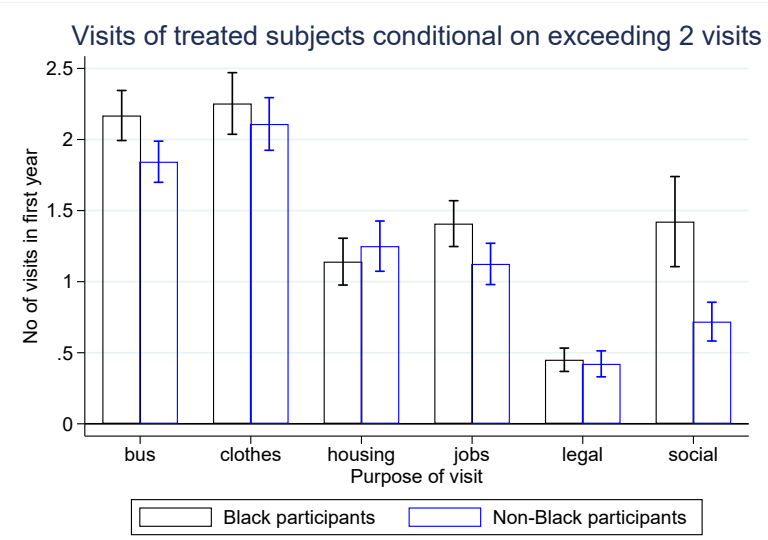


Figure A.5: Type of visit completed by race

Notes: The figure shows the number of visits by type and by race among treated participants who completed at least 3 visits. Visit types: bus is for a bus pass; clothes is for clothing; housing is for housing referral; jobs is for employment referral; legal is for legal referral, including getting an identification card; social is for peer or mentoring groups.

Table A.1: Services offered by the Aftercare Service Provider (ASP) and whether use of the service during a visit would count as a valid visit in the Easy and Hard treatments

	Services offer by Aftercare Service Provider	Counted as valid visit
1	Housing - search for options	yes
2	Food pantry and food voucher	yes
3	Clothing closet and clothing voucher	yes
4	Mentoring resources	yes
5	Family services	yes
6	Peer support group program	yes
7	Employment - resume, job applications, training and education searches	yes
8	Obtaining a state identification card	yes
9	Computer and phone use for housing or employment search	yes
10	Pick up bus pass (only reason for visit)	no
11	Personal use of phone or computer	no

Table A.2: Testing for interference using the Control group

	All	Non-Black	Black
Proportion in network assigned to Hard	0.059	-0.047	0.224
	(0.102)	(0.130)	(0.167)
Proportion in network assigned to Easy	-0.111	-0.094	-0.141
	(0.113)	(0.147)	(0.176)
Constant	0.198***	0.230***	0.191**
	(0.053)	(0.064)	(0.069)
Obs	366	202	164
R2	0.066	0.049	0.094
H0: no interference	0.530	0.762	0.311

Notes: Dependent variable is probability of re-arrest within 3 years of study enrollment. Controls include sex, race and whether the participant had an arrest in the three years prior to study enrollment. We test if arrest rates in the control group are affected by the assignment to treatment of other ASP clients in the participants' network. We represent a participant's network by an adjacency matrix. The adjacency matrix equals 1 if two participants visited the ASP on the same date at least once during the study and 0 otherwise. We use both qualified and non-qualified visits to create this matrix. Participants in the control group visited on the same date with 4.86 (std dev 3.80) participants during the study. The null hypothesis (H0) reports the joint significance test of the proportion of people assigned to both Easy and Hard treatments. While there is significant overlap on visit dates, we cannot reject the null hypothesis of no interference.

Table A.3: ITT estimates of the number of valid visits within a year of enrollment conditional on no arrest six months after enrollment

	All	Non-Black	Black
Hard	0.707	0.289	1.156
	(0.260)	(0.320)	(0.370)
Easy	1.297	0.476	2.099
	(0.326)	(0.275)	(0.662)

Notes: Number of observations is 490 for All, 225 for Black participants and 265 for non-Black participants. Standard errors in parentheses. ITTs are calculated using Stata `multe` command, which implements Goldsmith-Pinkham et al. (2024) multiple treatment contamination bias robust estimates. We interact treatment and race to expand the number of effective treatments. Heterogeneous treatment effects are estimated jointly but presented in separate columns. Regressions control for sex, race, residence status (temporary housing, rehab center, homeless) and whether the participant had an arrest in the three years before study enrollment.

Table A.4: Characteristics of Non-Black and Black participants

	Non-Black	s.d.	Black	s.d.	F-test	p-val.
Female	0.34	0.47	0.21	0.41	10.50	0.00
Age	41.20	9.79	44.26	12.00	9.64	0.00
Married/Partnered	0.05	0.22	0.09	0.28	2.07	0.15
Education (years)	12.41	1.83	12.03	1.39	6.46	0.01
Employed	0.24	0.43	0.27	0.45	0.91	0.34
Knew the aftercare	0.28	0.45	0.35	0.48	2.59	0.11
Number of children	1.81	1.93	2.31	2.31	6.81	0.01
Has other support	0.48	0.50	0.38	0.49	4.75	0.03
Year of last arrest	2,015.89	4.97	2,014.83	6.19	4.30	0.04
Duration of last incarceration	722.62	1,219.84	763.99	1,450.88	0.10	0.75
Arrested 36m prior to study	0.48	0.50	0.48	0.50	0.00	0.95
Institutional housing	0.66	0.48	0.45	0.50	22.35	0.00
Observations	265		225			

Notes: First and third columns are averages. Columns labelled s.d. are the standard deviation of the average. The last two columns report the F-statistic and p-value for a joint test of equality of the row variable across Non-Black and Black participants. Female is a dummy variable for female. Age is in years. Married/Partnered is a dummy variable for being in a relationship. Education is in years. Employed is a dummy variable for being currently employed at the time of enrollment into the study. Knew the aftercare is a dummy variable for having heard of our partner ASP while incarcerated. Number of children is number of children. Has other support is a dummy variable for having access to other support services. Year of last arrest is year. Duration of last incarceration is length of most previous incarceration in days. Arrested 36m prior is a dummy variable for having been arrested at least once during the 36 months prior to enrollment into the study. Institutional housing is a dummy variable for providing an institutional address at study enrollment, i.e. parole office, halfway house, rehabilitation facility, homeless shelter.

Table A.5: Treatment effects on 3-year re-arrest rate - not accounting for contamination bias and first-stage heterogeneity

	All		Non-Black		Black	
	ITT	ACR	ITT	ACR	ITT	ACR
Hard	0.031		0.026		0.030	
	(0.050)		(0.068)		(0.072)	
Easy	-0.019		0.114		-0.190	
	(0.050)		(0.068)		(0.072)	
Visits		-0.023		0.175		-0.125
		(0.044)		(0.124)		(0.054)

Notes: Number of observations is 490 for All, 225 for Black participants and 265 for non-Black participants. Standard errors in parentheses. ITT effects are estimates using OLS regression with dummies per treatment. ACR estimates are the effect of the number of visits on the probability of being arrested up to 36 after enrollment. The estimates implement are obtained assuming homogeneous response to treatment on visits. Regressions control for sex, race, residence status (temporary housing, rehab center, homeless) and whether the participant had an arrest in the three years before study enrollment. Missing information is handled by replacing missing items with a zero and adding a dummy for missing data.

Table A.6: Treatment effects on 3-year re-arrest rate - controlling for first-time visitor

	All		Non-Black		Black	
	ITT	ACR	ITT	ACR	ITT	ACR
Hard	0.014		-0.038		0.048	
	(0.049)		(0.059)		(0.061)	
Easy	-0.012		0.100		-0.103	
	(0.049)		(0.058)		(0.065)	
Visits		-0.075		0.173		-0.114
		(0.040)		(0.120)		(0.050)

Notes: Number of observations is 490 for All, 225 for Black participants and 265 for non-Black participants. Standard errors in parentheses. ITT effects are calculated using Stata `multe` command, which implements Goldsmith-Pinkham et al. (2024) multiple treatment contamination bias robust estimates. We interact treatment and race to expand the number of effective treatments. Heterogeneous treatment effects are estimated jointly but presented in separate columns. TOT estimates are the effect of the number of visits on the probability of being arrested up to 36 after enrollment. The estimates implement Abadie et al. (2024) approach to deal with first-stage heterogeneity. Regressions control for sex, race, residence status (temporary housing, rehab center, homeless), first-time visitor and whether the participant had an arrest in the three years before study enrollment. Missing information is handled by replacing missing items with a zero and adding a dummy for missing data.

Table A.7: Treatment effects on 3-year re-arrest rate - using Kolesár (2013)'s UJIVE method

	All	Non-Black	Black
ACR	-0.107	0.276	-0.148
	(0.059)	(0.268)	(0.072)

Notes: Number of observations is 490 for All, 225 for Black participants and 265 for non-Black participants. Standard errors in parentheses. We re-estimate the ACR specifications in Columns 2, 4 and 6 in Table 3 using Kolesár (2013)'s UJIVE method. Regressions control for sex, race, residence status (temporary housing, rehab center, homeless) and whether the participant had an arrest in the three years before study enrollment. Missing information is handled by replacing missing items with a zero and adding a dummy for missing data.

Table A.8: Complier characteristics

	Switched to 3-4 visits				Switched to 5+ visits			
height	All	Complier	NT	AT	All	Complier	NT	AT
Female	0.27	0.44	0.25	0.18	0.29	0.54	0.25	0.44
	(0.02)	(0.21)	(0.04)	(0.08)	(0.03)	(20.36)	(0.04)	(0.13)
Black	0.47	0.51	0.43	0.59	0.47	0.28	0.48	0.56
	(0.03)	(0.21)	(0.05)	(0.11)	(0.03)	(0.95)	(0.04)	(0.13)
Age	42.61	40.89	42.88	43.50	42.31	42.64	42.01	44.50
	(0.63)	(4.07)	(1.02)	(2.49)	(0.59)	(109.67)	(0.96)	(2.13)
Education(years)	12.23	11.93	12.34	12.05	12.20	12.55	12.12	12.56
	(0.09)	(0.70)	(0.15)	(0.25)	(0.09)	(3.06)	(0.14)	(0.33)
Employed	0.25	0.06	0.33	0.09	0.23	0.83	0.16	0.19
	(0.02)	(0.17)	(0.04)	(0.06)	(0.02)	(1.44)	(0.03)	(0.10)
Married/Partnered	0.06	0.04	0.08	0.00	0.07	0.05	0.06	0.13
	(0.01)	(0.08)	(0.03)	(0.00)	(0.01)	(0.29)	(0.02)	(0.09)
Knew the aftercare	0.32	0.10	0.35	0.50	0.30	0.05	0.33	0.31
	(0.03)	(0.20)	(0.04)	(0.11)	(0.03)	(14.23)	(0.04)	(0.12)
Number of children	1.99	2.11	1.94	2.10	2.07	1.73	2.07	2.47
	(0.12)	(0.74)	(0.20)	(0.61)	(0.12)	(4.96)	(0.20)	(0.32)
Has other support	0.42	0.42	0.46	0.20	0.45	0.60	0.41	0.60
	(0.03)	(0.16)	(0.05)	(0.09)	(0.03)	(0.94)	(0.04)	(0.13)
Year of last arrest	2,015.64	2,015.76	2,015.71	2,015.18	2,015.17	2,008.94	2,015.92	2,015.25
	(0.29)	(2.26)	(0.44)	(1.17)	(0.34)	(16.47)	(0.40)	(1.05)
Duration of last incarceration	762.99	254.55	857.26	1,174.24	722.21	45.18	790.83	959.87
	(78.02)	(438.48)	(144.65)	(456.56)	(78.63)	(2,644.24)	(124.68)	(393.71)
Arrested 36m prior to study	0.48	0.60	0.48	0.32	0.48	0.25	0.50	0.63
	(0.03)	(0.19)	(0.05)	(0.10)	(0.03)	(2.18)	(0.04)	(0.13)
Institutional housing	0.57	0.39	0.59	0.70	0.55	0.76	0.54	0.40
	(0.03)	(0.17)	(0.05)	(0.10)	(0.03)	(0.81)	(0.04)	(0.13)

Notes: The table reports average characteristics for each response type by compliance in the Easy (Switched to 3-4 visits) and Hard (Switched to 5+ visits) treatments. Standard errors in parentheses. Response types are Compliers, Never Takers (NT) and Always Takers (AT). Female is a dummy variable for female. Black is a dummy variable for black participant. Age is in years. Married/Partnered is a dummy variable for being in a relationship. Education is in years. Employed is a dummy variable for being currently employed at the time of enrollment into the study. Knew the aftercare is a dummy variable for having heard of our partner ASP while incarcerated. Number of children is number of children. Has other support is a dummy variable for having access to other support services. Year of last arrest is year. Duration of last incarceration is length of most previous incarceration in days. Arrested 36m prior is a dummy variable for having been arrested at least once during the 36 months prior to enrollment into the study. Institutional housing is a dummy variable for providing an institutional address at study enrollment, i.e. parole office, halfway house, rehabilitation facility, homeless shelter.

Table A.9: Treatment effect of one extra visit on criminal offenses, type and cost of crime

	Crim. Offenses		Arrests		Cost	
	All	Felony	Misdemeanor	Summary	Mean	Max
Visits	-0.061	-0.024	-0.056	-0.055	-2,092.016	-590.866
	(0.042)	(0.032)	(0.038)	(0.036)	(3,673.959)	(539.301)

Notes: Number of observations is 490. Standard errors in parentheses. UJS data set classifies offenses into summary, misdemeanors, and felonies. Misdemeanors and felonies have three subcategories each. To calculate costs, we attribute to each one of these categories the cost per crime estimated by Miller et al. (2021). In particular, we take the average of the crimes that account for at least fifty percent of offenses in such category. We recode criminal offenses without a category as zero. The total cost is calculated by adding all criminal counts listed thee years after recruitment. We winsorize the sum at the upper 1% due to the existence of extreme observations. Similarly, the maximum is calculated as the maximum cost over all criminal counts during the three years since recruitment. Regressions implement Abadie et al. (2024)'s approach to deal with first-stage heterogeneity. Regressions control for sex, race, residence status (temporary housing, rehab center, homeless) and whether the participant had an arrest in the three years before study enrollment. Missing information is handled by replacing missing items with a zero and adding a dummy for missing data.

Table A.10: Treatment effect of one extra visit by type of visit on arrests 3 years after recruitment

	Bus	Clothes	Housing	Jobs	Legal	Social
Visits	-0.116	-0.165	-0.167	-0.491	0.526	-0.267
	(0.110)	(0.105)	(0.126)	(0.340)	(0.448)	(0.148)

Notes: Number of observations is 490. Standard errors in parentheses. Each column estimates the TOT using different types of visits. TOT estimates are the effect of the number of visits on the probability of being arrested up to 36 after enrollment. The estimates implement Abadie et al. (2024) approach to deal with first-stage heterogeneity. Regressions control for sex, race, residence status (temporary housing, rehab center, homeless) and whether the participant had an arrest in the three years before study enrollment. Missing information is handled by replacing missing items with a zero and adding a dummy for missing data.

B Field experiment materials & study registration

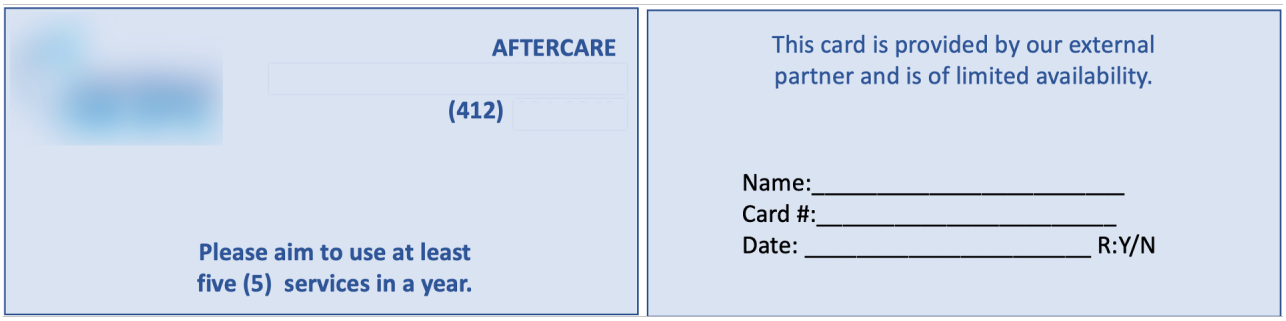


Figure B.1: Card for the Control group

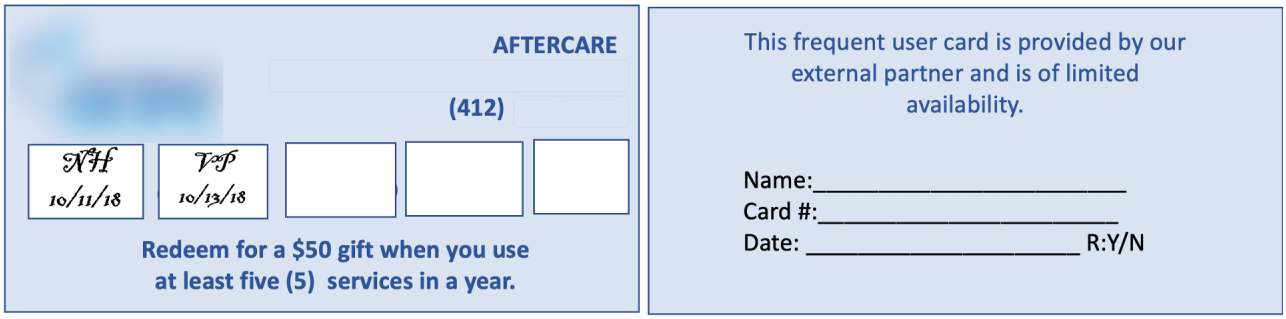


Figure B.2: Card for the Easy Treatment group

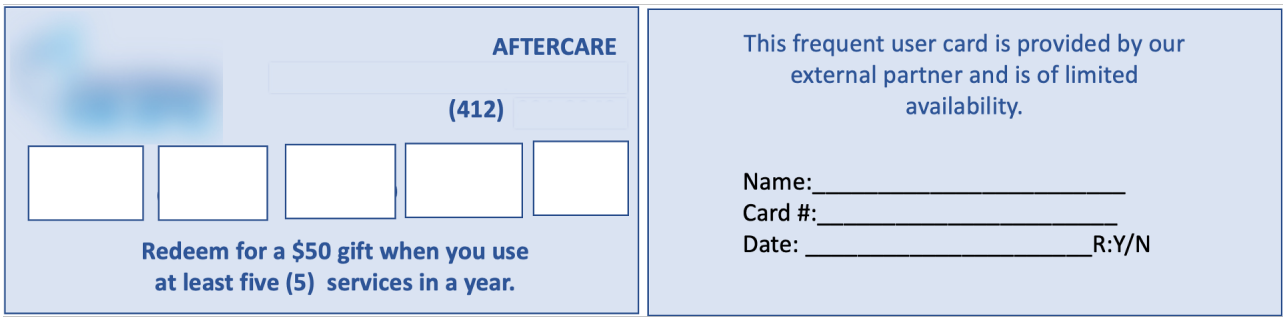


Figure B.3: Card for the Hard Treatment group

Intake survey questions

1. Name (first, last)
2. Date of birth
3. Address and zip code
4. Do you have a cell phone? If yes, what is phone number?
5. Do you have another contact person in case we cannot reach you? (name, phone number, relationship)
6. Highest level of education
7. Currently employed? Number of hours work per week
8. Marital status
9. Number of children
10. How did you get to our office today? (bus, drive, got a ride, took taxi/uber, walk, bike, other)
11. How long did it take you and how much did it cost?
12. Is this your first time coming to our office? If not, how long have you been coming to our office?
13. Did you know about us while you were still in jail? If so, how did you hear of us? what services did you think we provided?
14. Where were you last arrested (county)? What year?
15. How long was your last incarceration?
16. Was that your first arrest? If not, where was your first arrest (county) and year?
17. How long was your first incarceration?
18. Have you ever been arrested in other states outside of Pennsylvania?
19. Do you participate in other support programs outside of our services? If so, list them. How satisfied are you with them?

20. How many people can you think of that would help you out in the following situations? Name the top three persons (their relationship to you)? If you need cash, a job, life advice?

21. In a typical day, when do you get up, when do you go to sleep, how many hours do you spend outside the house?

22. How often do you attend church, chapel or other places of worship?

23. What is the biggest challenge you experienced after being released?

24. What service do you wish we provide to help you overcome the challenge?

25. Do you feel that our office can help you avoid future rearrests? Why or why not?

Analysis Plan - American Economics Association RCT Registry (AEARCTR-0003375)

Utilization of Social Services and Impact on Outcomes

October 2018

Analysis Plan

Our main outcome variables are re-arrests and usage of social services. For those two main variables, we will look at treatment effects (low and high usage intensity) relative to the control group in terms of frequency and types of crimes and services used.

Depending on the richness of data, we will also look at the duration of time before re-arrest, frequency of re-arrest, and type of crime re-arrested for. Similarly, for the usage of social services data, we will look at types of services used, timing of services used and time delay between service usage.

Depending on sample size and data availability, we will also try to control for potential confounders in our analysis, eg. type of crime, type of service used, age of participant, race, gender, etc.

C Marginal treatment effects

Our Easy and Hard treatments were designed to encourage more visits to our partner service provider and indeed increased visits compared to the Control group. The marginal treatment effects (MTE) analysis in this section uses this increase in visits, due to the treatment, to estimate the treatment effect on rearrests of these new individuals who visited our partner provider due to the incentives. Without incentives, these more "resistant" individuals are less likely to use services. We examine whether the treatment effect varies by resistance to treatment.

We exploit heterogeneous response to treatment to investigate whether selection into treatment is positive or negative. That is, we examine whether individuals who would benefit the most from using services are the ones who use them without encouragement (positive selection) or only use them with encouragement (negative selection). Variation in the propensity to complete a certain number of visits is used to estimate marginal treatment effects (Heckman and Vytlacil, 2007). Marginal treatment effects (MTE) are the derivative of the probability of being arrested within three years of enrollment with respect to the predicted probability of exceeding a threshold of visits. These can be interpreted as the treatment effect for individuals with different resistance to treatment (Heckman and Vytlacil, 2007).

We follow (Brinch et al., 2017) to estimate MTE. The authors show that, if potential outcomes are separable functions of observable and unobservable variables, discrete instruments interacted with covariates can be used to estimate treatment effects and the marginal response function over a larger support. To implement this approach, we create a dummy variable that equals one if at least three valid visits are completed within a year of enrollment. We use this definition of treatment completion to avoid potential exclusion restriction violations. This can happen if there are extensive margin effects on visits below such a defined threshold (Andresen and Huber, 2021). We augment the set of covariates by including distance between place of residence and the facilities of the ASP. This information is available only for the experimental sample, so the estimates are based on that population alone.

Figure C.1 presents our estimation results based on this approach. We present estimates of Black and Non-Black participants separately because these two subpopulations react differently to treatment. The top panel correspond to Black participants, and the bottom panel correspond to non-Black participants. Figures (a) and (c) present the distribution of propensity scores estimated using a Logit regression of completing at least three visits on covariates interacted with treatment indicators. Two distributions are shown. The distribution in blue is for treated individuals, i.e. those who completed three or more visits, and the distribution

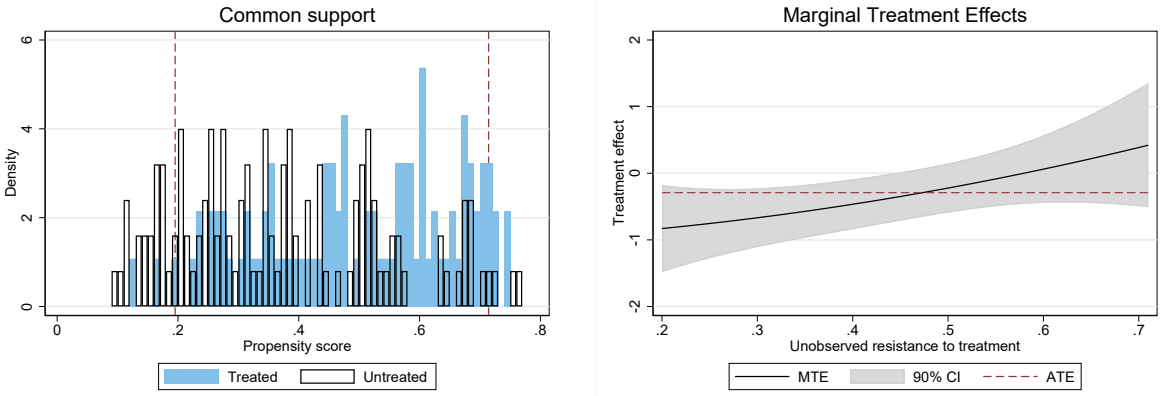
in grey is for untreated individuals, i.e. those who completed less than three visits. The overlapping distributions illustrate that the common support condition for MTE estimation is supported.[1] The Easy and Hard treatments increased the probability of completing three or more visits and thus brought in those who are more resistant to treatment. These individuals reside in the treated distribution. We use the propensity score support to estimate MTE. The support is trimmed by 2.5% on each tail, and we estimate the marginal response functions separately using a second-order polynomial. We use a parametric specification due to the small size of our sample.

Figures (b) and (d) show the MTE on rearrest by unobserved resistance to treatment. The interpretation is that those with higher resistance to treatment would not go to visit the service provider without incentives to do so. This is presented as a continuum on the x-axis in the figure. We see that the treatment effect on recidivism, for Black participants (Figure (b)), is more negative for those who are less resistant to treatment (i.e. between 0.20 and 0.45 of unobserved resistance). In other words, we observe positive selection into treatment for this group. We do not observe significant variation in marginal treatment effects for the subsample of non-Black participants (Figure (d)). The estimates reproduce the estimates imprecision levels reported in the main body of the paper.

In sum, these findings are consistent with positive selection into treatment for Black participants. Participants who might benefit the most from re-entry services are those who tend to use them more. Encouraging the use of services brings in participants with more to gain, and marginal treatment effects are the largest for this group.

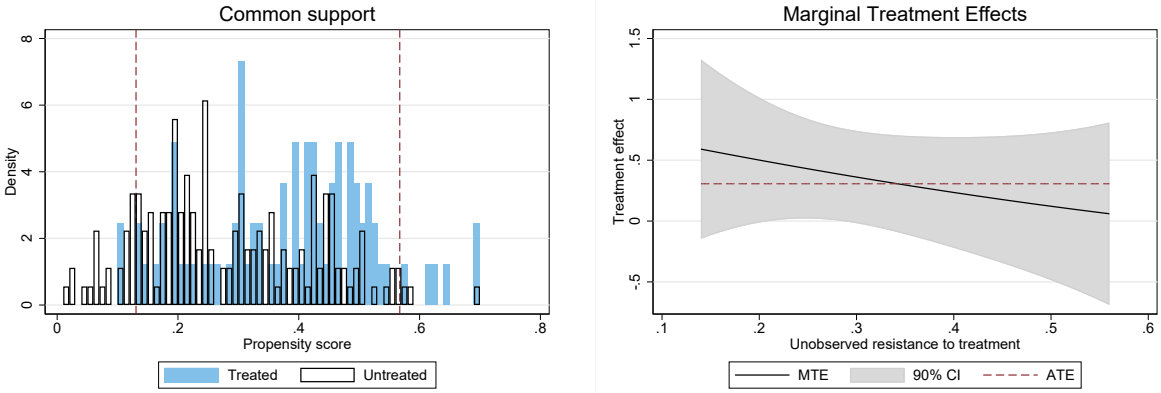
[1] The average probability of completing at least three visits is twenty four percent in the control condition, thirty-three percent in the Hard condition, and fifty percent in the Easy condition.

Black participants



(a) Propensity score (b) Marginal treatment effects

Non-Black participants



(c) Propensity score (d) Marginal treatment effects

Figure C.1: Marginal treatment effects of completing 3 or more visits

Note: The total number of observations is 490. The top two figures correspond to Black participants and the bottom two figures correspond to non-Black participants. In Panels (a) and (c), the dashed vertical lines represent the upper and lower bounds on the common support of the propensity score (based on 2.5% trimming) used to estimate the MTEs. Propensity scores are predicted with a Logit regression where the dependent variable is a dummy for three or more visits. Treated are those who completed three or more visits, and Untreated are those who completed fewer than three visits. In Panel (b) and (d), the MTE estimation is based on a quadratic polynomial specification in the sample with common support. The x-axis is the resistance to treatment. The y-axis is the estimated treatment effect, and the dotted horizontal red line is the average treatment effect. We follow Brinch et al. (2017) in assuming separability between observed and unobserved heterogeneity. We estimate marginal treatment response functions separately and calculate marginal treatment effects as the difference between them (Heckman and Vytlacil, 2007). All estimations were done via `mtefe` in Stata (Andresen, 2018).