Lessons from a partially controlled field trial

**Philip J. Cook · Mallory O'Brien · Anthony Braga ·
Jens Ludwig**

Abstract

Objectives Using the case of an on-going work-oriented prisoner-reentry experiment in Milwaukee, describe the challenges of organizing and sustaining a high-quality trial in the field in which only the randomization and data analysis are directly "controlled" by the evaluation team.

Methods The case study is of a randomized experiment involving youthful male prisoners with a history of violence and gang membership, scheduled for release into Milwaukee. The intervention included six months of pre-release services with a work-release opportunity, and intensive services and supervision following release. The case study describes the initial experimental plan and how much of that plan could be salvaged in the face of delays, administrative errors, and other problems.

Results The initial plan, when compared with the actual experiment, specified a larger and more homogeneous sample, more resources devoted to various aspects of the treatment, and more intensive supervision following release. These problems arose despite the best efforts of public officials. Randomization was preserved, and for that reason the results will still be of interest, although perhaps under-powered.

Conclusions The "gold standard" may become a bit tarnished in the field. It was crucial in this experiment to have a member of the experimental team engaged with the relevant state agencies at every step of the process to sustain this effort and to ensure that the treatment was delivered and relevant data generated. A newsletter and regular meetings with agents proved useful. The outcomes will have high internal validity.

Keywords Corrections · Rehabilitation · Gangs

P. J. Cook (✉) · M. O'Brien
Sanford School of Public Policy, Duke University, Box 90245, Durham, NC 27708, USA
e-mail: pcook@duke.edu

A. Braga
Rutgers University, Newark, USA

J. Ludwig
University of Chicago, Chicago, USA

Springer

Introduction

The high recidivism rate of the 700,000 prisoners released each year is costly to communities and government agencies. Indeed, the correctional system does little to prepare inmates for life after release, and more than two-thirds are rearrested for serious crimes within three years (Langan and Levin 2002). The problem is not just limited resources for rehabilitation-oriented programming, but lack of knowledge. There is scant scientific evidence to guide jurisdictions in developing reentry programs (Travis and Visher 2005).

When Milwaukee was awarded a federal grant through the US Department of Justice Comprehensive Anti-Gang Initiative, the Wisconsin Department of Corrections (WIDOC), in conjunction with several other agencies, developed an experimental program that was intended to shed light on what works in rehabilitating violent criminals. The Milwaukee Safe Streets Prisoner Reentry Initiative (PRI) was designed to provide a variety of reentry services to youthful gang-connected prisoners scheduled to be released into Milwaukee, with a focus on improving their employment prospects following release. A rigorous evaluation of this Initiative was made possible by WIDOC's decision to randomly select the group that was to receive special services from all the prisoners who were eligible and who volunteered. The Smith Richardson Foundation agreed to fund an external evaluation; that evaluation, with the authors of this paper as investigators, is ongoing.

WIDOC took the lead in developing the intervention, and has also managed the identification and recruitment of subjects, the delivery of the treatment, the data collection, and the transfer of data to the evaluation team. One member of the evaluation team, Mallory O'Brien, is based in Milwaukee, and has worked with state and local law enforcement agencies for many years. She controlled the random assignment process for the Safe Streets PRI, and has played an advisory and exhortative role in the implementation and monitoring of the treatment. But the PRI was created and administered by public agencies.

As it turns out, the PRI has evolved over time. Several changes from WIDOC's initial plan have been in the direction of dilution. The actual experimental sample is smaller, the subjects are more varied, and the resources devoted to the treatment group before and after release are somewhat less than planned. Inevitably some subjects have been lost to the experiment on both the treatment and control side. One consequence of these changes is to reduce the power of the experiment to detect effects of the treatment.

Nonetheless, the experimental treatment as actually delivered is still of interest. The unusual feature from the beginning, which has not changed, has been to combine pre-release programming with post-release follow-up services. The goal here is to create a relatively seamless transition to life on the outside whereby the offenders are ready to hold a job and have an opportunity to do so right away; thus perhaps reducing the high recidivism rate characteristic of the first few weeks following release. That remains an important feature of the PRI treatment. Furthermore, the randomization has not been compromised, and data on post-release arrests and revocations should be accurately recorded. At the end of the day, it will be reasonable to conclude that any significant differences between the treatment and control groups that are greater than chance are due to the treatment.

Motivation for this intervention

A simple, plausible idea animates the Milwaukee Safe Streets PRI: If released prisoners have a good chance to support themselves through legitimate means, they will be less inclined to return to criminal activity. Two mechanisms that may connect work to crime desistance seem particularly plausible. One is the effect of employment on social context: working in a legitimate job occupies time that might otherwise be spent on the street, and provides a group of associates who are presumably not (much) involved in criminal activity. A second is deterrence: improved legitimate opportunities provide the individual with something to lose if he is rearrested and sent back to prison, and hence sharpen the deterrent effect of that threat.

Prisoners are a highly disadvantaged group, whose employment prospects tend to be limited by their meager education, spotty work history, and unstable family life (Petersilia 2005; National Research Council 2008; Raphael 2011). Between 70 and 85 % of inmates have substance abuse problems (Petersilia 2000). Clearly this is a group that would have poor prospects even without the deleterious effects of criminal record and imprisonment. Survey data on employer preferences demonstrate widespread reluctance to hire ex-cons, particularly for positions that require substantial contact with customers (Holzer et al. 2006). Furthermore, employment regulations play a role: convicted felons are barred from a wide variety of sensitive occupations by federal, state, or local law (Bushway and Sweeten 2007).

In recent years, there has been an increase in interest and considerable policy action involving prisoner reentry issues, mostly at the state level (the traditional locus for corrections), but with financial support and some guidance from the federal level. Early work by the U.S. Department of Justice-sponsored Reentry Partnership Initiative informed a significant push to identify effective strategies for reentry (Travis 2005; Taxman et al. 2001a, b). The initiative encouraged state and local policymakers to develop multi-pronged strategies for successful reentry and continued community safety (Taxman et al. 2001b; Travis 2005). Subsequently, Congress has enacted a series of related programs, including the Offender Reentry and Community Safety Act of 2001, and the Second Chance Acts of 2004 and 2007.

In 2003, several federal agencies collaborated to establish the Serious and Violent Offender Reentry Initiative (SVORI), a large-scale program providing over $100 million to 69 grantees to develop programming, training, and state-of-the-art reentry strategies at the community level (Lattimore and Visher 2009). The SVORI programs were intended to reduce recidivism, as well as to improve employment, housing, and health outcomes of participating released prisoners. The 69 grantees operated 88 distinct programs for released prisoners. The services provided by these programs have focused on enhancing employment, community integration, family unification, substance abuse treatment, and skills building (Winterfeld and Lindquist 2005). RTI International and the Urban Institute conducted a quasi-experimental evaluation of SVORI in 16 selected sites (12 adult and 4 juvenile) in an attempt to estimate the impact of these reentry programs. The evaluation compared outcome measures between program participants and those who are in some sense comparable to SVORI participants, but has not used a randomized experimental design (www.svori-evaluation.org). The main conclusions, based on criminal record data and repeated interviews, were that SVORI programs were responsible for a modest increase in "the types and

amounts of needs-related services provided before and after release from prison (Lattimore and Visher 2009, p. vi)" but "for the adult men, there were no differences in arrest and reincarceration rates at 24 months (about 70 and 40 %, respectively)" despite some evidence of a decline in drug use. The results for women were rather odd, with lower arrest rate and a higher re-incarceration rate.

An extensive meta-analysis of the scientific literature on crime-control identified some more promising results on re-entry programs (Drake et al. 2009). In particular, vocational programs and basic-education programs, as well as cognitive-behavioral therapy, score well among the in-prison programs; workforce development efforts following release also reduced criminal offending. While all the 50-plus evaluations in the meta-analysis included some sort of control group, very few were based on randomized trials, and hence, keeping in mind Rossi's metallic laws of evaluation (Rossi 1987), may tend to provide a somewhat rosy set of conclusions (Drake et al. 2009; Raphael 2011).

One type of intervention that has been subjected to more rigorous evaluation is transitional employment for released prisoners (Visher et al. 2005; Raphael 2011). During the 1970s, the National Supported Work experiment evaluated the effect of providing transitional jobs to four at-risk groups (welfare recipients, released offenders, drug addicts, and high school dropouts). The initial evaluation found little impact of any sort for the ex-offender group, but a reanalysis based on all subjects with a criminal record found that while youthful subjects did not benefit, those aged 27 and over exhibited a substantial reduction (over 10 %) in arrests over three years in comparison to controls (Uggen 2000).

Promising results from transitional employment have also been generated by the ongoing randomized experimental evaluation of the Center for Employment Opportunities (CEO) in New York City. Ex-prisoners who showed up at the Center were placed in temporary, minimum-wage jobs with crews that worked under contract to city and state agencies, and then offered help in finding permanent jobs. (The control subjects in this experiment received basic job-search assistance.) The principal finding indicates reduced recidivism in the three years of follow-up with experimental subjects (Redcross et al. 2012), despite the fact that the treatment did not succeed in increasing employment after the transitional job ended. An analysis of heterogeneity in effects found that CEO caused reductions in recidivism for former prisoners who were at highest risk of recidivism (148 of the 568 subjects in the treatment group), for whom CEO reduced the probability of rearrest, the number of rearrests, and the probability of reconviction two years after random assignment (Zweig et al. 2010). Those in the medium and lowest-risk group had outcomes that were no better than the control group. The risk groups were defined by age, and prior arrest history; risk falls with age and increases with number of prior arrests. Seemingly at odds with Uggen's (2000) findings, then, younger subjects appear to have benefited more than older once arrest record is taken into account.

A similar experiment with transitional jobs is ongoing in four Midwestern cities (Chicago, Detroit, Milwaukee, and St. Paul), with the Joyce Foundation as the lead funder. In 2007–2008, more than 1,800 men who had recently been released from prison were assigned, at random, to a treatment group (offered transitional jobs) or a control group (offered basic job search assistance). The take-up rate by the treatment group was 85 % for the subsidized transitional jobs, but the treatment group was no more likely to

work in an unsubsidized job than the control group (Redcross et al. 2009, 2010). At the end of the first year, only about one-third of both groups was employed in the formal labor market, and there were no consistent impacts on recidivism during the first year of follow-up. (About one-third of each group was arrested.) These findings, while discouraging, are not inconsistent with the CEO findings, and it is possible that some positive effects may emerge in the second year of follow-up.

We can safely conclude that an offer of temporary low-wage jobs to released prisoners who volunteer to participate has the effect of increasing employment for this group so long as the jobs are available; and not thereafter. Whether the resulting increase in employment during the first couple of quarters following release reduces recidivism remains an open question. It is clear that low-wage employment is no magic bullet.

The Milwaukee Safe Streets PRI also focuses on improving employment opportunities, but takes a broader approach than these transitional-jobs experiments. The PRI treatment includes pre-release employment-related programming and relatively intensive post-release services, including but by no means limited to job creation. By the time they are released, offenders are prepared with basic job-finding and job-holding skills, and in many cases have been employed through a work-release program. The goal is a smooth transition to life on the outside, where barriers to working have been stripped away and the motivation to work is ensured by a mandate to work enforced through intensive parole supervision.

Origins and organization of the safe streets initiative

In 2006, the US Department of Justice initiated the Comprehensive Anti-Gang Initiative (CAGI) as a follow-on to Project Safe Neighborhoods. In a competition among the US Attorneys offices, the Eastern District of Wisconsin, which includes Milwaukee, was selected as one of the first six sites. Each winning District received $2.5 million for the initiative, with $1 million for *prevention*, $1 million for *enforcement*, and $500 K for *reentry* activities. The US Attorney's Office, in conjunction with many of the Milwaukee Project Safe Neighborhoods (PSN) partners, developed the implementation plan for the CAGI funds; this overarching umbrella was called Milwaukee Safe Streets. On the reentry front, WIDOC, from the Office of the Secretary, took the lead devising what has become PRI. Funds became available in early 2007.

The prisoner reentry component was sketched in the US Attorney's CAGI proposal as targeted on 100 "high-impact gang members returning to police districts 2 and 5 in the City of Milwaukee after a period of confinement….The reentry component of the program will include: pre-release reach-in activities that are designed to establish a comprehensive, individualized release plan which will prepare the offender for a seamless transition from confinement to the community; post-release services that include the use of a voucher program delivery system to address the offender's criminogenic needs; intensive case management along with high-risk supervision, working closely with law enforcement, prosecutors, treatment and service providers, faith based and community organizations, mentors and others with the community (Lipscomb 2006, p. 19)." The proposal, in short, was to concentrate a variety of available resources on high-risk prisoners released to high-crime areas of Milwaukee.

WIDOC leadership was committed to finding out whether this approach would help high-risk prisoners successfully navigate the challenges of reentry. As a member of the Advisory Group for Milwaukee PSN, Dr. O'Brien participated in the discussions that ultimately shaped the Milwaukee Safe Streets activities. Through these discussions, the US Attorney's Office, but more importantly WIDOC leadership, became convinced that since resources for the re-entry initiative were limited, they could serve only a subset of those inmates who were eligible, which created an opportunity for a rigorous evaluation of the initiative by selecting the treatment group randomly from the larger group of released prisoners who fit the criteria.

While the CAGI grant provided funding to support the delivery of the treatment, there were no funds to support an evaluation of this treatment. We applied for, and in 2008 received, a grant from the Smith Richardson Foundation to conduct this evaluation. The attractive features of the grant application included the strong experimental design, public interest in reducing recidivism by high-risk violent prisoners, and the unusually intensive treatment that included the reach-in feature as well as post-release programming.

At this point, the division of responsibility was clear. The Milwaukee Safe Streets PRI was a WIDOC program. The recruitment and supervision of prisoners, delivery of services before and after release, and data collection were all to be managed by WIDOC officials. We, the evaluation team, would be based at Duke University, and would be directly involved in the process only in managing the random assignment between treatment and control. Our main task was to be evaluating the effects of this initiative using data made available by WIDOC. This division of responsibility informed the procedures for human subjects protection that were worked out with the Duke IRB, which proceeded on the understanding that the evaluation team was conducting secondary analysis of the data produced by WIDOC, and that WIDOC itself was responsible for gaining informed consent and protecting the rights of the subjects while they were under correctional supervision. In practice we have had no contact with experimental subjects, and our contact with WIDOC service-delivery staff over the years of the PRI has focused on generating the necessary data (Table 1).

Implementing the field experiment

Researchers with experience in evaluating field experiments know that there are likely to be some bumps in the road (Berk 2005). Our experience has been no exception. We are happy to report that, while the process has not been smooth, the actual (as opposed to

Table 1 Control of various aspects of the PRI experiment

	WIDOC	DUKE
Design intervention	X	
Identify and recruit subjects	X	
Perform random assignment		X
Deliver services	X	
Generate data	X	
Analyze data		X

the planned) experiment is still of interest. Here, we recount some of the bumps and suggest some lessons.

Getting organized: a false start and eventual success

The PRI was conceived from the beginning as a multifaceted intervention involving a number of WIDOC units as well as several contractors. Inmates slated for release in Milwaukee were scattered across correctional facilities around the state. PRI entailed determining who among them were eligible for the intervention, gaining their consent to participate, randomly selecting those who would be exposed to the treatment among those who did consent (the PRI group), moving the PRIs to Racine six months prior to release, assessing their needs, organizing the provision of pre-release services, and managing a continuing flow of services following release.

In May 2007, a PRI coordinator was appointed by the Secretary's Office and supervised by Region 3 of WIDOC's Division of Community Corrections. The coordinator was given access to the WIDOC inmate records for identifying those eligible for the experiment. Her other task was to pull together the various pre-release and post-release services. The main thrust of these services was to prepare the offenders for reentry, specifically job preparedness. The coordinator facilitated monthly meetings with key stakeholders—WIser Choice for alcohol and other drug-abuse treatment, the US Attorney's Office, Division of Adult Institutions, and the Division of Community Corrections—to make the intervention operational. The planning process was slowed by the coordinator's maternity leave, but in March 2008 she began recruitment of subjects. Unfortunately, the Racine Correctional Institute (RCI) was not yet prepared to deliver the required services, and indeed the contracts had not yet been signed or RCI staff given the necessary training. As a result, the first subjects that were recruited did not receive services, and had to be dropped from the experiment. Recruitment was discontinued for over a year.

In August 2008, the US Attorney's Office contracted with an outside reentry workforce consultant to assist WIDOC in organizing service delivery in Racine and Milwaukee. Detailed pre- and post-release activities were planned (see Tables 2 and 3). Meanwhile, the evaluation team planned the data collection piece. We worked closely with WIDOC's Bureau of Technology Management (BTM) to determine data fields available from the various WIDOC records management systems and develop a plan for formatting and transferring them.

Table 2 Calendar of PRI events

Date	Event
2006	CAGI Award to Milwaukee$500,000 for proposed prisoner re-entry services
2008	Smith Richardson Foundation funds Duke team
2008	WIDOC coordinator begins recruiting subjects, then stops due to glitches in planning
2009 (April)	WIDOC re-commences recruitment
2010 (May)	Most of 236 subjects now released

Table 3 Proposed vs Actual PRI Activities

	Proposed as of Spring 2008	Actual or planned as of March 16 2011
Number of subjects	100 Es	106 Es
	200 Cs	130 Cs
Intervention Schedule	Intake: 5/08 to 4/09	Intake: 1/09 to 8/09
	Services: 6/08 to 4/10	Services: thru summer '11
Criteria for eligibility	Male	Male
	Age 18–30	Age 18–35
	Gang member	Violent record or gang member
	Risk to community	Risk to community
	>6 months supervision	>6 months supervision
	Connection with PD 2 or 5	Release to Milwaukee
	Informed consent	Informed consent
Randomization procedure	Evaluation team	Evaluation team
Intervention pre-release	Move to Racine CI	Move to Racine CI or Sturtevant Transitional Facility
	Work related training	Employment preparation through CCEP Job coach
		Cognitive-Behavioral Program & gang-diversion & AODA counseling
		Bi-weekly meeting with social worker. Risk and needs assessment
	Case worker to inmate ratio 1:20	Case worker to inmate ratio 1:40
	Work release	Work release with housing in Sturtevant Transitional Facility
	Employer recruitment (Rotary)	CCEP reach-in services
	Employer matching	
Intervention post-release	Intensive supervision by High Risk Unit	Same, but Agent to offender ratio 1:40
	Agent to offender ratio 1:20	
	Wrap around services	
	AODA needs (WIser Choice)	AODA needs (Wiser Choice)
	Team Meetings	CCEP services
		Team meetings: parole agent, job coach, social worker, AODA case worker

In February 2009, a new project coordinator was appointed, this time supervised directly by the WIDOC Director of Reentry for the state. This seemingly minor change of supervision, from the regional level to state level, facilitated implementation. The new coordinator could work directly with all divisions within WIDOC. Some of the main challenges were soon overcome: prisoner movement; the data-transfer plan; coursework development through the Division of Adult Institutions; and a plan for the pre- and post-release role for Community Corrections Employment Program. In April 2009, more than

a year after the false start, the effort to recruit inmates for the experiment re-commenced. This time all went well, and inmates who won the "coin toss" began transferring to Racine that summer.

Looking back, it is probably fair to say that the organizational complexities of implementing the multi-faceted PRI intervention were not fully anticipated. The difficulties cannot be attributed to the experimental overlay of PRI—rather, it was intrinsic to the intervention itself.

The missing subjects

The initial plan for PRI placed a variety of restrictions on eligibility. Recruitment of subjects was to be limited to male gang-connected inmates aged 18–30 who were scheduled for release into police districts 2 or 5 in Milwaukee with at least six months of community supervision. Various other restrictions applied as well. WIDOC officials had concluded from an initial analysis of their inmate population that there were 500 or more inmates who met all these criteria and would be released during the appropriate time period. We did not foresee any problems, since for purposes of the experiment we had decided to enroll just 300 inmates—a control group of 200 to go with the 100 members of the treatment group. (The size of the treatment group was determined by the Safe Streets PRI budget for the treatment services.) Yet when recruitment began in April 2009, a careful search of inmate lists did not produce nearly enough eligibles. WIDOC officials tried to address this problem by easing eligibility restrictions: the age range was extended to 35, the geographic restriction was eased to encompass the entire city of Milwaukee, and the gang-affiliation requirement was expanded to include all violent criminals. But there were still not enough eligible inmates, despite the near-100 % enrollment rate. When it became clear that there was no way to build up to the planned sample size, we decided to keep the experimental group near 100 and absorb all the shrinkage in the control group. That decision was based on statistical power considerations.[1]

The main effect of a smaller control group is to reduce the statistical power, which may prove to be a problem. On the other hand, we are not concerned about the easing of the eligibility conditions, since the sample still consists of high-risk offenders, as planned. The fact that the subjects in the experiment are more heterogeneous with respect to age, gang membership, and location of residence within Milwaukee changes the nature of the experiment somewhat, but does not detract from its validity as a test of the basic approach.

Random selection process

Inmates who were eligible for the experiment were informed by the prison social worker that, if they consented to participate, there was a one-in-three chance that they would be included in the treatment group. Dr. O'Brien selected the "winners" each week. WIDOC submitted to her the list of inmates that consented to participate. The

---

[1] Other things equal, and assuming that the treatment does not affect the variance of the outcome, then in an unbalanced experimental design, reallocation of given number of subjects in the direction of equality has the effect of increasing the precision of the estimated treatment effect.

list included WIDOC number, current institution, release date and termination date. The WIDOC number was entered into a computer-generated list randomizer and the numbers were randomized. From the randomized list, beginning with the first number, every third number was selected for treatment, thus providing the "winners"

When it became clear that we needed additional inmates in the treatment group, 16 inmates were chosen at random from the control group and reassigned to the treatment group. It is an interesting question whether the reassignment influenced behavior or attitudes of subjects who were first told, in effect, that they had lost the coin toss, and then that they had won after all.

Some dilution of "dose"

The multi-faceted treatment began with a transfer of PRI inmates to Racine six months prior to scheduled release. Racine is just 30 miles from Milwaukee, which provided a chance for work-release employment for the last few weeks of their sentence.

In Racine, PRI inmates were housed either in the Racine Correctional Institute, or in the Sturtevant Transitional Facility. In either case, they met regularly with a social worker and were assessed using a number of standard protocols. The social worker was responsible for handling pre-release activities, including obtaining necessary identification, Social Security numbers, and birth certificates. All PRI inmates were expected to participate in the Breaking Barriers Curriculum, and given access to reach-in services of the Community Corrections Employment Program (CCEP), alcohol and drug treatment (through WIser Choice), and remedial education. For minimum-security offenders who were placed in Sturtevant, work release was a possibility at the end of their term.

The coordinated-care team met with each PRI inmate 30 days prior to release, and continued to meet with him monthly following release. The PRIs continued to be eligible for AODA treatment in the community, as well as CCEP services, which often included job creation through subsidies to employers. They were subjected to enhanced supervision for six months.

In defining the experimental treatment, it is important to realize that the control-group inmates were in principle eligible to receive most of the same services as the PRI inmates. In practice, there was a rather sharp distinction based on the reality that these services are in short supply. Where the PRIs were guaranteed access, the controls were only able to get in line. The planned treatment was in effect to make available to PRIs everything that WIDOC had to offer, with the move to Racine thrown in as a bonus. It appears that this plan was implemented with some fidelity, although the case loads for the social worker in Racine and the parole agents were nearer 40 than the planned 20.

The main difference between the planned and actual treatment was the loss of the services of the Rotary Club. When Milwaukee won the CAGI grant in 2006, the local Rotary Club was part of the prisoner re-entry plan. Rotary members were committed to interviewing the Safe Streets inmates and hiring some of them even before they were released. Unfortunately, the long delays in the implementation of the initiative were too much, and the Rotarians turned their attention elsewhere. Job placement was left to CCEP, which has some funds for job creation. But the "dose" would have been stronger if Rotary had stayed in the mix.

Table 3 summarizes the various changes that occurred between the initial plan and the actual implementation of the PRI.

Social Interactions and SUTVA

The experimental process may have influenced the inmates who were recruited into the experiment in ways that were extraneous to WIDOC's intended treatment. The random assignment process may have affected the outlook of some inmates as they learned whether they were "winners" or "losers." That is a generic problem for experiments where "double blind" is not an option. In the Milwaukee PRI, there is also the possibility that the experiment created a new set of social interactions among inmates that had some influence on their behavior following release.

As noted, the PRI treatment group was relocated to Racine, and most of them ended up at one of two facilities there. The fact that the move brought most of them closer to home and gave them access to work-release jobs could be considered part of the treatment. Concentrating this group also created the opportunity for socializing among this experimentally created peer group. Some treatments were delivered to groups of inmates, including gang-diversion sessions. These group activities provided an opportunity for communication among members of the treatment group, including comparisons of expectations for participation. It is logically possible that the nature and strength of social influence might depend on the precise mix of individuals who happened to be assigned to the treatment group. For example, if several members of the same gang were assigned to a particular session, that prior relationship might influence the level of engagement with the gang-diversion programming. This could be considered a possible violation of what is known as the stable unit treatment value assumption (SUTVA) (Rubin 1990; Sampson 2010).

As an aside, it is interesting to observe that one effect of these prisoner-led discussions, when brought to the attention of RCI staff, was a closer adherence to the proposed treatment and increased effort by the staff to explain what was required of the inmates if they were to take full advantage of the treatment options.

It is not possible to determine whether this artificial feature of the experiment had meaningful influence on the behavior of the PRI treatment group. Our sense is that recidivism rates are not so easily changed. The rather discouraging history of systematic efforts to reduce recidivism supports this view.

Changing economic environment

Any field experiment is vulnerable to changing conditions in the field. Given the employment focus of the Safe Streets PRI, it is reasonable to presume that the condition of the Milwaukee labor market has been relevant. Ex-cons, who have limited employment opportunities in the best of times ("last hired, first fired"), are likely to be especially vulnerable to the weak labor market conditions associated with a recession.

When PRI was first proposed in 2006, Milwaukee County's unemployment rate was 5.6 %. A couple of years later, the local economy followed the national economy into the Great Recession. The unemployment rate jumped in early 2009 and averaged 9.3 % in that year, the first year that PRI subjects were released. The labor market has remained weak since then.

The concern from the evaluation viewpoint is that the employment-oriented treatment would be more effective in improving job prospects during good times than bad, in which case we will not observe the best case for the effectiveness of the PRI treatment. But, in fact, we do not know whether these services are more or less effective during recessions than during good times. That is, while we expect that released offenders will have more difficulty finding a job during a recession, the relative advantage conveyed to the treatment group in the Milwaukee PRI is not necessarily lessened.

Supplementing administrative data

The principal outcomes for this experiment are various measures of employment and of recidivism. Recidivism is readily measured by criminal-justice records on arrests, convictions, and revocations, all readily available to the evaluation team through WIDOC records. (WIDOC data are transferred to Duke on a regular basis by WIDOC.) Data on employment are more difficult to obtain, especially for members of the control group. For that group, parole agents may not have the time to check or record employment for each client, given their large caseloads. The alternative source of data on employment is the Employment Security system. We plan to request state employment records, but accessing those records requires Social Security numbers. Some of the subjects did not obtain Social Security numbers until after their release. Our staff has been successful in obtaining this information through repeated queries to the parole agents.

All the PRIs are supervised by one of two parole agents, and we are in touch with them. The control group parolees are divided among dozens of agents with whom we have no contact in the normal course of events. We have reached out to them by way of a colorful newsletter that provides updates on the experiment and the people involved. Distribution of the newsletter has provided a vehicle for being in touch with the parole agents, and many of them have responded and provided the necessary information on clients who are part of the PRI control group.

Enhancing power

A real advantage of this experiment is the access to extensive administrative data. These data, conveyed from WIDOC files, provide accurate measures of subjects' criminal histories, which are predictive of recidivism probability. (By prior agreement, WIDOC has transferred criminal-history information on all inmates in the experiment to Duke.) Controlling for these baseline covariates should not affect the estimated treatment effect of the intervention, since we expect treatment-group assignment to be uncorrelated with baseline characteristics by virtue of random assignment.[2] But controlling for baseline covariates ordinarily improves the statistical power of the study by reducing the

---

[2] David Freedman (2008) has argued that, if the assumptions of the OLS model are not met, then the inclusion of covariates may result in a biased estimate of the treatment effect, and that under some circumstances the adjustment can actually worsen the precision of the estimator. Under Freedman's influence, Berk et al. (2010) assert that "Random assignment does not justify any form of regression with covariates." Lin (2011) reviews the controversy and demonstrates that the statistical problems with regression adjustment tend to minor or readily fixed. All agree that a report of experimental results should begin with the difference in means between the experimental and control groups, and that if the inclusion of covariates causes a large change in the estimated treatment effect, that is cause for concern.

amount of residual variation in the post-randomization outcomes of interest, shrinking the standard errors of the impact estimates. Howard S. Bloom (2006) illustrates the increase in power resulting from use of covariates this way: "For example, an R_A^2 of 0.25 yields an effective sample that is one-third larger than that without covariates; an R_A^2 of 0.50 yields an effective sample that is twice as large; and an R_A^2 of 0.75 yields an effective sample that is four times as large (p. 12)." In this statement, R_A^2 is defined as "the proportion of pooled unexplained variation in the outcome within experimental groups predicted by covariates…(p. 12)."

The calculations reported in Table 4 further quantify the effect of covariates. With a study sample of our size, with no baseline covariates the minimum detectable effect size (MDES), or the smallest effect we could expect to detect, would be .37 standard deviations. If we had baseline covariates that had an R-squared in a regression against the outcomes of .33, the MDES would be reduced by around one-quarter, and if the R-squared on our baseline covariates was .71 the MDES would be shrunk about in half to 0.2 standard deviations. In that case, the study would be well-powered, since the convention (at least in the education research community) is to view effects of .2 standard deviations as "small."[3] By way of comparison, absent any baseline covariates, in order to improve our power by reducing MDES from .37 to .2 we would have to increase the size of our experiment by more than a factor of three. To the extent to which accessing administrative records that already exist is inexpensive, but expanding the study pool to provide the intervention to more study subjects is costly, enhancing the administrative data infrastructure that is available in jurisdictions can be a very cost-effective way to facilitate more informative randomized experimental evaluations of important government services.

Summing up

Like many another field experiment, the Milwaukee Safe Streets PRI has experienced some problems in implementation; the treatment and other facets of the experiment have evolved between the initial plan and the final execution. Depending on expectations, one could focus on the long delays in launching the PRI, or on the remarkable fact that when PRI finally did become operational, it was with a high degree of fidelity to the original concept. Surely part of the answer to the first issue relates to the complexity of the planned treatment, and the fact that the PRI was not a high priority for any of the operating agencies that were required to take on extra work in order to make it happen. Regarding the second matter, we note that PRI benefited from local leaders who did take the experiment seriously. Also important was the ability of the evaluation team to have some influence on the ground.

Some final reflections

In one image of the ideal controlled experiment, the researcher controls the design of the treatment, the selection of subjects and their assignment to treatment, the delivery

---

[3] Bloom et al. (2008) raise questions about the origins of this rule of thumb in education research, and suggest that the judgment of what is "small" or "large" should be made relative to the context of comparison.

Table 4 Some statistical power calculations

Number of study subjects (treatment plus control)	R-squared of baseline covariates in explaining variation in outcomes	Minimum detectable effect size (MDES), in standard deviation units
236	0	.37
236	.33	.30
236	.71	.20

By "baseline covariates" we exclude the indicator for treatment group assignment itself. We summarize the statistical power using the Minimum Detectable Effect Size (MDES), a widely-used statistic recommended by Bloom (1995) (see also Orr 1999) that is defined as the smallest true effect that has 80 % power for a two-tailed test of statistical significance at the .05 level, and is equal to 2.8 times the expected standard error of our impact estimate (where the outcome is standardized to be expressed in standard deviation units). Other ways of summarizing the expected statistical power of an experiment are also possible, of course, and are easily calculated using the Optimal Design software developed by Stephen Raudenbush; see http://www.wtgrantfoundation.org/resources/overview/research_tools

of the treatment, the data collection process, and, as much as possible, the environment. But for the Milwaukee Safe Streets PRI, most aspects have been controlled by public agencies. The only aspect that we researchers have controlled is the assignment of subjects to treatment or control groups. That narrow role is characteristic of "one off" field experiments (Bannerjee and Duflo, 2008). It should be said that the control of assignment is, by itself, of considerable scientific value. Assuming the process for recording administrative data on outcomes is not confounded by the experimental intervention, then random assignment is enough to ensure that if a difference in mean outcomes between the treatment and control groups is observed, the difference can be treated as an unbiased estimate of the causal effect of the treatment. Of course, providing a valid interpretation to this result requires having measures of the differential treatment received by the two groups.

This division of responsibility is a cost-effective arrangement, with the promise of considerable value relative to the size of the budget for evaluation. The leverage here derives from the fact that the intervention and data production are being provided at public expense by real-world agencies.

While randomized field experiments are considered the best way to estimate causal effects, these designs are limited by the fidelity with which they are implemented (Berk 2005). There is a long and ever-growing literature on field experiment implementation problems; many well-known criminal justice field experiments have experienced and successfully dealt with methodological difficulties. Our ability to detect, diagnose, and remedy implementation problems was greatly enhanced by having a member of the research team who was already well-embedded in local criminal justice professional networks. She was a steady advocate for the integrity of the research design and the provision of the promised services to the experimental group.

The fact that the intervention was designed by WIDOC staff oriented the experimental treatment to what seems an important question for WIDOC operations. The planned treatment incorporated available services delivered pretty much as usual, except the dose (and the expenditure per offender) was much larger than usual. In effect, the experiment became a test of whether the basic approach incorporated in WIDOC programming had promise for reforming a high-risk population, if enough resources were brought to bear.

If the PRI treatment group do better than the control group, that result will affirm WIDOC's basic approach. Of course, if there is no significant difference, then that may or may not suggest the failure of the approach, depending on the power of the experiment. If it turns out that the available covariates are highly predictive of post-release outcomes, then the experiment will be powerful enough to detect even small effects, and there will be a clear indication of whether the intervention would pass a cost-benefit test.

From the policy viewpoint, a potential drawback of the WIDOC-designed intervention is its complexity, although this issue is itself complex. The multi-faceted nature of PRI and other "kitchen sink" or "synergy" interventions can be motivated by the hypothesis that study samples that have multiple barriers to success need to have each of these barriers addressed in order to succeed. Or, put in statistical terms, the different services or mechanisms through which the intervention tries to help subjects may interact, and so have more-than-additive effects on outcomes. However, the assumption that a multifaceted treatment is required is not itself tested by a multi-faceted intervention. All that we can learn is the "black box" conclusion, namely that the treatment—in this case a combination of relocation prior to release, pre-release services, post-release services, and intensive supervision—was sufficient to reduce recidivism. It will not be possible to determine which of these aspects was really necessary, and which could be done without or with less. But of course a positive result from this experiment might be just what is needed to launch a series of narrower inquiries.

How might those future inquiries proceed? A scientific approach would entail creating an experiment or series of experiments that tests the basic ingredients individually or in different combinations. For the sake of argument, the ingredients in this experiment might be boiled down to pre-release preparation for job-finding and job-holding, pre- and post-release treatment for alcohol and drug problems, some temporary subsidies to create jobs, close supervision following release, and counseling. There are 32 logical combinations of these five dimensions (assuming that each of them is either "present" or "absent" in the treatment). One could imagine a mega-experiment with that many treatment arms, but it seems unlikely to say the least. The goal then would be to further reduce the treatment to still fewer essential mechanisms.

In any event, the Safe Streets PRI is not part of a scientific agenda, so much as it is designed to evaluate a particular operational approach to a real-world task. Will a concentrated dose of the sorts of reentry services offered in state corrections be enough to reduce crime involvement by high-risk offenders? That is surely an interesting question, and we look forward to being able to provide an answer.

References

Bannerjee, Abhijit V. and Esther Duflo (2008) "The experimental approach to development economics" Cambridge, MA: NBER Working Paper 14467.

Berk, R. (2005). Randomized experiments as the bronze standard. *Journal of Experimental Criminology, 4* (1), 417–433.

Berk, R., Barnes, G., Ahlman, L., & Kurtz, E. (2010). When second best is good enough: A comparison between a true experiment and a regression discountinuity quasi-experiment. *Journal of Experimental Criminology, 6*, 191–208.

Bloom, H. S. (1995). Minimum Detectable Effects: A Simple Way to Report the Statistical. Power of Experimental Designs. *Evaluation Review, 19*(5), 547–556.

Bloom, Howard S. (2006). *The Core Analytics of Randomized Experiments for Social Research."* MDRC: MDRC Working Papers on Research Methodology.

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). "Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions". *Journal of Research on Educational Effectiveness, 1*, 289–328.

Bushway, S., & Sweeten, G. (2007). Abolish lifetime bans for ex-felons. *Criminology and Public Policy, 6*, 697–706.

Drake, E. K., Aos, S., & Miller, M. G. (2009). Evidence-based public policy options to reduce crime and criminal justice costs: Implications in Washington State. *Victims and Offenders, 4*, 170–196.

Freedman, D. A. (2008). On regression adjustments to experimental data. *Advances in Applied Mathematics, 40*, 180–193.

Holzer, H. J., Raphael, S., & Stoll, M. A. (2006). Perceived criminality, criminal background checks and the racial hiring practices of employers. *Journal of Law and Economics, 49*(2), 451–480.

Langan, P. A., & Levin, D. J. (2002). *Recidivism of Prisoners Released in 1994*. Washington, DC: Bureau of Justice.

Lattimore, Pamela K. and Christy A. Visher (2009) *The multi-site evaluation of SVORI: Summary and synthesis*. RTI International.

Lin, Winston. (2011) "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique." http://polmeth.wustl.edu/media/Paper/agnostic20110901.pdf (accessed 3/19/12).

Lipscomb, William (2006) *Milwaukee Safe Streets Initiative: Project Safe Neighborhoods Anti-Gang Pilot Site* Application to U.S. Department of Justice, Comprehensive Anti-Gang Initiative.

National Research Council. (2008). *Parole, Desistance from Crime, and Community Integration. Committee on Community Supervision and Desistance from Crime*. Washington, DC: National Academies Press.

Orr, L. L. (1999). *Social experiments: evaluating public programs with experimental methods*. Thousand Oaks, CA: Sage Publications.

Petersilia, J. (2000). *When Prisoners Return to the Community: Political, Economic and Social Consequences*. Washington DC: US Department of Justice.

Petersilia, J. (2005). From cell to society: Who is returning home? In J. Travis & C. Visher (Eds.), *Prisoner Reentry and Crime in America*. New York: Cambridge University Press.

Raphael, Steven (2011). "Improving employment prospects for former prison inmates: Challenges and policy" in *Controlling Crime: Strategies and Tradeoffs,* edited by Philip J. Cook, Jens Ludwig, and Justin McCrary. Chicago: University of Chicago Press, forthcoming.

Redcross, C., Bloom, D., Azurdia, G., Zweig, J., & Pindus, N. (2009). *Transitional jobs for ex-prisoners: Implementation, two-year impacts, and costs of the Center for Employment Opportunities (CEO) Prisoner Reentry Program*. New York: MDRC.

Redcross, C., Bloom, D., Jacobs, E., Manno, M., Muller-Ravett, S., Seefeldt, K., Yahner, J., Young, A. A., Jr., & Zweig, J. (2010). *Work After Prison One-Year Findings from the Transitional Jobs Reentry Demonstration*. New York: MDRC.

Redcross, Cindy, Megan Millenky, Timothy Rudd, and Valerie Levshin (2012) *More Than a Job: Final Results from the Evaluation of the Center for Employment Opportunities (CEO) Transitional Jobs Program OPRE Report 2011–18* Office of Planning, Research and Evaluation (OPRE), Administration for Children and Families, U.S. Department of Health and Human Services.

Rossi, P. H. (1987). The Iron Law of Evaluation and Other Metallic Rules. *Research in Social Problems and Public Policy, 4*, 3–20.

Rubin, D. B. (1990). Formal Modes of Statistical Inference For Causal Effects. *Journal of Statistical Planning and Inference, 25*(1990), 279–292.

Sampson, R. J. (2010). Gold standard myths: observations on the experimental turn in quantitative criminology. *Journal of Quantitative Criminology, 26*(4), 489–500.

Taxman, F., Young, D., & Byrne, J. (2001a). *From prison safety to public safety: Best practices in offender reentry*. College Park, MD: Bureau of Governmental Research, University of Maryland.

Taxman, F., Young, D., Byrne, J., Hoslinger, A., Anspach, D., Thanner, M., & Silverman, R. (2001b). *The Eight Reentry Partnership Initiatives: Plans, early results & conceptual framework*. College Park, MD: Bureau of Governmental Research, University of Maryland.

Travis, J. (2005). *But They All Come Back: Facing the Challenges of Prisoner Reentry*. Washington, DC: The Urban Institute.

Travis, J., & Visher, C. (eds) (2005) *Prisoner Reentry and Crime in America* New York: Cambridge University Press.

Uggen, C. (2000). Work as a turning point in the life course of criminals: A duration model of age, employment, and recidivism. *American Sociological Review, 65*(4), 529–546.

Visher, C. A., Winterfield, L., & Coggeshall, M. B. (2005). Ex-offender employment programs and recidivism: A meta-analysis. *Journal of Experimental Criminology, 1*, 295–315.

Winterfeld, L., & Lindquist, C. (2005). *Characteristics of prisoner reentry programs. Reentry Research in Action*. Washington, DC: U.S. Department of Justice.

Zweig, J., Yahner, J., & Redcross, C. (2010). *Recidivism effects of the Center for Employment Opportunities (CEO) Program vary by former prisoners' risk of reoffending*. New York: MDRC.

Philip J. Cook is a professor of public policy and economics at Duke University. He is one of nation's leading economists with research specialty in criminal justice policy. His most recent books are *Controlling Crime: Strategies and Tradeoffs* (co-edited with Jens Ludwig and Justin McCrary: University of Chicago Press, 2011) and of *Paying the Tab* (Princeton University Press, 2007).

He has served on expert panels of the National Academies dealing with alcohol-abuse prevention; injury control; violence prevention; school rampage shootings; underage drinking; the prospects for a ballistics reference data base; and the deterrent effect of the death penalty. He is a member of the Institute of Medicine of the National Academy of Sciences, and a Fellow in the American Society of Criminology, as well as a research associate of the National Bureau of Economic Research. He serves as co-director of the NBER Working Group on the Economics of Crime.

Mallory O'Brien Ph.D. has led the development, implementation, and evaluation of the Milwaukee Homicide Review Commission since July 2004. Dr. O'Brien is trained as an epidemiologist and has worked in the field of violent injury prevention for the past 15 years. She is currently a member of the Project Safe Neighborhoods Taskforce for the Eastern District of Wisconsin, the Advisory Committee of the Safe Streets Initiative and the WI Attorney General's Crime Victim Council. Dr. O'Brien is a researcher with the Sanford School of Public Policy at Duke University and a Clinical Assistant Professor in the College of Nursing at Marquette University.

Anthony Braga Ph.D. is a Professor in the School of Criminal Justice at Rutgers University and a Senior Research Fellow in the Program in Criminal Justice Policy and Management at Harvard University. He is also a member of the University of Chicago Crime Lab and a Senior Fellow in the Chief Earl Warren Institute on Law and Social Policy at the University of California, Berkeley. He is currently the President and an elected Fellow of the Academy of Experimental Criminology. He received his M.P.A. from Harvard University and his Ph.D. in Criminal Justice from Rutgers University.

Jens Ludwig Ph.D. is the McCormick Foundation Professor of Social Service Administration, Law, and Public Policy at the University of Chicago, Director of the University of Chicago Crime Lab, and Co-Director of the University of Chicago Urban Education Lab. He is also research associate and co-director of the National Bureau of Economic Research working group on the economics of crime, non-resident senior fellow in economic studies at the Brookings Institution, and co-editor of the Journal of Human Resources. His research addresses urban policy challenges related to crime, schooling, health, housing and poverty. He is a member of the Board on Children, Youth and Families of the Institute of Medicine of the National Academy of Sciences, and recipient of the Association of Public Policy Analysis and Management's David Kershaw prize for contributions to public policy before age 40.