

# An Experimental Evaluation of the Impact of Intensive Supervision on the Recidivism of High-Risk Probationers

Crime & Delinquency  
2017, Vol. 63(1) 3–38  
© The Author(s) 2014  
Reprints and permissions:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/001128714555757  
cad.sagepub.com  


Jordan M. Hyatt<sup>1</sup> and Geoffrey C. Barnes<sup>1</sup>

## Abstract

This article reports the results of an experimental evaluation of the impact of Intensive Supervision Probation (ISP) on probationer recidivism. Participants, who were assessed at an increased likelihood of committing serious crimes and not ordered to specialized supervision, were randomly assigned to ISP ( $n = 447$ ) or standard probation ( $n = 385$ ). ISP probationers received more restrictive supervision and experienced more office contacts, home visitations, and drug screenings. After 12 months, there was no difference in offending. This equivalence holds across multiple types of crimes, including violent, non-violent, property, and drug offenses, as well as in a survival analysis conducted for each offense type. ISP probationers absconded from supervision, were charged with technical violations, and were incarcerated at significantly higher rates. Policy implications for these results are discussed.

## Keywords

community corrections, punishment, intensive probation, randomized trial, recidivism

---

<sup>1</sup>University of Pennsylvania, Philadelphia, PA, USA

## Corresponding Author:

Jordan M. Hyatt, Department of Criminology, University of Pennsylvania, 3718 Locust Walk, 483 McNeil Building, Philadelphia, PA 19104, USA.  
Email: [jhyatt@sas.upenn.edu](mailto:jhyatt@sas.upenn.edu)

Prison overcrowding and other concerns have led to an increased use of community-based supervision for serious offenders, and community corrections have become an increasingly common form of social control (Grattet, Lin, & Petersilia, 2011; Sampson, 1986). In response to increased demand, probation agencies have been forced to adopt increasingly differentiated supervision protocols for the most dangerous offenders in their caseloads. Intensive Supervision Probation (ISP), a control-based approach focused on small caseloads and increased reporting requirements, has been in use for decades, despite non-supportive findings in numerous evaluations. Advances in risk forecasting have left probation and parole agencies with a dilemma. These agencies are increasingly able to identify individuals who threaten public safety, but they have very few evidence-based options for managing these offenders. The result is that they continue to use the responses available and increase the intensity of traditional supervision methods. Updated and convincing research may better inform decisions and lead to policy change (see Lin, 2012).

The tailoring of supervision intensity to actuarially assessed levels of risk is a key component of the principles of risk-needs-responsivity (Andrews, Bonta, & Wormith, 2006; Taxman, Thanner, & Weisburd, 2006) and the implementation of evidence-based programming (EBP) in community corrections (McNeill, Farrall, Lightowler, & Maruna, 2012; Taxman & Belenko, 2012). In addition, recent promising developments with ISP programs combine treatment with enhanced surveillance, and some of these efforts are currently under evaluation across the country (Smith, Gendreau, & Swartz, 2009; Taxman et al., 2006). Because today's ISP often combines two or more different theoretical "levers" to modify offender behavior, there is an emerging need to replicate prior research on ISP so that evaluations can isolate the impact of the supervision component from that of any treatment elements. Understanding this relationship between supervision intensity and outcomes is necessary for the development of "effective and targeted interventions" (White, 2005, p. 12). In addition, identifying the downstream impact of ISP, including reincarceration and the prosecution of technical violations in court, illustrates the full costs of increasing supervision intensity today. We therefore replicate some of the ISP experiments of the past (Petersilia & Turner, 1990a, 1990b, 1993) to update our understanding about the effects of control-focused intensive supervision on serious offenders under modern constraints and conditions, including the use of advanced behavioral forecasting.

## **Background**

### *The Use of ISP for Community-Based Supervision*

In recent years, community corrections have been relied on with increasing frequency. Driven by both concerns about prison overcrowding and shifts in

sentencing trends away from incarceration, increasingly large numbers of offenders are being sentenced directly to non-custodial supervision, and this option is being used for a more diverse range of offenses (Austin, 2010; Petersilia, 2001). At the same time, reforms in indeterminate sentencing policies and parole eligibility have increased the frequency and nature of early release from incarceration (King, 2009; Tonry, 1999). A large majority of these individuals will be placed onto some form of probation or parole (Solomon, Kachinowski, & Bhati, 2005).

These forces have combined to make community corrections an increasingly common source of supervision within the criminal justice system (Clear, Reisig, & Cole, 2013; Pew Center on the States, 2009a). By 2012, approximately 1 in every 50 adults in the United States was under some form of community correctional supervision: 1 in 284 individuals were on parole and 1 in 61 were on probation (Maruschak & Bonczar, 2013). In Pennsylvania, relatively recent estimates of the prevalence of community correctional supervision have been as high as 1 in 37 residents (Pew Center on the States, 2009b). Our increasing reliance on probation and parole for an ever-widening range of offenders and offenses is not simply a temporary trend. This growth places enormous pressure on community corrections agencies, especially as historical data indicate that, at the agency level, staffing and budgetary increases have not kept pace with an exploding population (Gifford, 2002). If nothing else, the net effect of changes in community corrections and sentencing policy has been to increase the quantity and dangerousness of offenders being supervised on probation.<sup>1</sup>

Despite this increasing dependence on their services, community corrections agencies are often faced with criticism that their approach is “soft on crime” and cannot effectively prevent criminal conduct (Petersilia, 1999). This perception may also contribute to chronic under-funding of probation agencies, making it difficult to deliver effective supervision and protect public safety (Beto, Corbett, & Hinzman, 1999). At the same time, recidivism rates are generally high among probationers (Langan & Levin, 2002; Petersilia, 1987). These criminally active probationers and parolees are often reincarcerated (Petersilia & Turner, 1990a, 1993; Turner, Petersilia, & Deschenes, 1992), commonly for committing new offenses (Cohen, 1995; Solomon et al., 2005). These high rates have caused some to “. . . question the ability of community supervision to effect meaningful behavioral change in a direction favorable to public safety” (Lowencamp, Latessa, & Smith, 2006, p. 576).

One way that probation agencies have responded to these critiques is to intensify probation. *Intensive supervision probation* (ISP) is an umbrella term that encompasses many types of institutional responses characterized by stricter supervision protocols. These programs generally come in two types:

prison diversion and probation/parole enhancement (Petersilia & Turner, 1990b). Across both contexts, these programs often consist of increased office visits, more frequent drug testing, curfews, and a zero-tolerance policy toward minor infractions (Gill & Hyatt, in press). There is, however, significant between-program variation in supervision characteristics; there is no single, widely accepted definition of ISP.

ISP is, by itself, largely atheoretical and much of the related literature has treated it accordingly. The focus has remained on delivering higher levels of scrutiny and on increasing the frequency of contacts between officers and probationers; the specific mechanism for any assumed benefits is not often explored. As Petersilia and Turner (1993) noted, “[r]outine community supervision offers the weakest crime control. It often does not . . . deter . . . people from committing crimes, and it imposes relatively few punitive conditions” (p. 287). ISP was developed, in part, to address this febleness and increase the impact of probation. It was implied that, because traditional probation produced little deterrence due to its lack of punitive measures, increasing supervision intensity would remedy this. ISP protocols, featuring faster and more severe punishment along with higher levels of scrutiny, therefore adhere closer to the traditional principles of deterrence theory (Pratt, Cullen, Blevins, Daigle, & Madensen, 2006; Sherman, 1993) and discourage offending at the group and individual levels.

This ratcheting-up of supervision and sanctioning intensity for certain groups of offenders is also being used outside community corrections agencies. ISP and other types of intensive, control-focused supervision strategies have also thrived in light of the growth of specialty courts. These courts (e.g., drug, mental health, reentry, veterans) have addressed specific populations of offenders, focusing on the unique needs of each population (Dorf & Fagan, 2003; Marlowe & Kirby, 1999). Although not exclusively focused on control, and often using a meaningful treatment component, the programs run by these courts have placed a large number of offenders under ISP-like supervision. Given this combination of treatment with more intensive supervision, specialty courts provide yet another reason to understand the independent effects of both these elements, as well how they interact to influence subsequent offending. It is therefore crucial to isolate the impact of supervision intensity to demonstrate the separate effects of these more therapeutic elements. This study provides some of the necessary evaluative evidence on the control side of the supervision equation.

### *Evaluating ISP*

Researchers and policymakers have been assessing the potential of ISP programs for some time. Early, quasi-experimental evaluations of ISP were

completed in Georgia and Florida, and indicated that intensive probation had little impact on subsequent offending (Erwin, 1986; Nath, Clement, & Sistrunk, 1976). Regardless, these programs would become the model for many of the ISP programs that proliferated in the 1980s and early 1990s.

Efforts to better understand the effects of ISP increased along with practitioner-led demand for these programs. The largest of these early evaluations used a randomized design to evaluate 12 jurisdictions in which ISP was compared with some form of standard probation. Considered together, this multi-site evaluation found that ISP did not reduce multiple measures of recidivism, but rather increased rates of technical violations, and resulted in increased levels of incarceration. For example, across four of the larger sites (Houston, Los Angeles, Santa Fe, and Seattle), arrest rates were higher among the ISP group. However, the treatment group in three other sites (DeMoines, Macon and Ventura County) had lower reported arrest rates. Across the study, none of the comparisons reached statistical significance ( $p < .05$ ). At the conclusion of the 1-year follow-up period, and aggregated across all of the evaluations, about 37% of ISP and 33% of comparison offenders had been arrested (Petersilia & Turner, 1993). This evaluation, although now more than 20 years old, provided clear evidence of the ISP's impact on offending, at least in the specific context implemented at the sites of the RAND evaluation.

A recent meta-analysis considered the experimental and quasi-experimental evaluations of ISP conducted to date, examining both the impact of ISP on offending and the role of key moderator variables. Gill and Hyatt (in press), after reviewing 239 studies, assessed a total of 47 individual treatment-comparison contrasts—38 randomized trials and 9 quasi-experiments. Among the randomized controlled trials (RCTs), assignment to intensive supervision made no difference in the prevalence of rearrest (odds ratio = .93;  $p = .72$ ). Similarly, non-significant results were observed for quasi-experiments and for each of the policy-relevant program features considered in the meta-analysis. These results support the general conclusion regarding the effectiveness of ISP, but the relatively few true experiments, as well as a large degree of intra-program variation, make broad generalizations difficult.

Reductions in caseload size, a hallmark of many ISP programs, have also been shown to have little impact on offending. Latessa, Lawrence, Fulton, and Stichman (1998), using a randomized design, found no effect on arrest rates when caseloads were reduced as a component of ISP. Although officers who supervised fewer probationers had more time for administrative responsibilities (Taxman, 2002), and were therefore better able to cope with the increased contacts and stricter rules of ISP, this did not translate to reduced

offending. These general findings have been observed in numerous other studies (Farrington & Welsh, 2005; Gendreau, Goggin, Cullin, & Andrews, 2000; Taxman, 2002). Sherman and colleagues therefore classified control-only intensive probation as an approach to supervision that “doesn’t work” in preventing crime (Sherman et al., 1997). Aggregating the results of ISP evaluations across sites, outcomes, and populations does not challenge these verdicts. For example, a meta-analysis of more than 20,000 offenders enrolled in almost 50 studies of various types found that ISP, in the best cases, had no effect on recidivism, or, in the worst, increased offending by up to 6% (Gendreau et al., 2000).

Despite the weight of these discouraging findings, ISP in various forms continues to be used by many community supervision agencies, and would appear to be most common and easily implemented response when one group of offenders poses a higher risk of offending than normal. Given this prominence, evaluations of this particular approach to supervision should be updated to reflect the realities of community supervision today, including a small number of studies that challenge the earlier results. A matched sample comparison evaluation, conducted on a New Jersey ISP program, found that an intensive parole supervision program reduced new convictions by 28% and revocations by 21% within 12 months. At the same time, technical violations increased by 7%. However, this ISP program also fostered a more collaborative supervision relationship and, as the authors note, was “likely very different from surveillance-oriented ISPs” (Gendreau & Paparozzi, 2005, p. 462). When these results were analyzed according to organizational supportiveness and officer orientation, offenders under a non-supportive organization or from a law-enforcement focused officer performed significantly worse across all outcome measures. These components are, of course, the hallmarks of most standard ISP programs.

More recently, Jalbert, Rhodes, and Flygare (2010) conducted a multi-site study, examining the impact of decreased caseload sizes (and the associated increase in levels of supervision) as a component EBP using a regression discontinuity design (RDD). In this case, the Iowa Risk Assessment was used to both identify high risk offenders and create the RDD disjunction. The researchers found that, under those constraints and after 6 months, ISP reduced the likelihood of criminal recidivism by 25.5% ( $p = .037$ ) for all offenses. They replicate these findings within a quasi-experimental design focusing on caseload size (Jalbert & Rhodes, 2012). Although not causal evidence, these studies warrant a return to methodologically rigorous evaluations of the relationship between risk, improved forecasting techniques, ISP, and recidivism.

## Method

### *Setting*

The study was conducted in Philadelphia, Pennsylvania in conjunction with the Philadelphia Adult Probation and Parole Department (APPD). APPD's (2012b) mission is "to protect the community by intervening in the lives of offenders," with a specific focus on the prevention of violent crime by individuals under supervision. To focus supervision resources on those offenders likely to engage in serious offending, APPD has implemented a risk-stratified supervision structure that diverts the Department's resources away from individuals who may impose little or low-level risk to focus instead on those who are considered to present a high likelihood of violent recidivism (APPD, 2012; Hyatt, 2013; Barnes & Hyatt, 2012). The majority of the agency's offenders are managed within one of the risk-based units (high, moderate, and low). Offenders who are under a judicially mandated order to receive specialized types of supervision (e.g., house arrest, sex offenders, or domestic violence) are supervised within a fourth, mutually exclusive division.

Serious offenders comprise a large portion of the offenders under community supervision in Philadelphia. In February 2012, the department was responsible for the supervision and management of 43,676 offenders. This population included 3,819 offenders considered to be at a high risk of committing a serious or violent crime, as forecasted by the random forest model discussed below (APPD, 2012a, 2012b). The high risk population has increased in size over time; ISP covered 6,965 offenders by July 2014 and comprised 15.5% of all individuals under supervision at that time (APPD, 2014).

This risk-based supervision program has been evaluated, in stages, over the past several years. The first randomized trial, assessing the impact of highly reduced supervision on lower risk offenders, found no impact on recidivism when caseload sizes were increased to more than 400 offenders per officer and in-person contacts reduced to twice yearly. No significant differences in arrest rates were found after 12 months; 16% of the control group and 15% of the treatment group were charged with a new offense ( $p = .593$ ; Barnes et al., 2010). A follow-up evaluation found that this lack of difference persisted for up to 18 months ( $p = .874$ ; Barnes, Hyatt, Ahlman, & Kent, 2012).

### *Risk Forecasting, Eligibility, and Randomization*

The risk forecasting strategy used to identify offenders for enrollment into the research sample was a statistical procedure known as random forest forecasting. This method, a machine learning-based approach for prediction,

allows for the adaptation of the forecasting procedure to both the data that are available and the pragmatic needs of the agency. As Berk (2008) noted, the random forest approach controls for over-fitting, allows for the identification of non-linear relationships, and provides for the imposition of asymmetric costs for false positives versus false negatives. A complete discussion of the specifications and accuracy of this approach are beyond the scope of this article (see, Berk, 2012; Berk, Li, & Hickman, 2005; Breiman, 2001), although an analysis of this particular model can be found in Barnes and Hyatt (2012).

Outcomes for the model were each offender's likely conduct over the first 2 years of their term of supervision. The classification categories were mutually exclusive. All of the offenders involved in this research were classified as high risk, and were therefore predicted to commit a serious offense with this 2-year time frame. Serious, for these purposes, was defined as a murder, attempted murder, aggravated assault, robbery, or a sexual offense (e.g., rape, indecent assault). Moderate risk offenders were those forecasted to commit only offenses not classified as serious, including property and drug crimes. Finally, low risk offenders were those who were not predicted to commit any new offenses within the 2-year window.

A long-term follow-up of this model's forecasts (Barnes & Hyatt, 2012) showed that those placed into the high and moderate risk groups reoffended at very similar rates when crimes of any type were considered (54.8% of highs and 52.1% of moderates committed a new offense within 2 years). Not surprisingly, however, predicted high risks were much more likely (21.0%) than predicted moderates (11.0%) to have engaged in new serious offenses over this same period of time. The model identified a distinct population of high risk offenders for the present research. Although similar to moderate risk offenders in their overall likelihood of reoffending, the high risk probationers who participated in this experiment were noticeably more likely to commit new serious crimes.

The random forest model deployed by APPD during the course of this study used 48 different predictors to make these forecasts of future criminal behavior. These predictor variables, derived from administrative data sources, included measures of offender demographics, prior criminal history, the nature of the current offense, stays in the local jail system, prior sentences to both probation and incarceration, and neighborhood characteristics derived from census data (Barnes & Hyatt, 2012).

Risk forecasting is hardly an innovation in community corrections. However, the use of random forest forecasting models is uncommon. These models can be difficult to construct and rely on a vast amount of historical data (Berk, 2012). The development of more accurate forecasting methods,

however, is a key to effective program evaluation, especially for those interventions—such as ISP—that are intended only for serious offenders. In essence, because the risk of recidivism has been shown to correlate with the magnitude of effect sizes (see Lipsey, Landenberger, & Wilson, 2007), these “next generation” models allow for the identification of appropriately dangerous samples and for the development of evaluations that better reflect the true impact on the recidivism of serious offenders. They also may identify populations of offenders that are different, in unpredictable ways, from those on whom ISP has been tested in the past.

Both eligibility screening and random assignment for this RCT were conducted in a manner invisible to the end user and as part of APPD’s intake process. Because this process required that every new period of supervision be accompanied by a risk assessment, nearly every offender who began supervision during the enrollment period was screened for eligibility in the experiment. This screening took place automatically, using only machine-readable data, and required no direct effort by agency staff. Importantly, this allowed eligibility screening to be conducted uniformly throughout the entire experiment.

The most fundamental eligibility criterion for RCT eligibility was the offender’s forecasted risk category. Only offenders who were predicted as “high risk” were randomly assigned. In addition, those offenders placed into the experiment needed to be newly assigned to APPD’s forecasted high risk category. Offenders who were already on high risk supervision, or who had a previous high risk forecast within the previous year, were excluded from enrollment in the RCT. This ensured that our participants would have little prior experience with APPD’s normal procedures for high risk offenders, and that those assigned into ISP would be experiencing their randomly assigned treatment (in most cases) for the very first time. Along with a number of other eligibility criteria,<sup>2</sup> these rules were applied to the 27,196 forecasts run for the 19,998 offenders who began new cases<sup>3</sup> between May 1, 2010, and April 30, 2011. The resulting sample was composed of 832 male offenders<sup>4</sup> who were placed either into the ISP treatment group ( $n = 447$ ) or into a control condition ( $n = 385$ ).

All screenings were conducted, as noted above, by an automated computer program integrated into the agency’s case management system. This allowed all potential participants, regardless of officer preconceptions or demographics, to be screened consistently and universally. This system also recorded the reasons that an individual was considered ineligible. During the enrollment period, 4,203 high risk offenders (comprising 76.5% of high risk results<sup>5</sup>) did not meet the enrollment criteria and were excluded. As described in Table 1, each person could be ineligible for multiple reasons. The most common

**Table 1.** Reasons for RCT Ineligibility.

	Count	%
Already enrolled in RCT	458	8.4
CBT graduate	9	0.0
To be sent to specialized unit	1,203	22.0
Already in specialized unit	652	11.9
Eligible for YVRP	1,198	21.9
Already assigned to Anti-Violence unit	1,274	23.3
Less than 9 months remaining	908	16.6
Female	536	9.8
Non-Philadelphia resident	103	1.9
Previous high within last year	1,691	30.9

Note. CBT = cognitive-behavioral therapy; YVRP = Youth Violence Reduction Partnership.

reasons for ineligibility occurred when an individual was already supervised in one of the high risk units (23.3%), meaning that he or she had been exposed to the treatment condition prior to the RCT, or when an individual was judicially ordered into a specialized supervision unit (22.0%), meaning that he or she could not receive either of the randomly assigned treatments.

These criteria, representing the compromises and practical decisions necessary to implement an RCT of this scope, resulted in an evaluation sample that was likely to be slightly older (some younger offenders were diverted to a juvenile program), to have less experience with ISP (offenders already in ISP were ineligible), was all male (females were excluded by design), and which had longer sentences (because 9 months of supervision were required). As with most experiments, these qualifications should be considered when seeking to generalize these findings (Campbell, 1957; Weisburd, 2003).

### *Groups, Treatment Design, and Statistical Power*

The experimental ISP treatment, as defined by the department's written protocols, mandated an increased level of supervision and control across a number of dimensions. High risk offenders were required to report to APPD's centralized office location for a face-to-face meeting with their officer on a weekly basis. The protocol also mandated drug testing at least twice per month. The ratio of offenders to officers in the high risk units was intended to be 50:1, with the smaller caseload sizes allowing for monthly home visits and frequent follow-up contacts. Offenders under this protocol operated under a "zero-tolerance" policy for any rule violations, and all technical violations

were intended to be prosecuted fully. The ISP offenders were supervised in three geographically organized “Anti-Violence” units, each of which operated under identical protocols.<sup>6</sup>

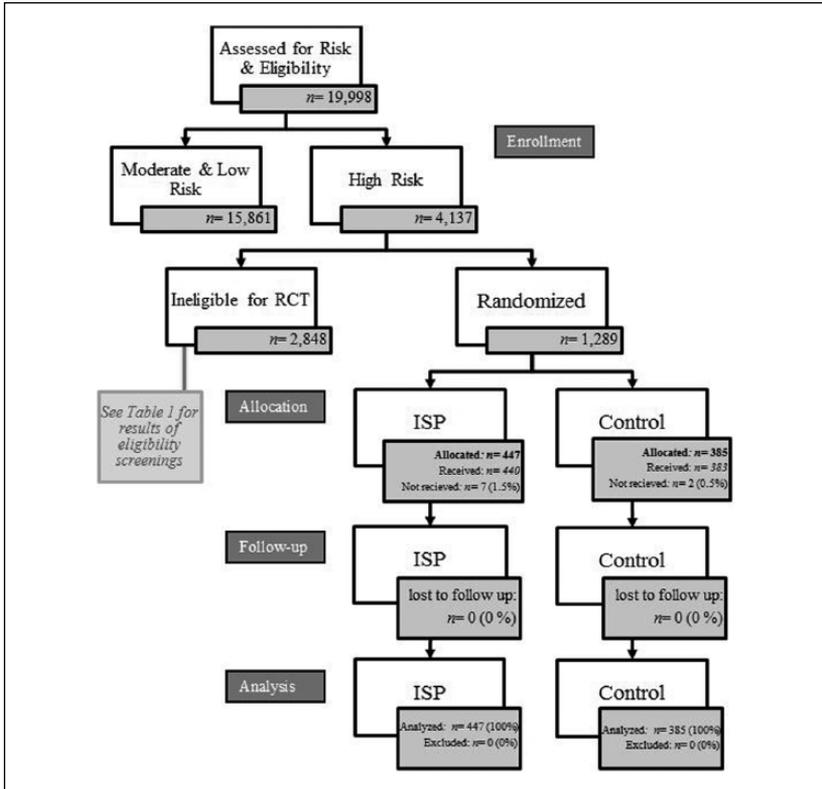
Within the control group, high risk offenders were assigned to the level of supervision traditionally reserved for offenders assessed as moderate risk. Delivering this treatment required these offenders to have the visible results of their risk forecasts changed during the assessment process. Although all of them were, in fact, forecasted as “high risk,” the result that appeared in the agency’s data systems instead labeled them as “moderate risk.” These altered forecast results allowed the offenders to be supervised within multiple “General Supervision” units, while also removing any labeling effects of being formally designated as high risk. Under this protocol, the offenders reported only once a month, and urinalysis screenings were administered only by judicial order or with cause. In addition, no out-of-office contacts, such as home visits, were permitted. Each officer in the moderate units was expected to manage approximately 150 probationers. This supervision plan closely mirrored the “one-size-fits-all” approach to supervision that was used in Philadelphia prior to the risk-based reorganization in 2009 (Barnes et al., 2010).

Because prior examinations of ISP have found little or no effect on recidivism, it was essential to design the evaluation with a strong likelihood of detecting even modest between-group differences. The random assignment of 832 offenders into the ISP ( $n = 447$ ) and control ( $n = 385$ ) groups produces a notable amount of statistical power when making these comparisons. With a “small” effect size of just  $d = 0.20$  (Cohen, 1988), a randomly assigned sample of this size will present no less than an 82% chance of detecting a statistically significant difference.

Figure 1 reports the flow of cases into and within the experiment and follows the Consolidated Standards of Reporting Trials (CONSORT; 2010) format. This format, in addition to encouraging transparency in the reporting of experiments, increases the descriptive validity of trials, a challenge in research of this nature (Mayo-Wilson et al., 2013; Perry, Weisburd, & Hewitt, 2010).

### *Sample, Participants, and Equivalence at Random Assignment*

Table 2 shows the results of a series of independent-sample  $t$  tests that compare the two treatment groups across a range of variables, measured at the moment that the offenders were randomly assigned.<sup>7</sup> It describes the types of offenders who were enrolled into the RCT, while also demonstrating that the randomization procedures successfully produced two statistically equivalent



**Figure 1.** CONSORT diagram.

Note. CONSORT = Consolidated Standards of Reporting Trials; ISP = Intensive Supervision Probation.

treatment groups. This holds true across several important variables types, including race, neighborhood-level socioeconomic status (SES), age, and criminal history. As Table 2 shows, the participants in the experiment were mostly African American, had extensive prior histories of criminal conduct (including violent offenses), and had spent previous time both on probation and in the local Philadelphia prison system. Although not available within the data, measures of gang involvement and substance abuse should also have been equivalent across the two conditions, in keeping with the strong assumptions that underlie randomized trials (see Sherman, 2003).

The identification of a treatment group and a control group which were uniformly assessed as high risk, prior to randomization, has been challenging

**Table 2.** Equivalence at Random Assignment.

Category	Variable	ISP	Control	p
		M (SD)	M (SD)	
Race	% Black	71.8 (0.45)	71.4 (0.45)	.903
	% White	21.0 (0.41)	21.6 (0.41)	.853
	% Other	0.72 (0.26)	0.7 (0.26)	.935
Age	Age at assignment	29.41 (9.482)	29.14 (9.32)	.686
SES	M household income	11,078 (16,084)	9,930 (15,480)	.298
	M home value	22,472 (21,077)	22,177 (20,724)	.804
Juvenile history	Any charge count	9.36 (11.53)	8.61 (11.99)	.362
	Serious charge count	0.94 (1.759)	0.96 (1.84)	.866
	Violent charge count	2.90 (4.80)	2.88 (4.93)	.950
	Sexual charge count	0.12 (0.676)	0.12 (0.90)	.916
	Property charge count	2.76 (4.76)	2.65 (4.66)	.739
	Drug charge count	1.23 (2.30)	0.96 (1.86)	.063
	Adult history	Any charge count	58.03 (47.35)	52.71 (40.39)
Adult history	Serious charge count	8.27 (8.17)	7.67 (7.24)	.260
	Violent charge count	19.23 (18.15)	17.81 (15.36)	.228
	Sexual charge count	0.79 (3.21)	0.87 (3.49)	.751
	Property charge count	15.55 (21.53)	13.62 (16.97)	.156
	Drug charge count	5.81 (6.32)	5.92 (6.41)	.804
Instant offense	Serious charge count	0.72 (1.35)	0.87 (1.62)	.144
	Violent charge count	1.44 (2.27)	1.67 (2.60)	.178
	Sexual charge count	0.05 (0.56)	0.06 (0.45)	.964
	Property charge count	0.97 (1.67)	1.00 (1.70)	.804
	Drug charge count	0.77 (1.14)	0.77 (1.12)	.948
Instant sentence	Probation sentence(s)	0.61 (0.97)	0.80 (1.21)	.015
	Incarceration sentence(s)	0.43 (0.80)	0.47 (0.88)	.456
Supervision history	% prior probation	62.6 (0.48)	67.2 (0.46)	.440
	% prior incarceration	94.4% (0.230)	95.5 (0.205)	.163

Note. ISP = Intensive Supervision Probation; SES = socioeconomic status.

in prior studies (e.g., Jalbert et al., 2010; Petersilia & Turner, 1993). The automated risk forecasting, eligibility screening program, and simultaneous data collection ensured that 100% of enrolled offenders were assessed, using the random forest model, as high risk. In addition, this process permitted the actual risk forecasting outcomes of the control group to be concealed from all APPD staff, with all of the control offenders instead being labeled and treated as moderate risk cases. In effect, this method of random assignment allowed

**Table 3.** Fidelity to Treatment Protocol.

Treatment event	ISP	Control	<i>p</i>
	<i>M (SD)</i>	<i>M (SD)</i>	
Scheduled office meetings	21.47 (15.03)	9.09 (5.62)	.000
Successful office meetings	18.67 (13.878)	7.35 (5.62)	.000
Scheduled home visits	8.82 (8.119)	0.12 (0.753)	.000
Successful home visits	5.32 (6.072)	0.08 (0.634)	.000
Scheduled phone contacts	8.45 (9.478)	4.08 (6.482)	.000
Successful phone contacts	5.5 (5.908)	2.69 (4.07)	.000
Drug tests administered	6.61 (5.849)	0.85 (2.063)	.000

Note. ISP = Intensive Supervision Probation.

this experiment to be double-blinded, a rarity in criminology. In both treatment groups, neither the offenders nor their supervising officers were aware that they were participating in a randomized trial, and there was effectively no way that these specific offenders could be isolated from the non-participating members of their officers' caseloads by anyone other than the research team.<sup>8</sup>

### *Supervision Intensity as Delivered*

As seen in Table 3, levels of treatment fidelity to APPD's written protocol were generally high. Probationers assigned to the ISP treatment group exhibited significantly higher levels of supervision and control when the two groups were compared using independent-sample *t* tests. This holds true for the number of face-to-face office meetings held ( $p = .000$ ) and the number of home visits ( $p = .000$ ), as well as for non-mandatory phone contacts ( $p = .000$ ). As was expected, the ISP group also was subjected to significantly more frequent urinalysis screenings ( $p = .000$ ). Exact measurement of treatment dosage is often lacking in ISP evaluations (Latessa et al., 1998). In this instance, these data were available and reliable, and all of the measured aspects of supervision indicate that the ISP group received the more intensive levels of supervision required.

The protocols governing supervision within both the high and moderate units do not provide for any therapeutic elements that directly address criminogenic needs, nor do they require that officers make or seek out referrals to such programs. Offenders assigned to ISP, who had far more frequent contact with their supervising officers, could have been more likely to have these needs discussed during the normal course of supervision. Information on

informal interventions, however, are largely invisible in these data; rates of referral to any form of external treatment programs are simply unavailable.

### **Outcome Measures**

Offense and criminal history data were collected for the 12 months following each participant's enrollment into the randomized trial. Although this means that the follow-up period for each probationer did not occur simultaneously due to the rolling RCT enrollment procedure, it ensures that each participant had equal time, post assignment, to engage in crime.

Recidivism is quantified as any charge for a new offense committed after random assignment. These data are limited to new criminal acts, and do not include technical violations of probation conditions. Charges are used in place of conviction data to better estimate underlying crime rates, as they are not confounded by systemic delays (Neithercutt, 1987). Although both measures are conservative and will undercount actual behavior, as Blumstein notes, "the errors of commission associated with truly false arrests are believed to be far less serious than the errors of omission that would occur if the more stringent standard of conviction were required" (Blumstein & Cohen, 1979, p. 565). Charges were grouped categorically: violent, serious, non-violent drug, property, and sexual offenses. These classifications, derived from a manual review of the full Pennsylvania Crimes code, were the same as those used to classify offenses for the risk forecasting process.

Information on new criminal offenses, including the date and nature of the offense, was extracted directly from the unified, computerized databases used by the police, courts, and correctional agencies in Philadelphia. Additional data, developed to more fully capture the impact of the ISP protocol on supervision compliance, were also collected during the same follow-up time frame. Data on absconding, drug test results, and supervision contacts were obtained directly from APPD's case management system. Imprisonment data were available from the local jail system by using the daily jail census files provided to the research team. In each case, the analysis period for all of these secondary measures was calculated in the same manner as the primary outcomes relating to criminal recidivism, covering 1 full year after each participant's enrollment into the RCT.

### **Results**

Prior research, discussed in some detail above, has found that ISP has a limited impact on the offending of participating offenders (Gottfredson & Gottfredson, 1985; MacKenzie, 2000; Petersilia & Turner, 1990a, 1993).

Some more recent work has challenged this contention (Jalbert et al., 2010). In this analysis, our goal is to assess the impact that ISP has on relatively short-term offending patterns and on compliance with probationary conditions.

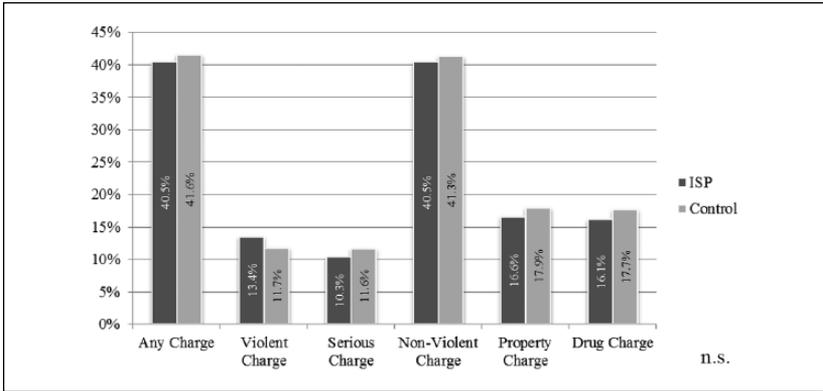
This analysis uses an Intention to Treat (ITT) design. This method requires the inclusion of all subjects, including those who fail to receive any treatment, who drop out of the trial, or who receive an intervention other than designated through the random assignment process (Hollis & Campbell, 1999). In this instance, we include probationers who, during the trial but after random assignment, were transferred out of their assignment unit. Deviations from the randomly assigned treatment were rare in both of the treatment groups. In total, the 832 offenders who participated in this research spent 272,222 days on active supervision<sup>9</sup> during their first year after random assignment. Their officer and unit assignments during this time show that they were supervised in accordance with their random assignment on 93.6% (254,704) of these supervision days.

With treatment integrity at these levels, an ITT approach is the optimal way to examine these findings. Although this method is relatively conservative, and may understate the magnitude of the observed effects (Aos, Miller & Drake, 2007; Gupta, 2011; Hollis & Campbell, 1999), it remains the best measure of the impact that implementing ISP for high risk probationers would have in a “real world” policy setting. This approach may be less than ideal for the identification of individual-level effects, but it is appropriate for specifying the types of “pragmatic estimate[s] of a change in treatment” that are essential in program evaluation and, in this case, for our agency partners (Hollis & Campbell, 1999, p. 673).

## Offending

The implementation of an ISP supervision strategy for high risk offenders had no significant effects on offending after 1 year. As indicated in Figure 2, roughly equal percentages of both the ISP treatment group (40.5%) and the comparison group (41.6%) were charged with any new offense ( $p = .756$ ). When these offenses are broken down by type, comparisons of violent offenses ( $p = .520$ ), serious offenses ( $p = .814$ ), non-violent offenses ( $p = .234$ ), property offenses ( $p = .603$ ), and drug offenses ( $p = .551$ ) all fail to reach customary levels of statistical significance.<sup>10</sup>

This pattern of non-significance observed in prevalence is mirrored in the measures of frequency, here quantified as the average number of charges lodged against members of each treatment group within 1 year of random assignment. Because the two groups spent roughly equivalent amounts of



**Figure 2.** Prevalence of offending within 12 months.

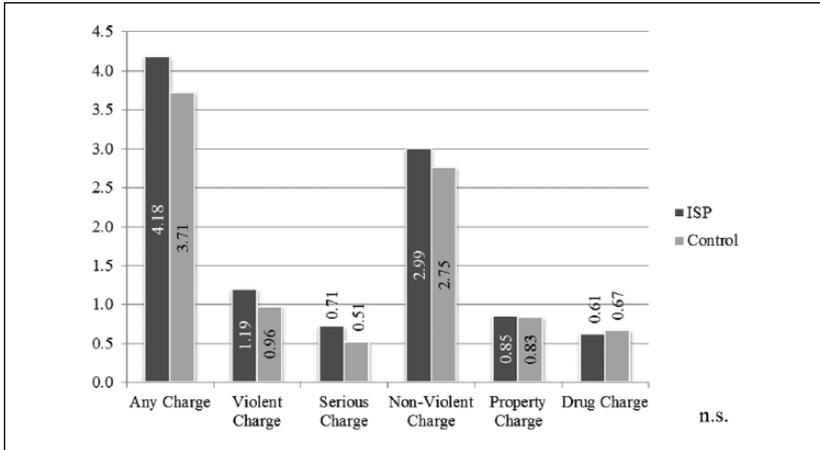
Note. ISP = Intensive Supervision Probation.

time in local jails during this period (see below), these frequency values stem from the raw offense counts over the entire year, and are not adjusted for the amount of time spent on the street.

The average numbers of offenses committed by the high risk offenders assigned to each of the two treatment groups were statistically indistinguishable. The mean differences are slight and average less than a single offense over the 1-year follow-up period. As shown in Figure 3, differences in the frequency of overall ( $M$  difference = .46,  $p = .535$ ), violent ( $M$  difference = .22,  $p = .484$ ), serious ( $M$  difference = .20,  $p = .334$ ), non-violent ( $M$  difference = .24,  $p = .611$ ), and drug offending ( $M$  difference = .02,  $p = .648$ ) suggest little practical difference between the two treatment groups.

### Time to Failure

A survival analysis was conducted to determine whether, even in light of a lack of overall differentiation in offending, ISP had an impact on how long probationers remained on supervision before offending. Because ISP is a control and surveillance focused approach to supervision, criminal misbehavior could have been detected earlier. We use a Kaplan–Meier survival analysis to study recidivism as a function of elapsed time. Each individual participant has exactly 1 full year of post-random assignment time included in this analysis. As noted in Table 4, below, differences in incarceration were not significantly different between the two groups, resulting in equivalent amounts of measurable “opportunity time” to offend. Results from



**Figure 3.** Mean number of charges, by offense, within 12 months.  
 Note. ISP = Intensive Supervision Probation.

Kaplan–Meier survival analyses for the time to first offense of any kind are presented in Figure 4. Analyses for the other offense categories were similarly not significant.

Among those probationers who committed a new offense, the two groups exhibited no significant differences in time between random assignment and their first charge of any type ( $p = .772$ ). This lack of a differential effect is also reflected in time elapsed until an offender's first charge for a serious ( $p = .551$ ), violent ( $p = .250$ ), property ( $p = .637$ ), non-violent ( $p = .814$ ), and drug ( $p = .492$ ) offense. Pragmatically and statistically, the impact of ISP on time until offending was minimal.

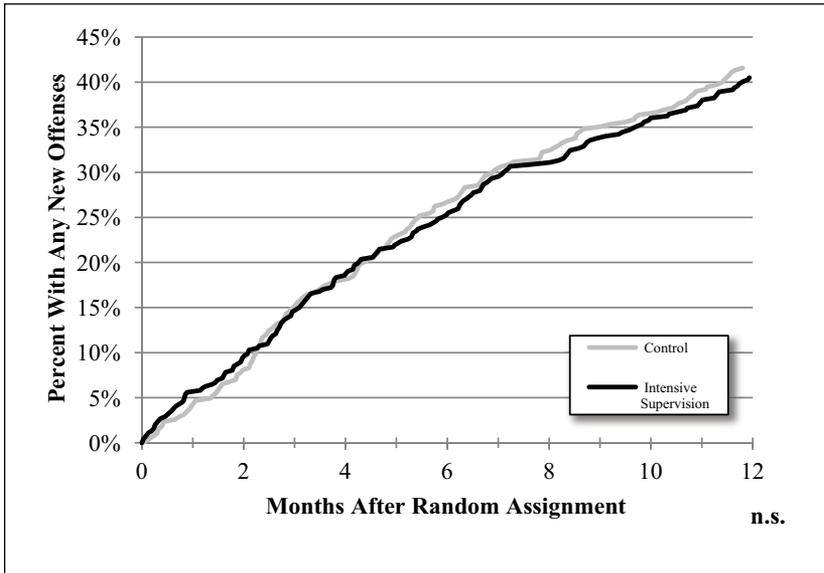
### *Absconding*

ISP had a clear impact on multiple measures of absconding. In Philadelphia, an offender was deemed to have potentially absconded after missing, without excuse or justification, two consecutive scheduled contacts. At that time, a contact notice was mailed to the address of record and additional attempts were made to locate the offender. If the offender did not get in touch with his or her officer, he or she was officially classified as having absconded, his or her case was transferred to a separate unit, and a warrant was requested. For the purposes of this evaluation, the absconding event was deemed to have taken place at the moment that the case was transferred into the absconding caseload.

**Table 4.** Absconding and Incarceration Within 12 Months.

	ISP	Control	p
<b>Incarceration</b>			
Percent incarcerated	67.6%	55.3%	.000
Number of incarceration incident	0.97	0.79	.003
Number of jail days	87.19	77.45	.181
<b>Absconding</b>			
Percent absconded	27.3%	16.10%	.000
M abscondings	0.365	0.18	.000

Note. ISP = Intensive Supervision Probation.



**Figure 4.** Time to failure, all offenses.

Differences in absconding persisted across measures of frequency and prevalence. Within 1 year of their assignment date, 11.2% more of the ISP group had absconded at least once (27.3%;  $p = .000$ ). Within that same time period, 16.1% of the comparison group absconded. On average, offenders in the experimental group engaged in 0.185 more individual absconding events ( $p = .000$ ) than those in the control group. These absconding rates are, across

both conditions, artificially low, as it is unlikely that an offender would pick up more than one absconding event during the observation period and, in both groups, the majority of probationers did not abscond.

### *Incarceration*

Offenders assigned to the ISP treatment condition were also incarcerated in the local jail system at significantly higher rates. Table 4 shows that, during the 1-year follow-up period, 12.3% more of the total ISP group was in custody, at least once, in the Philadelphia County Prison System ( $p = .000$ ). This count includes incarceration for any reason, including pre-trial detention, short sentences for technical violations, and any new judicial sentences of up to 24 months.<sup>11</sup> ISP offenders, on average, entered jail .20 more times in their first 12 months after random assignment ( $p = .003$ ). Despite being more likely to experience incarceration, however, it is important to note that ISP participants did not spend a significant number of additional days in the county correctional facility ( $M$  difference = 9.7 days,  $p = .181$ ) within the 12 month analysis period. Based on the available data, the two groups spent a statistically equivalent amount of time on the street during the first year after random assignment.

### *Violations of Probation*

An offender is charged with a technical violation when he or she does not adhere to the requirements of his or her supervision. These failures can take a number of forms, including failing to report, a positive drug screening, missing treatment or other court-mandated conditions, or not paying fines and court costs, among others. An arrest for a new offense is both a technical and a direct violation of probation, and gives rise to a new criminal matter that must be handled separately.

The violations process is handled in two stages. The first hearing after an individual is taken into custody for a violation is held via videoconference from the county jail (Gagnon 1). If the judge determines that there is sufficient evidence to proceed, a second, in-person hearing, referred to as a Gagnon 2, is subsequently held. At this hearing, which adopts many of the hallmarks of a standard court proceeding, the judge hears evidence about the alleged violation, renders a decision, and determines the appropriate sentence.

During the 12-month follow-up period, 43% of the ISP group was charged with a violation of probation and subject to a Gagnon 1 hearing. Only 27% of the comparison group had a violation hearing in that same time ( $p = .000$ ).

**Table 5.** Recorded Types for Violation Hearings and Prevalence of Justifications for Violation Hearings.

	ISP	Control	<i>p</i>
Type of violation			
Direct violation(s)	5.30%	4.40%	.618
Technical violation	29.30%	12.40%	.000
No reason given	1.10%	0.05%	.333
Justification for violation (prevalence)			
Drug test results	12.98%	1.30%	.000
Employment	0.22%	0.26%	.916
Failure to report	7.61%	4.94%	.111
Unpaid fines	3.58%	1.82%	.114
Misc. rules	4.25%	1.30%	.008
New arrest	14.32%	7.79%	.002
Treatment	1.82%	4.25%	.039

Note. ISP = Intensive Supervision Probation.

Not all of these violations were proven, as evidenced by the percentage that did not progress to a Gagnon 2 hearing. At this first stage, violations were dismissed at the same rate between groups. Seventy-seven (17.2%) ISP participants and 53 (13.7%) control offenders ( $p = .168$ ) were taken into custody and completed a Gagnon 1 hearing, but did not have a subsequent Gagnon 2 hearing within 1 year of random assignment.

Despite having equivalent rates of dismissal, violations were more prevalent in the ISP group. As a result, ISP offenders were significantly more likely to proceed on to a Gagnon 2 hearing, at 26% as compared with 13% ( $p = .000$ ) in control. The hearings in these two groups, however, took place for very different reasons. As shown in Table 5, this differentiation was driven by an increase in technical, not direct (i.e., a new arrest), violations. In addition to prevalence, ISP participants had, on average, twice as many hearings for technical violations (.29 compared with .12 hearings,  $p = .000$ ). There was no difference in direct violation counts ( $p = .618$ ). This is additional evidence of the implementation of a “zero-tolerance” policy within ISP, as well as the lack of an effect on underlying offending rates.

The full supervision file and court records for each violation hearing were manually reviewed, and the reason (or reasons) for every hearing was coded into mutually exclusive categories.<sup>12</sup> Counts were aggregated at the offender level. Table 5 also reports the percentage of probationers in each group that had at least one Gagnon 2 hearing in which each justification was recorded in the supporting documentation.

Unsurprisingly, given the accelerated rate and frequency of drug testing in the ISP group, positive drug tests were listed as a justification for significantly more ISP offenders ( $p = .000$ ). On a per-offender basis, new arrests were used as a reason to support a violation of probation for those assigned to ISP significantly more often ( $p = .002$ ). It is worth noting that the underlying rates of offending were the same between the groups, and that a new arrest could be presented as either a direct violation or a component of a technical violation with multiple causes. These measures are derived from the supervising officer's written justification for the violation and, therefore, reflect APPD's high risk protocol, not actual differences in criminal conduct. This evidence that officers were more likely to pursue violations under ISP provides additional empirical support for the successful implementation of the stricter ISP protocol during the evaluation.

Although the rate of missed appointments was higher for ISP offenders, there were no significant differences found in the percent of offenders who were violated for failing to report ( $p = .111$ ). This was surprising, as the ISP protocol required weekly (as opposed to monthly in the control group) reporting, and the supervision rules required less tolerance of missed reporting for offenders assigned to ISP. It may have been the case that, under the stricter guidelines, an ISP-supervised offender would have been classified as an absconder after missing several meetings and removed from active supervision, while someone in the less intensive comparison group would have been retained on supervision and eventually been referred for a violation hearing. Violation hearings for the absconders in ISP, which we know there were significantly more of, could not be held until the offender was located or rearrested, a series of events that were unlikely to occur within the 1-year follow-up period.<sup>13</sup>

ISP supervision also resulted in different outcomes for Gagnon 2 hearings. In this instance, a revocation of probation would result in the imposition of a new sentence, a continuation would leave the active probation sentences unchanged, and a termination of probation would end a probationer's supervision by APPD, at least for that specific case. All other sentences, including concurrent, active probation sentences would remain unmodified.

Overall revocation rates were significantly different. 14.9% of the ISP group and 8.3% of the comparison group had their probation revoked at some time during the evaluation ( $p = .002$ ). Although difficult to explain with the currently available data, offenders being supervised under ISP were also more likely to have their probation continued, leaving their sentence undisturbed ( $p = .002$ ), and terminated, thus ending their supervision on that case ( $p = .008$ ).

Lastly, we compared the prevalence of various sentencing outcomes for violation hearings between the two groups. A new sentence was required

**Table 6.** Prevalence of Outcomes for Violation Hearings and Sentences Resulting from Violation Hearings.

	ISP	Control	p
<b>Outcomes of violations hearings (prevalence)</b>			
Revoked	14.90%	8.30%	.002
Terminated	3.40%	0.80%	.008
Continued	10.30%	4.70%	.002
No sentence within 12 months	0.00%	0.05%	.158
<b>Sentences resulting from violation hearings</b>			
Probation	18.57%	10.13%	.001
Incarceration	9.40%	4.94%	.012
House arrest	0.00%	0.26%	.318
Parole	1.12%	0.00%	.025
No sentence recorded	1.30%	3.13%	.069

Note. ISP = Intensive Supervision Probation.

when the judge, at the Gagnon 2 hearing, revoked an individual’s active probation term. Given the series of conditional probabilities required to reach the sentencing phase of a violations hearing within 1 year, the proportion of each group that received each type of sentence is relatively small. However, significant and meaningful differences persist. As Table 6 also shows, differences in the prevalence of new sentences to further probation ( $p = .001$ ), incarceration ( $p = .012$ ), and parole ( $p = .025$ ) were significant, and more likely to occur in the ISP group.

## Discussion

Overall, after 12 months of supervision under an ISP supervision protocol, high risk offenders were not charged with significantly more (or less) offenses than those in the control group. This equivalence holds across multiple types of offending, including violent, non-violent, property, and drug offending, as well as for a survival analysis conducted for each offense type. Probationers receiving ISP supervision, however, absconded more frequently and were more likely to be incarcerated at least once during the 12-month follow-up period.

The observed increase in absconding in response to ISP supervision is, in many ways, unsurprising. ISP, at its inception, was designed to be a community-based supervision program as restrictive and invasive as full custody incarceration (Petersilia, 1990). For example, approximately 15% of all offenders who signed up for a voluntary New Jersey ISP program, designed

to encourage prisoners to take an early release option, withdrew their application when the program's requirements became clear (Pearson, 1988). High absconding rates may simply signal that, even with the increased consequences, the regularity of reporting and the intensity of control are too much for some high risk offenders to bear.

The significant differences in absconding and incarceration are problematic for agencies wishing to use an ISP strategy to increase their control over, or deliver therapeutic interventions to, serious offenders. Not only is it difficult to manage offenders who fail to maintain communication and report to their appointments, but once an offender has absconded, he or she cannot receive any treatment or access any reentry programming. In the long run, an increase in absconding probationers requires the expenditure of a significant amount of resources to locate these offenders, to incarcerate them on arrest, to hold violation hearings, and—in the vast majority of cases—to reintroduce them to a new (often longer) term of probationary supervision. From a cost-based perspective, this has the potential to offset any benefits of a risk-targeted ISP approach and, depending on the magnitude, overwhelm judicial and correctional systems already operating close to capacity.

It is clear that ISP failed to reduce offending and has negative implications for many other indicators of a successful supervision program. Although the reduction of crime is a key goal for probation agencies, it is not the only one. These results suggest that ISP is, in fact, a more severe, more invasive, and more restrictive protocol than traditional probation. The intensity and invasiveness of probation, which many offenders see as less appealing than prison (Crouch, 1993; Petersilia, 1990), allow for the scaling of punishment severity and the integration of probation into intermediate sanctioning and prison release systems. In addition, under these goals, ISP is not designed to reduce offending, but rather to serve as a mechanism through which infractions can be detected and non-complying offenders removed from the community (Tonry, 1999; Turner & Petersilia, 1992). Agencies may also, to satisfy the demands of policymakers and the general public, perceive a need to increase the intensity of supervision, especially for serious offenders released to the community. In this regard, the ISP program in Philadelphia met its goals.

Many jurisdictions are trying to move away from the control-only supervision strategies being evaluated here. Several recent meta-analyses have shown, across multiple treatment and intervention types, that the integration of a therapeutic component into supervision can have a positive impact on offending rates (Cullen, Wright, & Applegate, 1996; Dowden & Andrews, 2000; Lipsey & Landenberger, 2005; Smith et al., 2009). If

nothing else, ISP provides the opportunity and ability to levy sanctions for offenders who fail to comply with the requirements of their therapeutic programming. Increased contact requirements, a common characteristic of ISP, may also be necessary in creating the frequency and duration of interaction that has been shown to increase treatment effectiveness (Lipsey et al., 2007). For example, in Philadelphia, the delivery of cognitive-behavioral therapy to probationers requires, at a minimum, weekly contacts with offenders, an opportunity only available to high risk offenders who are supervised under ISP (Hyatt, 2013).

### *Limitations*

Despite a rigorous and well-implemented design, this study is subject to a number of limitations derived from the data that were accessible at the time of analysis and the sample studied. Like other studies relying on administrative data, our measures of recidivism are only as accurate as the records themselves and are limited in scope to information regularly collected by the agency. We therefore do not have access to many common measures associated with recidivism, including gang involvement, substance abuse history, and socioeconomic and marital statuses. Outcome data used here are also limited to conduct that took place in Philadelphia County, including information on new arrests and incarceration. Data on technical violations were developed through the manual review of supervision files and are subject to incomplete record keeping. However, an audit of the technical violations did not uncover systematic differences in record keeping between groups. Last, data on incarceration were derived from the prison system's daily census files and, as such, did not include an explanation of why an incarceration incident took place (i.e., revocation, new sentence, pre-trial detention, etc.).

The data used in this research are subject to certain qualifications. Conviction data were not used in this evaluation by design. We maintain, as others have, that arrests and new charges are the best proxy for offending patterns in the community (Neithercutt, 1987). However, using this as the sole outcome measure does impose some limitations on the policy implications of our findings. Many agencies are sensitive to the relationship between supervision strategies and the jail population. In Philadelphia, the probation agency has been singled out as a major contributor to overpopulation (Pew Center on the States, 2010). Therefore, even if ISP were to reduce offending but increase longer term returns to custody, the utility of the approach would need to be reconsidered.

These data also do not take the effects of external treatment services, which may reduce recidivism, into account (Lipsey, Chapman, &

Landenberger, 2001; Peters & Murrin, 2000). Our agency partners did not deliver any services directly, including drug treatment or mental health services, and so collected no data on participation in these programs. We cannot, however, discount that the provision of needs-focused treatment services, perhaps outside the scope and supervision of the agency, could have been distributed unequally among the treatment and control groups. At the same time, the null findings for subsequent offending make any effects from this potential difference in accessing treatment seem unlikely.

Our analysis has also focused only on the effects of ISP as a blanket policy for high risk offenders, and has not examined the possibility that ISP might be more effective for certain types of individuals. Determining these interaction effects would require a series of sub-group analyses within our sample of offenders, an approach that is fraught with potential challenges in both execution and interpretation (Wang, Lagakos, Ware, Hunter, & Drazen, 2007). It is also possible that the current sample, although large in aggregate terms, may become too small as it is repeatedly divided into a series of sub-groups. Nevertheless, differential effects for ISP (or other supervision strategies) are an intriguing possibility that has been little explored in the literature, and presents a promising avenue for further research.

This evaluation, like any randomized trial, also has clear limitations in external validity (Campbell, 1957; Weisburd, 2003), a constraint intensified by the stringent eligibility criteria used in the sample identification process during this experiment. We note, as many have before us (Farrington, 1983), the need for replication of these findings across other contexts. This study, if nothing else, may serve as an opportunity to revisit and reevaluate control-only ISP under more modern constraints, including the implementation of evidence-based policies in community corrections, and using the methods appropriate for causal identification.

### *Implications*

ISP programs remain, despite a wealth of contraindicating research findings, a prevailing model in community corrections. One meta-analysis of community corrections found that only 18% of such programs included a treatment component of any scope or quality (Gendreau et al., 2000). It is clear that, despite prior findings, ISP programs, such as the one in Philadelphia, remain in widespread use. By challenging this approach with modern and experimental evidence, these results open the door for the introduction of a therapeutic, hybrid model. The basic principles of EBP in community corrections require actuarial assessments of both risk and needs, as well as using interventions, including supervision, that target these identified criminogenic

factors (MacKenzie, 2000; Bogue, et al., 2011). The results reported here underscore the importance of integrated needs assessment and the delivery (or facilitation of) treatment that can address those specific criminogenic factors. In keeping with recent, meta-analytic findings, as well as the body of prior research (e.g., Gill & Hyatt, in press), these findings contribute to the conclusion that supervision alone is likely not an effective approach to crime reduction when these other factors are not directly and overtly addressed.

A policy of delivering an ISP protocol can still be evidence-based and remain a key component in managing high risk offenders when it creates the opportunities necessary to deliver treatment. As many other have suggested (Lowencamp, Flores, Holsinger, Makarios, & Latessa, 2010; Taxman, 2002; Thanner & Taxman, 2003), it is this integration of treatment into supervision that returns benefits. If nothing else, these null findings reinforce the claims that a hybrid approach reduces offending through exposure to therapeutic interventions, and not due to the increased intensity of supervision contacts. Associated increases in absconding, incarceration, and technical violations may encourage those wishing to deliver a longer term treatment to consider the implication of increased supervision requirements on potential program attendance.

Although there have been several notable evaluations of ISP in the past (Nath et al., 1976; Petersilia & Turner, 1990a, 1990b), research is an ongoing and iterative process. All too often, criminological research fails to explore issues of construct and external validity through replication (Farrington, 1983). This is especially important in experimental work given the intractable effects of context and implementation on RCT results. For example, a subsequent multi-site replication of the Minneapolis domestic violence experiment was necessary to better specify the relationship between police conduct and spousal abuse and, in some cases, these results were inconsistent (Sherman, 1992; Sherman, Schmidt, Rogan, & Smith, 1992).

This need for replication is especially apparent in light of more recent results that suggest that ISP may, contrary to the older, experimental literature, have a beneficial effect of offending rates (i.e., Gendreau & Paparozzi, 2005; Jalbert & Rhodes, 2012). An important question is why our findings failed to demonstrate a similar effect. These studies did not examine ISP in isolation, but instead used it as one component in broader array of evidence-based practices. They also used different procedures to identify their samples, and did not use the randomized experimental design needed to demonstrate a clear chain of cause and effect. A more detailed comparison of these findings (e.g., Gill & Hyatt, in press) may better shed light on how ISP can be leveraged to produce more beneficial results than it did here.

Finally, these results should be considered in light of the actual costs of running the ISP programs, not just in how ISP can be used to reallocate

resources. Drake, Aos, and Miller (2009), using a cost–benefit analysis, found that ISP is a net negative program, in that the costs far outweigh the value of the program. Each offender enrolled in a surveillance-based ISP program costs the system US\$3,869 and saves almost nothing in crime prevention or community benefits. Risk forecasting and stratification allow for resources to be spent according to the likelihood of offending, but this fails to take into account the opportunity cost of ISP itself. Based on these results, less intense and costly supervision alternatives can, without endangering the public, return the same benefits. The political pressures to implement intensive, ISP-like strategies for high risk offenders, in some jurisdictions, may make this consideration immaterial. Going forward, any evidence-based ISP supervision strategy should take these costs and results into account when determining the components of the protocol.

## **Conclusion**

Research on the impact of ISP has been largely consistent: Intensive probation, focused only on mechanisms of formal social control (e.g., Grattet et al., 2011), has little impact on recidivism. Our results here reinforce this conclusion using a rigorous methodology, an advanced risk forecasting technique, and under updated conditions of supervision. High risk offenders managed under a supervision policy that uses more restrictive protocols commit the same amount of crimes, but the increased supervision, including more frequent contacts, leads to absconding and could result in increased usage of law enforcement and correctional resources. Technical violations, a costly component of supervision, also are significantly more likely to occur. The efforts dedicated to increasing the intensity of supervision may be better allocated elsewhere, including treatment, especially for agencies operating under significant resource constraints.

These results underscore the notion that a policy of increasing the severity of supervision for high risk offenders is a potentially necessary, but certainly not sufficient, condition for the reduction of offending among a probation population. Supervision strategies focusing on the integration of treatment and control-focused characteristics represent an opportunity to utilize community corrections as a mechanism to reduce offending (Taxman, 2002). For many jurisdictions, however, therapeutic interventions are simply not within their budget. In those cases, these results should, at a minimum, serve as a catalyst for a reconsideration of how the intensity of supervision is allocated.

## **Author's Notes**

The authors would like to acknowledge and thank Dr. Richard Berk, for his efforts in building the forecasting model, and Dr. Lawrence Sherman, for his leadership in

developing the partnership that gave rise to this study. The authors also remain indebted to the staff and leadership of APPD and the First Judicial District of Pennsylvania, including Chief Robert Malvestuto, Chief Charles Hoyt and Dr. Ellen Kurtz.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was funded by the National Institute of Justice, 2008-IJ-CX-0024 and the Smith Richardson Foundation.

## Notes

1. Although probation and parole are, from a sentencing and procedural perspective, distinctly different sanctions, we refer to both options as simply probation for the duration of the article. In Philadelphia, probationers and (local) parolees are supervised by the same agency. They were not distinguished within the experiment.
2. The full set of eligibility criteria for this experiment required that (a) the offender started a new term of supervision at Adult Probation and Parole Department (APPD), (b) the new case resulted in a forecast of high risk using the APPD risk screening model, (c) the offender had no previous high risk forecasts within the past year, (d) the offender was not under current supervision by any of the units, which handled forecasted high risk offenders, (e) the offender was not currently under supervision within a specialized unit, (f) the offender was not already enrolled in the randomized control trial (RCT), (g) the offender was male, (h) the offender had a valid local police identification number, (i) the offender was a Philadelphia resident, (j) the offender was expected to remain under APPD supervision for at least the next 9 months, (k) there were no known court orders that required the offender to be supervised by a specific specialized unit at APPD (e.g., drug treatment, domestic violence), (l) the offender was not eligible for a targeted Youth Violence Reduction Partnership (YVRP) program in Philadelphia, and (m) the offender had not previously completed the cognitive-behavioral therapy (CBT) program during the pilot phase.
3. Each forecast applied to a criminal case. Individual offenders could, and often did, have multiple cases that were consolidated into concurrent sentences beginning on the same day. In that situation, the intake department would run a risk forecast for each case, and, once all forecasts were made, use only the highest score.
4. The experiment also included a third treatment group ( $n = 457$ ) that combined Intensive Supervision Probation (ISP) with a classroom-based CBT training program. This CBT group is excluded from this analysis.
5. A single person could have been, and in many cases was, screened for enrollment at the start of several new cases, all of which fell within the enrollment period. These

- probationers appear in analyses of enrollment and screening results multiple times (one for each new case). By design, an individual was rejected from enrolling in the RCT (as they were already an active participant) in all but the first instance.
6. Because the three "Anti-Violence" units operated under the same constraints, the ISP treatment group was, in fact, comprised of individuals from all of the units. When combined, these three units covered the entire city area, and there were no geographic limits on eligibility for the research.
  7. Measures of neighborhood-level socioeconomic status, including income and home values, are derived from the year 2000 census data, the most recent available at the time the research was being conducted.
  8. All of the offenders in this experiment were placed under APPD supervision, with the requirements and exact nature of their supervision to be determined by the agency. Each participant was identified as high risk, making the most intensive levels of supervision appropriate under agency protocols. It is worth noting that no participant had their supervision requirements increased as part of this research. The agency also had no requirement to inform the offenders about how these decisions were made.
  9. Active supervision, as used here, includes the time when the offender had an active sentence to supervision, was assigned to the caseload of a specific probation officer, and had not absconded from supervision. Note that APPD often retains an offender under this active supervision status, even when he or she is incarcerated for brief periods of time.
  10. These frequency analyses were conducted using independent-sample *t* tests on counts of new charges for various types of crime. All of these count distributions were over-dispersed, with variances far higher than the means. Although the large sample sizes in this study suggest that *t* tests should be adequate for these analyses, they were repeated using negative binomial regression to correct for this over-dispersion. The results of these regressions were identical to those produced by the *t* tests, and no significant between-group differences were found.
  11. In Pennsylvania, sentences of less than 2 years can be served within the county in which the offense took place. Offenders with a longer sentence are remanded to the state custody and serve their sentence in state prison.
  12. The "Misc. Rules" category includes all other conduct recorded in the files. These reasons included not providing a valid address, missing a status hearing, failing to receive a GED, refusing to submit to urinalysis, providing false information to APPD, tampering with drug testing process, failing to submit DNA as required, failing to attend status hearings, failing to report to immigration, not maintaining phone service while on house arrest, refusing to testify in a trial, fighting while in custody, impersonating an attorney, "improper suncreening," damaging ankle monitor, failing to register as a sex offender, escaping from custody, and attacking APPD security guards.
  13. Because ISP group participants reported 4 times more often, it was also possible for them to miss more meetings in a given period of time. This would have resulted in their being formally declared as an absconder (generally after failing

to make two consecutive appointments) in 2 weeks, a process that would have taken more than 2 months in the control group. Therefore, a violation hearing for failing to appear would have been redundant for a higher percentage of ISP participants as a bench warrant for absconding would already have been issued.

## References

- Adult Probation and Parole Department. (2012a). *Adult Probation and Parole Department Internal Report*. Philadelphia: First Judicial District of Pennsylvania, Court of Common Pleas, Trial Division.
- Adult Probation and Parole Department. (2012b). *2011 Adult Probation and Parole Department Annual Report*. Philadelphia: First Judicial District of Pennsylvania, Court of Common Pleas, Trial Division.
- Adult Probation and Parole Department. (2014). *Adult Probation and Parole Department Internal Population Report*. Philadelphia: First Judicial District of Pennsylvania, Court of Common Pleas, Trial Division.
- Andrews, D. A., Bonta, J., & Wormith, J. S. (2006). The recent past and near future of risk and/or need assessment. *Crime & Delinquency*, *52*, 7-27.
- Aos, S., Miller, M., & Drake, E., & Washington State Institute for Public Policy. (2007). Evidence-based public policy options to reduce future prison construction, criminal justice costs, and crime rates (Oct. 2006). *Federal Sentencing Reporter*, *19*, 275-290.
- Austin, A. (2010). *Criminal justice trends: Key legislative changes in sentencing policy, 2001-2010*. New York, NY: Vera Institute of Justice.
- Barnes, G. C., Ahlman, L., Gill, C., Sherman, L. W., Kurtz, E., & Malvestuto, R. (2010). Low-intensity community supervision for low-risk offenders: A randomized, controlled trial. *Journal of Experimental Criminology*, *6*, 159-189.
- Barnes, G. C., & Hyatt, J. (2012). *Classifying adult probationers by Forecasting Future Offending*. Washington, DC: National Institute of Justice.
- Barnes, G. C., Hyatt, J. M., Ahlman, L. C., & Kent, D. T. L. (2012). The effects of low-intensity supervision for lower-risk probationers: Updated results from a randomized controlled trial. *Journal of Crime & Justice*, *35*, 200-220.
- Berk, R. A. (2008). Forecasting methods in crime and justice. *Annual Review of Law and Social Science*, *4*, 219-238.
- Berk, R. A. (2012). *Criminal justice forecasts of risk: A machine learning approach*. New York, NY: Springer.
- Berk, R. A., Li, A., & Hickman, L. J. (2005). Statistical difficulties in determining the role of race in capital cases: A re-analysis of data from the state of Maryland. *Journal of Quantitative Criminology*, *21*, 365-390.
- Beto, D., Corbett, R., & Hinzman, G. (1999). Embracing key strategies for a rational probation system. In T. R. P. Council (Ed.), *Transforming probation through leadership: The "broken windows" model* (pp. 3-6). New York, NY: Center for Civic Innovation at the Manhattan Institute.
- Blumstein, A., & Cohen, J. (1979). Estimation of individual crime rates from arrest records. *Journal of Criminal Law and Criminology*, *70*, 561-585.

- Bogue, B., Campbell, N., Carey, M., Clawson, E., Faust, D., Florio, K., & ... Woodward, W. (2004). *Implementing evidence-based practice in community corrections: The principles of effective intervention*. National Institute of Corrections, Washington, DC.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5-32.
- Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin*, 54, 297-312.
- Clear, T. R., Reisig, M. D., & Cole, G. F. (2013). *American corrections*. Belmont, CA: Wadsworth.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Cohen, R. L. (1995). *Probation and parole violators in state prison*. Washington, DC: Bureau of Justice Statistics, U.S. Department of Justice.
- Consolidated Standards of Reporting Trials. (2010). *CONSORT: Flow diagram*. Retrieved from <http://www.consort-statement.org/download/Media/Default/Downloads/CONSORT%202010%20Flow%20Diagram.doc>
- Crouch, B. M. (1993). Is incarceration really worse? Analysis of offenders' preferences for prison over probation. *Justice Quarterly*, 10, 67-88.
- Cullen, F. T., Wright, J. P., & Applegate, B. K. (1996). Control in the community: The limits of reform. In A. T. Harland (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply* (pp. 69-116). Thousand Oaks, CA: Sage.
- Dorf, M. C., & Fagan, J. A. (2003). Problem-solving courts: From innovation to institutionalization. *American Criminal Law Review*, 40, 1501-1511.
- Dowden, C., & Andrews, D. A. (2000). Effective correctional treatment and violent reoffending: A meta-analysis. *Canadian Journal of Criminology*, 42, 449-467.
- 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 & Offenders*, 4, 170-196.
- Erwin, B. (1986). Turning up the heat on probationers in Georgia. *Federal Probation*, 50(109), 17-24.
- Farrington, D. P. (1983). Randomized experiments on crime and justice. *Crime and Justice*, 4, 257-308.
- Farrington, D. P., & Welsh, B. C. (2005). Randomized experiments in criminology: What have we learned in the last two decades? *Journal of Experimental Criminology*, 1, 9-38.
- Gendreau, P. A., Goggin, C., Cullin, F., & Andrews, D. (2000). The effects of community sanctions and incarceration on recidivism. *Forum on Corrections Research*, 12, 10-13.
- Gendreau, P. A., & Paparozzi, M. A. (2005). An intensive supervision program that worked: Service delivery, professional orientation, and organizational supportiveness. *The Prison Journal*, 85, 445-466.
- Gifford, S. L. (2002). *Justice expenditures and employment in the United States, 1999*. Washington, DC: Bureau of Justice Statistics, Office of Justice Programs.

- Gill, C. E., & Hyatt, J. M. (in press). Not just a numbers game: Is there hope for intensive supervision probation?
- Gottfredson, S. D., & Gottfredson, D. M. (1985). Screening for risk among parolees: Policy, practice, and method. In D. P. Farrington & R. Tarling (Eds.), *Prediction in criminology* (pp. 54-78). New York: State University of New York Press.
- Grattet, R., Lin, J., & Petersilia, J. (2011). Supervision regimes, risk, and official reactions to parolee deviance. *Criminology*, 49, 371-399.
- Gupta, S. (2011). Intention-to-treat concept: A review. *Perspectives in Clinical Research*, 2, 109-112.
- Hollis, S., & Campbell, F. (1999). What is meant by intention to treat analysis? Survey of published randomised controlled trials. *British Journal of Medicine*, 319, 670-674.
- Hyatt, J. M. (2013). *The impact of cognitive behavioral therapy on the recidivism of high risk offenders: Results from a randomized trial*. Philadelphia: University of Pennsylvania.
- Jalbert, S., & Rhodes, W. (2012). Reduced caseloads improve probation outcomes. *Journal of Crime & Justice*, 35, 221-238.
- Jalbert, S., Rhodes, W., & Flygare, C. (2010). Testing probation outcomes in an evidence-based practice setting: Reduced caseload size and intensive supervision effectiveness. *Journal of Offender Rehabilitation*, 49, 233-253.
- King, R. (2009). A change of course: Developments in state sentencing policy and their implications for the federal system. *Federal Sentencing Reporter*, 22, 48-52.
- Langan, P., & Levin, D. (2002). *Recidivism of prisoners released in 1994*. Washington, DC: Bureau of Justice Statistics, Office of Justice Programs.
- Latessa, E. J., Lawrence, T., Fulton, B., & Stichman, A. (1998). *Evaluating the prototypical ISP program*. Washington, DC: National Institute of Justice.
- Lin, J. (2012). Resisting "evidence" in challenging correctional environments. *Criminology & Public Policy*, 11, 421-428.
- Lipsey, M. W., Chapman, G. L., & Landenberger, N. A. (2001). Cognitive-behavioral programs for offenders. *The Annals of the American Academy of Political and Social Science*, 578, 144-157.
- Lipsey, M. W., & Landenberger, N. A. (2005). Cognitive-behavioral interventions: A meta-analysis of randomized controlled studies. In B. C. Walsh & D. P. Farrington (Eds.), *Preventing crime: What works for children, offenders, victims, and places* (pp. 57-71). Berlin, Germany: Springer.
- Lipsey, M. W., Landenberger, N. A., & Wilson, S. J. (2007). *Effects of cognitive-behavioral programs for criminal offenders*. Oslo, Norway: Campbell Systematic Reviews.
- Lowencamp, C. T., Flores, A. W., Holsinger, A. M., Makarios, M. D., & Latessa, E. J. (2010). Intensive supervision programs: Does program philosophy and the principles of effective intervention matter? *Journal of Criminal Justice*, 38, 368-375.

- Lowencamp, C. T., Latessa, E. J., & Smith, P. (2006). Does correctional program quality really matter? The impact of adhering to the principles of effective intervention. *Criminology & Public Policy*, 5, 575-594.
- MacKenzie, D. L. (2000). Evidence-based corrections: Identifying what works. *Crime & Delinquency*, 46, 457-471.
- Marlowe, D. B., & Kirby, K. C. (1999). Effective use of sanctions in drug courts: Lessons from behavioral research. *National Drug Court Institute Review*, 2(1), 1-32.
- Maruschak, L. M., & Bonczar, T. P. (2013). *Probation and parole in the United States, 2012*. Washington, DC: Bureau of Justice Statistics, U.S. Department of Justice.
- Mayo-Wilson, E., Montgomery, P., Hopewell, S., Macdonald, G., Moher, D., & Grant, S. (2013). Developing a reporting guideline for social and psychological intervention trials. *The British Journal of Psychiatry*, 203, 250-254.
- McNeill, F., Farrall, S., Lightowler, C., & Maruna, S. (2012). Reexamining evidence-based practice in community corrections: Beyond "a confined view" of what works. *Justice Research and Policy*, 14(1), 35-60.
- Nath, S. B., Clement, D. E., & Sistrunk, F. (1976). Parole and probation caseload size variation: The Florida Intensive Supervision Project. *Criminal Justice Review*, 1, 61-71.
- Neithercutt, M. G. (1987). Time to recidivism in probation/parole cohorts. *Journal of Contemporary Criminal Justice*, 3(2), 7-24.
- Pearson, F. S. (1988). Evaluation of New Jersey's Intensive Supervision Program. *Crime & Delinquency*, 34, 437-448.
- Perry, A., Weisburd, D., & Hewitt, C. (2010). Are criminologists describing randomized controlled trials in ways that allow us to assess them? Findings from a sample of crime and justice trials. *Journal of Experimental Criminology*, 6, 245-262.
- Peters, R. H., & Murrin, M. R. (2000). Effectiveness of treatment-based drug courts in reducing criminal recidivism. *Criminal Justice and Behavior*, 27, 72-96.
- Petersilia, J. (1987). Probation and felony offenders. *Federal Probation*, 51(2), 56-61.
- Petersilia, J. (1990). When probation becomes more dreaded than prison. *Federal Probation*, 54(1), 23-27.
- Petersilia, J. (1999). A decade of experimenting with intermediate sanctions: What have we learned? *Justice Research and Policy*, 1(1), 9-23.
- Petersilia, J. (2001). Prisoner reentry: Public safety and reintegration challenges. *The Prison Journal*, 81, 360-375.
- Petersilia, J., & Turner, S. (1990a). *Diverting prisoners to intensive probation: Results of an experiment in Oregon*. Santa Monica, CA: RAND Corporation.
- Petersilia, J., & Turner, S. (1990b). *Intensive supervision for high-risk probationers*. Santa Monica, CA: RAND Corporation.
- Petersilia, J., & Turner, S. (1993). Intensive probation and parole. *Crime and Justice*, 17, 281-335.
- Pew Center on the States. (2009a). *One in Thirty One: The long reach of American corrections*. Washington, DC: The Pew Charitable Trusts.

- Pew Center on the States. (2009b). *1 in 31: The long reach of American corrections, Pennsylvania*. Washington, DC: The Pew Charitable Trusts.
- Pew Center on the States. (2010). *Philadelphia's crowded, costly jails: The search for safe solutions*. Washington, DC: The Pew Charitable Trusts.
- Pratt, T. C., Cullen, F. T., Blevins, K. R., Daigle, L. E., & Madensen, T. D. (2006). The empirical status of deterrence theory: A meta-analysis. In F. T. Cullen, J. P. Wright & K. R. Blevins (Eds.), *Taking stock: The status of criminological theory* (pp. 367-396). New Brunswick, NJ: Transaction Publishers.
- Sampson, R. J. (1986). Crime in cities: The effects of formal and informal social control. *Crime and Justice*, 8, 271-311.
- Sherman, L. W. (1992). *Policing domestic violence: Experiments and dilemmas*. New York, NY: Free Press.
- Sherman, L. W. (1993). Defiance, deterrence, and irrelevance: A theory of the criminal sanction. *Journal of Research in Crime & Delinquency*, 30, 445-473.
- Sherman, L. W. (2003). Misleading evidence and evidence-led policy: Making social science more experimental. *The Annals of the American Academy of Political and Social Science*, 589, 6-19.
- Sherman, L. W., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P., & Bushway, S. (1997). *Preventing crime: What works, what doesn't, what's promising—A report to the United States Congress*. Washington, DC: Office of Justice Programs, U.S. Department of Justice.
- Sherman, L. W., Schmidt, J. D., Rogan, D. P., & Smith, D. A. (1992). The variable effects of arrest on criminal careers: The Milwaukee domestic violence experiment. *Journal of Criminal Law and Criminology*, 83, 137-169.
- Smith, P., Gendreau, P., & Swartz, K. (2009). Validating the principles of effective intervention: A systematic review of the contributions of meta-analysis in the field of corrections. *Victims & Offenders*, 4, 148-169.
- Solomon, A. L., Kachinowski, V., & Bhati, A. (2005). *Does parole work: Analyzing the impact of postprison supervision on rearrest outcomes*. Washington, DC: Urban Institute.
- Taxman, F. S. (2002). Supervision: Exploring the dimensions of effectiveness. *Federal Probation*, 66(2), 14-27.
- Taxman, F. S., & Belenko, S. (2012). *Implementing evidence-based practices in community corrections and addiction treatment*. New York, NY: Springer.
- Taxman, F. S., Thanner, M., & Weisburd, D. (2006). Risk, need, and responsivity (RNR): It all depends. *Crime & Delinquency*, 52, 28-51.
- Thanner, M., & Taxman, F. (2003). Responsivity: The value of providing intensive services to high-risk offenders. *Journal of Substance Abuse Treatment*, 24, 137-147.
- Tonry, M. (1999). *Reconsidering indeterminate and structured sentencing: Issues for the 21st century*. Washington, DC: National Institute of Justice.
- Turner, S., & Petersilia, J. (1992). Focusing on high-risk parolees: An experiment to reduce commitments to the Texas Department of Corrections. *Journal of Research in Crime & Delinquency*, 29, 34-61.

- Turner, S., Petersilia, J., & Deschenes, E. P. (1992). Evaluating Intensive Supervision Probation/Parole (ISP) for drug offenders. *Crime & Delinquency*, 38, 539-556.
- Wang, R., Lagakos, S. W., Ware, J. H., Hunter, D. J., & Drazen, J. M. (2007). Statistics in medicine—Reporting of subgroup analyses in clinical trials. *New England Journal of Medicine*, 357, 2189-2194.
- Weisburd, D. (2003). Ethical practice and evaluation of interventions in crime and justice: The moral imperative for randomized trials. *Evaluation Review*, 27, 336-354.
- White, T. F. (2005). *Re-engineering probation towards greater public safety: A framework for recidivism reduction through evidence-based practice*. Court Support Services Division, State of Connecticut, Judicial Branch. Hartford, CT.

### Author Biographies

**Jordan M. Hyatt** is the senior research associate at the Jerry Lee Center of Criminology of the University of Pennsylvania. His research interests include experimental evaluations, community corrections, sentencing and reentry.

**Geoffrey C. Barnes** is a research assistant professor in the Department of Criminology at the University of Pennsylvania. His research interests include risk forecasting, randomized experimentation, and restorative justice.