



The author(s) shown below used Federal funding provided by the U.S. Department of Justice to prepare the following resource:

Document Title: Assessing the Impact of a Graduated Response Approach for Youth in the Maryland Juvenile Justice System

Author(s): Jill L. Farrell, Sean Aaron Betsinger, Natalie Flath, John Irvine

Document Number: 254601

Date Received: March 2020

Award Number: 2016-JF-FX-0059

This resource has not been published by the U.S. Department of Justice. This resource is being made publically available through the Office of Justice Programs' National Criminal Justice Reference Service.

Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Assessing the Impact of a Graduated Response Approach for Youth in the Maryland Juvenile Justice System

February 2020

Jill L. Farrell, Sean Aaron Betsinger, Natalie Flath, & John Irvine*

The Institute for Innovation & Implementation
University of Maryland School of Social Work
306 West Redwood Street
Baltimore, MD 21201

*Maryland Department of Juvenile Services
120 West Fayette Street
Baltimore, MD 21201

This project was supported by Award No. 2016-JF-FX-0059, awarded by the Office of Juvenile Justice and Delinquency Prevention and managed by the National Institute for Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect those of the Department of Justice.

Table of Contents

Abstract..... 2

Introduction 4

 Overview of the Accountability and Incentives Management (AIM) System..... 6

 Theory and Research Review..... 9

 Prior Evaluations of Graduated Response Models 10

 The Present Study 12

Method 14

 Study Design 14

 Data Sources and Collection 15

 Study Groups..... 16

 Measures..... 18

 Propensity Score Matching..... 20

 Data Analysis..... 24

 Process Analysis 24

 Outcome Analysis 24

 Cost Analysis 25

Results..... 26

 Sample Characteristics..... 26

 Process Analyses 26

 Outcome Analyses 33

 Cost Analyses 37

Discussion..... 40

 Study Limitations 48

 Implications..... 50

References 52

Appendix A. AIM Sanctions and Incentives Grids 57

Appendix B. Additional Analyses 59

Abstract

Supervision violations are a major contributing factor to the incarceration of young people in this country, particularly for youth of color. Graduated response systems, which use a range of sanctions and incentives to respond to youth behaviors without relying on confinement, are one approach to achieving accountability, fairness, and recidivism prevention in the juvenile justice system. However, very little research has assessed the effectiveness of these approaches with juvenile offenders under community supervision.

This study evaluated the Accountability and Incentives Management (AIM) system, which was implemented by the Maryland Department of Juvenile Services (DJS) to promote supervision compliance and completion, reduce rates of supervision violations and recidivism, prevent detention and committed placements resulting from supervision violations, decrease lengths of stay under supervision and in placement, and address racial disproportionality in detention and committed placements. The study entailed a process evaluation of AIM's implementation; a two-group, quasi-experimental design to assess youth outcomes; and a cost savings analysis. The treatment group (or AIM group) included 1,983 adjudicated youths who started/completed probation or aftercare supervision with Maryland DJS between November 1, 2015 and October 31, 2017. Propensity score matching was used to create a statistically equivalent comparison group, comprised of youth who were supervised prior to AIM implementation (July 1, 2013-June 30, 2015).

Process analysis findings showed that 55% of youth under supervision received an AIM response, and youth were more likely to receive sanctions as opposed to incentives. Virtually all case managers adhered to the range of sanctions recommended through the structured AIM grids. Moreover, responses were applied within approximately 3 days of the identified behavior, on average. Outcome analyses indicated that youth supervised with AIM in place were significantly less likely to have a violation of probation filed with the court, be placed in a committed residential placement, and

commit an offense resulting in adjudication/conviction during supervision, relative to comparable youth supervised prior to AIM implementation. The effects for AIM did not vary by race or supervision type. Further, AIM has not had a substantial impact on placement costs. Study limitations, policy implications, and areas for future research are discussed.

Introduction

Juvenile justice system reform efforts have shifted tremendously over the past two decades, from policies and practices that primarily focused on punitive “lock ‘em up” responses to delinquency and criminal behaviors to those that are based on evidence from research and treating youth in the community whenever possible (National Research Council, 2013). Major reform initiatives have focused on reducing the use of detention and confinement of youth involved in the juvenile justice system (Mendel, 2011; National Research Council, 2013), and between 1997 and 2017 the number of youth held in juvenile residential facilities nationwide decreased 59% (OJJDP Statistical Briefing Book).

Despite this decline, troubling trends are still apparent. A substantial share of youth in juvenile residential placements are held for technical violations of parole or probation—19% of all detention placements and 13% of the committed population nationwide in 2017 (Puzzanchera & Hockenberry, 2019). Overall, the number of juvenile offenders in residential placement who had committed technical violations declined by 31% between 1997 and 2015; however, private facilities reported housing 12% more juvenile offenders for technical violations (Hockenberry, 2018). Technical violations typically involve noncompliance with conditions stipulated as part of probation or parole (e.g., not passing a drug test, missing appointments with probation officers, skipping school, or staying out past curfew) and not necessarily behaviors that present a risk to public safety (Leiber & Peck, 2013; Sedlak & Bruce, 2010). Detention and incarceration are costly (Austin, 1986) and potentially harmful responses (Fagan, 1996; Kupchik et al., 2003; Loughran et al., 2011; Mendel, 2011) for youth who do not pose a significant threat to the community.

Disproportionate minority confinement has been another long-standing problem for juvenile justice systems (e.g., Mendel, 2009; Development Services Group, 2014). In 2015, minority youth accounted for 69% of offenders in residential placement (Hockenberry, 2018), and the placement rate for minority youth in 2017 was more than twice the rate for white youth (OJJDP Statistical Briefing

Book). Further, approximately two-thirds of youth held in juvenile residential facilities for technical violations are youth of color (Sickmund et al., 2019), and research has shown that minority groups are more likely to be detained for technical violations relative to white youth (Leiber & Boggess, 2012; Rosay & Everett, 2006; Steinhart, 2001).

Increasingly, states and localities are incorporating the lessons learned from research on adolescent development and effective interventions into their juvenile justice reform strategies. In 2013, the National Research Council (NRC) released *Reforming Juvenile Justice: A Developmental Approach*, a report that summarized contemporary research and emphasized how knowledge about adolescent development aligns with the goals of the juvenile justice system, namely holding youth accountable, being fair, and preventing reoffending. The report specified several hallmarks of a developmental approach to juvenile justice reform, including (1) accountability without criminalization, (2) alternatives to juvenile justice involvement, (3) individualized response based on assessment of risks and needs, (4) confinement only when necessary for public safety, (5) a genuine commitment to fairness, (6) sensitivity to disparate treatment, and (7) family engagement. The authors urged policy makers at the federal, state, and local levels to incorporate these tenets into their juvenile justice reform efforts (also see NRC, 2014).

Graduated response systems are one approach to accomplishing the juvenile justice goals of accountability, fairness, and preventing recidivism. Graduated response systems typically entail an array of sanctions and incentives used with offenders under supervision to increase compliance with the conditions of supervision and promote public safety by deterring delinquent behavior without having to use detention or incarceration. Sanctions are tiered with increasing severity to match the increasing seriousness of the youth's behavior. Incentives serve as motivators and rewards for positive behaviors, affirmative changes, and compliance with supervision conditions. Supporters of these approaches contend that a "strong system of 'graduated responses' – combining sanctions for violations and

incentives for continued progress – can significantly reduce unnecessary incarceration, reduce racial and ethnic disparities, and improve successful probation completion rates and other outcomes for youth under supervision” (Center for Children’s Law and Policy, 2016, p. 4).

Overview of the Accountability and Incentives Management (AIM) System

In July 2015, the Maryland Department of Juvenile Services (DJS) implemented the Accountability and Incentives Management (AIM) system, a graduated responses approach designed to promote youth accountability and to reduce rates of supervision violations and recidivism for all youth under community supervision. The primary intent of AIM was to address the high rates of DJS detention and committed placements resulting from violations of probation (VOPs), and to address racial disproportionality in the deep end of the juvenile justice system.

Several studies and statistics related to Maryland DJS youth supported the intentions of the AIM system. For instance, more than one-third (35%) of DJS detention admissions in the first two months of 2013 resulted from the failure to comply with the conditions of either detention alternative programs or community-based supervision (Betsinger et al., 2013). In an analysis of more than 17,000 disposition decisions during State Fiscal Years (FY) 2013 and 2014, researchers found that DJS commitment was ordered in 47% of disposition decisions for which the most serious offense alleged or adjudicated was a VOP, whereas commitment was ordered in only 19% of disposition decisions adjudicated on crimes of violence. In addition, 59% of all DJS commitments involved low- and moderate-risk youth, suggesting that most of these youth did not pose a substantial risk to public safety that would warrant confinement/out-of-home placement (Annie E. Casey Foundation, 2015). Recidivism rates for DJS committed youth were also high—of youth released from a committed placement in FY14, 46% were rearrested within 12 months (Maryland DJS, 2015). Further, similar to the national statistics, while black youth comprised only 33% of Maryland’s youth population (ages 10-19) in 2013 (Maryland State Data

Center, 2015), 79% of DJS detentions and 69% of all DJS commitments were for black youth (Maryland DJS, 2013). Reducing this disproportionality was a major focus of the agency's efforts.

While DJS staff had been using sanctions and incentives with supervised youth for years, this practice was largely left to the discretion of the case managers, resulting in uneven application of these responses and an unknown impact on youth outcomes. With AIM, DJS sought to overcome the uncertainty associated with this informal approach by using a protocol that provides step-by-step guidance to case managers in appropriately responding to youth behavior and aligns with the developmental approach to juvenile justice reform. AIM was developed by DJS in partnership with the Center for Children's Law and Policy through a comprehensive process that included: a review of best practices and related research; a site visit to, and the development of tools from, a Juvenile Detention Alternatives Initiative (JDAI) model site in Santa Cruz, CA; feedback from line staff and managers about supervision challenges, needs, and best practices; and surveys to gauge staff knowledge of graduated responses and current supervision practices. The resulting graduated response system was initially piloted in three offices in January 2015, and statewide rollout followed in July of the same year. Training curricula were developed to support the rollout to supervision staff, and each office had an on-site AIM expert to provide ongoing consultation. As of November 1, 2015, staff had received further clarification on AIM procedures based on the first months of implementation, decision protocols were slightly revised based on feedback, and a larger pool of response options were made available (i.e., resources for incentives); at this point, AIM was considered to be fully implemented.

In the AIM training, graduated responses were defined as "holding youth accountable for misconduct using a wide array of sanctions implemented progressively in proportion to the noncompliant behavior and risk level of youth, and providing incentives to youth for making progress toward short and long term positive goals." The AIM protocol includes an *Infraction Determination Guide*, *Sanctions Grids*, an *Incentives Grid*, and an *AIM Tip Sheet*. AIM's *Infraction Determination Guide*

helps staff to classify the severity of a violation (i.e., minor, moderate, or serious) using five different assessment areas: relationship to the underlying offense, victim impact, frequency of the violation, compliance with conditions, and community safety. The *Sanctions Grids* were designed to match the resources available in each county. They outline a range of responses to a particular behavior based on the severity of the violation and the youth's probation supervision level. The *Incentives Grid* similarly provides a menu of options to reinforce positive youth behaviors (see Appendix A for an example of the Sanctions and Incentives Grids). Staff are encouraged to choose developmentally appropriate sanctions and incentives based on the youth's age, level of functioning, etc., and to work with supervisors and regional directors to develop tailored plans for special cases that may not fit within the standardized instruments. To increase youth and family perceptions of fairness, explanations are to be provided for the selection of particular sanctions, and VOPs may not be filed until less restrictive responses have been attempted. The *AIM Infraction Determination Guide, Sanctions Grids, and Incentives Grid* were programmed into DJS's case management information system to support implementation and monitoring of staff use of the system.

AIM is designed to improve DJS's community supervision practices and youth outcomes in several ways. Youth must meet certain terms and conditions as part of court orders and case plans, and AIM provides case management staff with effective tools to promote compliance by continually and appropriately responding to both positive and negative behaviors. AIM also helps case managers to convey that they want youth to be successful, and not just comply with the basic terms of probation. More specific intended outcomes include: fewer VOPs; reduced detention and committed admissions for supervision violations, particularly for low-risk offenders; higher successful supervision completion rates; shorter lengths of stay under supervision and in detention placements; lower recidivism rates; and greater consistency and equity in the use of violations, especially by race. It was also expected that AIM

would lead to reduced costs associated with residential care and shorter lengths of stay under supervision.

DJS completed an internal process evaluation of AIM implementation between July and October 2015 to examine the extent to which the practice was implemented as planned, explore possible barriers to implementation, and identify areas for improvement or refinement. Findings showed that sanctions accounted for a majority (91%) of the recorded entries. The low rate of recorded incentives (9%) may have been influenced by the lack of funds for incentives until late October, though field workers also reported that they under-recorded verbal praise in the AIM module. On average, the length of time between a youth's behavior and a response was 2.4 days ($SD = 4.9$), though response times ranged between 1 and 77 days. Case managers' responses were largely consistent with the responses prescribed via the AIM grids (i.e., there was a low rate of sanction grid overrides). Findings from this preliminary evaluation resulted in additional refinement to AIM definitions and responses to ensure more clarity for staff and improved implementation.

Theory and Research Review

AIM is primarily based on the principles of deterrence theory, which specifies that sanctions need to be certain, swift, and proportionate to the severity of the behavior to deter criminal activity (Beccaria, 1974; Bentham, 1789; Gibbs, 1975). Underlying these principles is the notion that humans are rational and weigh the costs and benefits of their actions. Theorists distinguish between two types of deterrence—general and specific. General deterrence is based in the notion that individuals learn from the experiences of others, and suggests clear laws with dependable enforcement will deter would-be offenders. On the other hand, specific deterrence is premised in experiential learning, suggesting an individual's own experiences shape assessment of risks and rewards for future behaviors. Thus, to be deterred, offenders must perceive a high likelihood of receiving a sanction (certainty). Further, the sooner a sanction is applied (celerity), the greater the effect it will have to deter future law violations,

with longer response times increasing perceptions that the sanction is unfair or questionable. Finally, per the severity principle, the sanction must be strong enough to deter the offense (i.e., the cost outweighs the benefit to the offender), though using strong sanctions such as detention and incarceration inappropriately or too quickly may lessen their deterrent effects. Moreover, any potential deterrence effect is likely dependent on both individual and situational characteristics, as opposed to uniform (Piquero et al., 2011), suggesting that responses should be individualized to the extent possible.

A highly consistent finding from the body of literature aimed at empirically testing deterrence theory is that the certainty, rather than the severity, of punishment yields the greater deterrent effect (Apel, 2013; Nagin, 2013). Research on celerity has provided less concrete evidence of a deterrent effect of swiftness (Paternoster, 2010). Despite mixed evidence, many criminal justice and correctional practices have been informed by deterrence theory. Summary evidence from reviews and meta-analyses of adult and juvenile correctional interventions have concluded that those based on deterrence are not effective and may even have harmful effects (MacKenzie & Farrington, 2015; Lipsey, 2009; Lipsey & Cullen, 2007). However, more recent interventions related to graduated sanctions have focused on increased certainty and celerity, and less on increased severity, with some reported success (e.g., Hawken & Kleiman, 2009; Hamilton et al., 2016; see below).

Prior Evaluations of Graduated Response Models

Over the past decade, the number of adult community supervision models based in deterrence theory, often referred to as “swift, certain, and fair” (SCF) approaches, have proliferated, largely due to the initial positive findings for Hawaii’s Opportunity Probation with Enforcement (HOPE) Program (Hawken & Kleiman, 2009). HOPE was an innovative program aimed at increasing swift and certain punishment of drug offenders through repeat testing and using sanctions less severe than sending violators to prison. Using a randomized control trial, Hawken and Kleiman (2009) demonstrated that HOPE resulted in reductions in positive drug tests, missed appointments, recidivism, revocations, and the number of days incarcerated for sanctions among adult drug offenders. However, subsequent

evaluations of similar approaches for adult offenders have yielded mixed evidence. An evaluation in Ohio found that community-based sanctions administered in a more certain, severe, and swift manner were effective in reducing adult parole violators' likelihood of recidivism and time to recidivism (Steiner et al., 2012), and Hamilton et al. (2016) found promising results such as reduced revocations and time in confinement for Washington State's Swift and Certain Program. On the other hand, Lattimore et al.'s (2016) multisite replication of HOPE and O'Connell et al.'s (2016) evaluation of Delaware's Decide Your Time did not find evidence supporting these programs.

Research also shows that incentives should be used in conjunction with sanctions to promote positive behavior. For instance, in a study of offenders who participated in Wyoming's Intensive Supervision Program between 2000 and 2003, Wodahl and colleagues (2011) found that the number of rewards applied had almost twice as strong of a relationship to success as the number of sanctions. Further, incentives and sanctions worked best when used together, and applying incentives at a ratio of four rewards to every one sanction continued to increase the chances of successful completion.

There are little to no studies focused on the use of graduated response approaches with juvenile offenders. Sanctions, incentives, and rewards (SIRs) were one component of the Juvenile Breaking the Cycle (JBTC) Program, a multi-modal intervention for drug-involved juvenile offenders designed to reduce substance use, recidivism, and other negative outcomes. Lattimore et al.'s (2005) evaluation of JBTC showed reductions in marijuana use and recidivism compared with youth who did not participate in JBTC, however, they were unable to assess the effects of individual model components (e.g., SIRs). Process findings highlighted lessons learned from SIRs implementation and indicated that program staff were receptive to the use of SIRs for securing compliance among juvenile offenders and found it to be effective.

Although some existing research evidence supports the use of graduated response approaches with justice-involved populations, there are no published rigorous evaluations of the effectiveness of

using graduated sanctions and incentives with a general population of juvenile offenders. Studies to date have focused on adult populations (e.g., Hawken & Kleiman, 2009) or juvenile drug offenders (e.g., Lattimore et al., 2005), and sanctions and incentives are often only one component of the model being evaluated (e.g., JBTC; Lattimore et al., 2005). Additional research is warranted to understand if these approaches are effective with general populations of juvenile offenders under community supervision.

Similarly, research is needed to determine the impact of graduated response systems on the overrepresentation of youth of color at the deep end of the juvenile justice system. Through its emphasis on the consistent and objective application of graduated sanctions and incentives, AIM was intended to address racial disproportionalities in DJS's detention and committed populations. To date, no peer-reviewed research has assessed whether a structured graduated response system of sanctions and incentives reduces racially disparate treatment of juvenile offenders.

The Present Study

This study assessed DJS's implementation of AIM and the effectiveness of this approach with respect to outcomes and costs with youth under community supervision. An important consideration when evaluating the effectiveness of these models, and justice programs and policies more generally, is whether the program was implemented as intended. Criminologists have called for more attention to implementation in evaluations (Mears & Kelly, 2002). Without full implementation, it is not possible to draw definitive conclusions about policy or program effectiveness (Corbett & Lennon, 2004). Indeed, Hawken (2010) drew attention to this notion as part of the HOPE experiment, citing the importance of implementing swift and certain supervision with high fidelity in order to achieve good outcomes. To that end, the research questions for the proposed study included:

1. To what extent did case managers implement AIM as intended?
 - a. How soon were graduated responses applied after the behavior was identified?

- b. How frequently did the case managers depart from *AIM Sanction Grid* recommendations?
 - c. Were there differences in the application of graduated responses by youth's gender, race, age, geographic location, supervision type, or risk level?
2. Did the implementation of AIM improve outcomes for youth under community supervision?
- a. Were there differences in AIM's effects by race?
 - b. Were there differences in AIM's effects by supervision type (probation and aftercare supervision)?
3. Did AIM implementation result in cost savings for the juvenile justice system?

The target population for this study included adjudicated youth who were under probation or aftercare supervision with Maryland DJS (i.e., community supervision). The study's relevance goes beyond Maryland, however. With over 155,500 delinquency cases resulting in probation in 2017 (Sickmund et al., 2019), and evidence for frequent and potentially unfair application of probation violations, this evaluation provides necessary evidence as to the effectiveness of graduated responses with this population and to support the implementation of these models more broadly.

Further, by quantifying program costs and conceptualizing the benefits as costs averted by reducing the use of detention and committed placements, policy makers will be able to weigh whether AIM is a good financial investment, and, ultimately, how to maximize their available resources so that placements are reserved for those youth who pose the greatest risk to public safety. In addition, this evaluation adds to a growing body of literature examining the financial benefits and costs associated with community supervision (e.g., Drake, 2018b; Hamilton et al., 2016; Roman & Harrell, 2001) by focusing specifically on potential cost savings associated with the implementation of graduated response systems in a juvenile justice context.

Method

This study included adjudicated youth under community supervision (i.e., either on probation or aftercare supervision upon reentry to the community from a committed placement) with the Maryland Department of Juvenile Services. DJS is an executive agency responsible for appropriately managing, supervising, and treating youth who are involved in the juvenile justice system in Maryland. The agency serves youth involved at all stages of juvenile justice case processing, providing intake, detention, probation, commitment, and aftercare services. In FY15 (just prior to AIM rollout), DJS handled 23,446 complaints at intake; 2,818 of these complaints resulted in an order to probation supervision and 928 resulted in a commitment to DJS (Maryland DJS, 2015).

Study Design

This study entailed process, outcome, and cost analyses, using retrospective and prospective data. Given that AIM was rolled out with all DJS case management staff at the same time (July 1, 2015) and prior to when evaluation efforts were initiated, a randomized control trial, the most rigorous experimental design, could not be considered for establishing a relationship, if any, between AIM and changes in the outcomes. Instead, a two-group, quasi-experimental design was employed. The treatment group, or AIM cohort¹, included youth who started and ended community supervision between November 1, 2015 and October 31, 2017. Although AIM was rolled out statewide on July 1, 2015 (i.e., staff were fully trained and the AIM module became available in the case management information system), a later start date for the treatment cohort was determined for a few reasons. First, staff received further clarification in the first few months of implementation, as they got used to the new system, and the decision protocols were slightly revised based on feedback. Further, monetary resources for incentives were not made available to staff until November 1, 2015. While case

¹ While we also refer to the treatment group as the “AIM cohort”, it is important to note that not every youth will exhibit a positive or negative behavior warranting an AIM response during supervision. Case managers had AIM available to them while supervising youth during this time period, however.

management staff had the ability to incentivize/reward youth with non-monetary incentives, the additional resources contributed to a larger pool of options. The comparison group, or pre-AIM cohort, included youth who started and ended community supervision prior to AIM implementation, between July 1, 2013 and June 30, 2015, excluding the three offices involved with the AIM pilot.

Data Sources and Collection

This study capitalized on the comprehensive data collected in DJS's data systems, including detailed data related to AIM implementation. The primary data sources were administrative data collected through DJS's management information systems, ASSIST (Automated Statewide System of Information Support Tools) and METS (Maryland Evaluation & Treatment Services). Information regarding a youth's demographic characteristics (e.g., gender, race/ethnicity, age, and county of jurisdiction), delinquency history (e.g., referrals to DJS, VOPs, adjudications, and commitment dispositions), placement history (e.g., detention and committed placements), and supervision history (e.g., aftercare and probation) is maintained in ASSIST. Case managers maintain information related to all behaviors, sanctions, and incentives in the METS AIM module; risk/needs assessment (risk level and treatment needs) and treatment planning information is also entered into this system. DJS's Budget and Finance Office maintains all cost data. These data are limited to per diem rates for detention and other residential placements; DJS was unable to quantify the costs specific to AIM implementation (e.g., training, IT programming, sanctions, incentives, etc.), and the agency does not calculate costs associated with supervision, which are independent of supervision caseloads.

DJS also receives adult criminal justice system data from the Maryland Department of Public Safety and Correctional Services (DPSCS) for DJS-involved youth. By routinely linking its data to DPSCS data, DJS is able to describe more accurately the level of recidivism for youth and to determine the extent to which youth penetrate the adult system (i.e., arrest, conviction, criminal sentence types).

DJS is a data-driven system and utilizes all data for management and decision-making purposes. Several systems are in place to ensure the quality and accuracy of data, including the use of

SafeMeasures® (an automated case management reporting service developed and maintained by the National Council on Crime & Delinquency), regular supervisory case reviews during which individual case data is closely reviewed for compliance with policy, and a rigorous audit system.

Data for study youth were collected through October 2017. The average length of stay under probation supervision was 13 months and 9 months for youth under aftercare supervision in FY15 (DJS, 2015), providing a 24-month window for the AIM cohort to start and end supervision—a sufficient time frame for many youth to complete a supervision episode. A two-year window was also used for the comparison group. DJS researchers de-identified the data files from ASSIST, METS, and DPSCS and transferred them to the University of Maryland researchers via UMB’s secure online data transfer system. The study was approved by the University of Maryland’s Institutional Review Board.

Study Groups

All youth who began and completed community supervision were eligible for sample inclusion, and the initial sample provided by DJS included all probation and aftercare supervision cases that opened during the study periods ($N = 11,562$). See Table 1 for steps used to determine the final study groups and the number of cases excluded at each step.

A primary goal of AIM is to support successful case closure; therefore, the measurement period covered time under supervision through case closure. Cases were included if supervision began and ended during the designated time frames for both treatment and comparison cohorts. Cases open for less than 30 days were excluded; these cases are relatively uncommon, and the first 30 days of supervision are generally devoted to case planning and provide an insufficient time frame to assess AIM practices and impacts. Two youths had duplicate supervision cases (i.e., two cases of the same type starting on the same day); the longer case was retained for each youth. Many youth had multiple supervision cases during the study window—the first was retained for analysis, as any subsequent cases would likely result from a failure event, measured as part of the outcomes. For those who were in a residential placement at the onset of their supervision, the supervision start date was revised to the

residential release date. Youth in residential placement during the entirety of supervision were excluded since the intent of this evaluation was to assess outcomes for youth under community supervision. Youth supervised for less than 30 days in the community once the start date was adjusted were also dropped. Youth served by the three offices that participated in the AIM pilot test in early 2015 were excluded from the comparison cohort. Very few cases had missing data on the study variables; these cases were dropped. Additionally, cases served both during the treatment and comparison cohorts were retained in the treatment cohort. The final sample size of the AIM ($n = 1,983$) and pre-AIM ($n = 2,329$) groups comprised of 4,312 youths.

Table 1. Case Selection for the Study Samples

	Pre-AIM	AIM	Total
Starting Sample of Cases	6,550	5,012	11,562
Excluded Cases (n):			
Missing Supervision Close Date	76	973	1,049
Supervision Close Date outside of cohort windows	2,830	1,641	4,471
Less than 30 days of supervision	128	104	232
Duplicate supervision cases, same Open Dates (longer case retained)	1	1	2
Multiple supervision cases (initial case retained)	625	236	861
In residential placement for the whole course of supervision	15	12	27
Less than 30 days post-residential supervision ¹	56	58	114
AIM pilot site	400	--	400
Missing data on study variables	19	4	23
Duplicate youth in both groups	71	--	71
Final Sample ²	2,329	1,983	4,312

¹ Supervision Start Date adjusted to Residential Release Date.

² Prior to propensity score matching.

The METS dataset comprised of youth behaviors and AIM responses (i.e., sanctions and incentives) was prepared to evaluate AIM implementation.² Behaviors and responses that did not occur during the youth's supervision episode were excluded. The final sample size comprised of 5,101 behaviors and responses³ among the AIM group of 1,983 youths.

² Note that a behavior cannot be recorded in METS without a response.

³ A small share of youth under community supervision (11%) also had AIM sanctions imposed by Community Detention staff. These responses are not included in the present analysis.

Measures

Process Variables

AIM-specific variables included youth behaviors (positive and negative) and responses (sanctions and incentives). The time lapse between behavior and response was measured as the difference between the corresponding dates. A departure from the sanctions grid recommendation (i.e., staff imposed a sanction either more or less severe than the recommended response), or AIM override, was indicated by a categorical measure (up, down, not applicable).⁴ Two variables representing the ratio of incentives to sanctions at the youth supervision level were generated. The first categorized the ratios into one of the following groups: more incentives to sanctions, same number of sanction and incentives, more sanctions to incentives, and no incentives or sanctions. The second categorized the incentive to sanction ratios more specifically (e.g., 4:1, 3:1, 2:1, etc.).

Outcome Variables

The outcomes for research questions (RQ) 2a-c were drawn from data collected by DJS and DPSCS. All outcomes include events that occurred during the supervision episode. The outcomes of interest for this study were: (1) VOP filed with the courts; (2) commitment due to VOP; (3) detention⁵; (4) committed residential placement; (5) successful supervision completion; (6) length of stay under supervision (number of days); (7) length of stay in detention (number of days); and recidivism as measured by (8) referral to DJS or adult arrest, (9) adjudication or adult conviction, and (10) DJS commitment or adult incarceration. For the adjudication/conviction and commitment/incarceration recidivism variables, the event is measured as the date of the associated referral to DJS or adult arrest. In addition to the presence or absence of each of the dichotomous outcomes, number of days between the start of supervision and the outcome were measured. Successful supervision completion was

⁴ AIM override procedures only apply to sanctions.

⁵ We initially planned to assess “detention due to VOP”, but these data were not systematically available at the time of the analysis.

defined as having a favorable case closure reason (i.e., successful termination or work completed)⁶ and not having a VOP, detention, committed residential placement, or recidivism event during supervision. Length of stay under supervision reflects the number of days from the adjusted supervision start date through case closure, thus measuring length of supervision in the community. For all outcome analyses, the treatment condition is defined as being under supervision once AIM was implemented (1) versus being supervised prior to AIM implementation (0).

Other Variables

Several variables were included in the analyses for matching, descriptive, and/or predictive purposes, including: age at the start of supervision, gender, race/ethnicity (African American/Black, Caucasian/White, or Other/Unknown)⁷, jurisdiction type (urban, suburban, or rural/large town)⁸, supervision type (probation or aftercare), most serious adjudicated offense type (against person, property, drug, or other⁹), and most serious adjudicated offense level (felony, misdemeanor, other¹⁰). Measures of risk for reoffending and criminogenic needs were derived from DJS's risk/needs assessment, the Maryland Comprehensive Assessment and Service Planning (MCASP) Assessment.¹¹ The MCASP Assessment is completed for all adjudicated youth by DJS case management staff prior to disposition, or within 30 days post-disposition for cases in which the case manager did not have an opportunity to complete it prior to disposition (i.e., adjudication and disposition hearings were held on the same day). Its two main purposes are to identify the youth's risk level (low, moderate, or high) and areas of need (total domain scores and categorized as low, moderate, or high¹²) to guide supervision

⁶ It is possible that case managers use inconsistent criteria to select case closure reasons, so additional criteria were added to specify a more conservative definition of "successful".

⁷ DJS's data system captures race and ethnicity in one field (i.e., Hispanic/Latino is an option in the drop-down list). For this analysis, Hispanic/Latino is grouped with Other/Unknown, due to relatively small numbers.

⁸ To operationalize jurisdiction type, each county was classified as urban, suburban, large town, or rural on the basis of its population size, population density, and proximity to metropolitan areas.

⁹ Other adjudicated offense types include unspecified felony and misdemeanor offenses, violations of probation, ordinance violations, status offenses, and citations.

¹⁰ Other adjudicated offense levels include ordinance violations, status offenses, and citations.

¹¹ The MCASP Assessment is based on the Washington State Juvenile Court Assessment.

¹² Only the categorical variables for each need domain were provided for the study.

and service plans. The assessment addresses delinquency history and nine need domains, including school, use of free time, employment, peers/relationships in the community, family, alcohol and drugs, mental health, anti-social attitudes, and aggression.

Propensity Score Matching

A quasi-experimental approach using propensity score matching (PSM; Guo et al., 2006; Rosenbaum & Rubin, 1983) was used to create statistically equivalent counterfactual groups for the comparative analyses (RQ2 and RQ3). PSM is a statistical matching method that allows for estimation of causal effects from observational data when a randomized trial is not possible. Given an assumption of selection on observables, PSM can reduce selection bias and threats to internal validity by creating a matched control group using covariates that are associated with treatment group membership, and thus eliminating them as confounders (Guo et al., 2006).

The treatment group (i.e., those who experienced AIM) was matched to youth who were under community supervision prior to AIM implementation. Matching characteristics included demographic covariates (age at start of supervision, gender, race, and jurisdiction type), supervision type, most serious adjudicated offense type, most serious adjudicated offense level, as well as risk level and treatment needs (school, peer, family, alcohol and drugs, mental health, and aggression). The anti-social attitude need domain was not included due to collinearity with aggression.

As an initial step, bivariate comparisons (t-tests and chi-square tests) and standardized differences (using the `pbalchk` command in Stata) were assessed for each matching variable to identify differences across groups prior to PSM. Propensity scores were estimated in Stata by regressing the likelihood of being in the treatment group on the set of covariates, which generated a set of predicted probabilities. Using a caliper of 0.001, 1:1 nearest neighbor matching without replacement was used to match the propensity scores with the treatment on comparison cases. Cases outside the common support and unmatched to the 0.001 caliper width were excluded in order to remove any unreliable

treatment effects. Standardized differences between the treatment and comparison groups indicated balance post-matching, defined as an absolute value less than $p < .10$ for each covariate. Mean bias within the matched groups was 1.6 as determined by the `pstest` command, falling within an acceptable range (below 5). Boxplots and kernel density plots further confirmed covariate balance (Figures 1 and 2). The final sample size was 3,180 with 1,590 treatment cases and 1,590 comparison cases.

Several other matching strategies were also employed. Using the previously established propensity scores, we conducted the PSM using 1:1 nearest neighbor without replacement and two different calipers (.03 and .001); 1:1 nearest neighbor with replacement and both calipers; 2:1 nearest neighbor with replacement and a caliper (.03); and inverse probability weighting, which retained all of the treatment and comparison cases. Table 2 shows summary statistics for each matching procedure. The selected matching strategy for all outcome analyses retained a substantial number of treatment and comparison cases and yielded the lowest mean bias, with almost identical propensity score distributions for each group. Table 3 shows the sample descriptives for all pre- and post-match measures; frequencies are shown for dichotomous measures and means and standard errors are shown for continuous measures. The relevant bivariate test statistics and percent bias show differences in the sample characteristics pre- and post-matching.

Figure 1. Boxplots of the Distributions of Group Propensity Scores, Pre- and Post-Matching

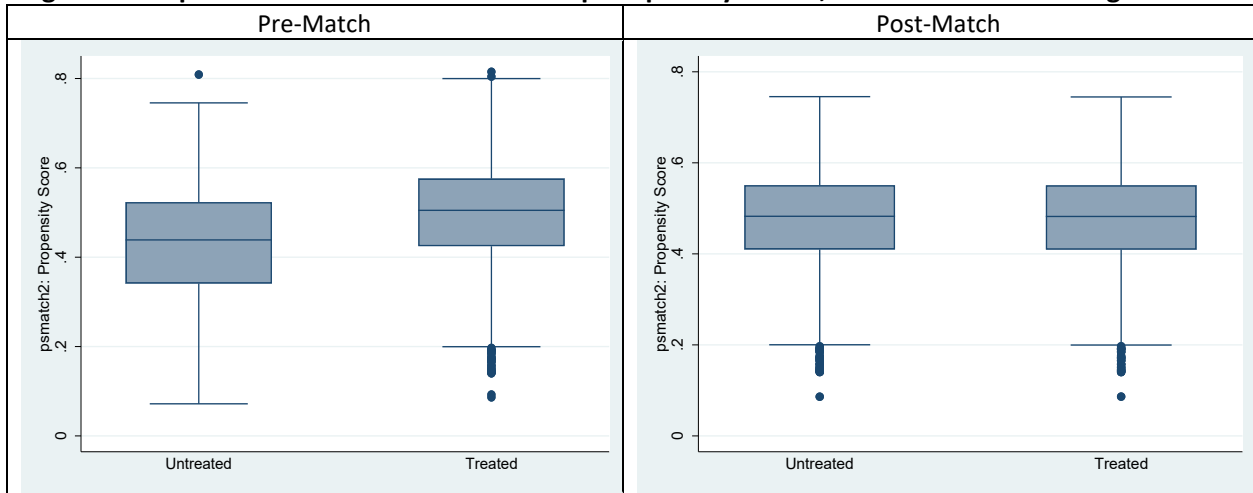


Figure 2. Kernel Density Plots of the Distributions of Group Propensity Scores, Pre- and Post-Matching

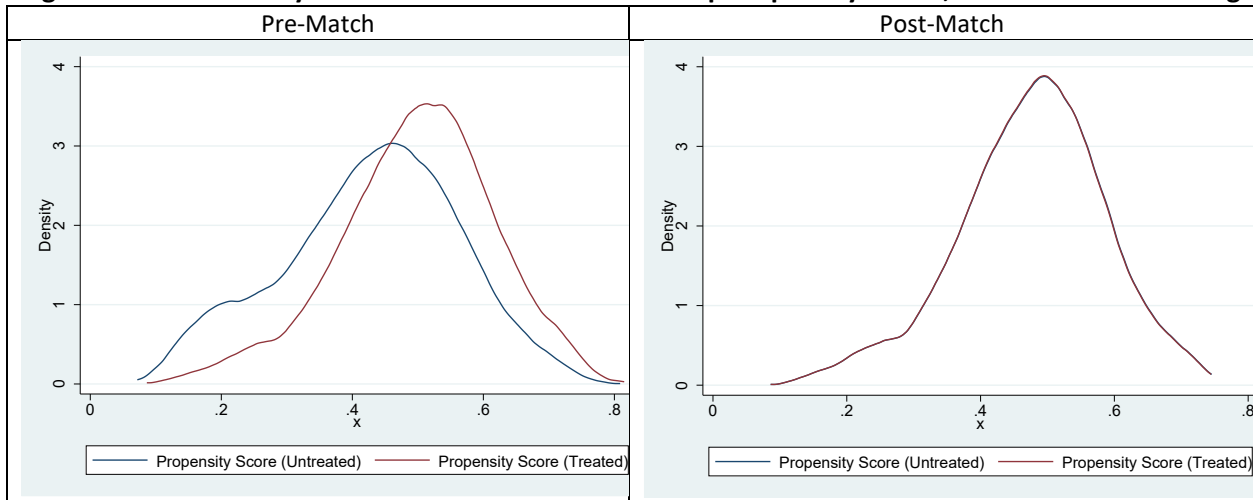


Table 2. Summary of Alternative Propensity Score Matching Specifications

	Sample Size		Mean Bias
	Pre-AIM	AIM	
Unmatched	2,329	1,983	7.9
1:1 No replacement, .03 caliper	1,679	1,679	1.6
1:1 No replacement, .001 caliper	1,590	1,590	1.6
1:1 With replacement, .03 caliper	1,981	1,981	2.1
1:1 With replacement, .001 caliper	1,941	1,941	2.2
2:1 With replacement, .03 caliper	1,981	1,981	1.8
Inverse probability weighting	2,329	1,983	4.3

Table 3. Sample Characteristics by Study Condition, Pre- and Post-Matching

Variable	Pre-Match				Post-Match			
	AIM Mean(SE)/%	Pre-AIM Mean(SE)/%	t-value/ χ^2	% bias	AIM Mean(SE)/%	Pre-AIM Mean(SE)/%	t-value/ χ^2	% bias
Total Sample	1,983	2,329	--	--	1,590	1,590	--	--
<i>Age at Start of Supervision</i>	16.3 (.03)	16.4 (.03)	1.61	-4.9	16.3 (.04)	16.3 (.04)	0.59	-2.1
<i>Gender</i>								
Male	80.7	77.8	5.75*	4.3	79.2	78.9	0.05	0.8
Female	19.3	22.2			20.8	21.1		
<i>Race/Ethnicity</i>								
African American/Black	68.9	65.0	27.99***	8.4	68.2	67.9	1.45	0.5
Caucasian/White	24.7	30.9		-13.9	26.6	27.7		-2.4
Other/Unknown	6.4	4.1		10.2	5.2	4.4		3.7
<i>Jurisdiction Type</i>								
Urban	22.4	29.5	33.22***	-16.1	25.3	24.7	0.41	1.4
Suburban	37.2	36.8		0.9	37.4	36.9		0.9
Rural/Large Town	40.4	33.7		13.8	37.4	38.4		-2.2
<i>Supervision Type</i>								
Probation	90.8	88.8	4.35*	6.4	90.0	90.0	0.00	0.0
Aftercare	9.2	11.2			10.0	10.0		
<i>Adjudicated Offense Type</i>								
Person	49.9	39.2	142.42***	21.8	46.4	47.8	1.32	-2.8
Property	26.5	28.6		-4.5	29.1	28.4		1.4
Drug	6.0	17.0		-35.1	6.9	7.3		-1.4
Other	17.5	15.2		6.2	17.7	16.5		3.2
<i>Adjudicated Offense Level</i>								
Felony	23.6	18.8	15.10**	11.5	20.6	20.5	0.05	0.2
Misdemeanor	69.0	74.0		-11.1	71.8	72.0		-0.6
Other	7.4	7.1		1.1	7.7	7.5		0.7
<i>Risk Level</i>								
High	13.3	10.6	16.08***	8.1	11.6	12.5	0.90	-2.7
Moderate	28.6	25.5		7.2	27.5	26.4		2.5
Low	58.1	63.9		-11.9	60.9	61.2		-0.5
<i>School Need</i>								
High	34.1	30.6	8.72*	7.6	32.8	32.7	0.02	0.3
Moderate	25.4	24.8		1.4	25.3	25.2		0.3
Low	40.5	44.7		-8.4	41.9	42.1		-0.5
<i>Peers/Relationships Need</i>								
High	29.0	28.9	1.36	0.2	29.1	29.2	0.52	-0.1
Moderate	54.1	52.9		2.4	52.9	53.8		-1.8
Low	16.8	18.2		-3.5	18.0	17.0		2.5
<i>Family Need</i>								
High	12.1	12.9	1.54	-2.5	12.7	13.0	1.02	-0.8
Moderate	30.0	31.0		-2.2	30.6	29.0		3.6
Low	57.9	56.1		3.7	56.7	58.1		-2.8
<i>Alcohol & Drugs Need</i>								
High	24.9	30.0	14.62**	-11.6	26.1	26.5	0.06	-0.8
Moderate	15.0	14.6		1.1	14.5	14.3		0.4
Low	60.1	55.3		9.7	59.4	59.2		0.5
<i>Mental Health Need</i>								
High	22.6	18.2	17.09***	10.9	20.1	21.3	1.32	-3.1
Moderate	16.3	14.9		3.8	15.8	16.5		-1.9
Low	61.1	66.9		-12.1	64.1	62.1		4.1
<i>Anti-Social Attitudes Need</i>								
High	40.3	39.4	1.35	1.9	40.9	41.3	0.16	-0.9
Moderate	20.6	22.1		-3.5	20.8	21.1		-0.6
Low	39.0	38.5		1.1	38.3	37.6		1.4
<i>Aggression Need</i>								
High	39.5	40.2	3.64	-1.3	39.9	41.0	1.05	-2.3
Moderate	26.3	28.2		-4.3	26.8	27.4		-1.3
Low	34.2	31.6		5.4	33.3	31.6		3.6

Notes: SE = Standard error. *p<.05, **p<.01, ***p<.001

Data Analysis

Multiple analyses were conducted to address each research question. The pre-matching treatment group was used to evaluate AIM implementation (RQ1a-c). The matched treatment and comparison samples were employed for analysis of the main outcome effects (RQ2a), subgroup interaction effects (RQ2b, c), and cost-savings (RQ3).

Process Analysis

For RQ1, descriptive analyses examined whether AIM was implemented as intended by case management staff. First, types of behaviors and responses were described with frequency measures at the behavior level. Evaluation staff worked with DJS staff to classify the behaviors and responses into categories to facilitate assessment of the data. Aggregated measures for responses were also assessed at the youth level. Following, average number of days between youth behavior and case management response was evaluated to examine how swiftly responses were applied following the youth's behavior (RQ1a). Next, we examined how closely case managers followed the AIM protocols to determine the appropriate sanction by assessing rates of departure from the recommendations resulting in more or less severe sanctions (i.e., overrides) and documented reasons (RQ1b). Further analyses assessed the data by subgroups to determine whether there were differences in the application of AIM based on youth characteristics, supervision type, risk level, needs, and/or geographic location types (RQ1c). In addition to summary statistics (i.e., frequencies, means), analyses related to each of the questions outlined in RQ1 employed ANOVA and chi-square tests to examine group-based differences in the implementation outcomes.

Outcome Analysis

For RQ2, using the matched treatment and comparison group samples, we first calculated percentages/means for each outcome. Next, multivariate analyses, including Cox regression, logistic regression, and linear regression models, were conducted for each outcome, as appropriate, to measure the impact of AIM. For most outcomes, each individual in the sample was at risk of experiencing a failure

event between the start and end of his/her supervision episode, which can vary widely in length, and survival analysis techniques involve statistical controls to account for different times at risk (Hosmer & Lemeshow, 2003). To account for any residual imbalance after propensity score matching, we estimated regression models with all of the covariates included in the propensity score models, otherwise known as doubly robust regression models. Results presented are based on the regression-adjusted models using 1:1 nearest neighbor matching without replacement and a caliper of .001, though we conducted multiple sensitivity tests to check the consistency of our findings with different specifications of the propensity score match. To assess if there were differences in AIM effects by race, we included an interaction term between race and treatment group in the regression models that included all covariates. To assess if there were differences by probation or aftercare supervision, we interacted supervision type with study condition and evaluated outcomes in similar manner. For all analyses, we use a minimum α level of .05 to identify a statistically significant effect.

Cost Analysis

Cost-offset (also known as cost-savings) analysis, a simplified form of cost-benefits analysis, was originally chosen as the approach for examining whether the costs associated with implementing the AIM model in Maryland would be offset by reductions in costs associated with secure detention and committed placements (RQ3). Though it was anticipated that the marginal costs associated with AIM could be identified to include only those that were accrued beyond the normal costs of DJS case management for youth under supervision (e.g., programming METS to support implementation, training DJS case managers, providing incentives to promote positive behavior), DJS was unable to isolate and quantify these costs. Accordingly, the approach to the cost analysis was modified to focus primarily on potential benefits, including decreased lengths of stay in detention and committed placements. In addition to the data used for RQ1 and RQ2, data used for this analysis included DJS per diem rates for detention and committed placements. Because the analysis took the perspective of DJS rather than program recipients or society more broadly, it did not address opportunity costs to individual

participants, nor did it attempt to monetize broader savings to the public (e.g., averted pain and suffering caused by subsequent delinquency). Although it was anticipated that estimates based on these analyses would likely be conservative due to the relatively short time frame used (i.e., placements during supervision), these estimates were still expected to be beneficial to DJJ (and juvenile justice agencies more broadly), given the extent to which they are accustomed to making decisions based on short-term outcomes, in line with legislative cycles (Maher et al., 2012; Chamberlain et al., 2011).

Results

Sample Characteristics

Table 3 (above) illustrated the descriptive statistics for the treatment group (prior to matching). Of the 1,983 youths in the treatment condition, 80.7% were male, 68.9% were *African American/Black* and 24.7% *Caucasian/White*, and their average age at the start of supervision was 16.3 years old. The majority were under supervision in jurisdictions classified as *rural/large town* (40.0%) and *suburban* (37.2%). Most youth (90.8%) were under probation supervision and 9.2% were under aftercare supervision. Their most serious adjudicated offenses were primarily person-related (49.9%; e.g., assault, robbery) and misdemeanors (69.0%). Per the MCASP Assessment, most youth were classified as *low* risk for reoffending (58.1%), and needs were variable—40.3% rated as *high* need for anti-social attitudes, compared with 39.5% for aggression, 34.1% for school, 29.0% for peers/relationships, 24.9% for alcohol and drugs, 22.6% for mental health, and 12.1% for family.

Process Analyses

An initial analysis assessed the behaviors and responses for all youth at the behavior level. There were a total of 5,101 behaviors with an AIM response among youth under community supervision from November 1, 2015 to October 31, 2017; of these, 71.5% were negative behaviors that resulted in a sanction and 28.5% were positive behaviors with associated incentives/rewards.

Table 4 shows the frequencies for the 3,649 negative behaviors reported for AIM. The most common behaviors were related to supervision adherence (e.g., *missed office appointment*, 20.7%), school (e.g., *unexcused school absence*, 17.3%), and curfew (e.g., *curfew violation*, 13.6%), and the least frequent was related to restitution/fine payments (0.6%). One-quarter (25.2%) of the behaviors were indicated using the “other” option. Analysis of the text responses for “other” suggested that this response option was often selected by staff to input more than one behavior resulting in the response; to provide more qualitative information about the behavior, which frequently could have been selected from the drop-down; and to insert behaviors that were not available in the drop-down options.¹³

Table 4. Negative Behaviors Identified During Community Supervision (N=3,649 behaviors)

Behavior Category	% of Total	Behavior	% of Total
Supervision	20.7	Missed office appointment	20.7
School	17.3	Unexcused school absence	17.3
Curfew	13.6	Curfew violation	13.6
Substance Use	11.4	Positive drug results - Marijuana	9.6
		Positive drug results - Other CDS	1.4
		Positive drug results - Alcohol	0.4
Program/ Services	11.3	Missed treatment or program meeting	11.3
Restitution/ Fines	0.6	Failure to pay restitution/fine	0.6
Other	25.2	Other	25.2

The most common response to negative behaviors was *verbal warning* (47.2%), followed by *program/services* (13.8%) and *make up missed appointment* (11.8%; Table 5). The least frequent response involved community service. Preliminary analysis of the text field associated with the option *programs/services* showed the most frequent entry did not actually specify a program or service (29%), but often referenced another sanction type (e.g., detention, VOP, or another court-related action). Frequently described program types included substance use disorder services and behavioral health services, including family therapy.

¹³ The “other” responses were analyzed and categorized in a separate project to inform the addition of drop-down options.

Table 5. Sanctions Applied During Community Supervision (N=3,649 sanctions)

Sanction Category	% of Total	Sanction	% of Total
Verbal Warning	47.2	Verbal warning	47.2
Program/ Services	13.8	Programs/services	13.6
		Community conferencing	0.1
Make Up Missed Appointment	11.8	Make up missed appointments	11.8
Meeting	9.1	Youth & parent meeting (3 bus. days)	4.7
		Youth & parent meeting (1 bus. day)	2.2
		Administrative meeting with youth, family, supervisor (1 bus. day)	1.9
		Administrative meeting with youth, family, supervisor (2 bus. days)	0.2
Increase Restrictions/ Decrease Privileges	7.2	Restrict activity (short-term)	4.9
		Modify curfew (extended term)	1.5
		Modify curfew (short-term)	0.5
		Increase drug screenings	0.4
		Restrict activity (extended-term)	--
Increase Supervision	3.4	Increase supervision contacts	2.6
		GPS for 30 days (VPI only)	0.5
		Increase supervision level	0.2
		Increase supervision level to moderate	0.1
VOP	3.3	File VOP	3.3
Reading/ Writing Assignment	2.2	Written essay or oral report	1.9
		Assigned reading or video viewing	0.3
Community Service	2.0	Community service (up to 4 hours)	0.8
		Community service (up to 12 hours)	0.7
		Community service (up to 8 hours)	0.3
		Community service (up to 10 hours)	0.2
		Community service (up to 6 hours)	--
		Community service (up to 14 hours)	--

The most common positive behaviors recorded in the AIM system were related to compliance (76.5%), followed by pro-social involvement (16.2%) and self-advocacy (7.3%; see Appendix B, Table 1).¹⁴ Positive behaviors included attendance and participation in program/services (25.1%); school attendance, behavior, and performance (18.2%); and supervision adherence (15.6%). The least common positive behaviors were related to paying restitution (1.9%) and peers (1.4%; Table 6).

¹⁴ These categories are part of the AIM data entry screen.

Table 6. Positive Behaviors Identified during Community Supervision (N=1,452 behaviors)

Behavior Category	% of Total	Behavior	% of Total
Program/ Services	25.1	Attends and/or completes court-ordered cognitive programs/ treatment services	23.9
		Participates in additional services	1.2
School	18.2	Improves school attendance	4.7
		Improved grades in school	4.5
		Improved behavior in school	2.7
		Receives promotion in grade level or employment	2.3
		Attends school and all classes as scheduled	1.7
		Enrolls in an academic or vocational program	1.4
		No suspensions/expulsions	0.5
		Maintains grade improvements	0.4
Supervision	15.6	Keeps office appointments with reminders	5.8
		No incidents of whereabouts unknown	3.6
		Keeps office appointments without reminders	3.1
		Takes initiative to meet at least one probation condition	2.6
		Notifies CMS of change in contact information	0.6
Employment	7.3	Obtains/retains employment	7.3
Change Initiative	6.5	Handles a difficult situation well	1.6
		Makes progress toward or completes at least one personal goal	1.0
		Manages time well for at least one important task	1.0
		Acknowledge and/or articulate needs	0.9
		Seeks help	0.9
		Establishes at least one personal goal	0.6
		Begins to acknowledge and/or articulate Needs	0.3
		Completes at least one personal goal	0.2
Substance Use	6.3	Clean urine screen/drug free	6.3
Family/Home	5.9	Abides by set family/house rules	5.9
Curfew	4.9	Abides by curfew	4.9
Community Service	4.6	Locates and/or begins court-ordered community service	1.9
		Completes at least 50% of court-ordered community service	1.4
		Completes at least 75% of court-ordered community service	0.8
		Completes 100% of court-ordered community service	0.6
Extracurricular Activity	2.3	Joined/Participates in at least one extra-curricular activity (school/community)	2.3
Restitution	1.9	Makes consistent payments and/or pays off restitution	1.6
		Begins paying restitution	0.3
Peers	1.4	Remains detached from negative peer groups	0.7
		Detaches from negative peer groups	0.8

Table 7 shows the frequencies for incentives applied during the course of supervision for the study sample. The most salient response among AIM incentives was *verbal praise* to the youth or to the parent (66.9%). Tangible items were the second most common response (23.8%), with the most

frequent item being *food/candy/treat* (19.9%). Supervision incentives, increases of privileges/decreases of restrictions, recognition (beyond verbal praise), and tickets for events/activities were less than 5% of the total responses. None of the responses involved other pro-social trips/activities or opportunities.

Table 7. Incentives Applied During Community Supervision (N=1,452 incentives)

Incentive Category	% of Total	Incentive	% of Total
Verbal Praise	66.9	Verbal praise to youth	61.3
		Verbal praise to parent	5.6
Tangible Item	23.8	Food/candy/treat	19.9
		School supplies	1.5
		Personal hygiene supplies	1.0
		Meal coupons/meal for youth's family	0.8
		Books/magazines	0.4
		Media card (iTunes, Play Store)	0.2
		College/university gear	--
		Magazine subscription	--
		Recognition/ Certificate	2.9
Certificate of achievement	0.6		
Invitation to monthly/annual recognition ceremony	0.5		
Positive letter home from case manager	0.3		
Positive letter home from Supervisor or Director	--		
Publicly display work	--		
Decrease Supervision	2.4	Reduced level of supervision (major)	1.8
		Reduced level of supervision (moderate)	0.3
		Request removal of electronic monitoring	0.3
Terminate Supervision	1.9	Petition of termination of case	1.9
Decrease Restrictions/ Increase Privileges	1.4	Extend curfew	1.4
		Allow previously restricted activity	--
		Request reduction in community service hours (major)	--
		Request reduction in community service hours (moderate)	--
		Restoration of driving privileges (major)	--
		Restoration of driving privileges (moderate)	--
Tickets/Passes	0.7	Tickets to sporting events	0.6
		Tickets to other activities (amusement park, museums, etc.)	0.1
		Membership in athletic/art/other classes	--
		Movie tickets	--
		Recreation center passes/classes	--
Opportunity/ Trip/Activity	--	Field trips with staff (e.g. hiking, rafting)	--
		Invitation for youth to serve on agency leadership council/serve in come leadership role	--
		Job shadowing/apprenticeship opportunity	--
		Lunch with case manager/judge	--
		Manicure/pedicure or other self-care activity	--
		Meal/meeting with elected official or other well-known individual	--
		Tour of local college/university	--
		Tour of other local business of interest	--
		Tour of stadium or athletic facility	--

Just over half (55.3%, $n = 1,097$) of youth had at least one response during their supervision episode—42.3% of youth had at least one sanction and 28.9% had at least one incentive (Table 8). Of those with at least one response, the average number of responses per youth was 4.65 ($SD = 5.96$), with a range of 1 to 64. Of those with sanctions applied, the average number per youth was 4.35 ($SD = 5.58$); and of those with incentives, the average number was 2.53 ($SD = 3.54$).

Table 8. Responses During Supervision per Youth (N=1,983)

	% Youth with No Responses (n)	% Youth with Responses (n)	Youth with Responses			
			Mean	Standard Deviation	Minimum	Maximum
Incentive	71.1 (1,409)	28.9 (574)	2.53	3.54	1	49
Sanction	57.7 (1,145)	42.3 (838)	4.35	5.58	1	59
Incentive and/or Sanction	44.7 (886)	55.3 (1,097)	4.65	5.96	1	64

Chi-square tests of independence were performed to examine the relationships between several factors and whether youth had *any* responses applied during supervision. Youth with responses were significantly more likely to be ages 14-16 at the start of supervision, relative to 13 and under or 17 and over [$\chi^2(2, N = 1,983) = 8.35, p = .015$]. They were also more likely to be supervised in rural/large town jurisdictions relative to urban and suburban [$\chi^2(2, N = 1,983) = 37.06, p = .000$]. Youth classified as moderate and high risk [$\chi^2(2, N = 1,983) = 30.08, p = .000$], and generally those having moderate or high needs, relative to those classified as low, were also more likely to have any responses. There were no significant associations between having a response and gender, race/ethnicity, or supervision type. See Appendix B, Table 2 for a full summary of these results.

Of youth with responses, we also assessed the ratio of incentives to sanctions applied during supervision. Overall, just under two-thirds (63.1%) of youth had *more sanctions to incentives*, and just under one-third (31.3%) had *more incentives to sanctions* (Table 9). Again, chi-square tests were performed to assess the relationship of the ratio with factors included in the study. Significant relationships were identified between the response ratio category and race/ethnicity [$\chi^2(4, N = 1,097) = 21.27, p = .000$], jurisdiction type [$\chi^2(4, N = 1,097) = 75.10, p = .000$], risk level [$\chi^2(4, N = 1,097) = 66.21, p$

= .000], and the need domains, with the exception of mental health. Youth were more likely to receive *more sanctions to incentives* if they were *African American/Black*, supervised in an *urban* jurisdiction, classified as *high risk*, and classified as having *high needs*, relative to their respective counterparts. There were no significant associations between response ratio category and age, gender, or supervision type. See Appendix B, Table 3 for a full summary of these results. Further, prior research has suggested that a ratio of 4 incentives to 1 sanction is most effective to support behavior change among offenders. In the AIM sample, only 2.0% ($n = 22$) of youth with responses (or 1.1% of the full AIM sample) had the respective ratio of 4:1.

Table 9. Application of Incentives to Sanctions During Supervision per Youth (N=1,097)

Ratio Category	Percent (n)
More Incentives to Sanctions	31.3 (343)
Same Number of Incentives and Sanctions	5.7 (62)
More Sanctions to Incentives	63.1 (692)

Time to response was examined using two approaches upon further review of the data—at the behavior/response level and at the youth level (of youth with responses). Summary statistics are reported in Table 10. The mean number of days to response for all behaviors was 2.79 days ($SD = 6.79$). The average time to response was lower for incentives ($M = 2.05$, $SD = 8.16$) relative to sanctions ($M = 3.08$, $SD = 6.14$). While the response time ranged from 0 to 133 days across all behaviors, the modal response time was 0, or the same day (50.5%; 79.0% for incentives and 39.1% for sanctions). The average response time during supervision was also calculated for all youth with any responses. Of youth with responses, the mean response time was 3.24 days ($SD = 7.60$), ranging from 0 to 133 days.

Table 10. Time from Behavior to Response (Number of days)

	N	Mean	Std. Deviation	Minimum	Maximum
Behavior/Response Level					
Incentive	1,452	2.05	8.16	0	133
Sanction	3,649	3.08	6.14	0	92
Total	5,101	2.79	6.79	0	133
Youth Level					
Average Response Time during Supervision	1,097	3.24	7.60	0	133

Differences in the time to any response were assessed using one-way ANOVA tests with the main study factors at both the response and youth levels. At the response level, there were statistically significant differences in response time by race/ethnicity [$F(2, 5098) = 9.99, p = .000$], jurisdiction type [$F(2, 5098) = 16.73, p = .000$], school need [$F(2, 5098) = 3.30, p = .037$], and alcohol and drug need [$F(2, 5098) = 4.78, p = .008$]. Youth identified as *African American/Black* had longer response times relative to *Caucasian/White* youth, and youth supervised in *urban* jurisdictions had longer response times relative to their counterparts. Youth identified as *high* need for school (relative to *moderate* need) and *high* need for alcohol and drugs (relative to *low* need) had lower response times. There were no statistically significant differences between group mean response times by age, gender, supervision type, most serious adjudicated offense type and level, risk level, and the remainder of needs. At the youth level, of youth with responses, jurisdiction type was the only factor that had a statistically significant difference between groups [$F(2, 1094) = 4.29, p = .014$], with youth in *urban* jurisdictions having longer average response times ($M = 4.58, SD = 10.18$) than youth in *suburban* ($M = 2.82, SD = 6.79$) and *rural/large town* ($M = 2.98, SD = 6.78$) locations. See Appendix B, Tables 4 and 5 for a full summary of these results.

Per AIM procedures, overrides are only applicable for applying sanctions. Staff overrode the AIM recommended responses in only four (0.1%) cases among the 3,649 community supervision sanctions imposed. All cases involved overrides up to more serious sanctions. Due to the very small number of relevant cases ($n = 4$), planned analyses to examine differences in the use of overrides by study factors were not conducted.

Outcome Analyses

The effects of AIM were assessed for multiple outcomes, using the matched pre-AIM and AIM sample; findings are summarized in Table 11. For each outcome, means are shown for the treatment and comparison groups, as well as the relevant effect estimate. The results show that AIM youth were less likely to have a VOP filed with court during supervision (16.2%) compared with pre-AIM youth

(20.2%). Study condition was significant in the Cox regression model, indicating that AIM youth had a 25% lower risk of having a VOP filed compared to pre-AIM youth ($HR = 0.75, p = .001$). AIM youth were also significantly less likely to have a new committed residential placement ($HR = 0.63, p = .009$) and have a new referral to DJS/arrest that results in adjudication/conviction ($HR = 0.82, p = .046$) during supervision, relative to pre-AIM youth. AIM youth were also less likely to have a new commitment due to a VOP and a new detention, and slightly more likely to have a new referral to DJS/adult arrest and one that resulted in commitment or incarceration; however, these differences were not statistically significant in the Cox regression models. The rates of successful supervision closure were similar for the treatment (48.0%) and comparison (49.1%) groups, and study condition was not significant in a logistic regression model.

Table 11. Summary of Outcomes during Supervision, Bivariate and Regression Model Estimates

Outcome	% / Mean		Treatment Effect in Regression ¹ with Covariates ²			
	Comparison (Pre-AIM)	Treatment (AIM)	Hazard Ratio	SE	Sig.	95% CI
VOP Filed with Court	20.2	16.2	0.75	0.06	.001	0.63 – 0.89
VOP Commitment	10.1	9.4	0.89	0.11	.335	0.71 – 1.12
Detention	29.2	27.2	0.91	0.06	.187	0.80 – 1.04
Committed Residential Placement	5.7	4.2	0.63	0.11	.009	0.44 – 0.89
Referral to DJS/Adult Arrest	31.9	32.5	1.01	0.06	.845	0.89 – 1.15
Adjudication/Adult Conviction	13.8	11.5	0.82	0.08	.046	0.67 – 1.00
Commitment/Incarceration	8.6	8.7	1.01	0.13	.941	0.79 – 1.29
			Odds Ratio	SE	Sig.	95% CI
Successful Supervision	49.1	48.0	0.96	0.07	.581	0.82 – 1.11
			Coeff.	SE	Sig.	95% CI
Number of Days in Detention	11.23	12.34	0.90	1.18	.445	-1.41 – 3.21
Number of Days under Supervision	263.10	269.43	5.64	4.77	.237	-3.72 – 15.00

Notes: SE = Standard error. CI = Confidence interval.

¹ Cox regression models were conducted for dichotomous outcomes with varying times at risk; logistic regression models were conducted for dichotomous outcomes measured as of discharge; linear regression models were conducted for continuous outcomes.

² Covariates included: age at the start of supervision; gender (male/female); race/ethnicity (black, white, other/unknown); jurisdiction type (urban, suburban, rural/large town); supervision type (probation, aftercare); most serious adjudicated offense level (felony, misdemeanor, other); most serious adjudicated offense type (person, property, drug, other); risk level (low, moderate, high); and school, peer relationships, family, alcohol and drug, mental health, and aggression needs (low, moderate, high).

Two continuous outcome measures were also assessed—number of days in detention during the supervision episode and length of supervision. Although it was expected that AIM would result in shorter lengths of time for both outcomes, on average, AIM youth had slightly more days in detention ($M = 12.34$) compared with pre-AIM youth ($M = 11.23$), and slightly longer lengths of supervision ($M = 269.43$ and $M = 263.10$, respectively). Study condition was not statistically significant in the linear regression models, however. As previously noted, multiple sensitivity tests were conducted to check the consistency of our findings with different specifications of the propensity score match. These results are summarized in Appendix B, Table 6; substantive findings were largely consistent with the results shown here.

To assess if there were differences in AIM effects by supervision type (probation and aftercare supervision), interaction terms were computed and each regression model (with covariates) was re-estimated. Table 12 shows the outcome means by study condition for probation ($n = 2,862$) and aftercare ($n = 318$) youth. While there are some differences apparent in the failure rates by supervision type (e.g., a larger share of aftercare youth and a smaller share of probation youth had a new detention during supervision post-AIM implementation), the interaction terms were not statistically significant in each model (findings not shown), suggesting that the effects of AIM did not vary by supervision type.

A similar approach was employed to assess if there were differences in AIM effects by race. Table 13 shows the outcome means by study condition for *African American/Black* ($n = 2,164$), *Caucasian/White* ($n = 863$), and *Other/Unknown* race ($n = 153$) youth. Once again, there are some differences apparent in the failure/success rates by group, but none of the interaction terms were statistically significant, though the interaction term in the model estimating the effects on length of supervision approached significance ($p < .064$). Bivariate findings show that youth identified as *Caucasian/White* had significantly longer lengths of stay under supervision post-AIM implementation, while the length of supervision slightly decreased for *African American/Black* youth.

Table 12. Differences in Outcomes by Supervision Type

	% / Mean		t-value/χ ²
	Comparison (Pre-AIM)	Treatment (AIM)	
VOP Filed with Court			
Aftercare	9.4	6.9	0.67
Probation	21.4	17.2	8.08**
VOP Commitment			
Aftercare	2.5	2.5	0.00
Probation	10.9	10.1	0.45
Detention			
Aftercare	34.6	37.1	0.22
Probation	28.6	26.1	2.28
Committed Residential Placement			
Aftercare	16.4	19.5	0.53
Probation	4.5	2.5	8.80**
Referral to DJS/Adult Arrest			
Aftercare	32.7	30.8	0.13
Probation	31.8	32.7	0.27
Adjudication/Adult Conviction			
Aftercare	12.6	10.7	0.28
Probation	13.9	11.6	3.42
Commitment/Adult Incarceration			
Aftercare	8.8	8.8	0.00
Probation	8.6	8.7	0.00
Successful Supervision			
Aftercare	43.4	32.1	4.34*
Probation	49.8	49.8	0.00
Days in Detention			
Aftercare	17.30	23.04	-1.03
Probation	10.56	11.15	-0.49
Days under Supervision			
Aftercare	254.91	252.79	0.13
Probation	264.02	271.28	-1.45

*p<.05, **p<.01

Table 13. Differences in Outcomes by Race/Ethnicity

	% / Mean		t-value/χ ²
	Comparison (Pre-AIM)	Treatment (AIM)	
VOP Filed with Court			
African American/Black	19.8	16.1	5.21*
Caucasian/White	20.7	15.6	3.74
Other/Unknown	22.9	20.5	0.13
VOP Commitment			
African American/Black	9.2	9.2	0.00
Caucasian/White	12.1	9.0	2.14
Other/Unknown	11.4	13.3	0.12
Detention			
African American/Black	31.2	28.6	1.75
Caucasian/White	25.0	22.7	0.63
Other/Unknown	24.3	31.3	0.93
Committed Residential Placement			
African American/Black	5.4	4.3	1.25
Caucasian/White	6.6	3.6	4.13*
Other/Unknown	4.3	4.8	0.02
Referral to DJS/Adult Arrest			
African American/Black	34.7	35.8	0.27
Caucasian/White	27.3	25.8	0.25
Other/Unknown	17.1	24.1	1.11
Adjudication/Adult Conviction			
African American/Black	15.5	13.0	2.67
Caucasian/White	10.2	8.5	0.75
Other/Unknown	10.0	7.2	0.38
Commitment/Adult Incarceration			
African American/Black	9.7	9.7	0.00
Caucasian/White	6.6	6.4	0.02
Other/Unknown	4.3	7.2	0.59
Successful Supervision			
African American/Black	46.3	43.5	1.77
Caucasian/White	54.1	58.4	1.62
Other/Unknown	61.4	54.2	0.81
Days in Detention			
African American/Black	13.26	13.72	-0.28
Caucasian/White	6.57	7.72	-0.70
Other/Unknown	9.30	17.92	-1.40
Days under Supervision			
African American/Black	271.95	269.00	0.50
Caucasian/White	244.22	266.33	-2.47*
Other/Unknown	245.36	290.82	-1.93

*p<.05

Cost Analyses

To assess whether AIM was associated with cost savings to DJS, we examined whether the costs associated with placements during supervision were reduced for the AIM group compared to the

comparison group. Though we originally proposed assessing data related to costs of AIM implementation, including programming changes to METS, training provided to DJS case managers, and sanctions and incentives used to address behavior, these costs could not be not be quantified by DJS. Therefore, the analyses that follow rely on per diem rates for placements in group homes and in DJS-operated facilities.

Costs were estimated using two different sets of analyses. First, average detention placement costs for the AIM and comparison groups were estimated based on average detention lengths during supervision as well as average detention per diem rates. For this analysis, per diem rates were calculated based on each DJS-operated facility's total expenditures and number of beds.¹⁵ An average per diem rate was then calculated across the seven detention facilities for FY16 through FY18, covering the implementation period for AIM and thereby allowing comparability across the AIM and comparison groups. The results of this analysis suggest that adjusted detention costs for the two groups were very similar, with average costs per stay for the AIM group being \$6,982 and those for the comparison group being \$6,358. The higher cost for the AIM group stems from the fact that, on average, youth in the AIM group spent slightly longer periods, albeit not statistically significant, in detention during supervision than did members of the comparison group.

A second method for estimating costs focuses on the subsample of youth in the two groups who were placed in staff secure, hardware secure, or other committed placements during supervision. This analysis uses the total length of stay for the placement, which extends beyond the supervision period.¹⁶ As with the previous analyses, per diem costs for DJS-operated placements (i.e., staff and hardware

¹⁵ Though DJS has traditionally published per diem rates based on average daily populations, these rates are variable depending on the number of youth placed, even though the overall cost of operating facilities has remained more stable over time. The current approach using the number of beds, which DJS will also employ for future cost reporting, attempts to overcome the variability associated with the ADP approach and provides more stables estimate over time.

¹⁶ For purposes of this analysis, only the length of the first placement during the supervision period is considered; subsequent transfers to different placements and other placements that originated after the supervision period are not considered.

secure facilities) were calculated based on the number of beds in each. However, costs for group homes, treatment foster care, independent living, residential treatment centers, and diagnostic units are based on the number of youth actually placed with each private provider agency; therefore, per diems rates for these placements were based on average daily populations (ADP).

Table 14. Estimated Placement Costs, Youth Placed during Supervision

	Comparison			Treatment (AIM)		
	N	Average Length of Stay	Average Cost Per Placement	N	Average Length of Stay	Average Cost Per Placement
Staff Secure ¹	12	142.3	\$58,104	10	108.3	\$44,211
Hardware Secure ¹	0	--	--	1	186.0	\$107,534
Group Home ²	9	200.6	\$46,647	10	110.3	\$25,655
Treatment Foster Care ²	12	128.9	\$20,423	12	126.5	\$20,040
Independent Living ²	1	48.0	\$6,536	2	120.0	\$16,341
Residential Treatment Center ²	5	174.4	\$71,212	2	249.0	\$101,673
Diagnostic Unit ²	7	90.7	\$27,981	1	92.0	\$28,377

¹ Per diem rates for staff secure and hardware secure facilities were calculated based on the number of beds in each facility and averaged by placement type and across the years covered by AIM.

² Per diem rates were calculated based on ADP and averaged across the years covered by AIM. Group home placements include therapeutic group homes.

A relatively small proportion of youth in the sample is represented by these analyses. A slightly larger number of youth in the comparison group ($n = 12$ vs. $n = 10$) were placed in DJS-operated staff secure facilities during their supervision periods, and they spent longer periods, on average, in these placements than did youth in the AIM group. Accordingly, the estimated costs per staff secure placement are higher for the comparison group (\$58,104) than for the AIM group (\$44,211). On the other hand, only one youth in the AIM group was placed in a hardware secure setting while under supervision, for an estimated total placement cost of \$107,534.

Cost estimates for group home and treatment foster care placements generally favor the AIM group. While more youth in the AIM group ($n = 10$) than in the comparison group ($n = 9$) were placed in group home settings, youth in the comparison group spent longer periods in the placements on average; this translates into higher estimated costs per placement for the comparison group (\$46,647) than for the AIM group (\$25,655). Similarly, though both groups had 12 youth who were placed in treatment

foster settings, youth in the comparison group averaged slightly longer spells in these placements, which again translates in to a slightly higher estimated cost per placement (\$20,423 for youth in the comparison group vs. \$20,040 for youth in the AIM group). On the other hand, relative to the comparison group, youth in the AIM group averaged longer placements in independent living programs, residential treatment centers, and diagnostic units, resulting in higher estimated placement costs for each these settings.

Discussion

This evaluation examined the implementation and outcomes of Maryland DJS's graduated response system, AIM. Results showed that just over half (55%) of youth supervised with AIM in place received at least one sanction or incentive during their supervision, and sanctions were more commonly applied than incentives. Responses to positive and negative behaviors were applied relatively swiftly—within 3 days on average—and staff followed the structured guidance for sanctions in almost 100% of their decisions. The study included an assessment of several youth outcomes, showing some favorable results among youth who received AIM. Specifically, youth supervised with AIM in place were less likely to have a VOP filed with the court, a new committed residential placement, and a new DJS referral/arrest that resulted in an adjudication/conviction during supervision. Finally, a cost analysis yielded mixed evidence related to potential cost savings attributable to AIM.

AIM is a tool for case managers to use with youth with the goals of encouraging supervision compliance and reducing unfavorable outcomes. Training materials drew largely from deterrence theory, calling for responses to be certain, swift, and proportionate to the behavior to achieve positive results. To that end, this study could not assess whether responses were applied when they were warranted (certainty). Complicating this assessment is the fact that not every youth under supervision will require an AIM response to be successful. Indeed, our analysis showed that just over half of the youth had a behavior and response documented as part of AIM. There are three possible explanations

for youth who did not have an AIM response indicated in their case record: 1) they did not exhibit any behaviors that warranted a response or demonstrate a need to be incentivized; 2) a response was warranted but the case manager did not apply it; or 3) a response was applied but it was not recorded in the data system. Regarding the latter groups, DJS has a fairly robust supervision and quality assurance model in place, but we cannot look past the possibility that some responses were not applied or not logged.

Bivariate analyses showed that youth who did not have an AIM response were more likely to be low risk for reoffending, lending some support to the premise that in many cases a response was probably not necessary. Indeed, particularly with low-risk offenders, juvenile justice agencies must be cautious with applying graduated responses that may lead to unintentional net-widening (Goldstein et al., 2016). Conversely, youth classified as moderate/high risk and needs were more likely to have sanctions and incentives applied. To the extent these youth have more conditions and interventions as part of their supervision, it is also conceivable they would have more responses to address their adherence and engagement in required and needed services. Further inquiry into the application of graduated responses by risk levels—and associated outcomes—is warranted to ensure the approach is achieving its intended effects.

Case managers were also more likely to use sanctions as opposed to incentives with youth, countering the recommended practice based on prior research (Wodahl et al., 2011). This finding is consistent with DJS's initial assessment of AIM, which showed approximately 90% of all responses were sanctions and that lack of resources for incentives and under-documenting verbal rewards in the system contributed to at least some of this imbalance. Additional barriers may include lack of line staff buy-in regarding the use of incentives and system-related barriers to utilization of incentives that involve reduced supervision levels and contacts. A deeper assessment of barriers to use of incentives is warranted. Findings from this study also suggest that staff may benefit from additional training and

coaching from supervisors to reinforce the use of incentives. Although AIM was not explicitly grounded in a contingency management (CM) framework in training materials, principles from this approach could be useful for strengthening the theoretical basis for use of incentives. CM is a behavioral approach for incentivizing changes in behaviors and attitudes, using principles of operant conditioning and behavioral strategies (see Trotman & Taxman, 2011), with evidence supporting its use among adolescents with substance use disorders (Stanger et al., 2016).

AIM guidance specifies that responses must occur soon after a behavior so youth learn the connection between the behavior and the response (celerity). The average times to AIM response for recorded behaviors were 3 days and 2 days for sanctions and incentives, respectively, and in half of all responses, it was applied on the same day. There is currently no guidance on optimal response time beyond as soon as possible, but these findings fell within DJS's expectations and were viewed as favorable. Further research is warranted to determine if there is a target response period for shaping adolescent offenders' behaviors and at what point responses may have diminishing effects.

Ensuring responses proportionate to the behavior (severity) was accomplished through structured AIM sanctions and incentives grids. An important aspect of AIM is that it allows case managers flexibility to tailor responses to the individual youth, rather than assuming a single type would be effective for all youth. Beyond offering a range of responses to select through the structured decision-making grids, DJS built in override procedures so that staff could select sanctions outside of the range when they felt it was necessary. In this study, there were just four overrides for all sanctioned behaviors. While this finding was somewhat surprising, in recent years, DJS has identified low override rates with its other structured decision tools (e.g., Detention Risk Assessment Instrument). Taken together, on the one hand, staff may be reluctant to deviate from the recommendations indicated by these tools. Another possibility may be that the range of sanctions available for use within each cell is too wide, making it unnecessary for staff to override the recommended array. The procedures for AIM

overrides are also rigorous, requiring approval of the Regional Director. While the intent of this procedure is to prevent overuse of VOPs, the process may be too onerous to implement, especially in the daily practice of case management with the goal of swift identification and response. Some staff may manipulate the *Infraction Determination Guide* scoring to ensure the behavior would yield a response they already had in mind to prevent the need for override approval. Prior research suggests that some correctional staff may not use structured decision-making tools with fidelity, instead influencing scores to fit their subjective judgments (Murphy & Turner, 2009; Miller & Maloney, 2013). The finding in this study suggests that the sanction grids and override procedures warrant review to ensure intended outcomes.

Despite evidence to suggest that AIM implementation was not optimal (e.g., more use of sanctions than incentives), outcome analyses using quasi-experimental methods revealed some positive impacts. Youth who were supervised post-AIM implementation were 25% less likely to have a VOP filed with court, and less likely to have a new committed residential placement or new referral to DJS/arrest that resulted in adjudication or conviction, during supervision. There were also null findings for several outcomes, including having a new detention; having a new referral to DJS/arrest, as well as one resulting in commitment/incarceration; successful supervision completion; number of days spent in detention; and number of days under supervision. Notably, none of the primary analyses indicated that AIM youth were significantly more likely to have a negative outcome, and several of the null results still trended in a direction that favored AIM. Further, we assessed supervision type separately due to knowledge of other reform efforts related to aftercare that could independently improve outcomes, though no differences were observed based on these analyses. Nor did the effects of AIM vary by race. Taken together, these findings offer promising support for this model. In addition to reassessing these outcomes in other samples, future research should evaluate the impact of AIM post supervision.

Another primary goal of AIM was to address racial disproportionalities in the deep end of the system by ensuring more fairness and objectivity in the application of responses, and reducing the rate at which youth—particularly African American youth—are detained or committed due to VOPs. An assessment of AIM implementation measures and outcomes by race revealed some concerns and some promising findings. While race was not related to the likelihood of a youth receiving *any* response, African American/Black youth were more likely to receive sanctions over incentives, and had a longer average time to response, as compared with Caucasian/White youth. This result is potentially related to differences in the application of AIM by jurisdiction type, given African American/Black youth comprised most youth in the urban jurisdiction. On the other hand, bivariate analyses of outcomes showed that African American/Black youth were significantly less likely to have a VOP filed with AIM in place, and other outcomes were similar or lower than their counterparts in the comparison group, indicating that AIM is a promising strategy to address racial disproportionalities. Notably, however, there was a significant increase in supervision length for Caucasian/White youth post-AIM implementation. This finding runs counter to an intention of AIM—to reduce supervision lengths by increasing compliance—and will need a deeper review to understand the drivers of this result.

While this study did not assess the effectiveness of response types, it was notable that verbal warning and verbal praise were the most prominent responses. In particular, the high number of instances of verbal praise to youth calls attention to the notion that this response was likely used as an ad hoc reward for an identified positive behavior or accomplishment as opposed to a response that was used to incentive the behavior. In Lattimore et al.'s (2005) process evaluation of the Juvenile Breaking the Cycle Program, they noted that staff found it easier to administer rewards over incentives because rewards were more clearly defined. Attention to the distinction and use of these mechanisms may promote improved use of rewards and incentives by DJS staff. However, while both mechanisms support positive behavior change, the relative importance of one versus the other for supervising juvenile

offenders is unknown. Youth participation in future AIM development may also provide insight into effective and rewards incentives from their perspectives.

Additional factors may impact the effectiveness of graduate response approaches. For one, deterrence theory would suggest that youth need to be made aware of the sanctions they may face without compliance to supervision conditions. While the training specified the importance of reviewing AIM with youth under supervision and an AIM flyer was created for dissemination to families, it is unknown the extent to which these actually happened in practice and when it was discussed. Youth also have to be responsive to the sanction or incentive applied. Again, as part of AIM, case managers are instructed to use an individualized approach in the selection of responses, but we did not assess the “appropriateness” of the response for the youth. As noted by O’Connell et al. (2016), effective responses likely vary by offender types (e.g., gender, age, etc.). Further, the AIM training emphasized that responses should remain proportionate to the behavior and not necessarily escalated, especially if the response was previously effective with the youth. Research indicates that youth learn through repetition, and while it may seem counterproductive to use the same response over and over, this approach may be the most effective for behavior change.

This study also did not address staff factors that may affect AIM implementation. It is highly likely that AIM is used more frequently, and more consistently, by some staff more than others. Prior research shows that case managers have different orientations to their role, where some take the perspective of counselor/social worker and others a more law enforcement approach (e.g., Shearer, 2002). AIM may be perceived and utilized differently depending how they view their roles. Use of AIM may also vary by how case managers view the reform. Steiner et al.’s (2011) study of parole officers revealed resistance and cynicism to a new graduated sanctions grid, citing concern about restrictions on their decision-making capacity. Data for this analysis did not include the youth’s case manager, but future analyses should assess this variation and impact on outcomes.

Ultimately, case managers work with other stakeholders to effectively manage and serve youth on their caseloads—court personnel, service providers, community members, caregivers, and others fulfill roles that contribute to the goals of supervision, and their actions may support or impede AIM. For example, judges receive systematic reports of youth progress on community supervision and/or conduct regular review hearings, in which they may impose sanctions (or incentives) to youth behavior. To the extent that responses do not align with the case managers’ recommendations and/or intentions (e.g., the judge orders detention when the case manager recommended a less severe response), AIM-related outcomes could be negatively impacted. Similarly, DJS’s electronic monitoring staff also use a version of AIM to sanction youth (incentives are not part of their model), and a small share of youth in this study were concurrently supervised. This study did not account for, or assess the impact of, responses contemporaneously imposed by others.

An assessment of sanctions imposed as part of AIM also raises additional questions for consideration. DJS purposefully separated program and service options from the sanctions list in the jurisdiction-based graduated response grids so they would not be perceived as punishments. A review of text responses associated with programs/services revealed several entries involving detention, electronic monitoring, and other more punitive programs that were not necessarily intended for this field. Shifting language away from “sanction” and to something that encompasses both punitive and supportive responses to negative behavior may be more consistent with the underlying approach and goals of the model.

Experts from the Center for Children’s Law and Policy (2016) who provide guidance on graduated response systems have warned that behavior labels may not be specific enough to capture the severity of the violation. In this study, the most frequent negative behavior option selected was “*other*.” *Other* was intentionally added to the drop-down list to capture behaviors not considered in the initial design. An assessment of related text responses suggests that case managers selected this option

primarily to specify a behavior not available in the drop-down list, as intended, but also to indicate multiple behaviors for the response and/or to provide more information regarding the behavior and/or response. Deeper assessment of these responses will be instructive for future AIM technical revisions to ensure that AIM implementation can be easily monitored. Again, attention should be paid to relatively low-level behaviors (e.g., missing curfew by a short time frame) resulting in more severe sanctioning over time, as well as behaviors that do not fall under the scope of AIM (e.g., school-specific behaviors that are addressed by the school).

Findings related to potential cost savings are mixed. Though outcome analyses indicate that the AIM and comparison groups are statistically similar with respect to both their likelihood of experiencing a new detention as well as their lengths of stay in detention while under supervision, the slightly longer average detention length for the AIM group (12 days versus 11 days) translated into slightly higher estimated costs per stay in detention (\$6,982 for the AIM group versus \$6,358 for the comparison group). Additional analyses examining the estimated costs for members of both groups who were placed in staff secure, hardware secure, and other committed placement settings during supervision suggest the AIM group may have lower average costs per placement for some settings (i.e., staff secure, treatment foster care, and group home placements), while the comparison group may have more favorable costs for others. Although this set of findings is based on a small number of placed youth and must be interpreted cautiously, additional follow-up analyses are warranted to investigate whether AIM has benefits beyond those that may occur during the supervision.

It should be noted that, while intended to address high placement rates, AIM was not designed with the goal of providing significant cost savings related to detention or committed placements. The overall operational costs for individual DJS-operated facilities generally fluctuate very little over time regardless of the number of occupants or their lengths of stay. In order for true cost savings to be realized, the number of youth placed in detention or committed placements would have to be reduced

to such a degree that the number of correctional staff could be reduced or facilities/units could be closed. Our findings do not suggest that AIM has had such an effect, and, again, that was not one of its intended goals.

Study Limitations

While this study had many strengths with regard to the availability and comprehensiveness of data to assess the primary research questions, it also has limitations. For one, while we expect that the quality of data was high given DJS's use of quality assurance mechanisms, it is possible that staff did not record all youth behaviors and corresponding sanctions/incentives into METS. Dates and other data entry fields rely on workers having an accurate understanding of how to code all circumstances, and it is possible staff drifted from DJS guidance over time.

Second, while a quasi-experimental research design using propensity score matching is generally considered a strong method for evaluating outcomes, PSM cannot control for unobservable variation outside of the measured confounders. Of highest concern was that parallel reform efforts implemented between July 2013 and October 2017 may have influenced outcomes in each cohort (comparison and treatment). Prior to evaluating AIM, DJS's Reentry Strategic Plan was the primary system change identified that might have impacted outcomes for the treatment group. DJS finished developing and started to implement the plan in January 2016, and strategies were implemented over the course of the year. Because this effort consisted of initiatives designed to improved reentry outcomes, including supervision completion and recidivism, outcomes for youth under aftercare supervision were assessed separately (again, no significant differences in outcomes were identified). Changes were also made to DJS's most intensive level of supervision, VPI (Violence Prevention Initiative), as of July 1, 2016, but these changes largely involved revising the supervision label and condensing it into one level of supervision, as opposed to three levels, during the study time frame. Additionally, a new version of the Detention Risk Assessment Instrument (DRAI) was rolled out in July 2017, but the DRAI is used primarily

to drive decision-making for pre-adjudicated youth. Therefore, it seems unlikely that the DRAI had an appreciable impact on outcomes for youth under probation and aftercare supervision during AIM implementation.

Cost analyses were limited due to the unavailability of costs related to AIM implementation. The analyses presented here assessed whether AIM had potential fiscal benefits in the form of reduced detention and placement stays, but they did not address broader questions about whether AIM is financially “worth the investment” by DJS, since costs for programming, training, and using sanctions and incentives to address behavior could not be estimated. It is worth noting again that the most frequently used incentive was verbal praise, which comes at no cost to DJS, as do many of the other incentives and sanctions utilized, including changes in restrictions and supervision levels. However, tangible incentives comprised nearly one-quarter (24%) of the incentives used, and programs/services accounted for approximately 14% of the sanctions used, and each comes with their own costs. A national scan conducted by the Vera Institute of Justice (2012) indicated minimal costs to states implementing graduated response grids, though some states experienced cost increases when new community-based programming had to be installed. Conversely, some states witnessed reduced costs when they were able to reserve more costly resources for those offenders who represent the highest public safety risk. As such, it may be particularly valuable for DJS to assign costs to sanctions and incentives as the agency digs deeper into questions about which may work best for their supervised population.

As with most cost analysis studies, our findings were also limited by the fact that all cost data presented were based on estimates rather than actual expenditures. In addition, some analyses relied on very small numbers of youth and utilized lengths of stay in committed placements extending after the supervision period, which were not a primary focus of the analyses covered in this report. Additional analyses, which further examine outcomes beyond the supervision period, may be instructive for

examining whether costs savings may be more fully realized after supervision ends. In addition, because we limited the scope of our analysis to costs and potential benefits to those experienced by DJS, and not the juvenile justice system or society more broadly, we did not assess other potential cost savings associated with implementation, including outcomes of interest beyond the juvenile justice system (see Drake, 2018a).

Implications

Overall, this study contributes to our understanding of the implementation, outcomes, and potential cost efficiencies of using graduated responses systems with juvenile offenders. Since the reported success of Project HOPE (Hawken & Kleiman, 2009), at least 161 jurisdictions of adult correctional systems have adopted similar models (Bartels, 2016). Toolkits are currently available to assist juvenile justice agencies with implementing graduated response systems (e.g., Center for Children’s Law and Policy, 2016), and several agencies have contacted DJS to inquire about use of the AIM system. Given the apparent interest and the relatively straightforward adoption and implementation requirements of AIM and related systems, additional research on these approaches is sorely needed to back their widespread use in the field, particularly with juvenile offenders. Specifically, this evaluation provided insight into: (a) how well a juvenile justice agency implemented a new graduated response system and where drift may occur; (b) how effective the system was in addressing a multitude of important outcomes for community supervision; (c) how effective the system might be in assisting juvenile justice agencies to address racial disparities in confinement; (d) whether there were differential effects for youth under probation versus aftercare supervision; and (e) whether AIM might help juvenile justice systems to save on costs related to negative outcomes.

The evaluation results have been instructive for DJS. During the evaluation process, as findings became available, the AIM Steering Committee reconvened in the Fall 2017 to focus on using the results to inform revisions to training, policies and procedures, quality assurance and supervision protocols, and

the data system. While encouraged by some positive outcomes, the low use of incentives was surprising, and efforts are underway to conduct refresher trainings that emphasize the intent and use of incentives with youth. The application of AIM responses may yield more even favorable outcomes if incentives were used more frequently.

Another important takeaway from this study is that implementation is ongoing, and even tightly designed interventions may not be carried out as expected. Implementation teams are a recommended approach to managing programs and interventions—to get started with a new effort and for ongoing oversight (Fixsen et al., 2017; Metz et al., 2015). DJS designed and initially implemented AIM using a team approach; however, once it was viewed that full implementation was achieved, the central implementation team disbanded. The reconvening of this team is a positive development, which should enhance current implementation and future outcomes.

Taken as a whole, findings from this study support graduated response systems as a promising strategy for improving youth outcomes under community supervision. The results point toward the importance of continuing the improvement and evaluation of these models, including additional areas for future research, as noted throughout this discussion. The results also contribute more generally to the growing body of work that establishes evidence-based practices for community supervision of juvenile offenders. This study highlights the importance of designing and implementing interventions with evaluation in mind—we were able to conduct a rigorous assessment of AIM retrospectively due to the wealth of administrative data collected by DJS, and specifically to AIM implementation. Practitioners and policy makers should take care to measure program implementation and outcomes wherever possible to inform their efforts and the field of juvenile justice more broadly.

References

- Annie E. Casey Foundation. (2015). *Doors to DJS commitment: What drives juvenile confinement in Maryland?* Baltimore, MD.
- Apel, R. (2013). Sanctions, perceptions, and crime: Implications for criminal deterrence. *Journal of Quantitative Criminology*, 29, 67–101.
- Austin, J. (1986). Using early release to relieve prison crowding: A dilemma in public policy. *Crime & Delinquency*, 32, 404-502.
- Bartels, L. (2016). Looking at Hawaii’s Opportunity with Probation Enforcement (HOPE) program through a therapeutic jurisprudence lens. *QUT Law Review*, 16(3), 30-49.
- Beccaria, C. (1764). *On crimes and punishments*. Indianapolis, IN: Hackett.
- Bentham, J. (1780). *The principles of morals and legislation*. Oxford, UK: Clarendon Press.
- Betsinger, S., Strickman, S., Farrell, J., Irvine, J., Stables, M., Meyers, V., Garry, L., & Walters, S. R. (2013). *Doors to detention: Statewide utilization study*. Baltimore, MD: The Institute for Innovation & Implementation.
- Center for Children’s Law and Policy. (2016). *Graduated responses toolkit: New resources and insights to help youth succeed on probation*. Washington, DC.
- Chamberlain, P., Snowden, L. R., Padgett, C., Saldana, L., Roles, J., Holmes, L., . . . Landsverk, J. (2011). A strategy for assessing the costs of implementing new practices in the child welfare system: Adapting the English cost calculator in the United States. *Administration and Policy in Mental Health and Mental Health Services Research*, 38, 24-31.
- Corbett, T., & Lennon, M. C. (2003). Implementation studies: From policy to action. In M. C. Lennon & T. Corbett (Eds.), *Policy into action: Implementation research and welfare reform* (pp. 1-14). Washington, DC: Urban Institute Press.
- Development Services Group, Inc. (2014). *Disproportionate minority contact. Literature review*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention.
- Drake, E. K. (2018a). Benefits and costs of “Swift, Certain, and Fair” supervision: Is a bottom-line estimate really the bottom line? *Criminology & Public Policy*, 17(4), 865-874.
- Drake, E. K. (2018b). The monetary benefits and costs of community supervision. *Journal of Contemporary Criminal Justice*, 34(1), 47-68.
- Fagan, J. (1996). The comparative impacts of juvenile and criminal court sanctions on adolescent felony offenders. *Law and Policy*, 18(1), 77-119.
- Fixsen, D. L., Blasé, K. A., & Fixsen, A. A. M. (2017). Scaling effective interventions. *Criminology & Public Policy*, 16(2), 487-499.

Gibbs, J. (1975). *Crime, punishment, and deterrence*. New York: Elsevier Scientific.

Goldstein, N. E. S., NeMoyer, A., Gale-Bentz, E., Levick, M., & Feierman, J. (2016). "You're on the right track!" Using graduated response systems to address immaturity of judgment and enhance youths' capacities to successfully complete probation. *Temple Law Review*, *88*, 805–836.

Guo, S., Barth, R. P., & Gibbons, C. (2006). Propensity score matching strategies for evaluating substance abuse services for child welfare clients. *Children and Youth Services Review*, *28*, 357-383.

Hamilton, Z., Campbell, C.M., van Wormer, J., Kigerl, A., & Posey, B. (2016). Impact of swift and certain sanctions: Evaluation of Washington State's policy for offenders on community supervision. *Criminology & Public Policy*, *15*(4), 1009-1072.

Hawken, A., & Kleiman, M. (2009). *Managing drug involved probationers with swift and certain sanctions: Evaluating Hawaii's HOPE*. (National Institute for Justice No. 2007-IJ-CX-0033). Washington, DC: National Institute for Justice.

Hawken, A. (2010). HOPE for probation: How Hawaii improved behavior with high-probability, low-severity sanctions. *Journal of Global Drug Policy and Practice*, *4*(3).

Hockenberry, S. (2018). *Juveniles in residential placement, 2015*. *Juvenile Justice Bulletin*. Washington, DC: U.S. Dept. Of Justice. Office of Juvenile Justice and Delinquency Prevention.

Hosmer, D. W., & Lemeshow, S. (2003). *Applied survival analysis: Regression modeling of time to event data*. New York: Wiley.

Kupchik, A., Fagan, J., & Liberman, A. (2003). Punishment, proportionality, and jurisdictional transfer of adolescent offenders: A test of the leniency gap hypothesis. *Stanford Law & Policy Review*, *14*(1), 57-83.

Lattimore, P. K., Krebs, C. P., Graham, P., & Cowell, A. J. (2005). *Evaluation of the Juvenile Breaking the Cycle program*. Washington, DC: National Institute of Justice.

Lattimore, P. K., MacKenzie, D. L., Zajac, G., Dawes, D., Arsenault, E., & Tueller, S. (2016). Outcome findings from the HOPE demonstration field experiment: Is swift, certain, and fair an effective supervision strategy? *Criminology & Public Policy*, *15*, 1103–1141.

Leiber, M. J., & Boggess, L. N. (2012). Race, probation violations and structured secure detention decision making in three jurisdictions. *Youth Violence and Juvenile Justice*, *10*(4), 333-353.

Leiber, M. J., & Peck, J. H. (2013). Probation violations and juvenile justice decision making: Implications for Blacks and Hispanics. *Youth Violence and Juvenile Justice*, *11*, 60–78.

Lipsey, M. (2009). The primary factors that characterize effective interventions with juvenile offenders: a meta-analytic overview. *Victims and Offenders*, *4*, 124–147.

Lipsey, M. W., & Cullen, F. T. (2007). The effectiveness of correctional rehabilitation: a review of systematic reviews. *Annual Review of Law and Social Science*, *3*, 297–320.

Loughran, T. A., Mulvey, E. P., Schubert, C. A., Chassin, L.A., Steinberg, L., Piquero, A. R., Fagan, J., Cota-Robles, S., Cauffman, E., & Losoya, S. (2010). Differential effects of adult court transfer on juvenile offender recidivism. *Law and Human Behavior, 34*(6), 476-488.

MacKenzie, D. L., & Farrington, D. P. (2015). Preventing future offending of delinquents and offenders: What have we learned from experiments? *Journal of Experimental Criminology, 11*, 565-595.
and meta-analyses?

Maher, E. J., Corwin, T. W., Hodnett, R., & Faulk, K. (2012). A cost-savings analysis of a statewide parenting education program in child welfare. *Research on Social Work Practice, 22*, 615-625.

Maryland Department of Juvenile Services. (2013). *Data resource guide: Fiscal Year 2013*. Online. Available: [http://www.djs.maryland.gov/drg/Full DRG With Pullouts 2013.pdf](http://www.djs.maryland.gov/drg/Full_DRG_With_Pullouts_2013.pdf)

Maryland Department of Juvenile Services. (2015). *Data resource guide: Fiscal Year 2015*. Online. Available: http://www.djs.maryland.gov/drg/2015/2015_Full_DRG.pdf

Maryland State Data Center. (2015). *Total population estimates by race and Hispanic origin for July 1, 2013*. Online. Available: http://www.mdp.state.md.us/msdc/Pop_estimate/estimate_10to14/CensPopEst10_14.shtml

Mears, D. P., & Kelly, W. R. (2002). Linking process and outcomes in evaluating a statewide drug treatment program for youthful offenders. *Crime and Delinquency, 48*, 99-115.

Mendel, R. A. (2009). *Two decades of JDAI: From demonstration project to national standard*. Baltimore, MD: Annie E. Casey Foundation.

Mendel, R. A. (2011). *No place for kids: The case for reducing juvenile incarceration*. Baltimore, MD: Annie E. Casey Foundation. Retrieved from http://www.aecf.org/~media/Pubs/Topics/Juvenile%20Justice/Detention%20Reform/NoPlaceForKids/J_NoPlaceForKids_Full.pdf

Metz, A., Bartley, L., Ball, H., Wilson, D., Naoom, S., & Redmond, P. (2014). Active Implementation Frameworks (AIF) for successful service delivery: Catawba County child wellbeing project. *Research on Social Work Practice, 25*(4), 415-422.

Miller, J., & Maloney, C. (2013). Practitioner compliance with risk/needs assessment tools: A theoretical and empirical assessment. *Criminal Justice & Behavior, 40*(7), 716-736.

Murphy, A. and Turner, S., (2009). *Parole Violation Decision-Making Instrument (PVDMI) process evaluation*. Irvine, CA: Center for Evidence-Based Corrections.

Nagin, D. S. (2013). Deterrence: A review of the evidence by a criminologist for economists. *Annual Review of Economics, 5*, 83-105.

National Research Council. (2013). *Reforming juvenile justice: A developmental approach*. Committee on Assessing Juvenile Justice Reform, R. J. Bonnie, R. L. Johnson, B. M. Chemers, and J. A. Schuck, Eds.

Committee on Law and Justice, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Research Council. (2014). *Implementing juvenile justice reform: The federal role*. Committee on a Prioritized Plan to Implement a Developmental Approach in Juvenile Justice Reform, Committee on Law and Justice, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

OJJDP Statistical Briefing Book. Online. Available:

https://www.ojjdp.gov/ojstatbb/special_topics/qa11801.asp?qaDate=2017. Released on April 23, 2019.

O'Connell, D. J., Brent, J. J., & Visher, C. A. (2016). Decide Your Time: A randomized trial of a drug testing and graduated sanctions program for probationers. *Criminology & Public Policy*, *15*, 1073–1102.

Paternoster, R. (2010). How much do we really know about criminal deterrence? *Journal of Criminal Law & Criminology*, *100*(3), 765-824.

Piquero, A. R., Paternoster, R., Pogarsky, G., & Loughran, T. A. (2011). Elaborating the individual difference component of deterrence theory. *Annual Review of Law and Social Science*, *7*, 335–360.

Puzzanchera, C., & Hockenberry, S. (2019). *Trends and characteristics of youth in residential placement, 2017*. Pittsburgh, PA: National Center for Juvenile Justice.

Roman, J., & Harrell, A. (2001). Assessing the costs and benefits accruing to the public from a graduated sanctions program for drug-using defendants. *Law & Policy*, *23*(2), 237-268.

Rosay, A., & Everett, R. (2006). *Quantitative analysis of disparities in juvenile delinquency referrals*. Anchorage: Justice Center, University of Alaska.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*, 41–55.

Sedlak, A. J., & Bruce, C. (2010). *Youth's characteristics and backgrounds: Findings from the survey of youth in residential placement*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention. Retrieved from http://ncfy.acf.hhs.gov/sites/default/files/docs/19912-Youths_Characteristics_and_Backgrounds.pdf

Shearer, R. A. (2002). Probation strategies of juvenile and adult preservice trainees. *Federal Probation*, *66*, 33–37.

Sickmund, M., Sladky, T. J., Kang, W., & Puzzanchera, C. (2019). *Easy access to the census of juveniles in residential placement*. Available: <https://www.ojjdp.gov/ojstatbb/ezacirp/>

Stanger, C., Lansing, A. H., Budney, A. J. (2016). Contingency management approaches for adolescent substance use disorders. *Child and Adolescent Psychiatric Clinics of North America*, *25*(4), 645-659.

Steiner, B., Makarios, M. D., Travis, L. F., & Meade, B. (2012). Examining the effects of community-based sanctions on offender recidivism. *Justice Quarterly*, *29*(2), 229-257.

Steiner, B., Travis, L. F., & Makarios, M. D. (2011). Understanding parole officers' responses to sanctioning reform. *Crime & Delinquency*, 57(2), 222-246.


Steinhart, D. (2001). *Special detention cases: Strategies for handling difficult populations pathways to juvenile detention reform*. Baltimore, MD: Annie E. Casey Foundation.

Trotman, A. J., & Taxman, F. S. (2011). Implementation of a contingency management-based intervention in a community supervision setting: clinical issues and recommendations. *Journal of Offender Rehabilitation*, 50(5), 235-251.

Vera Institute of Justice. (2012). *Supervision response guidelines*. Online. Available: <https://www.appa-net.org/eWeb/Resources/SPSP/State-Response-Guide.pdf>

Wodahl, E. J., Garland, B., Culhane, S. E., & McCarthy, W. P. (2011). Utilizing behavioral interventions to improve supervision outcomes. *Criminal Justice and Behavior*, 38, 386-405.

Appendix A. AIM Sanctions and Incentives Grids

 MARYLAND Department of Juvenile Services				
County Graduated Responses Grid 2015				
Probation Supervision Level				
	Low	Moderate	High/VPI	
INFRACTION SEVERITY LEVEL	Minor	CMS Options: <ul style="list-style-type: none"> Verbal Warning Make up Missed Appointment Program/Service Interventions <ul style="list-style-type: none"> Anger Management (Walden) Substance Abuse Assessment-(Walden) Cove program (Walden) Ladies of Excellence program (SMCPS) Victim awareness (DIS) Basketball Leadership (SMCPS) Youth Empowerment (TCYSB) Mentoring program (IFCS) FFT (Center for Children) 	CMS Options: <ul style="list-style-type: none"> Verbal Warning Make up Missed Appointments Written Essay or Oral Report Assigned Reading or Video Viewing Community Service (up to 4 hours) Program/Service Interventions <ul style="list-style-type: none"> Mental Health Assessment (Walden) FFT (Center for Children) Drug Court Substance Abuse Assessment (Walden) Mentoring program (IFCS) Youth intervention specialist (TCYSB) Cove program (Walden) Ladies of Excellence program (SMCPS) Victim awareness (DIS) Basketball Leadership (SMCPS) Youth Empowerment (TCYSB) Evening Counseling Center (SMCPS) 	CMS Options: <ul style="list-style-type: none"> Verbal Warning Youth & Parent Meeting (1 bus. day) Make up Missed Appointments Written Essay or Oral Report Assigned Reading or Video Viewing Community Service (up to 10 hours) Restrict Activity (Short-Term) Modify Curfew (Extended Term) Program/Service Interventions <ul style="list-style-type: none"> Mental Health Assessment (Walden) FFT (Center for Children) Drug Court Substance Abuse Assessment (Walden) Mentoring program (IFCS) Youth intervention specialist (TCYSB) Cove program (Walden) Ladies of Excellence program (SMCPS) Victim awareness (DIS) Basketball Leadership (SMCPS) Youth Empowerment (TCYSB) Evening Counseling Center (SMCPS)
	Moderate	CMS Options: <ul style="list-style-type: none"> Verbal Warning Youth & Parent Meeting (3 bus. days) Community Service (up to 4 hours) Program/Service Interventions <ul style="list-style-type: none"> Youth intervention specialist (TCYSB) Anger Management (Walden) Substance Abuse Assessment-(Walden) Cove program (Walden) Ladies of Excellence program (SMCPS) Victim awareness (DIS) Basketball Leadership (SMCPS) Youth Empowerment (TCYSB) Mentoring program (IFCS) FFT (Center for Children) 	CMS Options: <ul style="list-style-type: none"> Verbal Warning Youth & Parent Meeting (3 bus. days) Make up Missed Appointments Written Essay or Oral Report Assigned Reading or Video Viewing Community Service (up to 8 hours) Restrict Activity (Short-Term) Modify Curfew (Short-Term) Increase Supervision Contacts Program/Service Interventions <ul style="list-style-type: none"> Mental Health Assessment (Walden) FFT (Center for Children) Drug Court Substance Abuse Assessment (Walden) Mentoring program (IFCS) Youth intervention specialist (TCYSB) Cove program (Walden) Ladies of Excellence program (SMCPS) Victim awareness (DIS) Basketball Leadership (SMCPS) Youth Empowerment (TCYSB) Evening Counseling Center (SMCPS) 	CMS Options: <ul style="list-style-type: none"> Administrative Meeting with Youth, Family, Supervisor (1 bus. day) Make up Missed Appointments Written Essay/Oral Report Assigned Reading or Video Viewing Community Service (up to 12 hours) Restrict Activity (Short-Term) Modify Curfew (Extended Term) Increase Supervision Contacts Increase drug screenings (if applicable) GPS for 30 Days (VPI Only) File VOP Program/Service Interventions <ul style="list-style-type: none"> Mental Health Assessment (Walden) FFT (Center for Children) Drug Court Substance Abuse Assessment (Walden) Mentoring program (IFCS) Youth intervention specialist (TCYSB) Cove program (Walden) Ladies of Excellence program (SMCPS) Victim awareness (DIS) Basketball Leadership (SMCPS) Youth Empowerment (TCYSB) Evening Counseling Center (SMCPS)
	Serious	CMS Options: <ul style="list-style-type: none"> Youth & Parent Meeting (1 bus. day) Community Service Hrs. (6 hours) Modify Curfew (Short-Term) Increase Supervision Level to Moderate Program/Service Interventions <ul style="list-style-type: none"> Anger Management (Walden) Substance Abuse Assessment-(Walden) Cove program (Walden) Ladies of Excellence program (SMCPS) Victim awareness (DIS) Basketball Leadership (SMCPS) Youth Empowerment (TCYSB) Mentoring program (IFCS) FFT (Center for Children) 	CMS Options: <ul style="list-style-type: none"> Administrative Meeting with Youth, Family, Supervisor (2 bus. days) Make up Missed Appointments Written Essay or Oral Report Assigned Reading or Video Viewing Community Service (up to 10 hours) Restrict Activity (Extended Term) Modify Curfew (Extended -Term) Increase Supervision Contacts Increase drug screenings (if applicable) Increase Supervision Level to High Program/Service Interventions <ul style="list-style-type: none"> Mental Health Assessment (Walden) FFT (Center for Children) Drug Court Substance Abuse Assessment (Walden) Mentoring program (IFCS) Youth intervention specialist (TCYSB) Cove program (Walden) Ladies of Excellence program (SMCPS) Victim awareness (DIS) Basketball Leadership (SMCPS) Youth Empowerment (TCYSB) Evening Counseling Center (SMCPS) 	CMS Options: <ul style="list-style-type: none"> Administrative Meeting with Youth, Family, Supervisor (1 bus. day) Make up Missed Appointments Written Essay/Oral Report Assigned Reading or Video Viewing Community Service (up to 14 hours) Restrict Activity (Short-Term) Modify Curfew (Extended Term) Increase Supervision Contacts Increase drug screenings (if applicable) GPS for 30 Days (VPI Only) File VOP Program/Service Interventions <ul style="list-style-type: none"> Mental Health Assessment (Walden) FFT (Center for Children) Drug Court Substance Abuse Assessment (Walden) Mentoring program (IFCS) Youth intervention specialist (TCYSB) Cove program (Walden) Ladies of Excellence program (SMCPS) Victim awareness (DIS) Basketball Leadership (SMCPS) Youth Empowerment (TCYSB) Evening Counseling Center (SMCPS)

DJS COMMUNITY SUPERVISION INCENTIVES FOR POSITIVE BEHAVIORS

		Achieve/Sustain Behavior 0-30 days	Achieve/Sustain Behavior 31 – 90 days (1-3 mos.)	Achieve/Sustain Behavior 91-180 days (4-6 mos.)	Achieve/Sustain Behavior 181 to Indefinite (>6 months)
TYPES OF POSITIVE BEHAVIORS	COMPLIANCE	<ul style="list-style-type: none"> - Enrolls in an academic or vocational Program - Improves school attendance - Attends court-ordered Cognitive Programs/treatment services - Keeps office appointments with reminders - Abides by curfew - Improved behavior in school - Improved grades in school - Abides by set family/house rules - Locates and/or begins court-ordered community service - Begins paying restitution - No incidents of whereabouts unknown - Notifies CMS of change in contact information - No suspensions/expulsions 	<ul style="list-style-type: none"> - Enrolls in an academic or vocational program - Improves school attendance - Attends and/or completes court-ordered Cognitive programs/treatment services - Keeps office appointments without reminders - Abides by curfew - Improved behavior in school - Improved grades in school - Abides by set family/house rules - Clean Urine Screen/Drug Free - Completes at least 50% of court-ordered community service - Makes consistent payments and/or pays off restitution - No incidents of whereabouts unknown - Notifies CMS of change in contact information - No suspensions/expulsions 	<ul style="list-style-type: none"> - Enrolls in an academic or vocational program - Attends school and all classes as scheduled - Attends and/or completes court-ordered Cognitive programs/treatment services - Keeps office appointments without reminders - Abides by curfew - Improved behavior in school - Maintains grade improvements - Abides by set family/house rules - Clean Urine Screen/Drug Free - Completes at least 75% of court-ordered community service - Makes consistent payments and/or pays off restitution - No incidents of whereabouts unknown - Notifies CMS of change in contact information - No suspensions/expulsions 	<ul style="list-style-type: none"> - Enrolls in an academic or vocational program - Attends school and all classes as scheduled - Attends and/or completes court-ordered cognitive programs/treatment services - Keeps office appointments without Reminders - Abides by curfew - Improved behavior in school - Maintains grade improvements - Abides by set family/house rules - Clean Urine Screen/Drug Free - Completes 100% of court-ordered community service - Makes consistent payments and/or pays off restitution - No incidents of whereabouts unknown - Notifies CMS of change in contact information - No suspensions/expulsions
	PROBATION INVOLVEMENT	<ul style="list-style-type: none"> - Obtains/Retains Employment - Receives promotion in grade level or employment - Joined/Participates in at least one extra-curricular activity (school/community) - Participates in additional services - Detaches from negative peer groups - Establishes at least one personal goal 	<ul style="list-style-type: none"> - Obtains/Retains Employment - Receives promotion in grade level or employment - Joined/Participates in at least one extra-curricular activity (school or community) - Participates in additional services - Remains detached from negative peer groups - Makes progress toward or completes at least one personal goal 	<ul style="list-style-type: none"> - Obtains/Retains Employment - Receives promotion in grade level or employment - Joined/Participates in at least one extra-curricular activity (school or community) - Participates in additional services - Remains detached from negative peer groups - Makes progress toward or completes at least one personal goal 	<ul style="list-style-type: none"> - Obtains/Retains Employment - Receives promotion in grade level or employment - Joined/Participates in at least one extra-curricular activity (school or community) - Participates in additional services - Remains detached from negative peer groups - Completes at least one personal goal
	SELF-REGULATORY	<ul style="list-style-type: none"> - Acknowledge and/or articulate needs - Seeks help - Manages time well for at least one important task - Handles a difficult situation well - Takes initiative to meet at least one probation condition 	<ul style="list-style-type: none"> - Begins to acknowledge and/or articulate Needs - Seeks help - Manages time well for at least one important task - Handles a difficult situation well - Takes initiative to meet at least one probation condition 	<ul style="list-style-type: none"> - Begins to acknowledge and/or articulate needs - Seeks help - Manages time well for at least one important task - Handles a difficult situation well - Takes initiative to meet at least one probation condition 	<ul style="list-style-type: none"> - Begins to acknowledge and/or articulate needs - Seeks help - Manages time well for at least one important task - Handles a difficult situation well - Takes initiative to meet at least one probation condition
AVAILABLE INCENTIVES					
RANGE OF INCENTIVES	MINOR ACHIEVEMENTS	<ul style="list-style-type: none"> - Verbal praise to youth - Verbal praise to parent - Positive letter home from case manager - Positive letter home from Supervisor or Director - Publicly display work - Food/candy/treat - Certificate of achievement - Invitation to monthly/annual Recognition Ceremony 	<ul style="list-style-type: none"> - Verbal praise to youth - Verbal praise to parent - Positive letter home from case manager - Positive letter home from Supervisor or Director - Publicly display work - Food/candy/treat - Certificate of achievement - Invitation to monthly/annual Recognition Ceremony - School supplies 	<ul style="list-style-type: none"> - Verbal praise to youth - Verbal praise to parent - Positive letter home from case manager - Positive letter home from Supervisor or Director - Publicly display work - Food/candy/treat - Certificate of achievement - Invitation to monthly/annual Recognition Ceremony - School supplies 	<ul style="list-style-type: none"> - Verbal praise to youth - Verbal praise to parent - Positive letter home from case manager - Positive letter home from Supervisor or Director - Publicly display work - Food/candy/treat - Certificate of achievement - Invitation to monthly/annual Recognition Ceremony - School supplies
	MODERATE ACHIEVEMENTS	<ul style="list-style-type: none"> - School supplies - Books/Magazines - Personal hygiene supplies 	<ul style="list-style-type: none"> - School supplies - Books/Magazines - Personal hygiene supplies - Magazine subscription - Extend curfew - Tour of local college/university - Tour of other local business of interest - Tour of stadium or athletic facility - Meal coupons/meal for youth's family - Allow previously restricted activity - Court recognition 	<ul style="list-style-type: none"> - Books/Magazines - Personal hygiene supplies - Magazine subscription - Extend curfew - Tour of local college/university - Tour of other local business of interest - Tour of stadium or athletic facility - Meal coupons/meal for youth's family - Allow previously restricted activity - Court recognition - Job Shadowing/Apprenticeship Opportunity 	<ul style="list-style-type: none"> - Books/Magazines - Personal hygiene supplies - Magazine subscription - Extend curfew - Tour of local college/university - Tour of other local business of interest - Tour of stadium or athletic facility - Meal coupons/meal for youth's family - Allow previously restricted activity - Court recognition - Job shadowing/apprenticeship opportunity - Reduced level of supervision - Restoration of driving privileges - Request reduction in community service - College/university gear
	MAJOR ACHIEVEMENTS	<ul style="list-style-type: none"> - Meal coupons/meal for youth's family 	<ul style="list-style-type: none"> - Recreation center passes/classes - Tickets to sporting events - Tickets to other activities (amusement park, museums, etc.) - Movie tickets - Request removal of Electronic Monitoring - Media card (iTunes, Play Store) - Field trips with staff (e.g., hiking, rafting) 	<ul style="list-style-type: none"> - Recreation center passes/classes - Tickets to sporting events - Tickets to other activities (amusement park, museums, etc.) - Movie tickets - Request removal electronic monitoring - Media card (iTunes, Play Store) - Field trips with staff (e.g., hiking, rafting) - Reduced level of supervision - Restoration of driving privileges - Request reduction in community service hours - College/university gear - Membership in athletic/art/other classes - Petition for termination of case - Invitation for youth to serve on agency leadership council/serve in some leadership role - Manicure/pedicure or other self-care activity 	<ul style="list-style-type: none"> - Recreation center passes/classes - Tickets to sporting events - Tickets to other activities (amusement park, museums, etc.) - Movie tickets - Request removal of electronic monitoring - Media card (iTunes, Play Store) - Field trips with staff (e.g., hiking, rafting) - Reduced level of supervision - Restoration of driving privileges - Request reduction in community service hours - Membership in athletic/art/other classes - Petition for termination of case - Invitation for youth to serve on agency leadership council/serve in some leadership role - Manicure/pedicure or other self-care activity - Meal/meeting with elected official or other well-known individual - Lunch with case manager/judge

Appendix B. Additional Analyses

Table 1. Positive Behaviors Identified During Community Supervision (N=1,452 behaviors)

Behavior Category	% of Total	Behavior	% of Total
Compliance	76.5	Attends and/or completes court-ordered cognitive programs/ treatment services	23.9
		Clean urine screen/drug free	6.3
		Abides by set family/house rules	5.9
		Keeps office appointments with reminders	5.8
		Abides by curfew	4.9
		Improves school attendance	4.7
		Improved grades in school	4.5
		No incidents of whereabouts unknown	3.6
		Keeps office appointments without reminders	3.1
		Improved behavior in school	2.7
		Locates and/or begins court-ordered community service	1.9
		Attends school and all classes as scheduled	1.7
		Makes consistent payments and/or pays off restitution	1.6
		Completes at least 50% of court-ordered community service	1.4
		Enrolls in an academic or vocational program	1.4
		Completes at least 75% of court-ordered community service	0.8
		Completes 100% of court-ordered community service	0.6
		Notifies CMS of change in contact information	0.6
		No suspensions/expulsions	0.5
		Maintains grade improvements	0.4
Begins paying restitution	0.3		
Pro-Social Involvement Behaviors	16.2	Obtains/retains employment	7.3
		Joined/Participates in at least one extra-curricular activity (school/community)	2.3
		Receives promotion in grade level or employment	2.3
		Participates in additional services	1.2
		Makes progress toward or completes at least one personal goal	1.0
		Detaches from negative peer groups	0.8
		Remains detached from negative peer groups	0.7
		Establishes at least one personal goal	0.6
Self-Advocacy Behaviors	7.3	Completes at least one personal goal	0.2
		Takes initiative to meet at least one probation condition	2.6
		Handles a difficult situation well	1.6
		Manages time well for at least one important task	1.0
		Acknowledge and/or articulate needs	0.9
Seeks help	0.9		
Begins to acknowledge and/or articulate Needs	0.3		

Table 2. Differences in Youth With and Without AIM Responses (N=1,983)

Variable		% With No Responses	% With Any Responses	Pearson Chi-Square	Sig.
<i>Age at Start of Supervision (years)</i>	13 and Under	48.6	51.4	8.35	.015
	14-16	41.6	58.4		
	17 and Over	48.0	52.0		
<i>Gender</i>	Male	44.6	55.4	0.02	.880
	Female	45.0	55.0		
<i>Race/Ethnicity</i>	African American/Black	44.3	55.7	0.33	.849
	Caucasian/White	45.5	54.5		
	Other/Unknown	46.0	54.0		
<i>Jurisdiction Type</i>	Urban	51.2	48.8	37.06	.000
	Suburban	49.7	50.3		
	Rural/Large Town	36.5	63.5		
<i>Supervision Type</i>	Probation	45.0	55.0	0.81	.368
	Aftercare	41.5	58.5		
<i>Most Serious Adjudicated Offense Type</i>	Person	44.6	55.4	8.64	.035
	Property	47.0	53.0		
	Drug	52.1	47.9		
	Other	38.8	61.2		
<i>Most Serious Adjudicated Offense Level</i>	Felony	46.0	54.0	0.48	.785
	Misdemeanor	44.2	55.8		
	Other	44.9	55.1		
<i>Risk Level</i>	High	39.5	60.5	30.08	.000
	Moderate	36.6	63.4		
	Low	49.8	50.2		
<i>School Need</i>	High	38.7	61.3	24.64	.000
	Moderate	42.3	57.7		
	Low	51.2	48.8		
<i>Peers/ Relationships Need</i>	High	37.8	62.2	23.32	.000
	Moderate	45.4	54.6		
	Low	54.2	45.8		
<i>Family Need</i>	High	38.8	61.3	23.60	.000
	Moderate	38.2	61.8		
	Low	49.3	50.7		
<i>Alcohol & Drugs Need</i>	High	40.0	60.0	17.21	.000
	Moderate	37.6	62.4		
	Low	48.4	51.6		
<i>Mental Health Need</i>	High	41.1	58.9	9.22	.010
	Moderate	39.6	60.4		
	Low	47.4	52.6		
<i>Anti-Social Attitudes Need</i>	High	38.4	61.6	21.64	.000
	Moderate	48.4	51.6		
	Low	49.2	50.8		
<i>Aggression Need</i>	High	41.6	58.4	6.23	.044
	Moderate	44.9	55.1		
	Low	48.1	51.9		

Table 3. Differences in Response Ratio per Youth with Responses (N=1,097)

Variable		% Incentives > Sanctions	% Incentives = Sanctions	% Sanctions > Incentives	Pearson Chi-Square	Sig.
<i>Age at Start of Supervision (years)</i>	13 and Under	40.7	3.3	56.0	8.93	.063
	14-16	28.5	5.2	66.2		
	17 and Over	33.3	6.8	59.8		
<i>Gender</i>	Male	27.1	8.1	64.8	4.27	.118
	Female	32.2	5.1	62.7		
<i>Race/Ethnicity</i>	African American/Black	28.2	4.7	67.1	21.27	.000
	Caucasian/White	39.7	8.6	51.7		
	Other/Unknown	32.4	4.4	63.2		
<i>Jurisdiction Type</i>	Urban	8.8	4.6	86.6	75.10	.000
	Suburban	32.3	5.4	62.3		
	Rural/Large Town	40.1	6.3	53.6		
<i>Supervision Type</i>	Probation	29.0	5.6	65.4	0.30	.859
	Aftercare	31.5	5.7	62.8		
<i>Most Serious Adjudicated Offense Type</i>	Person	30.3	6.2	63.5	15.00	.020
	Property	25.8	5.7	68.5		
	Drug	36.8	0.0	63.2		
	Other	39.4	5.6	54.9		
<i>Most Serious Adjudicated Offense Level</i>	Felony	31.3	4.0	64.7	7.38	.117
	Misdemeanor	30.2	5.9	63.9		
	Other	40.7	8.6	50.6		
<i>Risk Level</i>	High	15.1	3.8	81.1	66.21	.000
	Moderate	21.9	7.5	70.6		
	Low	41.5	5.0	53.5		
<i>School Need</i>	High	18.8	4.6	76.6	71.19	.000
	Moderate	30.3	6.6	63.1		
	Low	45.2	6.1	48.7		
<i>Peers/ Relationships Need</i>	High	19.0	5.9	75.1	60.47	.000
	Moderate	33.1	6.1	60.8		
	Low	52.9	3.3	43.8		
<i>Family Need</i>	High	18.4	4.8	76.9	40.56	.000
	Moderate	24.2	4.9	70.9		
	Low	39.0	6.4	54.6		
<i>Alcohol & Drugs Need</i>	High	20.3	7.8	72.0	37.36	.000
	Moderate	24.7	7.0	68.3		
	Low	38.5	4.2	57.2		
<i>Mental Health Need</i>	High	26.5	6.4	67.0	7.59	.108
	Moderate	28.7	7.7	63.6		
	Low	34.0	4.7	61.3		
<i>Anti-Social Attitudes Need</i>	High	22.5	5.1	72.4	45.44	.000
	Moderate	30.3	5.7	64.0		
	Low	42.7	6.4	50.9		
<i>Aggression Need</i>	High	24.0	5.2	70.7	38.73	.000
	Moderate	28.9	4.5	66.6		
	Low	42.6	7.1	50.3		

Table 4. Differences in Time to Response, Days: Response Level (N=5,101)

Variable		Mean	F	Sig.
<i>Age at Start of Supervision (years)</i>	13 and Under	2.16	2.12	.120
	14-16	2.89		
	17 and Over	2.76		
<i>Gender</i>	Male	2.63	0.73	.393
	Female	2.83		
<i>Race/Ethnicity</i>	African American/Black	3.05	9.99	.000
	Caucasian/White	2.06		
	Other/Unknown	2.75		
<i>Jurisdiction Type</i>	Urban	3.54	16.73	.000
	Suburban	2.10		
	Rural/Large Town	2.95		
<i>Supervision Type</i>	Probation	2.82	1.17	.279
	Aftercare	2.47		
<i>Most Serious Adjudicated Offense Type</i>	Person	2.73	0.41	.747
	Property	2.78		
	Drug	2.68		
	Other	3.01		
<i>Most Serious Adjudicated Offense Level</i>	Felony	2.68	0.37	.690
	Misdemeanor	2.84		
	Other	2.61		
<i>Risk Level</i>	High	2.52	2.56	.078
	Moderate	2.64		
	Low	3.02		
<i>School Need</i>	High	2.62	3.30	.037
	Moderate	3.20		
	Low	2.69		
<i>Peers/ Relationships Need</i>	High	2.62	1.10	.332
	Moderate	2.92		
	Low	2.83		
<i>Family Need</i>	High	2.57	1.24	.291
	Moderate	2.70		
	Low	2.95		
<i>Alcohol & Drugs Need</i>	High	2.45	4.78	.008
	Moderate	2.56		
	Low	3.07		
<i>Mental Health Need</i>	High	2.71	1.37	.255
	Moderate	2.53		
	Low	2.92		
<i>Anti-Social Attitudes Need</i>	High	2.88	1.31	.270
	Moderate	2.47		
	Low	2.81		
<i>Aggression Need</i>	High	2.90	2.57	.077
	Moderate	2.97		
	Low	2.44		

Table 5. Differences in Time to Response, Days: Youth Level (N=1,097)

Variable		Mean	F	Sig.
<i>Age at Start of Supervision (years)</i>	13 and Under	2.85	0.40	.671
	14-16	3.42		
	17 and Over	3.06		
<i>Gender</i>	Male	3.14	0.94	.333
	Female	3.70		
<i>Race/Ethnicity</i>	African American/Black	3.56	2.17	.115
	Caucasian/White	2.47		
	Other/Unknown	2.79		
<i>Jurisdiction Type</i>	Urban	4.58	4.29	.014
	Suburban	2.82		
	Rural/Large Town	2.98		
<i>Supervision Type</i>	Probation	3.29	0.32	.572
	Aftercare	2.85		
<i>Most Serious Adjudicated Offense Type</i>	Person	3.34	0.49	.693
	Property	3.11		
	Drug	2.08		
	Other	3.21		
<i>Most Serious Adjudicated Offense Level</i>	Felony	2.70	0.96	.382
	Misdemeanor	3.36		
	Other	3.80		
<i>Risk Level</i>	High	3.53	0.63	.534
	Moderate	2.88		
	Low	3.39		
<i>School Need</i>	High	2.89	1.19	.305
	Moderate	3.79		
	Low	3.21		
<i>Peers/ Relationships Need</i>	High	3.26	0.13	.876
	Moderate	3.31		
	Low	2.95		
<i>Family Need</i>	High	3.44	0.11	.898
	Moderate	3.11		
	Low	3.28		
<i>Alcohol & Drugs Need</i>	High	3.07	0.50	.609
	Moderate	2.88		
	Low	3.44		
<i>Mental Health Need</i>	High	3.38	0.39	.679
	Moderate	3.59		
	Low	3.08		
<i>Anti-Social Attitudes Need</i>	High	3.35	1.58	.206
	Moderate	2.43		
	Low	3.55		
<i>Aggression Need</i>	High	3.34	0.08	.922
	Moderate	3.12		
	Low	3.21		

Table 6. Differences in Outcomes Using Alternative Propensity Score Matching and Model Specifications

	Frequency/Mean		Treatment Effect in Regression ¹ without Covariates		Treatment Effect in Regression ² with Covariates ³	
	Pre-AIM	AIM	HR	Sig.	HR	Sig.
VOP Filed with Court						
Unmatched	20.1	16.9	0.81	.003	0.76	.000
1:1 No replacement, .03 caliper	20.9	16.0	0.73	.000	0.72	.000
1:1 No replacement, .001 caliper	20.2	16.2	0.77	.002	0.75	.001
1:1 With replacement, .03 caliper	22.8	16.9	0.70	.000	0.68	.000
1:1 With replacement, .001 caliper	22.5	16.6	0.70	.000	0.67	.000
2:1 With replacement, .03 caliper	22.5	16.9	0.71	.000	0.70	.000
Inverse probability weighting	20.7	16.3	0.75	.000	0.75	.000
VOP Commitment						
Unmatched	10.2	9.5	0.91	.308	0.78	.018
1:1 No replacement, .03 caliper	10.7	8.8	0.79	.031	0.80	.053
1:1 No replacement, .001 caliper	10.1	9.4	0.91	.405	0.89	.335
1:1 With replacement, .03 caliper	11.7	9.4	0.79	.014	0.75	.005
1:1 With replacement, .001 caliper	11.4	9.4	0.80	.027	0.75	.007
2:1 With replacement, .03 caliper	11.6	9.4	0.80	.021	0.79	.017
Inverse probability weighting	10.9	8.7	0.78	.014	0.78	.015
Detention						
Unmatched	28.6	28.3	0.97	.630	0.95	.404
1:1 No replacement, .03 caliper	29.6	27.2	0.89	.086	0.91	.182
1:1 No replacement, .001 caliper	29.2	27.2	0.91	.167	0.91	.187
1:1 With replacement, .03 caliper	31.9	28.3	0.85	.007	0.85	.008
1:1 With replacement, .001 caliper	31.8	28.1	0.85	.007	0.84	.004
2:1 With replacement, .03 caliper	30.9	28.3	0.89	.057	0.91	.093
Inverse probability weighting	29.4	27.7	0.93	.224	0.95	.450
Committed Residential Placement						
Unmatched	5.7	4.6	0.74	.035	0.68	.013
1:1 No replacement, .03 caliper	5.7	4.5	0.72	.041	0.68	.024
1:1 No replacement, .001 caliper	5.7	4.2	0.65	.014	0.63	.009
1:1 With replacement, .03 caliper	6.8	4.6	0.62	.001	0.59	.000
1:1 With replacement, .001 caliper	6.9	4.4	0.58	.000	0.55	.000
2:1 With replacement, .03 caliper	6.1	4.6	0.70	.017	0.68	.008
Inverse probability weighting	5.7	4.6	0.75	.055	0.70	.025
Referral to DJS/Adult Arrest						
Unmatched	31.5	33.4	1.05	.377	1.02	.782
1:1 No replacement, .03 caliper	32.0	32.5	0.99	.898	1.01	.888
1:1 No replacement, .001 caliper	31.9	32.5	1.00	.945	1.01	.845
1:1 With replacement, .03 caliper	32.6	33.4	1.00	.992	1.01	.875
1:1 With replacement, .001 caliper	32.8	33.2	0.99	.824	0.99	.853
2:1 With replacement, .03 caliper	32.7	33.4	1.00	.953	1.03	.630
Inverse probability weighting	31.8	32.4	1.00	.935	1.02	.705
Adjudication/Adult Conviction						
Unmatched	13.3	12.2	0.88	.142	0.84	.048
1:1 No replacement, .03 caliper	13.7	11.5	0.80	.023	0.82	.050
1:1 No replacement, .001 caliper	13.8	11.5	0.81	.032	0.82	.046
1:1 With replacement, .03 caliper	14.9	12.1	0.78	.004	0.79	.006
1:1 With replacement, .001 caliper	15.0	12.0	0.77	.003	0.77	.004

Table 6. Differences in Outcomes Using Alternative Propensity Score Matching and Model Specifications

	Frequency/Mean		Treatment Effect in Regression ¹ without Covariates		Treatment Effect in Regression ² with Covariates ³	
	Pre-AIM	AIM	HR	Sig.	HR	Sig.
2:1 With replacement, .03 caliper	14.3	12.1	0.82	.023	0.83	.041
Inverse probability weighting	13.7	11.6	0.82	.029	0.84	.060
Commitment/Adult Incarceration						
Unmatched	8.4	8.7	1.00	.982	0.99	.950
1:1 No replacement, .03 caliper	8.5	8.5	0.96	.754	1.02	.880
1:1 No replacement, .001 caliper	8.6	8.7	0.98	.883	1.01	.941
1:1 With replacement, .03 caliper	10.2	8.7	0.81	.047	0.83	.086
1:1 With replacement, .001 caliper	10.3	8.7	0.80	.035	0.82	.063
2:1 With replacement, .03 caliper	9.4	8.7	0.90	.313	0.93	.517
Inverse probability weighting	8.6	8.4	0.95	.607	0.99	.951
	Pre-AIM	AIM	OR	Sig.	OR	Sig.
Successful Supervision						
Unmatched	49.0	47.0	0.92	.202	0.98	.800
1:1 No replacement, .03 caliper	48.7	48.0	0.97	.704	0.95	.532
1:1 No replacement, .001 caliper	49.1	48.0	0.96	.523	0.96	.581
1:1 With replacement, .03 caliper	46.9	47.0	1.00	.949	1.02	.773
1:1 With replacement, .001 caliper	47.1	47.1	1.00	1.000	1.02	.759
2:1 With replacement, .03 caliper	46.8	47.0	1.01	.903	0.99	.936
Inverse probability weighting	48.4	48.4	1.00	.988	1.00	.992
	Pre-AIM	AIM	Coeff.	Sig.	Coeff.	Sig.
Days in Detention						
Unmatched	10.56	12.93	2.36	.022	1.58	.131
1:1 No replacement, .03 caliper	11.21	12.61	1.40	.239	1.48	.199
1:1 No replacement, .001 caliper	11.23	12.34	1.11	.367	0.90	.445
1:1 With replacement, .03 caliper	12.66	12.91	0.26	.823	-0.07	.948
1:1 With replacement, .001 caliper	12.59	12.60	0.01	.995	-0.43	.694
2:1 With replacement, .03 caliper	11.36	12.91	1.56	.203	1.48	.215
Inverse probability weighting	11.06	12.69	1.63	.152	1.67	.128
Days under Supervision						
Unmatched	262.95	271.56	8.61	.039	6.72	.118
1:1 No replacement, .03 caliper	262.54	272.05	9.51	.043	9.09	.050
1:1 No replacement, .001 caliper	263.10	269.43	6.32	.189	5.64	.237
1:1 With replacement, .03 caliper	262.76	271.60	8.84	.041	8.39	.050
1:1 With replacement, .001 caliper	263.42	271.12	7.70	.076	7.37	.087
2:1 With replacement, .03 caliper	265.18	271.60	6.42	.215	6.76	.185
Inverse probability weighting	264.03	270.49	6.45	.139	6.54	.130

Note: The matching strategy/models in bold font were selected as the primary evaluation models for interpretation.

¹ Cox regression models were conducted for dichotomous outcomes with varying times at risk; logistic regression models were conducted for dichotomous outcomes measured as of discharge; linear regression models were conducted for continuous outcomes. HR = Hazard Ratio. OR = Odds Ratio. Coeff. = Coefficient.

² Covariates included: age at the start of supervision; gender (male/female); race/ethnicity (black, white, other/unknown); jurisdiction type (urban, suburban, rural/large town); supervision type (probation, aftercare); most serious adjudicated offense level (felony, misdemeanor, other); most serious adjudicated offense type (person, property, drug, other); risk level (low, moderate, high); and school, peer relationships, family, alcohol and drug, mental health, and aggression needs (low, moderate, high).