DISCUSSION PAPER SERIES

IZA

Causal Effects of Mental Health Treatment on Education Outcomes for Youth in the Justice System

Alison Cuellar
Dhaval M. Dave

Causal Effects of Mental Health Treatment on Education Outcomes for Youth in the Justice System

Alison Cuellar
*George Mason University
and NBER*

Dhaval M. Dave
*Bentley University,
NBER and IZA*

ABSTRACT

Causal Effects of Mental Health Treatment on Education Outcomes for Youth in the Justice System[*]

This study assesses whether mental health interventions can improve academic outcomes for justice-involved youth. Only a limited number of studies have linked justice policies to outcomes beyond crime, particularly education, which carries large monetary and non-monetary benefits. The current study relies on detailed administrative data and unique policy rules under which youth are assigned to behavioral treatment programs. The administrative data allow for a rich set of controls for observed family- and youth-specific heterogeneity. In addition, the treatment assignment rules create discontinuous thresholds among youth who are deemed eligible or not eligible for treatment, rules which the study exploits empirically to address the non-random selection bias in estimating plausibly causal effects of treatment eligibility and treatment receipt. Estimates indicate that certain types of intensive mental health intervention can lower dropout and increase high-school completion for justice-involved youth. Effects on grades are negative or not significant, possibly due to the greater retention of less academically-skilled students. We also assess heterogeneity in the treatment effects, and find that the effects on dropout tend to be greater among youth believed to be less academically engaged prior to treatment. The positive effects on education, and the earnings returns alone, are large enough to produce benefits in excess of the cost for the more intensive of the mental health interventions.

JEL Classification: I1, I21, H75, K42

Keywords: mental health, treatment, juvenile justice, crime, education, high school, grades, GED, dropout

Corresponding author:

Dhaval Dave
Department of Economics
Bentley University
175 Forest Street
Waltham, MA 02452
USA
E-mail: ddave@bentley.edu

I. Introduction

In 2010 juveniles accounted for 1.6 million non-traffic related arrests, ranging from curfew violations to violent crimes (Sickmund and Puzzanchera, 2014). Youth in the juvenile justice system typically are at least 10 years old and most are age 16 or 17. They have, by definition, exhibited antisocial behaviors that lead them to the attention of the justice system. A variety of adolescent intervention programs have been created which target criminal and antisocial behaviors (Greenwood, 2008; Cuellar 2015). These programs stand in contrast to broad, primary prevention approaches which target young children often before significant problems emerge. In part, the debate surrounding the most effective policies centers on whether adolescence is too late to improve the trajectory of antisocial behaviors or academic achievement (Cook, Dodge, Farkas, et al., 2014).

For youth with relatively significant mental health and behavior problems, who may have a lengthy history of serious behavior and family troubles, particular treatment programs have been developed called multi-systemic therapy (MST) and functional family therapy (FFT). These intervention programs are relatively intensive and costly, as much as $7,000 per treated youth. Yet they have the potential to reduce social costs across a variety of dimensions, not only those related to violent and property crimes, but also those associated with substance abuse, early teen births, and low educational outcomes.

Although treatment programs such as MST and FFT can reduce criminal recidivism, their high cost has impeded widespread availability in the justice system, even to youth who would meet the relatively stringent eligibility criteria based on individual mental health and social factors. One possible reason is that the full extent of societal benefits from these programs has not been explored. Broader and durable societal benefits could be achieved if these treatment

programs affected outcomes such as high school dropout and completion rates. Such improvements would have long-lasting effects on employment, income and participation in social programs and, thus, modify any cost-benefit calculus of the interventions.

Multi-systemic Therapy (MST), Functional Family Therapy (FFT) and to a lesser extent Aggression Replacement Training (ART) are among the most studied treatment programs for youth in the justice system. With these programs justice agencies seek to reduce recidivism and potentially criminal behavior that could persist into adulthood. Thus, the focus of research on MST and FFT has been largely and expectedly on criminal outcomes. Indeed, there is some evidence discussed below that these programs achieve reductions in recidivism (Littell, et al, 2005, Lucenko, et al 2011). However, less well studied is their impact on academic success.

This study assesses the short and medium-run academic benefits of mental health treatment among a population of youth in the juvenile justice system, using rich administrative data from Washington State and unique program rules under which youth are assigned to these treatment interventions. We have extensive data on each youth allowing us to control for observed differences in family and youth characteristics. In addition, the treatment assignment rules in Washington provide plausibly exogenous variation among sub-populations of youth who are deemed eligible or not eligible for treatment, rules which we exploit empirically to address the non-random selection bias in estimating causal effects of treatment eligibility. We are then able to deploy our rich data and compare secondary education outcomes for these high-risk youth. Among education outcomes we are able to observe secondary school completion, transfers, dropouts, suspensions, and average grades, allowing us to assess the impact of mental health treatment across a range of important measures of human capital acquisition.

We find consistent evidence that FFT and MST, the more intensive mental health treatment programs, have broader spillover benefits among high-risk youth. The intent-to-treat and the treatment-on-the-treated effects point to lower dropout rates and higher secondary school completion. Assessment of heterogeneity in these treatment effects further confirms that these programs improve educational outcomes even among the higher-risk individuals – those who had weaker school engagement and had a history of mental health problems prior to their treatment screening.

The rest of the study is laid out as follows. Section 2 reviews the previous literature regarding mental health treatment for youth in the justice system and experiments to improve their educational outcomes. Section 3 discusses our data and Section 4 outlines our empirical approach in exploiting the natural experiment afforded by Washington's eligibility guidelines for treatment diversion. The results are discussed in Section 5, followed by some policy implications in the concluding section.

2. Background

Measuring the causal link between mental health treatment and education is a challenging task, as a result of which many researchers have resorted to small-scale randomized trials of treatments. We review relevant strands from two literatures, the first on several treatments that have focused on youth in the justice system. In addition, we include studies that have focused on educational outcomes among high-risk adolescents where behavior problems are also prevalent.

Aggression Replacement Training, Functional Family Therapy and Multi-Systemic Therapy

Aggression Replacement Training (ART) is the least intensive of the three interventions examined here. It relies on a group therapy format and lasts only 10 weeks (Washington State

Institute for Public Policy, 2004a).[1] ART teaches youths how to control their angry impulses and take perspectives other than their own through moral reasoning with the goal of reducing aggressive behavior and violence (Office of Juvenile Justice and Delinquency Prevention). There is no explicit academic component or educational goal, although improved peer and teacher relationships in school could conceivably lead to greater school success. ART also has less empirical support than either MST or FFT. One ART study among justice-involved youth used a waiting list comparison and found a 16% lower rate of felony recidivism among those treated with ART. A second study of 65 youth examined ART in school settings and found that parents and teachers reported improved social skills, including such dimensions as self-control and cooperation, as well as improvements in problem behaviors (Gundersen and Svartdal, 2006).

Another intervention for youth with behavioral problems, conduct disorder, delinquency and substance abuse is FFT, which although multi-faceted it is not as intensive as MST (Littell, et al 2007). The Blueprints for Violence Prevention estimate for the year-one cost of FFT is $1,679,000 or $2,800 per youth or family served.[2] Any public mental health system or private clinic seeking to newly deliver FFT services would contact the national coordinating entity, FFT LLC in Seattle, for technical assistance and training.[3] FFT includes up to 30 hours of family-based intervention and is typically delivered by teams of 3-8 therapists with relatively low caseloads of 10-12 families. The FFT intervention is highly structured and typically lasts 3 months. It is designed to prevent further delinquency and violence by promoting communication and enhancing support within the family. Families are trained to modify communication that

---

[1] The cost to implement Aggression Replacement Training in Washington State was estimated at $745 per youth. See https://www.crimesolutions.gov/ProgramDetails.aspx?ID=254.
[2] These estimates program descriptions were accessed at http://www.blueprintsprograms.com/programCosts.php?pid=0a57cb53ba59c46fc4b692527a38a87c78d84028
[3] Additional information on FFT implementation is available at http://fftllc.com

contributes to delinquency and dysfunction, to negotiate more effectively, and to set clear rules and responsibilities. While longer term goals include lower reduced recidivism, substance abuse and sibling criminal involvement, school outcomes are not explicitly part of the formal intervention model[4] although one might posit that greater problem-solving and communication skills could lead to improvements in at least some education outcomes. While earlier trials of FFT found reductions in reoffending, a more recent trial found no differences between FFT and the control group with respect to reoffending (Sexton and Turner, 2010).

Separately, a natural experiment-based study of FFT was conducted in Washington State which examined a specific sub-population of justice-involved youth, those who had been incarcerated (Lucenko, et al 2011). A policy change occurred under which youth were released from commitment facilities and received FFT in one year, but in the following year youth who were released no longer received FFT due to budget cuts. Both groups were observed for 12 months after release from incarceration. The study used propensity score matching to compare incarcerated youth before and after the program cut and finds that arrest rates declined, while employment rates increased. The change in arrest and employment is attributed to FFT. However, the study is not able to control for secular trends in youth crime or employment. Potentially, a recession that led to both the JRA programmatic cut and lower availability of jobs could confound the employment results while cuts in police or probation workforces could have led to less crime detection over time and confounded the arrest results. Education outcomes were not measured.

Among the most intensive and costly programs for justice-involved youth is Multi-Systemic Therapy. The Blueprints estimate for the year-one cost of two MST teams serving 132

---

[4] A logic model is provided by the Blueprints for Violence Prevention group at
http://www.blueprintsprograms.com/factSheet.php?pid=0a57cb53ba59c46fc4b692527a38a87c78d84028

families is $933,000 or $7,068 per family.[5] MST is delivered by trained masters-level therapists who have caseloads of 4 to 6 families and are available to families 24 hours a day, 7 days per week. MST is not delivered by justice system probation staff, but rather by MST-trained clinical staff. MST implementation and replication is coordinated through a national entity, MST Services in South Carolina.[6] The MST model combines family-based and individual cognitive-behavioral therapies with a range of individualized support services. MST therapists intervene with the family, peers, schools, and other treatment agencies. Treatment involves an initial period of intensive, potentially daily home visits, which are tapered over 3 to 5 months to weekly visits. Therapists work with youth and their families to address parent-focused behavior management and communication skills training as well as collaborative links between families and school (Henggeler, et al, 1998). Fidelity to the MST treatment model is maintained by weekly supervision meetings involving 3 to 4 therapists and a clinical supervisor. There is some experimental evidence that MST reduces arrests and convictions relative to treatment as usual (Littell, et al, 2005).

All three mental health interventions, ART, FFT, and MST use mental health therapy to address youth conduct problems, but it is not clear whether they are sufficient to remediate academic problems specifically. Other interventions for at-risk adolescents attempt to directly target school achievement, for example, through school-based interventions or tutoring. In some cases the intervention combines mental health treatment and academic components. Often the academically focused interventions target adolescents in high-risk schools, not specifically youth in the justice system.

---

[5] These estimates and program descriptions were accessed at
http://www.blueprintsprograms.com/programCosts.php?pid=cb4e5208b4cd87268b208e49452ed6e89a68e0b8
[6] Additional information on MST implementation is available at http://mstservices.com

Other Intervention Studies of Education for High-Risk Youth

Using an exclusively school-based approach, a study conducted by Fryer (2011) introduced instructional change with a "high-dosage" math tutoring component for all students in specific grades, as well as removal of poor-performing teachers and administrators, longer instructional time, data-driven instruction, and a culture of high expectations. These instructional changes were implemented in elementary, middle and high schools in Houston public schools that served low income students, 88% of whom were black or Hispanic. The study finds significant improvements on math achievement scores from the intervention, highlighting the possibility of significant improvements in math skills at relatively late stages, i.e., among 6th and 9th graders.

In contrast, the Pathways to Education Program was broader (Oreopoulos, Brown, and Lavecchia, 2014). It was offered in a housing project in Toronto and spanned case management, intensive tutoring, group activities, and financial support for school, college, and transportation expenses. Although it is not possible to unpack the effect of these different components, the Pathways study found improved grades, large increases in high school graduation rates (rising from 38% to 58%) as well as higher college entry.

Another school intervention took place with males in 7th through 9th grade in Chicago Public Schools and focused on crime and school outcomes (Heller, et al, 2013). The study randomized over 2,700 youth to school as usual, a group-based mental therapy program called "Becoming a Man (BAM)" offered in school, BAM offered with after-school programming, or after-school programming only. There was no tutoring group in this particular study. The BAM groups performed better on their grade point averages, but not on drop-out by year's end or on average days of school attendance.

Finally, Cook, et al (2014) conducted a study among 10th and 11th grade inner-city boys, using both group therapy and academic remediation in the form of math tutoring. The group therapy was based on BAM, while the academic intervention was daily two-on-one math tutoring for one hour based on the model developed by Match Education. While the pilot reports no statistically significant differences in disciplinary incidents in school or in days suspended out of school between those who received treatment and those who did not, it does find significant effects on math achievement.

Contributions of Study

Our study addresses key knowledge gaps and makes contributions along several dimensions. First, our study bridges the two literatures reviewed above, the first studying the effects of mental health interventions on crime outcomes among justice involved-youth and the second studying the effects of academically-focused interventions on educational outcomes among high-risk youth. Virtually all of the prior work on non-academic behavioral interventions, targeted at justice-involved youth, have understandably focused on first-order effects – studying criminal and conduct outcomes, which these interventions were intended to improve. Furthermore, virtually all of the literature on adult and juvenile crime has also understandably focused on the effects of justice policies and interventions on crime.[7] We add to the very sparse literature linking crime-based policies to domains other than crime. In the process, we provide some of the first causal evidence on whether, and the extent to which, these often-endorsed behavioral and mental health treatment programs, some of which do not have an explicit academically-focused component, nevertheless have broader spillover benefits on human

---

[7]One recent exception in the juvenile justice literature is Aizer and Doyle (2015), who use random judge assignment in Chicago to study the effects of juvenile incarceration on schooling and future adult crime. They find that juvenile incarceration results in large decreases in the likelihood of completing high school and large increases in the likelihood of being incarcerated as an adult.

capital development through educational acquisition. Given the vast literature linking education

to improvements in health behaviors, morbidity, and mortality, reduction in future crime and

delinquency, and other non-monetary benefits, an unbiased cost-benefit calculus of any

intervention would need to account for any positive (or negative) effects on educational

attainment.

Second, the school-based studies above included students attending school and by design

excluded youth who had dropped out. In our study, we start from a population of youth in the

justice system and examine whether mental health intervention improves their academic

achievement using several measures, two of which are dropping out of secondary school and

completion since our sample is not selected on current students. Third, the focus on justice-

involved youth – an inherently high-risk sample of adolescents – is particularly salient to the

debate on primary versus secondary prevention, and our study contributes evidence on whether

behavioral interventions can be effective in affecting positive change among older youth who

have already started to exhibit delinquent behaviors. Fourth, much of the juvenile justice debate

has also centered on the effectiveness of harsher sentencing such as detention and confinement.

However, most juvenile offenders are not incarcerated, and states differ in terms of how to deal

with these offenders who are released back into the community.[8] Our study thus adds to the

sparse literature on the causal effects of alternatives targeted at youth who are not sentenced to

confinement, and provides valuable policy guidance on whether behavioral interventions provide

any benefits over other less-costly local sanctions comprising probation, small fines, or some

community restitution.

---

[8] Among 112,766 offenders entering the juvenile justice system in WA between 2004-2009, about 4,211 unique
offenders were confined in a Juvenile Rehabilitation Center at some point over this period (authors' calculations
from the administrative data).

3.   Data

Administrative Data from WA

In order to perform detailed analyses of how mental health treatment in the justice system affects education outcomes, data were received from two agencies in Washington State, the statewide juvenile justice system and the statewide school system. The Washington state juvenile justice agency uses a computerized survey tool to measure a youth's risk of reoffending and eligibility for mental health services (Washington State Institute for Public Policy, 2004b). The survey instrument has three parts, a pre-screen assessment, full assessment, and a reassessment and is administered by trained probation officers and other staff. Youth rating moderate or high risk (see Figure 1; described below) on the pre-screen complete the full assessment, whereas those rating low risk do not get a full assessment. The instrument spans 12 domains including criminal history, demographics, free time, school, employment, relationships, mental health, substance use, family, attitudes, aggression, and skills.

The justice-system instrument captures criminal and social risk factors, as well as ratings of aggressive behavior and problems with family and living situation. Together these elements define eligibility for the treatment programs as described below. In addition, the justice data capture whether a youth was deemed eligible for treatment by treatment type (MST, FFT, or ART), whether they were referred for treatment, and reasons why the services may not have been received, such as lack of available openings or refusal to participate. Our data indicate that the average time from assessment to treatment planning was 164 days for MST, 119 days for FFT, and 131 days for ART. The average time between assessment and treatment completion for youth who completed treatment was 306 days for MST, 228 days for FFT, and 189 days for ART.

11

The justice data in our study include all youth who were screened or assessed for treatment services from January 2004 to April 2009. Our observation is a person and screening date pair. We exclude youth who ever appear in the data as deceased. We also exclude youth who ever report being incarcerated in long-term commitment facilities because these youth would receive education in those facilities and not through the regular education system.

The school data capture enrolled youth for the 2004/5 through 2008/9 academic years. These monthly data include information on the school district code[9], demographics (age, gender, race/ethnicity), details on participation in school programs (free lunch, special education, reading or math assistance), enrollment status (graduation or transfer), as well as grade point averages for youth in higher grades. Because the data include detailed monthly school information, we are able to construct multiple outcomes each measured over the 12 months after the youth was screened for treatment eligibility. We constructed four dependent variables: 1) indicator for whether a given youth dropped out of school; 2) an indicator for whether a youth transferred schools, which captures general disruptions in the youth's academic experience; 3) a broad indicator for being absent, which captures whether the youth dropped out of school, was suspended, or just did not show up to school; and 4) grade point average. In order to assess whether any potential benefits realized from not dropping out of school are persistent, we also assess high school or GED completion over the 24 months after the youth was screened for treatment eligibility.[10]

---

[9] There are 295 school districts in Washington.

[10] Given that our data right-censored as some juvenile offenders age out of high school eligibility, we chose to estimate 12- and 24-month outcomes as an alternative to estimating hazard models. By estimating the outcomes over this fixed duration, we ensured that all offenders are observed for at least that period of time.

These justice and education data were linked by the state agencies using last name, first initial, and date of birth.[11] After merging the screening and treatment data with the school data, our sample includes 35,020 observations over the 2004-2009 period.

Treatment Diversion Program Rules

Our study relies on a unique natural experiment afforded by program rules governing treatment diversion, which we exploit along with the rich information in the administrative records to identify the causal effects of treatment. In   Washington State treatment services are delivered to youth in the juvenile justice system based on fixed criteria. Youth are assessed after referral to the justice system, which is analogous to an adult arrest. Based on their assessment scores youth are deemed eligible for MST, FFT, or ART.[12] Although the agency refers designated justice-involved teens for treatment, the treatment itself is delivered in the community. Even when youth are found to be eligible, not all youth are assigned to treatment as they may refuse to participate or treatment slots may not be available.

Eligibility for each treatment type is determined by scores calculated from the youth assessment instruments in several key areas. The key dimensions on the assessment instrument are broader categorizations of criminal risk and social risk, the component domains of family living situation, and aggression-related domains. Criminal risk and social risk together capture a youth's recidivism risk level. Criminal risk scores (ranging from 0-31) are calculated from a

---

[11] The agency-matched files were de-identified and then provided to the research team. While we are enormously grateful to the agencies for conducting the matching and providing de-identified files, we are unable to explicitly test the sensitivity of our results to the matching procedure. In other work with county education and justice data (Cuellar and Markowitz, 2015) we were able to conduct our own matching, converting names to sound codex groupings using methods detailed in Trajtenberg et al., 2006). In that study the results were not sensitive to the matching method, which is reassuring.

[12] Some low-risk youth also are offered a low intensity program called Coordination of Services (COS) which we do not study here. Because our focus is on high risk youth, we exclude youth who receive COS services from our analyses. Furthermore, there is very limited evidence base for COS, even with respect to the first-order effects on crime and re-offending.

series of questions related to prior offending, including the age at which the youth first committed an offense, the types of previous offenses, and any prior incarcerations. Social risk scores (ranging from 0-18) are calculated from questions about difficulties in school, history of voluntary or court-ordered placement out-of-home, such as in foster care, relationships with peers, whether a parent is serving time in jail, parent control problems, and history of abuse, neglect, and mental health problems. Criminal and social risk are combined to determine whether a youth is considered low, moderate or high risk, as shown in Figure 1.

One component of social risk, family living arrangements, summarizes questions on whether the child is living with his or her parents, whether the family is supportive of the youth, whether there are sibling problem behaviors, whether the parents exercise appropriate supervision and control over the youth and whether reward and punishment behavior is appropriate and consistent. Deficits in these areas signify family instability and are fundamental to the appropriate allocation of FFT and MST therapies.

Another component domain, aggression, includes specific questions related to aggression, including tolerance for frustration, hostile interpretations of actions and intentions of others, belief in verbal and physical aggression as a legitimate means to resolve conflicts, and reports of past aggression problems. In addition, there are questions about pro-social attitudes and behaviors, such as attitudes about past crimes, optimism, empathy, impulsiveness, and respect for authority figures and others. Finally, there are questions related to skills, including goal setting, consequent thinking, basic social skills and dealing with emotions. Deficits in these areas are indicative of a need for ART.

Given that MST, FFT and ART each follow specific intervention models, it is not possible for a youth to receive multiple mental health treatments concurrently. Furthermore, the

14

high resource intensity of the FFT and MST partly reflects the individualized, coordinated, and targeted nature of these interventions, which are being tailored to treat the specific youth in question. While the specific treatment may differ from youth to youth, we are interested in the mean treatment effect of each program as implemented in practice. Moreover, as we examine the effect of one treatment, for example, FFT, we exclude youth who have received MST or ART.

Figure 2 summarizes how scores in these domains result in treatment eligibility. To be eligible for treatment, youth must first meet the appropriate criterion for recidivism risk, and second must meet certain score thresholds for either family instability or aggression (Figure 2). For MST eligibility a youth must be rated high-risk and have a score of at least an 8 out of 34 on family instability. For FFT, the family instability situation criterion is the same (8 out of 34), but the youth must be at moderate recidivism risk or greater. For ART, the youth must be of at least moderate recidivism risk, in addition to having aggression problems (2 on a scale of 13), or attitude problems (5 on a scale of 23), or skills problems (4 on a scale of 18). We utilize these eligibility thresholds, essentially comparing youth who are otherwise similar on all other observed measures of risk but differ in their eligibility for treatment assignment, in order to assess the causal effects of the treatment interventions, as detailed below.

4.   Analytical Framework

The objective of this study is to assess the causal impact of behavioral interventions on educational outcomes for justice-involved youth. We frame this question, and the empirical methodology, within an intent-to-treat analysis. Consider the following education production function relating eligibility for a particular treatment to post-treatment educational outcomes.

(1) $E_{ist+k} = \alpha + \beta\, EligThreshold_{it} + X_{ist+k}\Phi + \text{District}_k\Psi + Time_{t+k}\Omega + \mu_{it} + \varepsilon_{it+k}$

15

Specifically, educational outcome (E; for instance, dropping out of school, GPA, or high school completion), for the ith juvenile offender in school district s, realized over some period k beyond their intake at time t, is a function of their meeting the eligibility threshold (*EligThreshold*) for treatment subsequent to their offense and arrest. This threshold is a dichotomous indicator for whether the offender has met the relevant score threshold (Figure 2) for the treatment being studied. Thus, with respect to ART, this is an indicator for whether the offender has a high-enough score in the aggression component, or the attitudes/behavior component, or the skills component. For FFT and MST, this is an indicator for whether the offender has a high enough score in the family instability component.

The vector X includes gender, age at the time of the assessment, and race. It also includes subsets of social and criminal history measures further described below. The outcomes are related to the three mental health behavioral interventions described above: ART, FFT, and MST. All models control for district-level fixed effects (*District*), to account for any stable heterogeneity across the 295 school districts in WA (for instance, related to services offered, district-specific policies and enforcement, district-level demographics and peer groups, and quality of educational inputs), and year fixed effects (*Time*), to account for statewide trends in schooling outcomes (for instance, related to state-level policy changes or general economic conditions affecting all school districts). We estimate models for dichotomous outcomes (for instance, dropout, transfer, school completion) via probit regression and models for continuous outcomes (for instance, GPA) via OLS.[13] Standard errors are adjusted for arbitrary correlation in the error term (ε) across individuals in a given school district.

---

[13] Results are not sensitive to alternate estimation methods and specifications, including logit and OLS models for dichotomous outcomes, and log-linear models for GPA.

The parameter of interest is β, which captures the causal intent-to-treat (ITT) effect of making an offender eligible for a given treatment program on their schooling outcomes. The comparison is relative to no treatment and usual sanctions comprising probation, small fines, or some community restitution. This parameter expressly identifies the reduced-form impact of the initial treatment eligibility, and as such is distinguished from the structural effect of the "treatment on the treated" (TOT), which would capture the impact of actually receiving the treatment. The ITT effect provides valuable information on the likely direct effects of an important policy tool. State juvenile justice programs can manipulate the criteria determining treatment diversion and thus expand eligibility. However, eligibility does not necessarily translate into treatment receipt due to attrition and also because the treatment program may be unavailable in the district or there are no slots available. Thus, the ITT is informative in gauging the "real world" policy response to expanded treatment diversion after accounting for offenders who may fail to start the program for various reasons or dropout prior to completion.

A key empirical challenge in deriving a consistent estimate of β stems from potential non-random selection into treatment diversion. Eligibility for treatment may be correlated with unobserved offender-specific heterogeneity (μ), which may confound the estimated effect of treatment assignment. With detailed information on the offenders' social and criminal history, differences in eligibility around the various thresholds within fine homogeneous cells can be exploited for credible identifying variation to bypass this selection bias.

The empirical models proceed in a stepwise manner in order to address and inform these and other specific issues. First, we estimate a parsimonious version of equation (1), controlling only for basic offender demographics along with district and year fixed effects. These models establish the potential selection bias at play. Given that eligibility for treatment diversion is

limited to relatively higher-risk youth who have significant family difficulties and problems with their aggression, attitude, and coping skills – factors which would also predict poor school performance and enrollment – these naïve estimates of the ITT are expected to be biased downwards (upwards) for positive (negative) measures of school performance.

We therefore expand this specification to further control for a rich set of covariates. These include lagged schooling outcomes prior to the treatment screen, capturing the offender's GPA, whether he was suspended, was eligible for accommodations and services under Section 504, and their special education status (Model 2).[14] Such pre-treatment measures may partly account for selection bias from unobserved student- and family-specific factors (Angrist and Pischke 2009; Dave and Colman 2012). These models also control non-parametrically for the criminal history score (ranging from 0-31) and the social history score (ranging from 0-18), and also add linear and quadratic forms of the respective component scores that drive eligibility for ART, FFT, and MST. Also included in these models are indicators for various key component measures that form the criminal history and social history scores (other than those in the component score used to determine treatment eligibility),[15] such as age of first offense, prior misdemeanor and felony arrests, prior detention and confinement, mental health and substance abuse history, peer relationships, and history of other abuse and neglect.

---

[14] Section 504 is a part of the Rehabilitation Act of 1973 that prohibits discrimination based upon disability. Each district generally addresses students' needs and services under Section 504 differently, which would be captured by the district-specific fixed effects. Examples of such accommodations include highlighted textbooks, extra textbooks for home use, rearranged schedules, extended time on tests and assignments, preferred seating in the classroom, frequent feedback, peer assistance with note taking, oral exams in lieu of written exams, taping class lectures, and computer-aided instruction.

[15] For instance, the family instability score (which is part of the offender's broader social history) is based on measures such as the offender's current living arrangements, household income, problem history with parents and siblings, family support, parental supervision, and criminal history of other household members. This score governs eligibility into FFT or MST treatment (see Figure 2). Models cannot control for these components separately as they are subsumed in the family instability score (which we control for through a second-order polynomial).

Identification in this specification is driven by the jump in eligibility around the respective eligibility component score, within groups of offenders who alternately have either the same total criminal history risk score or the same total social risk history score. While this identifying variation can be framed in the context of a regression discontinuity design (RDD), it differs in some respects from a conventional RDD wherein the discontinuity exists on a single forcing variable. With respect to the treatment diversion programs that we are studying, several discontinuous thresholds are relevant. Thus, the contention is not that individuals with a family instability risk factor of 8 or 9 (and thus eligible for FFT treatment assignment) would be near-identical to those with a family instability risk factor of 6 or 7 (and thus ineligible for FFT), because this risk factor score itself has several components, and there are other first-tier thresholds (relating to overall social and criminal history risk, as shown in Figure 1).[16] However, once we control for all of the observable components of all the social and criminal history scores, then any differences between both sides of treatment assignment would be (and are indeed) minimized. That is, within a very fine cell where individuals have the exact same social and criminal history profiles and thus are observationally similar, large differences in eligibility are driven by small shifts in specific components (for instance, a 7 vs. an 8 on family instability due to a change in some sub-component that defines family instability).[17]

The next two specifications exploit variation across the eligibility thresholds within increasingly finer sub-cells of offenders who are more homogeneous in terms of their

---

[16] Indeed, there are some differences on observables and predetermined measures of schooling around the family instability risk score threshold. For instance, offenders with a family instability score of 8 (a threshold for FFT eligibility) are about 3.5 percentage points more likely to have dropped out of school prior to their screening and are 4.1 percentage points more likely to be classified as moderate or high risk, relative to those offenders with a family instability score of 7.

[17] While in principle it may be possible go within a relatively fine cell such that only a single discontinuity would apply, we invariably run into power issues and limited sample sizes which makes this infeasible. Furthermore, this would also substantially limit the external validity of any estimates based on such a narrow range.

observables and presumably also on their unobservables. Model 3 is a fully-saturated model; it adds indicators for interactions between each score for social history and criminal history risk (31 x 18 = 558 indicators). Note that not everyone who achieves the score thresholds in Figure 2 is eligible; eligibility also requires first-tier criteria (Figure 1), being classified to a high-enough risk group. Models 1-3 are estimated for all offenders including risk groups who are not eligible, and thus the estimates represent the ITT effect averaged over the entire juvenile justice population in WA.

Model 4, which is our preferred specification, excludes those risk-groups who would not be eligible for a particular behavioral intervention; for instance, low-risk offenders are not eligible for diversion into ART and FFT, and both low- and moderate-risk offenders are not eligible for MST. Thus, excluding these groups, as appropriate to the intervention being studied, should increase the magnitude of the estimated effect, in the spirit of a dose-response check. The identification in this model is driven by individuals who are identical in terms of both their social and criminal histories (that is, they have exactly the same combination of social *and* criminal history scores), have similar pre-treatment educational outcomes, have similar histories with respect to peer and gang relationships, mental health, substance abuse, age at first offense, prior misdemeanors and felony arrests, prior detention and confinement, and various other component measures noted above, but differ only in being on either side of the eligibility threshold. The thought experiment underlying this model is to compare two youths from the same school district at the same period, with both offenders being identical along their social and criminal history scores, as well as their prior educational outcomes, mental health, substance abuse, and offending history. The only observable difference between these youths is that they are on either side of the eligibility threshold, thus making one eligible for treatment and the other ineligible.

Comparison of the estimates across models 1-4 provide valuable information on the direction of the selection bias, and whether this bias is being attenuated as the identifying variation is arguably becoming more credible.

We next assess heterogeneity in these effects across gender and age as well as across prior school performance measures and their mental health history as reflected in the offenders' initial screening assessment. It should be noted that these measures are predetermined and predate the treatment. These analyses are important towards understanding whether the treatment interventions can improve schooling among those who are at the greatest academic disadvantage.

Not everyone who is eligible for treatment actually receives treatment. Hence, we supplement the ITT analyses with an instrumental variables (IV) estimation of the effects of the "treatment on the treated" (TOT), which is the actual causal effects of treatment take-up on educational attainment. Eligibility for treatment diversion, in the context of the thresholds noted above, would constitute a plausibly valid IV for treatment take-up. Additionally, many offenders are unable to commence their assigned treatment program because the particular program is not available in their area at the time or because all slots in the program are filled.[18] These supply-side constraints can also be used as an additional IV in the structural estimation of the TOT.

Finally, we implement several specification checks to validate our identification strategy. These include assessing whether our estimates are sensitive to outliers or alternate polynomial transforms of the scores, further sample restrictions to smaller neighborhoods surrounding the thresholds, assessing that the covariates and predetermined schooling outcomes are balanced

---

[18] Lack of treatment slots is a function of timing and available staff. Mental health clinics serve youth who are involved in the justice system and those who are not; thus, a given referral may be declined because trained staff were lacking at that particular time or because trained staff did not have openings.

across the eligibility thresholds in our preferred specifications, implementing various placebo

analyses, and assessing treatment effects at alternate eligibility thresholds relating to social

history risk and criminal history risk (from Figure 1). These are detailed in the next section.

5.   Results

Effects of Treatment Eligibility on Schooling Outcomes

Table 1 presents means for key variables over our analysis sample. The average juvenile

offender is 16 years of age, male (76% of the sample), and White (60% of the sample). Given

the high-risk nature of this sample of justice-involved youth, school performance measures are

expectedly low. Over the one-year period after screening in the juvenile justice system, the

school dropout rate is relatively high at 27.6%, with about 51.5% suspended or leaving school

this period, and the average GPA is 1.53. Among offenders ages 17 or above, only about 23.4%

complete high school or their GED over the two-year period after screening. Columns 2 and 3

present means for offenders stratified across their eligibility status for treatment with FFT, and

the final two columns present means based on eligibility for MST. Youths who are eligible for

these intensive treatment programs are not a random subset of all justice-involved youths. About

68% of FFT-eligible youth had low grades prior to their treatment diversion and only 18.1%

attended school regularly. This compares with FFT-ineligible youth, among whom 43.1% had

low grades and 52.7% attended school regularly. Furthermore, youth who are eligible for

treatment services are relative higher-risk in terms of both their criminal history as well as their

social history, and more likely to have experienced family problems. These systematic

differences are statistically significant and are expected to confound post-treatment schooling

outcomes, which on average are generally worse among the eligible youth. For instance, the

one-year school dropout rate is 32.8% among those who are FFT-eligible, compared with 23.9%

among those who are ineligible. The multivariate models presented below address this non-random selection bias.

Table 2 presents estimates of the ITT effect of ART eligibility on schooling outcomes, measured over a period of 12 months subsequent to treatment screening. Column 1 presents estimates from a parsimonious version of equation (1), controlling only for the offender's socio-demographics in conjunction with district and year fixed effects. Across all outcomes, these estimates suggest that eligibility for diversion into ART is associated with worse school performance (significantly higher dropout and lower GPA) and bring to light the substantial selection bias at play. Observed and unobserved risk factors underlying the youth's social and criminal history, which are key drivers of eligibility into the behavioral therapy program, are also predictive of poor school outcomes. Thus, these naïve estimates are biased towards showing that the intervention is less effective in improving schooling than may be the case. Models 2-4 add richer controls to account for this selection, while exploiting more selective variation surrounding the eligibility cut-offs within increasingly finer cells of alike youth. Comparing estimates across specifications 1-4 suggests that the bias is attenuated in the expected direction as we employ more credible identifying variation. However, estimates from our preferred specification (model 4) do not show any significant effects on these measures of education. There is some suggestive decrease in transfers, which may be indicative of a reduced disruption to schooling, though the estimate is statistically insignificant.

It should be noted that ART is implemented through group therapy format, and is the least intensive of the three mental health interventions. While ART teaches youths to control their angry impulses and aggression, and develop their moral reasoning skills, there is no explicit

educational goal, and we do not find that ART leads to any discernible academic improvements at least over a one-year follow-up.

Tables 3 and 4 present the estimated ITT effects for the more intensive of the treatment programs, FFT and MST respectively. Specifically, Panel A of Table 3 assesses how meeting the second tier eligibility threshold for FFT affects a broad measure of school absentee status, a measure which includes both dropping out of school as well as suspensions. As before, model 1 suggests that eligibility is associated with a higher likelihood of dropping out, an association that is largely driven by the positive selection bias. This bias diminishes, and effect sizes shift towards becoming more negative as successive models exploit more credible variation.

Our preferred specification (model 4) controls for all criminal and social risk factors non-parametrically, along with the rich set of covariates, and is limited only to those youth who are classified as moderate or high risk. Here, we find that the ITT effect of a 100% increase in FFT eligibility is to reduce the 12-month absentee rate by 4 percentage points (about 5% relative to the baseline mean), though this effect is imprecisely estimated (p-value of 0.2) due to reduction in sample size and the variation being exploited.[19] It is validating that excluding the low-risk individuals from the sample, who are not eligible for FFT, causes the magnitude to increase from 1.5 (model 3) to 4.0 percentage points (model 4). We do not find any significant or meaningful effects of FFT eligibility on the probability of being suspended.[20] Thus, all of the effects on the broad measure of absentee status are driven by decreases in the likelihood of dropping out of school. Panel B of Table 3 confirms this, suggesting that a 100% increase in FFT eligibility

---

[19] As models become progressively richer, we lose observations due to missing information on the pre-screen. The observations with missing information on the pre-screen are not significantly different (beyond a Type I error) with respect to all other observable characteristics relative to the full sample.

[20] Results are not reported for economy of space, and available upon request.

reduces the propensity to drop out of school over the next 12 months by a statistically significant 5.7 percentage points (about 18% relative to the sample mean).

Panels C and D suggest slight negative effects on both GPA (by about 0.04 index points) and the likelihood of transferring to another school (by about 1.1 percentage points). Though these estimates are generally insignificant in the richer models, they are consistent with the reduction in dropout. The lower likelihood of transfers is potentially less disruptive and may promote the likelihood of school retention. Prior work on other educational inputs has also generally been hard pressed to find positive beneficial effects on GPA and academic achievement (Cullen, et al 2013). In our case, the greater selection of higher-risk and lower-achieving students staying in school may be responsible for the slight negative effects on GPA.

The ITT effects for MST, which is the most intensive of the behavioral mental health interventions, are presented in Table 4. Generally, these effects mirror those for FFT (in Table 3), suggesting significant decreases in the likelihood of dropping out of school. It is again validating that the effect magnitude increases substantially between models 3 and 4, when we omit risk groups (low and moderate-risk youth) who are not eligible for MST treatment diversion despite meeting the threshold on the family instability score.

In comparison with FFT, the effect magnitudes for MST are larger. Model 4 suggests an 8-10 percentage points decrease in the probability of dropping out of school due to a 100% increase in MST eligibility. This higher ITT effect may reflect both the greater intensity of the intervention as well as a differentially larger effect due to the more higher-risk individuals being targeted for this treatment. In later stratified analyses, we further explore this question of heterogeneous treatment effects based on schooling-related risk measures.

Figure 3 shows the discrete reduction in the 12-month dropout rate for FFT and MST around the threshold determining eligibility, based on local linear regression estimation. It is validating that these estimates also suggest a reduction on the order of about 5-9 percentage points, though they are less precise (p-value of 0.22 for FFT and 0.08 for MST) due to the non-parametric estimation.

The analyses thus far suggest that the intensive interventions, FFT and MST, increase the likelihood of staying in school over the short term. Table 5 assesses whether these effects are persistent and translate into a greater likelihood of completing secondary school or its equivalent through the GED. As before, model 1 suggests that there is strong negative selection bias leading to higher eligibility being associated with a lower likelihood of graduation/GED completion. This selection bias is attenuated in models 2-4, however, which exploit the discontinuities and fully saturate the non-parametric controls for interactions between social risk and criminal history risk. These models suggest that all three interventions increase the probability of completing secondary school on the order of 5-10 percentage points.

Heterogeneous Effects of Treatment Eligibility on Schooling Outcomes

Justice-involved youth are predominantly males, about 75% of the sample. Furthermore, a large literature suggests differences in problem behaviors and coping mechanisms across boys and girls. Thus, Table 6 explores heterogeneity in the ITT effects across gender (columns 1 and 2). In general, we find that both FFT and MST eligibility are effective in reducing dropout for males and females; however, due to inflated standard errors, we cannot rule out that there is no difference in the ITT effects, though the ITT effect for FFT is an order of magnitude higher for females than for males. A larger ITT effect may reflect either an increase in take-up rates and/or an increase in the effectiveness of actually receiving the treatment (TOT). Over the sample

period, the take up rate for MST is not significantly higher among females, and thus the suggestively higher ITT effect for MST for females may reflect a greater effectiveness of the program for this group.

We also assess heterogeneity in the effects by age of screening (results not reported in the table), comparing those who were screened when 16 or younger vs. those who were screened when 17 or older.[21] Differential effects by age may capture differences in the short-term vs. longer-term responses; for instance, a 17-year old who is being screened and treated may not dropout since he may still be monitored by an MST therapist, but the student who is younger when screened may have relatively poorer educational outcomes once the supervision and therapy fades. These results indicate that both FFT and MST reduce 12-month dropout for both age groups; the effect however is significantly stronger (the magnitude is about twice as large) for the older offenders ages 17+. This may reflect a stronger beneficial response in the short-term, though also validates the point that treatment programs targeted at older high-risk offenders can be effective in raising educational attainment.

Given that FFT and MST diversion appears to be effective in raising school retention and graduation in general among those who are made eligible, an important question is whether these interventions are effective for those youth whose educational prospects are among the weakest – those youth who are at higher risk of dropping out of school and recidivating. Models 3-10 assess heterogeneity in the ITT effects based on predetermined schooling outcomes and assessment, and whether the offender has any history of mental health problems, all measured prior to the treatment screening. In general, higher eligibility for both FFT and MST impart a

---

[21] Due to sample size limitations, we are not able to stratify the sample by age at screening. Hence, we explore differential effects based on an interaction term between the eligibility cutoff and an indicators for ages 17+ vs. ages 16 and younger.

positive benefit for students, regardless of their prior school performance and attendance. Even among students who previously were performing poorly in school (low grades and/or poor attendance record), the ITT effect for both interventions is to reduce the likelihood of dropping out of school.

The policy relevance of these effects is particularly prominent in columns 7 and 8. The prescreen included the interviewer's assessment, based on the various domains and their own impressions of the youth, his/her family, and the surroundings, on how likely is the youth to graduate from high school (columns 7 and 8). There is evidence that these interventions can impart important positive educational spillovers to at-risk youth whose prospects of graduating from high school were particularly poor. We qualify this discussion by noting that in stratifying the analyses to assess heterogeneity across these margins, many of these estimates are imprecise due to reduced sample sizes, though the effect magnitudes and patterns remain robust and are in line with those reported in Table 3-4.

Effects of Treatment Take-up on Education

We note that the effects discussed thus far are intent-to-treat effects, measuring the effects of eligibility rather than the effects of actual treatment receipt (TOT). The two diverge since take-up rates of the treatment interventions are far less than 100%. As shown in Table 1, 63.3% of the sample is eligible for ART, 40.4% is eligible for FFT, and 29.5% if eligible for MST. While we cannot directly observe actual referral, we can impute referral based on those who have a treatment plan and begin services and those who could not commence treatment due to supply-side constraints (treatment was not offered in their area at the time, or slots were filled up, for instance).[22] Thus, among those who are eligible for each of the three programs and do

---

[22] The imputed referral figures are not used in any of the analyses and are only noted here for descriptive purposes.

not start any other program, 32% are referred to ART (among those who are ART-eligible), 28% are referred to FFT (among those who are FFT-eligible), and 19% are referred to MST (among those who are MST-eligible). Among those who are referred, not everyone is able to start and receive treatment services due to the supply-side constraints; thus, among those who are referred about 65% actually take up and receive ART services, 64% receive FFT, and a little over 15% start MST treatment.

Figure 4 shows the discrete jump in the FFT and MST take-up around the threshold determining eligibility, based on local linear regression estimation conditioning on observed covariates. These estimates are statistically significant and are very similar to the unconditional differences in mean take-up rates between eligible and non-eligible offenders. They suggest about a 25% increase in take-up from becoming eligible for FFT, and about a 4-5% increase in take-up from becoming eligible for MST.

In Table 7, we directly estimate the effects of treatment take-up on schooling outcomes, based on an IV methodology. [23] In the fully-saturated model, eligibility for the particular treatment program under study and an indicator for whether the youth did not start the program due to supply-side constraints (it was not offered in the area at the time, or there were no available spots) are plausible IVs for program take-up. Indeed, the first-stage estimates reported in Table 7 confirm this, and show that eligibility significantly and positively predicts program take-up while supply-side constraints significantly reduce take-up. Particularly for ART and FFT, the IVs strongly predict take-up, with the F-statistic on the excluded instruments ranging

---

[23] An alternate approach to deriving the TOT is based on the Wald estimator (ratio of the effect of eligibility on education to the effect of eligibility on take-up), comparable to an estimator that uses the thresholds as an IV in a just-identified model. The take-up effects are shown in Figures 4. Dividing the ITT effects by these take-up effects yields estimates similar to those reported in Table 7, though they are less precise. We exploit the supply-side constraints as an additional IV in Table 7 for added statistical power, and also to draw policy implications in the discussion relating to the supply-side constraints driving some of the low take-up rates.

from 97 to 305. The IVs are jointly significant at the 1-percent level for MST take-up, though the F-statistic is lower (ranging from 6 to 8) due to the weaker take-up rate of the program among eligible youth.[24] The over-identification test also supports the orthogonality of the IVs from the structural error; while this is statistically validating, we note that the test should be interpreted with caution since the two instruments (eligibility and supply constraints) are likely working at different margins with heterogeneous treatment effects across these margins.[25]

Columns 1-3 present the TOT effects for ART take-up on dropout and high school completion (over 12- and 24-months post screening). Consistent with the weak to null ITT effects, we do not find any significant or meaningful effects of ART take-up on drop-out or graduation over the short-term. Column 3 is suggestive of some positive cumulative effect on completion rates over a 24-month period. Columns 4-6 report the TOT effects of FFT take-up, which are fully in line with the earlier-report ITT effects. Take-up of FFT reduces the likelihood of dropping out of school by about 9 percentage points and raises the likelihood of graduating by about 5-8 percentage points. The final three models present the TOT estimates for MST participation. These also confirm that MST take-up reduces the likelihood of dropping out of school by about 24 percentage points and increases the likelihood of completing secondary school by between 25-53 percentage points. That is, while most youths who are eligible for MST do not enter treatment, those who do enter treatment significantly benefit from it in the form of greater educational attainment.

Specification Checks

---

[24] This is due to the reasons noted in the text. There is a higher rate of refusal because the MST program is intensive and requires a significant level of participation and effort from the youth and their families; also, supply constraints are more binding given the intensive resources required for offering MST.

[25] The Hausman test also confirms that program take up is expectedly endogenous.

We implemented several additional checks to assess the plausibility of our identifying assumptions, and to verify that the estimates are robust to alternate specifications and estimation issues. First, we confirm that covariates and predetermined outcomes are balanced around the eligibility thresholds. Table 8 presents estimates wherein we predict the demographics and predetermined schooling outcomes, which predate the screening, using the eligibility thresholds and our preferred specification. If the youths being compared in these models are alike, then we should not observe any meaningful effects of these thresholds on their demographics or prior schooling outcomes. It is validating that other than a small marginally significant difference on predetermined GPA, we find no statistically or economically significant effects on schooling outcomes measured prior to the arrest and screening.[26]

Second, we confirm that our estimates are not driven by functional form. Model 5 in Tables 3 and 4 control for up to a third-order polynomial in the respective scores, yielding very similar treatment effects to those that control for a second-order polynomial. Estimates are also not sensitive to linear or fourth-order polynomials[27], and we also obtain similar magnitudes of the treatment effects with non-parametric local linear regression estimates (Figure 3).

Our preferred specifications always limit the sample to those risk levels who are potentially eligible for treatment, and then exploit program rules within these risk levels that determine eligibility. Third, we assess that our estimates are robust to restricting the sample to a narrow range (+/- 5; relative to an unrestricted potential range from 0-34) around the program rule cutoff scores for FFT and MST, in order to identify an even more homogeneous group of

---

[26] The small difference in pre-arrest GPA is in the opposite direction to be driving any beneficial effects on school completion.

[27] Gelman and Imbens (2014) recommend against controlling for higher-order polynomials in such regression discontinuity-based models.

offenders. Model 6 in Tables 3 and 4 presents these results.[28] While standard errors inflate somewhat, these estimates are also in line with our main results. Additionally, the local linear regression estimates are also not sensitive to alternate bandwidths (+/-5, +/- 8, optimal bandwidth that minimizes MSE, 200-500% of optimal bandwidth), though the variance increases as the bandwidth gets smaller compounding our power issue. The optimal bandwidth for the local linear regression-based estimates (based on minimum MSE) is generally +/- 2 of the instability score, and while the effect magnitudes are highly robust (similar to those presented in Table 1), the standard errors are larger and render most of these effects to be statistically insignificant at conventional levels.

Fourth, we assess whether screeners could be manipulating the scores affecting eligibility to divert certain offenders into treatment. The survey is completed by probation officers, though the ultimate decision to deliver services rests with the mental health providers who deliver MST, FFT or ART. If the scores are being manipulated, we would observe discontinuous densities at scores around the threshold. We do not find this to be the case. See Figure 5, for instance, below which normalizes the relevant score so that the eligibility cutoff is at 0. We statistically tested for any breaks in the density at the relevant thresholds, based on the test proposed by McCrary (2008), and in all cases fail to reject the null of any significant break in the density at these thresholds.

Fifth, we conducted falsification checks wherein the eligibility thresholds should not have any effects on outcomes. The first tier of eligibility criteria depends on certain risk levels and

---

[28] We do not report estimates for ART, since Table 2 suggests no significant effects. However, we confirm that these estimates are also not sensitive to functional form (estimate for 12-month dropout is 0.0225 with a third-order polynomial) or to a limited bandwidth (estimate for 12-month dropout if 0.0127 when limiting the sample to a narrower bandwidth around the thresholds: +/- 2 on Aggression as the threshold occurs at 2; +/- 5 on Attitudes and Behaviors as the threshold occurs at 5; and +/- 4 on Skills as the threshold occurs at 4).

suggests a natural placebo check. For instance, in order to be eligible for FFT-based mental health treatment, the offender must have a minimum risk level of moderate or high, and then attain a score above some threshold. Thus, among low-risk offenders who are not eligible for treatment, being above these thresholds should not have any beneficial effects on educational attainment. Similarly, the most intensive mental health treatment (MST) requires that offenders be classified as high-risk, and then attain further thresholds on family instability. Thus, among low or moderate risk offenders, achieving these thresholds would not make one eligible for treatment and should not have any effects on dropping out of school. Model 7 (Tables 3 & 4) limits the sample to low risk offenders who would not be eligible for FFT treatment, and to low and moderate risk offenders who would not be eligible for MST treatment. They do not show any significant effects on schooling outcomes at the thresholds among these groups.

All of the preferred estimates are limited to fixed categories of social and criminal history risk, and then use the second-tier domains (family instability; aggression, attitude, or social skills problems; see Figure 2) to define discrete thresholds in eligibility for treatment. We do so in order to preserve our sample and maximize statistical power. Sixth, , we respectively exploit shifts in eligibility due to other first-tier thresholds (Figure 1) in the social history risk score and the criminal history risk score. Specifically, when studying FFT and MST, we restrict the sample to those with a family instability score ≥ 8 (which is one of the criteria for treatment eligibility) and then further restricted the sample based on fixed categories of social risk, and then using only the criminal history risk thresholds to identify treatment eligibility. For instance, among those offenders with a family instability score ≥ 8 and who have a social history risk score between 6-9, a criminal history risk ≥ 3 would make this offender eligible for FFT, and a criminal history risk ≥ 8 would make him eligible for MST. Alternately, we also restrict

offenders with a family instability score ≥ 8 to fixed categories of criminal history risk, and use only movements in social history risk to define eligibility. All of these models utilize exogenous identifying variation at other margins of social and criminal history risk. The estimates are highly imprecise due to inflated standard errors and reduced sample sizes from the multiple sample restrictions. Nevertheless, they are generally consistent with the results presented in Tables 3-4, and suggest a reduction in dropout associated with FFT and MST treatment eligibility on the order of 5-10 percentage points.[29]

6.   Conclusions

The study adds to the sparse literature linking justice policies to outcomes other than crime – most notably, education acquisition – which has important benefits both monetarily and non-monetarily. Our analyses suggest that intensive behavioral interventions, such as FFT and MST, can have positive impacts on academic achievement for troubled teens. The evidence is consistent with improvements in school completion rates for justice-involved youth. Moreover, the results suggest that the effects may be larger for girls and for youth who are believed to have worse education prospects. The evidence to support ART, a less intensive intervention, is not as strong.

According to the Blueprints Initiative, FFT costs $2,800 and MST costs $7,068 per youth. Our estimates suggest that actual FFT take-up reduces drop-out by about 10 percentage points, and MST take-up reduces drop out by about 24% (based on the IV estimates). The average high-school dropout earns $20,241 (U.S. Census) annually. The return to an extra year

---

[29] We also replicated the analyses for ART based on these alternate eligibility thresholds at the margins of social and criminal history risk. These effects are statistically insignificant, as those reported in Table 2, though the magnitudes (-3.1 to -4.9) are suggestive of some reduction in dropout. This would also be consistent with the improvement in 24-months secondary school completion that we find with respect to ART, both in terms of the ITT and the TOT analyses.

of education, even if this person does not finish high school is 7-8% (Angrist and Krueger, 1991). Hence, an extra year of schooling is predicted to raise earnings by about $1,500. FFT participation will therefore, on average, raises earnings by about $150/year ($1,500*0.075) and MST participation will raise earnings by about $375 ($1,500*0.24). In present value terms, the lifetime increase would more than compensate for the cost of both FFT and MST, even at the conservative benefit levels of just reducing dropout (not even counting high school completion).

The results of this study are important because they suggest that interventions for youth in the justice system have broad societal benefits outside the health care and juvenile justice sectors. Strictly speaking, the estimates we identify apply only to compliers, the specific subpopulation surrounding the thresholds and whose eligibility is discontinuously impacted by just meeting these thresholds. While the compliers are a relevant subpopulation, our population of interest, and our sample, consisting of juvenile offenders who have come into contact with the justice system, is relatively high risk (in terms of reoffending, engagement in other risky behaviors, and poor schooling outcomes) to begin with. Any treatment programs, which could be targeted within this population and which are shown to be effective, would be highly valuable for policy makers. While technically the added educational benefits are indicated only for those individuals on the margin of treatment assignment, the subgroup analyses and the robustness of the effects at multiple margins provide some suggestive indication that the responsive population who may benefit from these treatment programs may be broader. Consequently, a strong argument could be made for making these interventions much more widely available through education-system investments.

Currently, only limited funding is provided by justice systems, including in Washington state where services are frequently not available even when youth are eligible. To launch MST

and FFT programs, mental health clinics must invest in training up-front of clinical teams on each of the intervention methods. In addition they also must make an ongoing commitment to maintaining treatment fidelity and low caseloads. Designing funding schemes would likely necessitate joint planning across education, justice and Medicaid agencies.

The estimated effects indicate that mental health treatment interventions offered by the justice system can have important impacts on human capital development. Consistent with other recent studies, we find evidence that intervention programs can make an academic difference for adolescents, suggesting that adolescents is not too late to improve the course of anti-social behaviors or academic achievement. Our results show that these interventions are effective even among the highest-risk youth with poor schooling prospects and a history of mental health problems. Future work should focus on interventions modified for adolescents with other problems, such as substance abuse and family history of mental health.

References

Aizer, A. and J. Doyle Jr. "Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges," *Quarterly Journal of Economics* vol. 130 (2015): 759-803.

Angrist, J. D. Krueger, A.B. "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics* vol. 106 (1991): 979-1014

Angrist, J.D. and J.S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton and Oxford.

Acemoglu, D. and J. Angrist. 2000. How large are human-capital externalities? Evidence from compulsory-schooling laws. *NBER Macroeconomics Annual*, Vol. 15 (2000), pp. 9-59.

Colman, G. and D. Dave "Physical Activity and Health," *International Journal of Arts and Sciences*, 5, No. 6 (2012), 29-45.

Cook, Philip J., K. Dodge, G. Farkas, et al. 2014. *The (Surprising) Efficacy of Academic and Behavioral Intervention with Disadvantaged Youth from a Randomized Experiment in Chicago.* Institute for Policy Research, Northwestern University.

Corman, H., D. Dave, and N. Reichman. "Effects of Welfare Reform on Women's Crime," *International Review of Law and Economics*, 40, (2014), 1-14.

Cuellar, Alison "Preventing and Treating Child Mental Health Problems" *Future of Children* issue on *Policies to Promote Child Health*, forthcoming

Cullen, JB, Levitt, S, D. et al. (2013) "What Can Be Done To Improve Struggling High Schools?" *Journal of Economic Perspectives*, no. 27: 133–152

Fryer, Roland G. (2014) "Injecting Charter School Best Practices into Traditional PublicSchools:Evidence From Field Experiments." *Quarterly Journal of Economics* 2014;129(3):1355-1407

Gelman, A and G. Imbens, "Why High-order Polynomials should not be used in Regression Discontinuity Designs," *National Bureau of Economic Research Working Paper Series*, No. 20405, August 2014.

Greenwood, P. (2008). "Prevention and intervention programs for juvenile offenders." *The Future of Children, 18*(2), 185-210

Gundersen, Knut K., and Frode Svartdal. 2006. "Aggression Replacement Training in Norway: Outcome Evaluation of 11 Norwegian Student Projects." *Scandinavian Journal of Education Research* 50(1):63–81.

Heller, Sara B. and colleagues. "Preventing youth violence and dropout: A randomized field experiment." Cambridge, MA: NBER Working Paper #19014 (2013)

Henggeler, S.W., Schoenwald, S.K., Borduin, C.M., Rowland, M.D., and Cunningham, P.B. (1998). Multisystemic treatment of antisocial behavior in children and adolescents. New York : Guilford Press.

Lipsey, M. W., Howell, J. C., Kelly, M. R., Chapman, G., and Carver, D. (2010). *Improving the effectiveness of juvenile justice programs: A new perspective on evidence-based practice*. Center for Juvenile Justice Reform, Georgetown University

Littell, Julia H, Margo Campbell, Stacy Green, and Barbara Toews. "Multisystemic Therapy for social, emotional, and behavioral problems in youth aged 10-17" *Cochrane Database Systematic Review*. 19, no. 4 (2005)

Littell, Julia H, Aina Winsvold, Arild Bjørndal, Karianne T Hammerstrøm. 2007. Functional Family Therapy for families of youth (age 11-18) with behavior problems. *Cochrane Database of Systematic Reviews* Issue 2.

Lucenko, Barbara A, Lijian He, David Mancuso, and Barbara Felver, *Effects of Functional Family Parole on Re-Arrest and Employment for Youth in Washington State: Technical Appendix* RDA Report 2.24A (Olympia, WA: Department of Social and Health Services, September 2011)
https://www.dshs.wa.gov/sites/default/files/SESA/rda/documents/research-2-24.pdf

McCrary, J. "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 142(2008): 698-714.

OJJDP Model Programs Guide: Aggression Replacement Training (ART)," U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention (OJJDP) accessed at http://www.crimesolutions.gov/ProgramDetails.aspx?ID=254

Oreopoulos, P, R.S. Brown, and A.M. Lavecchia. "Pathways to Education: An Integrated Approach to Helping at-risk High School Students," NBER Working Paper 20430 (2014).

Sexton, Thomas and Charles W. Turner . "The effectiveness of functional family therapy for youth with behavioral problems in a community practice setting." *Journal of Family Psychology*. 24, no. 3(2010):339-348

Sickmund, Melissa, and Puzzanchera, Charles (eds.). 2014. Juvenile Offenders and Victims: 2014 National Report . Pittsburgh, PA: National Center for Juvenile Justice

Washington State Institute for Public Policy. 2004a. *Outcome Evaluation of Washington State's Research-Based Programs for Juvenile Offenders*. Olympia, Wash.: Washington State Institute for Public Policy. http://www.wsipp.wa.gov/rptfiles/04-01-1201.pdf

Washington State Institute for Public Policy. 2004b. Washington State Juvenile Court Assessment Manual Version 2.1 Olympia, Wash.: Washington State Institute for Public Policy. http://www.wsipp.wa.gov/ReportFile/873/Wsipp_Washington-State-Juvenile-Court-Assessment-Manual-Version-2-1_Manual.pdf

Wolff, Jennifer C. and Thomas H. Ollendick. 2006 "The Comorbidity of Conduct Problems and Depression in Childhood and Adolescence," *Clinical Child and Family Psychology Review* 9 (3/4): 201-220

Figure 1

Recidivism Risk-level definitions based on Criminal History and Social History Risk Scores			
Criminal History Risk Score	**Social History Risk Score**		
	0 – 5	**6 – 9**	**10 – 18**
0 – 2	Low	Low	Moderate
3 – 7	Low	Moderate	High
8 – 31	Moderate	High	High

Figure 2

Treatment Eligibility Determination		
Treatment Type	**Risk-level**	**Other**
Multi-systemic Therapy (MST)	High	Family Instability: Risk Factor Score of at least 8 out of 34 on specific measures of family instability within the Social History component
Family Functional Therapy (FFT)	Moderate or High	Family Instability: Risk Factor Score of at least 8 out of 34 on specific measures of family instability within the Social History component
Aggression Replacement Training (ART)	Moderate or High	Aggression, Attitude, or Social Skills Problems: Risk Factor Score $\geq$ 2 (out of 13) on the Aggression component domain of Social History risk Risk Factor Score $\geq$ 5 (out of 23) on the Attitudes/Behavior component domain of Social History risk Risk Factor Score $\geq$ 4 (out of 18) on the Skills component domain of Social History risk

Figure 3

12-month Dropout & FFT	12-month Dropout & MST
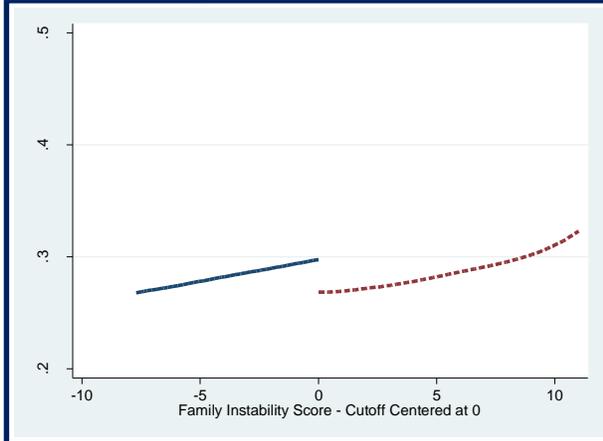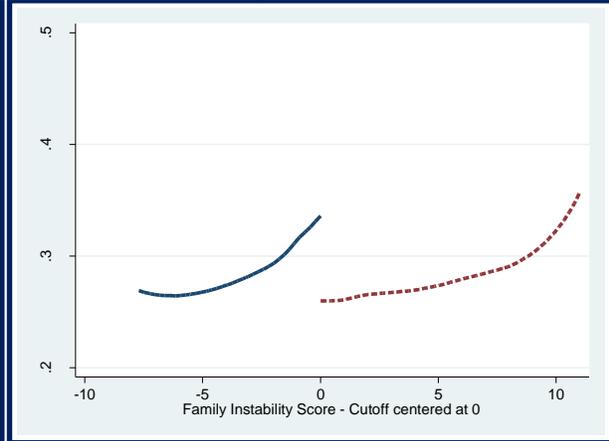
Notes: For FFT, sample limited to moderate or high-risk individuals, and youth who receive ART, MST or COS are excluded. For MST, sample is limited to high-risk individuals, and youth who receive ART, MST or COS are excluded. Outliers (Family instability score of 20 or more) are excluded. Estimates are based on 250% of optimal bandwidth. Estimates control for socio-demographics, covariates listed in Tables 2 & 3, and interactions between social history risk and criminal history risk.

Figure 4
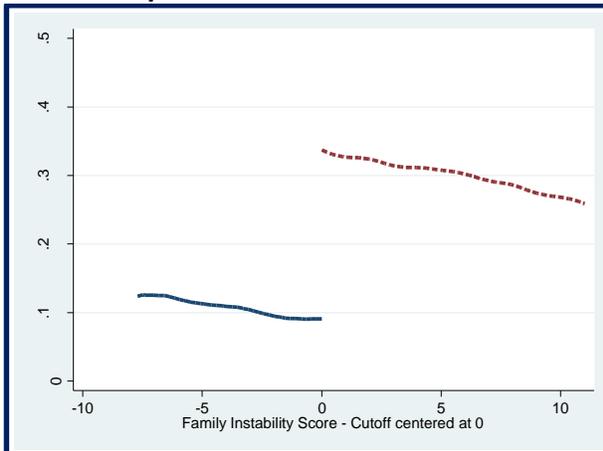
FFT Take-up	MST Take-up
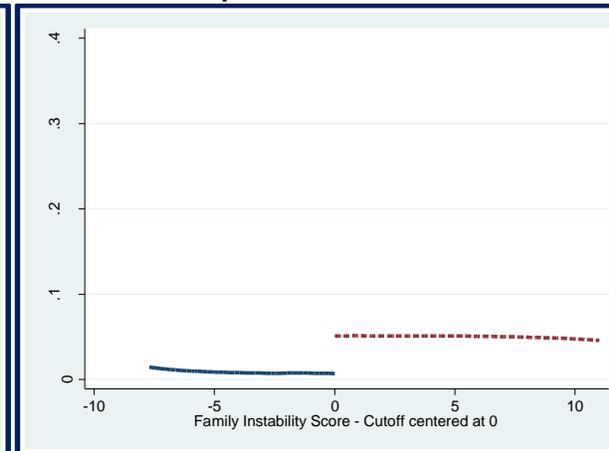
Notes: See Figure 3.

Figure 5
Density at Relevant Score Thresholds: Moderate or High Risk Offenders

Aggression
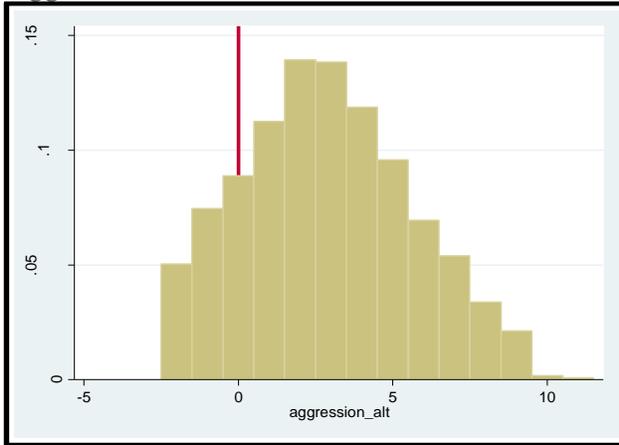
Attitudes/Behaviors
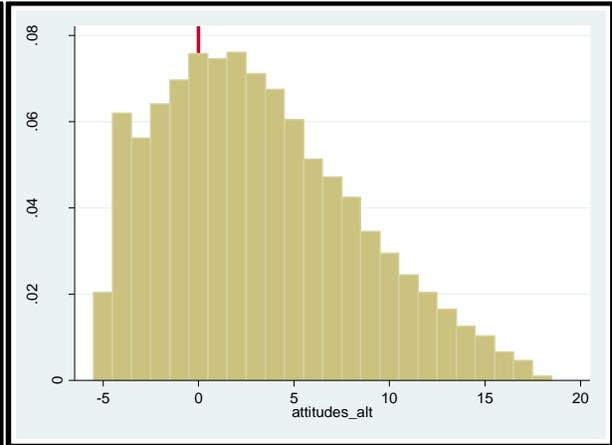
Skills
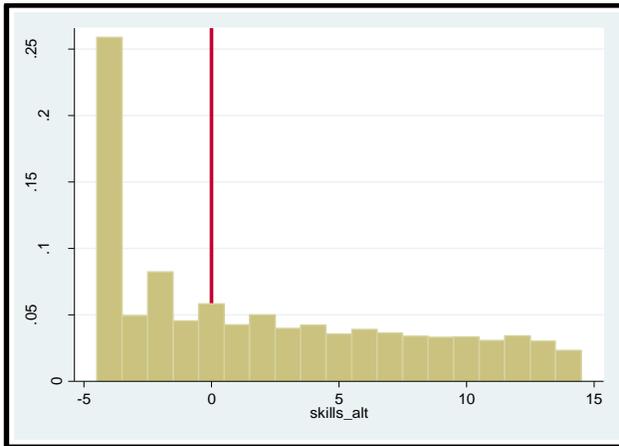
Family Instability
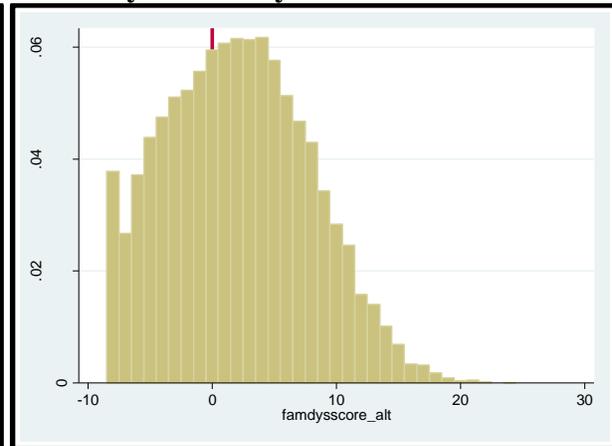
Table 1
Sample Means
Matched Intervention-Education Records from WA, 2004-2009

|	1	2	3	4	5
	All	FFT	FFT	MST	MST
Variable		Eligible	Not Eligible	Eligible	Not Eligible
Age	15.92	15.80***	16.00	15.79***	15.97
Male	0.7557	0.7576	0.7545	0.7765***	0.7470
White	0.6004	0.5702***	0.6208	0.5523***	0.6205
Black	0.1145	0.1313***	0.1032	0.1440***	0.1022
Asian	0.0342	0.0260***	0.0397	0.0244***	0.0382
Hispanic	0.1574	0.1637***	0.1531	0.1649**	0.1542
Race/Ethnicity missing	0.0936	0.1087***	0.0833	0.1144***	0.0849
Absent (Dropout, Suspended or No-show) (1-year)	0.6362	0.7614***	0.5483	0.7848***	0.5721
Dropout (1-year)	0.2756	0.3276***	0.2391	0.3336***	0.2505
Transfer (1-year)	0.8844	0.9135***	0.8640	0.9184***	0.8697
Suspended or Left (1-year)	0.5147	0.6464***	0.4222	0.6783***	0.4440
GPA (1-year)	1.5319	1.2676***	1.7041	1.2483***	1.6450
High School / GED Completion (2-year) (all ages)	0.1013	0.0560***	0.1342	0.0495***	0.1248
High School / GED Completion (2-year) (ages 17+)	0.2338	0.1170***	0.2985	0.0979***	0.2804
ART Eligible	0.6332	0.9831***	0.3963	0.9905***	0.4839
FFT Eligible	0.4037	1	0	1	0.1545
MST Eligible	0.2948	0.7302***	0	1	0
Social History Score	6.6363	9.2985***	4.8337	10.2517***	5.1250
Criminal History Score	7.7280	9.9231***	6.2417	11.0925***	6.3215
Risk Level – Low	0.3396	0	0.5696	0	0.4816
Risk Level – Moderate	0.2543	0.2698***	0.2437	0	0.3605
Risk Level – High	0.4061	0.7302***	0.1867	1	0.1579
Family Dysfunction Score	7.0578	13.3640***	2.7878	13.9146***	4.1914
Family Dysfunction Score ≥ 8	0.4245	0.9865***	0.0440	0.9887***	0.1887
Low Grades – GPA ≤ 2.0 (Pre-Treatment)	0.5326	0.6823***	0.4312	0.7103***	0.4583
Attended School Regularly – Few or No Unexcused Absences (Pre-Treatment)	0.3874	0.1809***	0.5272	0.1386***	0.4914
Interviewer's assessment: Very likely that the youth will stay in and graduate from high school or equivalent (Pre-Treatment)	0.1694	0.0957***	0.2193	0.0593***	0.2155

Notes: Maximum number of observations is 35,020; sample size is lower for some variables due to missing information (see text). Asterisks denote statistical significance of the difference between columns 2 & 3 and between columns 4 & 5, as follows: *** p-value ≤ 0.01; ** 0.01 < p-value ≤ 0.05; * 0.05 < p-value ≤ 0.10

Table 2
Effects of Aggression Replacement Training (ART) Eligibility on 12-month School Outcomes

Model	1	2	3	4
Panel A	Outcome: Absent (Dropout, Suspended, or No-show)			
Eligibility Threshold – ART	0.1942***	0.0419	0.0345	0.0146
	(0.0122)	(0.0265)	(0.0272)	(0.0328)
Observations	23389	9708	9484	7872
Panel B	Outcome: Dropout			
Eligibility Threshold - ART	0.0722***	0.0382*	0.0315	0.0306
	(0.0090)	(0.0212)	(0.0222)	(0.0319)
Observations	23311	9657	9493	7894
Panel C	Outcome: GPA			
Eligibility Threshold – ART	-0.3071***	0.0111	0.0466	0.0404
	(0.0307)	(0.0535)	(0.0531)	(0.0701)
Observations	19728	8234	8234	6782
Panel D	Outcome: Transfer			
Eligibility Threshold – ART	0.0375***	-0.0227*	-0.0194	-0.0139
	(0.0093)	(0.0109)	(0.0122)	(0.0154)
Observations	23117	9361	8758	7150
District & Year indicators	Yes	Yes	Yes	Yes
Socio-demographic Covariates	Yes	Yes	Yes	Yes
Extended Covariates	No	Yes	Yes	Yes
Crim. History Score indicators	No	Yes	Yes	Yes
Social History Score indicators	No	Yes	Yes	Yes
Interactions: Social History Score * Criminal History scores (558 cells: see Fig. 1)	No	No	Yes	Yes
Sample	All	All	All	Moderate or High Risk

Notes: Marginal effects from probit regression models are presented for absent, dropout, and transfer; coefficients from OLS models are presented for GPA. Standard errors are clustered at the school district level, and reported in parentheses. Youth who receive FFT, MST or COS are excluded from the sample. All models control for socio-demographics (age, gender, race/ethnicity), and district and year fixed effects. Extended covariates include pre-screening school outcomes (GPA, suspension, eligibility for Section 504 services, and special education), linear and quadratic measures of the relevant scores (aggression, attitudes/behavior, social skills, and family instability) which determine eligibility for the particular treatment program, as denoted in Figure 2, and detailed components of criminal and social history (indicators for age of first offense, categorical indicators for the number of prior misdemeanor referrals and for prior felony referrals, indicators for the number of prior sexual misconduct misdemeanor referrals and for prior such felony referrals, categorical indicators for prior detention and for prior Juvenile Rehabilitation Administration-JRA confinement episodes, history of and current pro-social/anti-social friends/peers, history of and current involvement in gangs, history of truancy/running away, history of neglect, history of court-ordered out-of-home placement, and history of mental health problems, alcohol and drug problems, and abuse). Asterisks denote statistical significance as follows: *** p-value ≤ 0.01; ** 0.01 < p-value ≤ 0.05; * 0.05 < p-value ≤ 0.10.

Table 3
Effects of Functional Family Treatment (FFT) Eligibility on 12-month School Outcomes

Model	1	2	3	4	5	6	7
Panel A	Outcome: Absent (Dropout, Suspended, or No-show)						
Eligibility Threshold - FFT	0.1943***	-0.0245	-0.0153	-0.0398	-0.0380	-0.0430	0.0340
	(0.0108)	(0.0291)	(0.0289)	(0.0319)	(0.0328)	(0.0427)	(0.0692)
Observations	22179	9890	9659	5823	5823	3347	3664
Panel B	Outcome: Dropout						
Eligibility Threshold - FFT	0.0897***	-0.0362*	-0.0331	-0.0573*	-0.0657**	-0.0551	0.0206
	(0.0076)	(0.0207)	(0.0217)	(0.0307)	(0.0318)	(0.0401)	(0.0533)
Observations	22107	9773	9542	5807	5807	3373	3471
Panel C	Outcome: GPA						
Eligibility Threshold - FFT	-0.3320***	-0.0188	-0.0233	-0.0419*	-0.0382	-0.0640	0.0305
	(0.0282)	(0.0179)	(0.0188)	(0.0252)	(0.0249)	(0.0410)	(0.0488)
Observations	18739	9696	9696	5968	5968	3604	3728
Panel D	Outcome: Transfer						
Eligibility Threshold - FFT	0.0336***	-0.0116	-0.0173	-0.0114	-0.0029	0.0107	-0.0073
	(0.0069)	(0.0200)	(0.0207)	(0.0195)	(0.0201)	(0.0290)	(0.0634)
Observations	21921	9684	9096	5239	5239	2930	3536
District & Year indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Socio-demographic Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Extended Covariates	No	Yes	Yes	Yes	Yes	Yes	Yes
Crim. History Score indicators	No	Yes	Yes	Yes	Yes	Yes	Yes
Social History Score indicators	No	Yes	Yes	Yes	Yes	Yes	Yes
Interactions: Social History Score * Criminal History scores (558 cells: see Fig. 1)	No	No	Yes	Yes	Yes	Yes	Yes
Relevant scores - Cubic	No	No	No	No	Yes	No	No
Sample	All	All	All	Moderate or High Risk	Moderate or High Risk	Moderate or High Risk	Low Risk
						Limited Bandwidth	

Notes: See Table 2. Youth who receive ART, MST or COS are excluded from the sample. Limited Bandwidth restricts the sample to +/-5 surrounding the eligibility threshold on the family instability score (family instability score from 3-12).

Table 4
Effects of Multi-systemic Therapy (MST) Eligibility on 12-month School Outcomes

Model	1	2	3	4	5	6	7
Panel A	Outcome: Absent (Dropout, Suspended, or No-show)						
Eligibility Threshold - FFT	0.2021***	-0.0227	-0.0147	-0.0834*	-0.0830*	-0.1029*	0.0364
	(0.0111)	(0.0294)	(0.0298)	(0.0444)	(0.0442)	(0.0587)	(0.0334)
Observations	20299	9236	9010	3060	3060	1475	5715
Panel B	Outcome: Dropout						
Eligibility Threshold - FFT	0.0872***	-0.0488**	-0.0467**	-0.0991**	-0.0975**	-0.0243	-0.0844
	(0.0078)	(0.0197)	(0.0206)	(0.0458)	(0.0464)	(0.0257)	(0.0633)
Observations	20197	9120	8890	3130	3130	5503	1541
Panel C	Outcome: GPA						
Eligibility Threshold - FFT	-0.3320***	-0.0176	-0.0206	-0.0562	-0.0574	-0.0844	0.0037
	(0.0292)	(0.0188)	(0.0196)	(0.0437)	(0.0440)	(0.0570)	(0.0267)
Observations	17227	9059	9059	3314	3314	1789	5745
Panel D	Outcome: Transfer						
Eligibility Threshold - FFT	0.0332***	-0.0067	-0.0118	0.0064	0.0145	-0.0028	0.0261
	(0.0076)	(0.0201)	(0.0210)	(0.0293)	(0.0310)	(0.0300)	(0.0331)
Observations	20036	9035	8456	2693	2693	5489	1264
District & Year indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Socio-demographic Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Extended Covariates	No	Yes	Yes	Yes	Yes	Yes	Yes
Crim. History Score indicators	No	Yes	Yes	Yes	Yes	Yes	Yes
Social History Score indicators	No	Yes	Yes	Yes	Yes	Yes	Yes
Interactions: Social History Score * Criminal History scores (558 cells: see Fig. 1)	No	No	Yes	Yes	Yes	Yes	Yes
Relevant scores - Cubic	No	No	No	No	Yes	No	No
Sample	All	All	All	High Risk	High Risk	High Risk	Low or Moderate Risk
						Limited Bandwidth	

Notes: See Table 2. Youth who receive ART, FFT or COS are excluded from the sample. Limited Bandwidth restricts the sample to +/-5 surrounding the eligibility threshold on the family instability score (family instability score from 3-12).

Table 5
Effects of ART, FFT & MST Eligibility on 24-month High School / GED Completion

Model	1	2	3	4
Panel A				
Eligibility Threshold - ART	-0.0858***	0.0422**	0.0603**	0.0460***
	(0.0103)	(0.0119)	(0.0203)	(0.0106)
Panel B				
Eligibility Threshold - FFT	-0.0645***	0.0190	0.0256	0.0512
	(0.0067)	(0.0222)	(0.0372)	(0.0389)
Panel C				
Eligibility Threshold - MST	-0.0685***	0.0133	-0.0029	0.1052***
	(0.0070)	(0.0255)	(0.0413)	(0.0258)
District & Year indicators	Yes	Yes	Yes	Yes
Socio-demographic Covariates	Yes	Yes	Yes	Yes
Extended Covariates	No	Yes	Yes	Yes
Crim. History Score indicators	No	Yes	Yes	Yes
Social History Score indicators	No	Yes	Yes	Yes
Interactions: Social History Score * Criminal History scores (558 cells: see Fig. 1)	No	No	Yes	Yes
Sample	All	All	All	ART & FFT: Moderate or High Risk MST: High Risk

Notes: See Tables 2-4. Observations range from 10137 to 1541 as models include more detailed information on social and criminal history.

Table 6
12-month Dropout
Heterogeneity across Gender and Prior (Pre-screening) School Performance

Model	1	2	3	4	5	6	7	8
Sample	Males	Females	Low Grades	High Grades	Irregular Attendance	Regular Attendance	Assessment: Uncertain / Unlikely to Graduate	Assessment: Likely to Graduate
Panel A								
Elig. Threshold-FFT	-0.0742**	-0.1136	-0.0535	-0.0841	-0.0584	-0.0692	-0.0409	-0.0153
	(0.0323)	(0.0769)	(0.0413)	(0.0631)	(0.0392)	(0.0635)	(0.0364)	(0.0754)
Panel B								
Elig. Threshold - MST	-0.0860***	-0.0655	-0.0688*	-0.1212*	-0.0839**	-0.1082	-0.0582*	-0.0459
	(0.0317)	(0.0642)	(0.0392)	(0.0638)	(0.0380)	(0.0665)	(0.0352)	(0.0833)
District & Year indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Socio-demographic Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Extended Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Crim. History Score indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Social History Score indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Interactions: Social History Score * Criminal History scores (see Fig. 1)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: See Tables 2-4. Sample stratified across school performance utilize predetermined measures which refer to the period prior to the treatment screening. Samples are restricted to moderate or high risk (for FFT) and high risk (for MST). Sample sizes range from 480 to 6042.

Table 7
Effects of Treatment Take-up on Dropout and Completion
IV Estimation

Treatment Intervention	ART			FFT			MST		
Outcome	Dropout – 12 months	School Completion - 12 months	School Completion – 24 months	Dropout – 12 months	School Completion - 12 months	School Completion – 24 months	Dropout – 12 months	School Completion - 12 months	School Completion – 24 months
Treatment Take-up	0.0345	-0.0008	0.1151**	-0.0901*	0.0479	0.0761	-0.2390	0.2469	0.5250*
	(0.0498)	(0.0297)	(0.0553)	(0.0500)	(0.0324)	(0.0608)	(0.5838)	(0.2180)	(0.2748)
First-stage: Instruments									
Eligibility	0.4214***	0.4456***	0.4390***	0.3136***	0.3229***	0.3090***	0.0383***	0.0426***	0.0761***
	(0.0277)	(0.0244)	(0.0460)	(0.0188)	(0.0187)	(0.0226)	(0.0132)	(0.0117)	(0.0207)
Treatment not available / Capacity constraint	-0.1518***	-0.1582***	-0.1893***	-0.0942***	-0.1022***	-0.1223***	-0.0148***	-0.0167	-0.0251***
	(0.0088)	(0.0072)	(0.0178)	(0.0076)	(0.0074)	(0.0150)	(0.0044)	(0.0041)	(0.0083)
F-statistic on Excluded IVs	198.77***	304.72***	103.46***	146.72***	152.62***	96.60***	5.66***	8.51***	7.80***
District & Year indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Socio-demographic Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Extended Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Crim. History Score indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Social History Score indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Interactions: Social History Score * Criminal History scores (558 cells: see Fig. 1)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8240	11026	2919	11401	14937	3540	11401	14937	3540

Notes: Coefficients from models estimated via two-stage least squares are reported. Standard errors are clustered at the school district level, and reported in parentheses. See Table 2 for the full list of control variables.

Table 8
Testing Covariates Balance across Eligibility Thresholds
Demographics and Schooling outcomes measured 1-year prior to Arrest and Screening

	Age	White	Dropout	GPA	Transfer	Suspended	Section 504 Particip.	National School Lunch Particip.
Elig Threshold - FFT	0.0517 (0.0449)	-0.0141 0.0212	0.0128 (0.0151)	-0.1051* (0.0549)	0.0158 (0.0249)	0.0088 (0.0236)	0.0056 (0.0067)	-0.0217 (0.0196)
Elig Threshold - MST	0.0500 (0.0478)	-0.0138 (0.0228)	0.0191 (0.0177)	-0.1132* (0.0584)	0.0214 (0.0265)	0.0048 (0.0252)	0.0155 (0.0098)	-0.0152 (0.0209)

Notes: See Tables 3-4. Marginal effects from probit regression are reported for all outcomes except for GPA, wherein OLS coefficients are reported. Standard errors are clustered at the school district level, and reported in parentheses. Models for FFT follow specification 4 in Table 3, and models for MST follow specification 4 in Table 4.