Longer-term Impacts of a Youth Behavioral Science Intervention: Experimental Evidence from Chicago *

Nour Abdul-Razzak† Brandon Domash‡ Kelly Hallberg§ Cristobal Pinto Poehls¶

June 11, 2025

Abstract

We conduct a large-scale, randomized controlled trial of a six-month intervention combining intensive mentoring and group cognitive behavioral therapy (CBT) for youth in Chicago, following study participants for up to five years. The program was designed to engage young people at higher risk of engagement with the criminal justice system, and successfully did so with a take-up rate of 62%. Over 24 months, youth offered the program experienced an 18% reduction in the probability of being arrested, with no impact on number of arrests. We find a significant impact on violence engagement, with a 23% reduction in the probability of a violent-crime arrest within 24 months. We find the program's impact in preventing any arrest persists into adulthood, up to four years post randomization. The program moderately improves school engagement in the first year as well. Sub-population analyses suggests that all youth are benefiting from the program, but that the program may be moving different outcomes for different groups of youth in ways related to baseline risk of engaging in the justice system or disengaging from school. We conclude that programs that combine CBT and mentoring can serve as a model to engaging a harder-to-reach population of youth, predominantly outside of school, and be cost effective in reducing criminal justice contact in the longer run. *JEL codes:* C53, C93, I12, K40, K42

*This is a University of Chicago Crime Lab research project, with generous support by the Sports Alliance, CDC, and AbbVie. We'd like to thank Brightpoint (formerly Children's Home and Aid) and Youth Advocate Programs in making the intervention possible. We'd like to thank Lydia Jessup, Christina Leon, Max Lubell, Heather Bland, Lucia Delgado Sanchez, Emily He, Luke Karner, Chelsea Hanlock, Kyle Pinder, and Mariah Van Ermen for excellent research assistance and program management assistance. Special thanks to Jens Ludwig, Aurelie Ouss, Chris Blattman, Leonardo Bursztyn, Max Kapustin, Sarah Heller and seminar participants at the Transatlantic Workshop on the Economics of Crime, the Virtual Crime Economics (ViCE) seminar, Texas Economics of Crime Workshop, and the America Latina Crime and Policy Network Conference (AL CAPONE) for helpful comments and feedback. Nour Abdul-Razzak gratefully acknowledges support from the National Science Foundation (NSF) Graduate Research Fellowship Program. Research reported in this publication was additionally supported by the Centers for Disease Control (CDC) National Center for Injury Prevention and Control under award number 1-R-01-CE002971 and of the National Institutes of Health under award number 5P01HD076816. The content is solely the responsibility of the authors and does not necessarily represent the official views of the CDC, the National Institutes of Health, NSF, other funders, data providers, or implementing agencies.

†University of Chicago Inclusive Economy Lab and Harris School of Public Policy. Email: abdulrazzak@uchicago.edu, Address: 111 W Washington St, Suite 1023, Chicago, IL 60602.

‡University of Chicago Crime Lab. Email: bdomash@uchicago.edu

§University of Chicago Inclusive Economy Lab and Harris School of Public Policy. Email: hallberk@uchicago.edu

¶University of Chicago Crime Lab. Email: capinto@uchicago.edu

1 Introduction

Frequent exposure to violence and justice system contact can affect youth mental health, academic engagement, future earnings, and political participation (Aizer and Doyle, 2015; Ang, 2021; Ang and Tebes, 2020; Owens, 2017; Hjalmarsson, 2008; Flannery et al., 2004; Cloitre et al., 2009; Margolin and Gordis, 2000; Ford et al., 2008). These costs are disproportionately born by youth of color, and in particular Black youth (Western, 2006; Legewie and Fagan, 2019; Desmond et al., 2016; Lerman and Weaver, 2014; Gelman et al., 2007; Weaver and Geller, 2019). Fortunately, in the last decade, there has been significant progress in identifying programs that reduce violence *and* criminal justice contact. Many of these successful programs seek to address the psychological impacts of repeated trauma exposure by providing mental health resources that use behavioral science principles, such as cognitive behavioral therapy (CBT) (Heller et al., 2017; Landenberger and Lipsey, 2005; Bhatt et al., 2023; Barnes et al., 2017; Blattman et al., 2017; Dinarte-Diaz and Egana-delSol, 2024; Blattman et al., 2023; Arbour, 2021; Batistich et al., 2024). These interventions typically help individuals by slowing down their reactions and providing alternative responses to stressful or challenging situations.[1]

However, to be effective at scale, this kind of programming must enroll young people at risk for engagement in the criminal justice system. Many of the programs that have been studied to date have been implemented in institutional settings (e.g. schools and detention centers) where reaching young people is relatively straightforward. However, the same factors that cause youth to be at elevated risk for violence or criminal justice engagement make them less likely to engage in school (Aizer and Doyle, 2015; Hjalmarsson, 2008; Lochner, 2020). This creates an important challenge for the field: finding a way to get young people who are beginning to disconnect from or are no longer enrolled in school to participate in potentially beneficial programming, without having to wait until they become deeply engaged in the criminal justice system and are detained in an institutional setting. The salience of this challenge has been heightened by the rise in chronic absenteeism following the COVID-19 pandemic. Post-pandemic, nearly a third of students in the

[1] Framed within the standard economic theory of crime model, these interventions target the decision making process itself or the underlying skills, preferences or identities that can impact the types of choices or alternatives that come to mind in a given situation rather than attempting to change the incentives people face when deciding whether to engage in criminal activity (Kahneman, 2011; Ross and Nisbett, 2011; Shah and Ludwig, 2016; Becker, 1968).

U.S. are missing more than 10% of school days (Dee, 2024; OCO, 2025). This growth in absenteeism is not only putting youth at increasing risk of engaging in the criminal justice system, but might have downstream impacts on educational attainment and employment (Liu et al., 2021; Cattan et al., 2023; Lochner, 2004; Lochner and Moretti, 2004; Lochner, 2011). In this context, is it possible to design programming that engages young people in CBT programming without relying on school attendance as a pathway to participation?

A separate, but related challenge to scale is whether CBT-based programs can achieve sustained effects. Most of the successful CBT-based programs in the U.S. find large, but short-lived effects. Helping youth successfully navigate the critical juncture they face between the ages of 16 and 18, with age-crime profiles peaking during this time period, could lead to longer-term positive impacts on employment and human capital accumulation (Heckman and Kautz, 2013; Lochner, 2020; Farrington, 1986).

We study a program in Chicago that aims to meet these challenges. Choose to Change (C2C) is a six month program that combines group-based CBT with intensive mentorship and wraparound supports. The key innovation of the C2C program is the use of an intensive "advocate" or mentor to recruit and retain the youth to participate in and *practice* the CBT programming. This unique combination of services is designed to better engage youth who are marginally connected to school and at elevated risk of criminal justice involvement. The majority of programming occurs outside of the school building and mentors utilize a strength-based approach to engage with program participants, identify their needs, and develop an individualized support plan. Once a stable relationship has developed, the mentors and youth begin attending group CBT sessions specifically designed to help young people who have experienced trauma build the social and emotional skills needed to cope with stressful situations and reduce conflict. Beginning in 2015, we designed and implemented a large-scale randomized controlled trial (RCT) to study the impact of the C2C program, assigning successive rolling cohorts of youth to be offered C2C services (treatment) or not (control) with a total study sample of 2,074 youth. Critically, we track outcomes up to five years post randomization to examine the longer-term impacts of the program as youth transition into adulthood.

The program was successful in serving a range of youth at heightened risk for criminal justice system engagement, including a sizable number of young people who would likely not have been reached by a strictly school-based program. All study youth for the RCT were drawn from neighborhoods in Chicago experiencing the highest levels of violence and youth involvement in the criminal justice system. Half of program referrals came from non-traditional school sources, including alternative high schools, the district's student outreach and re-engagement center which works with students who have dropped out of school, and juvenile probation. Prior to program referral, 50% of youth had been a victim of a crime, 35% had a prior arrest, and 30% were disengaged from school, with two-thirds of the sample falling into at least one of these risk factors.

Of the 1,052 youth offered the program, C2C was able to successfully engage 62%, a take-up rate higher than many in-school programs in Chicago. Even among youth who are are marginally or completely disconnected to school, we observe a nearly 50% take-up rate. The program engaged youth for an average of eight hours a week of mentoring and wraparound supports designed to address each young person's unique needs and goals (e.g. food security, consistent transportation to school) and provide opportunities for social outings with other youth. The average participant in the program attended nine CBT sessions focused on helping them understand how prior trauma can influence their decision-making and develop the skills to regulate emotions and resolve conflicts. Program staff and youth alike cite the relationship developed with the mentor as critical to driving CBT engagement. Importantly, the mentor attends the group CBT sessions with program participants and provides opportunities for youth to *practice* the new tools they are building on a regular basis outside of the formal CBT sessions. The "learn by doing" focus of the program is designed to help participants build and retain the confidence, skills, and habits to navigate stress and conflict safely in the longer-term after the program ends.

Looking at our pre-registered primary criminal justice outcome, total arrests, we find that C2C has a measurable and sustained impact on whether a young person was ever arrested post-randomization yet no statistically significant impact on the intensive margin, number of arrests.[2] We find that being offered C2C substantially reduces the probability of any arrest by 6.3 percentage points (18% of the control mean) during the 24 months after randomization (p-value < 0.01 after adjusting for multiple hypotheses). This translates to a treatment-on-the-treated effect of 10.3 percentage points (31% of the control complier mean). Point estimates for number of arrests, are negative but noisy and not statistically significant at conventional benchmarks during any time period. The large and significant extensive margin impacts continue to persist up to 48 months post

[2]Please see `https://www.socialscienceregistry.org/trials/933` for the pre-analysis plan.

randomization, even after adjusting for multiple hypotheses. At 48 months post randomization, C2C continues to reduce the probability of any arrest by 5 percentage points (12% of the control mean). Notably, over the same time period, we see no change in program participants' likelihood of being stopped by the police. This suggests that the change in likelihood of arrest is reflective of a real change in behavior rather than a change in their ability to avoid arrest or in where they are spending time.

Importantly, the program significantly decreased arrests for violent offenses. During the course of programming, being offered C2C reduces the number of arrests for a violent offense by roughly 36% of the control mean, resulting in roughly 2.3 fewer violent arrests per 100 youth after six months (p-value < 0.05 and $< .10$ after adjusting for multiple hypotheses). We also see these reductions in arrests for violent offenses persist in the years after the program ends. In the 24 months after randomization, we find that being offered C2C reduces the probability of any arrest for a violent offense by 3.7 percentage points (23% of the control mean, with p-value $< .05$ after adjusting for multiple hypotheses). Along the intensive margin at 24 months, the longer-term estimates for number of arrests for a violent offense are slightly noisier (18% reduction of control mean, with p-value < 0.10, $< .38$ after adjusting for multiple hypotheses).

We also find evidence that the program led to a modest increase in engagement in school, as measured through a pre-registered index that combines the number of days attended, an indicator of whether a student was enrolled in coursework or had graduated, the number of misconduct incidents, GPA, whether the student was enrolled in a school within a juvenile justice facility, and the number of schools attended. We find that being offered C2C leads to an increase of 0.03 and 0.05 standard deviations in this combined school index in the first and second semesters post-randomization, respectively. Looking at the effects on the individual components, we find that most of this is driven by positive behavior change in school (e.g. a reduction in the number of misconducts) rather than through improved academic performance (e.g. GPA). The program had no effect on participants' likelihood of graduating from high school. This pattern of results is consistent with the program's focus on building participants' social and emotional skills, rather than a focus on academic remediation.

The program's aim to serve a different group of youth than those who are typically reached by school-based programs raises questions about how treatment effects differ for these groups of youth and whether the program's effectiveness could be increased through better targeting given the range of youth the program served. To answer these questions we conducted several exploratory subgroup analyses to look at treatment effect heterogeneity. We find suggestive evidence that all youth are benefiting from the program, but that the program may be moving different outcomes for different groups of youth in ways that may be related to baseline propensity to experience a given outcome. We find evidence that the program is more effective in reducing violence for youth at elevated risk for criminal justice engagement: youth with prior arrests, males, and youth not engaged in school. Conversely, we find the program is effective in reducing number of arrests for youth with no prior arrest and youth with fewer baseline prior arrests. For education outcomes, we find larger school engagement improvements on the academic margins (including increasing graduation rates) for youth who were more engaged in school at baseline.

To assess the net social value of the C2C program, we assign monetary values to each observed C2C arrest and victimization event, weighting each by the estimated cost they impose on society. We find that C2C maximizes savings to society after three years, and saves between \$6,800 and \$20,200 per participant ($p < .05$ and $p < .10$, respectively) in the form of reduced contact with the criminal justice system. This alone is enough to offset the costs of the program, although these estimates likely represent a conservative lower bound, as they exclude educational gains and rely solely on observed arrest and victimization data.

Taken as a whole, our paper makes three main contributions. First, this study extends the existing CBT literature to demonstrate this kind of programming can engage and impact young people outside of institutional settings. The program was effective in enrolling some young people who were disconnecting or disconnected from school and in reducing the likelihood that they would be arrested at all and that they would be arrested for a violent offense. Qualitative data collected from both program participants and program staff suggest that the mentors play a critical role in in overcoming youth hesitation to participate in the program overall and in the therapy sessions in particular (Lubell et al., 2024). Given the mixed evidence on the value of wraparound programs, this study also shows the promise of wraparound support services when thoughtfully used as they likely acted as an incentive to join the program and reduced barriers to both program participation and school attendance (Doleac, 2019).

Our findings also highlight the consistency of the effectiveness of CBT-based programming. Our short-term effect sizes (specifically violent-related arrest impacts) are similar in magnitude to what previous studies have found (Heller et al., 2017). While at least a sizable portion (30%) of the C2C study population are youth disconnected from school or would be difficult to reach by traditional school programs, one could also describe the C2C population as a linear combination of past-studies: school engaged youth, disconnected youth, youth in detention, and adults at high risk of engaging in violence (Bhatt et al., 2023; Heller et al., 2017; Blattman et al., 2017; Arbour, 2021). In this respect, highlighting the replicability of CBT-based programming (with a different program structure) helps move this type of intervention from "promising" to a more proven practice that leads to the behavior changes needed to tackle difficult issues like violence and crime (Heller, 2022; Ludwig, 2025).

Second, unlike other CBT programs studied in the U.S. context, we look at longer-term outcomes and find sustained effects up to four years after randomization. The program served young people when they are at or near their peak risk for being arrested, and decreased this risk for a sustained period of time, playing a critical role in helping many young people make the transition into adulthood without criminal justice system involvement.[3] The longer-term impacts of this program, with all study youth within the study outcome window becoming adults, are especially promising as many youth programs see fade-out shortly after program end and few studies have been able to track youth into adulthood (Heckman and Kautz, 2013; Heller, 2014; Davis and Heller, 2020).

We see these sustained effects, long past when young people are actively engaged in the program, coupled with the finding that the program did not change the likelihood that young people were stopped by police, as evidence that the program led to sustained behavioral changes. Based on the qualitative data we collected from program participants and staff, we hypothesize that this sustained behavioral change may have been facilitated by the opportunity to practice the skills gained in CBT – both through engagement with the advocate who can emphasize these skills outside of group, but also by the wraparound supports that help meet young people's basic needs giving them the bandwidth for practice. This is consistent with literature in other contexts that

[3]C2C engaged youth at a time of elevated risk for involvement in the criminal justice system: in the 12 months prior to randomization, youth in the control group had experienced an average of 63.8 arrests per 100 youth, the highest yearly average we see among youth in the study up to five years post randomization.

have found that practice facilitates sustained behavior change (Blattman et al., 2023; Wilson, 2004; Beck and Beck, 2011).

Finally, these findings deepen our understanding of the promise and limitations of CBT programming in improving educational outcomes. The program modestly increased participants' engagement in school, primarily by reducing disciplinary incidences but did not impact high school graduation on average. Taking into account some heterogeneity in our education impacts, these findings contribute to a growing literature that finds that for young people who are very behind in school, behavioral interventions alone may be insufficient to make-up the ground necessary to succeed academically (Guryan et al., 2021).

The paper is organized as follows: first we discuss the details of the Choose to Change intervention and the theory behind how the intervention was designed to support young people. Next, we discuss the study population and the data used in the project. We then describe the methods behind the RCT and the main RCT findings. Lastly, we dive into some heterogeneity analysis and a discussion of findings.

2 Choose to Change: Program Model

In response to a 2015 Design Competition, an initiative launched by the University of Chicago Crime and Education Labs to crowdsource ideas for youth violence prevention interventions from across the city of Chicago, two local Chicago nonprofits, Brightpoint (formerly Children's Home & Aid) and Youth Advocate Programs (YAP), Inc. created Choose to Change (C2C): Your Mind, Your Game.[4] C2C combines individualized wraparound services, intensive mentoring, and group-based, trauma-informed therapy. This combination support is designed to help youth understand how past traumatic experiences and chronic stress can impact their thinking and behavior and how this, in turn, affects their emotional response to situations they encounter in daily life.

Over the course of the six-month program, mentors (known as advocates in the program) from YAP meet with youth individually for at least eight hours a week, building strong interpersonal bonds and offering wraparound and mentoring services focused on meeting each person's goals and addressing their specific needs. In addition, therapists from Brightpoint lead trauma-informed

[4]The 2015 Design Competition was held in partnership with GET in Chicago, a local philanthropic organization and the MacArthur Foundation.

CBT sessions called SPARCS (Structured Psychotherapy for Adolescents Responding to Chronic Stress), which helps youth process prior trauma and develop a new set of decision-making skills, helping youth reduce their reliance on maladaptive responses and behaviors. Program staff identify three core program components: relentless, strength-based engagement and strong relationships; developing decision making skills; and applied learning.

The bundled treatment costs about $4,500-7,600 per participant at the time the program was implemented, with the cost increasing over time in response to increasing wages. Because we study the complete bundle of program services provided together, we are only able to identify the effect of the program as a whole and cannot disentangle the relative contribution of the CBT or mentoring components, or/and synergies generated by implementing these program components together. We briefly describe each of these program components below.

2.1 Strength-based engagement and strong relationships

Youth are referred to the program from either a community-based organization, public agency, or their middle or high school based on the referral partners' assessment of youth risk for engagement in violence or the criminal justice system or at risk for school disengagement. C2C advocates then connect with the youth and their family members to introduce the program and obtain consent to participate. Advocates are persistent in their recruitment and engagement of young people, and try multiple tactics to locate youth including using their connections in the community. Youth are often told they have been referred to this program based on their *potential* for change and future success. Their past behavior issues are not highlighted, and instead program staff hope this re-framing allows youth to see this program as an opportunity, not a form of punishment.

Services formally begin with the advocate convening a meeting that includes the youth, their parents, and other family members or individuals who might be able to support the youth. Through this meeting, the advocate develops a picture of the youth's life and how they can advocate on the youth's behalf to help them reach the best version of themselves. The advocate, with the cooperation and input of the family and youth, then draws up a service plan (including action steps and goals) based on the improvement areas and strengths identified. These goals could include a young person's desire to get employment, address legal challenges, help secure stable housing, or realize educational aspirations.[5] Because the wraparound services are tailored to the individual, which services are provided is dependent on the needs of the youth and their family.

This strengths-based approach continues throughout a young person's engagement in the program. C2C employs a "No Reject, No Eject" policy that dictates that youth are not denied or discontinued from services in response to new challenges, or for non-compliance with program policies. Rather, C2C staff work to adapt services to meet their specific needs and situations. Throughout the intervention, the advocate engages in one-on-one meetings with the youth, family meetings, and weekly group recreational activities. The group activities tend to be fun experiences, such as playing basketball at a local gym, going to see movies, or going out to eat.[6] These interactions also serve as an opportunity for advocates to get to know their clients and build trust. Advocates provide around-the clock support to the young people they serve, communicating frequently via text message or phone calls between formal outings. Advocates and youth share similar backgrounds and often come from the same neighborhoods. This helps youth develop functional and trusting relationships that are key to helping them continue to engage in the program and the CBT sessions. Advocates serve between five and 15 youth at a time to support relationship building.[7] The elements of C2C provided by the advocates/mentors are summarized in Table A.I in the Appendix.

2.2 Developing decision making skills

After a month has passed, advocates introduce youth to the therapy component of the program. This iterative roll-out of program components ensures that trust has developed between the young person and the program staff before they are asked to participate in CBT sessions. The C2C program utilizes the SPARCS curriculum to support youth in developing the skills needed to resolve conflicts and regulate emotions (De Rosa et al., 2004). SPARCS is trauma-informed and aims to help youth understand and change the way stress or past traumatic experiences can influence

[5] The model is rooted in the belief that all youth, adults, and families have strengths that can and should be developed. The principle of strength-based services encourages teams to create goals that reflect building family and youth assets, capacities, and resilience.

[6] During some cohorts, YAP also referred and encouraged many participants to engage in summer job programs ("One Summer Chicago") sponsored by the City.

[7] Program staff highlighted that even after the program ends officially, some youth continue to reach out to their mentor/advocate for support if crises arise.

decision-making. Typically, groups of eight to 10 youth will attend 12-16 sessions, 45-60 minutes each. Once sessions start, they are held once a week and typically scheduled in school (to facilitate attendance) or at Brightpoint's community facilities. SPARCS targets four domains of functioning: regulating emotions and behavior, attention and awareness of self, relationships, and hopefulness and sense of purpose in life. For example, C2C sessions often start with a check-in or "temperature-check" where youth are asked to slow down and take a deep breath, look around and be present in the moment, and then self-check how they are feeling (how distressed they feel on a scale of 1 to 10, and how in control they feel on a scale of 1 to 10). These practices help youth become aware of how their thoughts and feelings might affect their reactions and choices.

Many skills are taught through role-playing a conflict, real or hypothetical. In these sessions, therapists highlight existing coping skills that youth might use that could temporarily make them feel better, but later lead to more issues like using drugs, isolating from others, or getting into fights. During sessions, therapists highlight how "emotional leftovers" from traumatic or highly stressful experiences can make youth more willing to engage in these short-term coping strategies that perpetuate a cycle of distress or escalate conflict (Van Dijk, 2013). Youth learn how to articulate what their goals are in a given situation and what options they have that could support them in a positive way. Through various exercises and conversations, youth learn to better regulate their emotions, engage helpful coping strategies, and build problem-solving and communication skills. SPARCS combines elements from traditional CBT and Dialectical Behavioral Therapy (DBT), a form of CBT that incorporates more mindfulness and acceptance techniques.[8] See Table A.II in the Appendix for a more detailed summary of the different topics covered in these sessions.

2.3 Applied learning

One of the unique aspects of the C2C program is the ability and focus on practicing the CBT/SPARCS lessons in real life. The program model creates many synergies that help youth reinforce their new skills. First, during the program, advocates are required to attend the SPARCS sessions with the

[8]The curriculum incorporates core elements of treatments for youth experiencing trauma, including psychoeducation (developing an understanding of how people react to trauma), relaxation and emotion regulation skills, mindfulness, and cognitive skill building (Santiago et al., 2018). An important element of SPARCS is removing judgment of current behaviors that youth are relying on to manage their emotions (Van Dijk, 2013). The program manual reminds the therapists that even maladaptive coping skills are coping skills, and youth are doing the best they can with the skills they have.

youth. Advocates then practice the SPARCS skills in their individual and group interactions with youth out in the community. This allows for "learning by doing" and creates opportunities for youth to develop new habits that will benefit them after the program ends. Furthermore, since the advocates interact regularly with program participants in various settings, they can highlight key moments when CBT techniques may be useful. Youth often call or text their advocate if a stressful or safety issue arises. Advocates can use these opportunities to remind the youth of helpful coping strategies and the different choices they have in how they react. This is strengthened as the therapists and advocates try to maintain a close relationship to track the progress of the youth. Second, the CBT sessions are consistently presented to youth in a non-clinical framework to destigmatize therapy. For example, group-based therapy sessions are led by Masters-level clinicians, but they are referred to as "coaches" rather than therapists to ensure that they are approachable to youth who may have misgivings about participating in therapy.

3 Eligibility and Study Population

Between November 2015 and December 2019, the study team worked closely with program staff to identify and recruit eligible youth who could benefit from C2C. Through these efforts, we identified a study sample of $2,074$ young people across eight cohorts.[9] Youth were drawn from neighborhoods on the south and west sides of Chicago. Because C2C advocates frequently travel to participants' homes and schools, geographic boundaries were drawn at the start of the program to limit the distance advocates had to drive to reach youth. Figure A.I shows a map of where study youth reside based on their address of residence at the time of referral.

The neighborhoods from which the study youth are drawn have been exposed to challenging conditions, including but not limited to segregation, poverty, disinvestment, and experience some of the highest rates of violence and police contact in the city. Figure A.II shows a map of the 5-year average number of shootings by community area (2015 to 2019), and Figure A.II shows the number of complaints filed against the Chicago Police Department (CPD) in the last three decades (using data from the Invisible Institute). Both maps show high concentrations in the south and westside neighborhoods where our study youth reside.

[9] 6 youth never attended Chicago Public Schools and an additional 3 youth had already graduated prior to randomization, and thus are removed from all educational analyses.

For all cohorts, in the months prior to randomization, C2C staff conducted outreach to community service providers and schools in the neighborhoods they served to identify young people who were eligible to participate in the program. Because the purpose of the program was to reach young people at risk of dropping out of school, interacting with the criminal justice system, or engaging in violence, referral sources were asked to identify young people between the ages of 13 and 18[10] who were actively affiliated with gangs or at risk or gang engagement; on juvenile probation; previously found guilty of weapons offenses; seriously disruptive at school through chronic truancy, serious misconduct and/or frequent suspensions; and/or direct victims of or witness to traumatic violence. The program providers hoped to engage youth who may not fully benefit from existing in-school services, either because the young person was not in school consistently or because they had other (often basic) needs that need to be met first. Referrals came from Chicago Public Schools (CPS) - including neighborhood and charter schools (50%), alternative schools (18%), the Student Outreach and Re-engagement Centers (20%) which focus on re-engaging chronically truant and out-of-school youth, and the CPS Office of Safety and Security (6%) - as well as Cook County Juvenile Probation (5%). The administrators identified young people in need of services based on their internal data and personal knowledge of whether the young people met the referral criteria. Program providers preferred this approach to a data-driven approach (such as a risk score) because it took advantage of the knowledge and relationships that school administrators had with youth.

Randomization and enrollment occurred on a rolling cohort basis over four years, with all youth ensured a minimum length of programming. For the first four cohorts, services lasted five months. This expanded to six months in cohort 5 as more funding became available. Randomization and subsequent enrollment continued until a cohort had reached capacity.[11] We use the date of randomization as the start of the outcome-tracking period, but recognize that treatment youth may have actually enrolled in the program in subsequent weeks or even months. The majority of the youth in the last cohort of programming (cohort 8) were receiving services during the early days

[10] Although the program targeted 13-18 year-olds with respect to the program letters and advertisement sent out to the referral agencies, occasionally youth outside of this age range were served if age was not available prior to randomization.

[11] Typically 100 participants made up one cohort, although this varied cohort to cohort depending on funding and decreased over time with the program serving about 50 to 60 youth per cohort in later years of the study.

of the COVID-19 pandemic.[12] Across all eight cohorts, 1,052 youth were randomized to treatment and 1,022 youth were randomized to control. Randomization was stratified by referral source and cohort.

Table I shows that the program succeeded in identifying youth in the target population. Referred youth were, on average, 16 years old and in the first few years of high school. About 95% of the youth were Black and almost all youth qualify for free or reduced lunch. A significant percentage had been previously arrested (35%), and 20% had a gap in school enrollment at some point during the prior three semesters. In the prior semester, the attendance rate was about 73%, highlighting that study youth are missing on average more than a quarter of school days. GPA is hovering is the below C range – with average GPA around 1.94. Approximately half of the youth in the study had experienced a previous victimization incident. Comparing these baseline characteristics with the average high school student in CPS, we find that C2C youth are at significantly increased risk of engaging in the criminal justice system. On average, only 3% of high school students in CPS attending traditional schools have a prior arrest and 12% have a prior victimization (Lab, 2021)[13].

To better understand the sample's risk of criminal justice system engagement in the absence of intervention, Figure A.III displays the control means of total arrests in six month intervals leading up to and following randomization. We see that the risk of being arrested peaks in the year prior to randomization, with a mean of approximately 0.3 or 30 arrests per 100 youth in the six-months interval just before randomization. After randomization, we see that the risk of arrest remains relatively stable for about 12 months (mean of 0.25 arrests or 25 arrests per 100 youth) and then begins to gradually drop off. Within five years, almost 44% of the control group has an arrest with an average of 1.6 cumulative arrests (or 160 arrests per 100 youth). For violence (see Figure A.IV in Appendix), we see a similar pattern with arrest risk also peaking in the six-month interval prior

[12]Given that most of the cohort 8 randomization and recruitment happened prior to the pandemic, we do not see significant impacts on the take-up rate during this cohort. However, programming changed dramatically with most services moving online. Necessities such as groceries were still delivered to families in person, but mentoring and therapy services were primarily conducted virtually, with some outdoor, in-person services offered to supplement the online engagement.

[13]While this comparison is limited to a public report that uses data from traditional high schools in CPS during the school year 2018-2019, only about 10% of CPS students attend non-traditional or alternative schools. Another comparison that can be done is compare directly the C2C male study population to (Heller et al., 2017) given the similarities in program structure and geographic coverage (Becoming a Man or BAM only worked with males). While juvenile arrest rates have fallen by more than 50% in the last decade, we still find on almost every measure, the C2C males are more disconnected from school and have greater prior involvement in the criminal justice system compared to the BAM study (and compared to the females in the C2C study), see Table A.XVI for more information

to randomization; while post-randomization risk peak happens within 18 months. Combining the baseline characteristics of the sample with this projection of "realized risk" suggests that not only did the program successfully engage youth in high need of this program, they also did so at a time of heightened risk for criminal justice system engagement.

Randomization was successful in balancing the treatment and control groups across almost all baseline characteristics (see Table I). An F-test for joint significance shows we cannot reject the null hypothesis that treatment and control groups are equal. We recognize a few characteristics are marginally significant in the difference between treatment and control groups (number of prior violent arrests, any prior drug arrest, and prior attendance rate). Given the large number of baseline characteristics we look at, it's very possible to find a few characteristics that are not perfectly balanced by chance. Given the results of the F-test, we are confident in the comparability of the two groups, but we control for these baseline characteristics in our regressions to account for any residual group differences.[14]

Lastly, for approximately 44% of the youth who participated in the program, we have baseline data from mental health assessments the program providers collected during the first CBT/SPARCS session. Given we do not have these statistics for the full study population, these characteristics are not included in the balance tests or later in our impact analysis. However, they still help us understand the characteristics of the young people the program was able to engage. Of the 289 youth who filled out the baseline mental health assessment, 92% of the youth reported experiencing at least one traumatic event in their life. The type of traumas that were most often shared by youth fit in the category of "community and family violence", with 50% of youth noting that they have been told about someone's killing or injury, 40% of youth having seen or heard someone being badly hurt, 45% having been hurt or injured themselves and 39% having seen someone dead or dying or watched or heard someone being killed. These sobering statistics highlight the need for the trauma-informed supports in the C2C program.

[14]We specifically control for the characteristics for which we have non-missing baseline data for all youth – demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations.

4 Data and Outcomes Measured

4.1 Criminal justice outcomes

To understand how C2C impacts contact with the justice system, we matched our study youth to CPD arrest and victimization data using a probabilistic matching algorithm based on first name, last name, date of birth, and home address.[15] Our pre-registered primary criminal justice outcome is criminal activity as measured by the number of arrests for the full study up to 24 months post randomization, looking at outcomes at 6 month intervals.[16] We show both the number of arrests after randomization, and whether or not a youth had any arrest after randomization. Given that the number of arrests variable is skewed with a mass at zero and the high variation in this measure for this study population, an analysis of the average change might miss important changes in whether youth had any engagement in the criminal justice system through being arrested.[17] We believe this slight deviation from the pre-analysis plan is necessary as both margins are important to examine given any arrest can have negative and costly consequences for youth and society (Aizer and Doyle, 2015; Western, 2006; Legewie and Fagan, 2019; Desmond et al., 2016; Lerman and Weaver, 2014; Gelman et al., 2007; Weaver and Geller, 2019).

To unpack the criminal justice results further, we examine our pre-registered secondary outcomes by examining the impact of the program on arrests by the charge associated with each arrest. To understand the types of arrests impacted, we follow the existing literature and break down number of arrests and whether a young person has an arrest by charge type (violent, property, drug, or other), with the "other" arrest category consisting of arrests for violations such as trespassing, disorderly conduct, weapons violations, vandalism, warrants, etc.[18] Violent arrests comprise about

[15]The Appendix discusses details of the matching procedure. Our project through the University of Chicago Crime Lab has a master data sharing agreement with the Chicago Police Department and Chicago Public Schools. The CPD data was provided by and belongs to the CPD. Any further use of the data must be approved by CPD. Points of view or opinions contained within this document are those of the authors and do not necessarily represent the official position or policies of the Chicago Police Department.

[16]AEA's RCT registry: `https://www.socialscienceregistry.org/trials/933`.

[17]The standard deviation of the number of arrests outcome in 24 months post-randomization is 1.75. Later in the paper, we also highlight that we have some outliers in our sample that have accumulated a lot of arrests (both at baseline and in the post period) that increase the variance of our point estimates for number of arrests.

[18]Arrests are classified using FBI codes and statute descriptions. Violent arrests comprise of homicides, sexual assault, robbery, aggravated assault, aggravated battery, simple assault, simple battery, sexual misconduct, and some violent disorderly conduct and violent miscellaneous statutes. Property arrests comprise of burglary, larceny, motor vehicle theft, arson, and stolen property. Other arrests are all arrests that do not fall under violent, property, or drug classification. Each arrest category is mutually exclusive.

24 percent of the outcome study arrest sample (within 24 months post-randomization), property arrests make up about 26 percent, drug arrests 7 percent and other arrests make up almost 43 percent of the arrests in the sample.[19] Treatment and control group differences are assessed every six months post randomization. At present, enough time has elapsed to track the entire sample of 2,074 youth up to 60 months post randomization. Examining these time intervals provides insight into whether the program primarily succeeded due to incapacitation–by keeping the youth busy during the first six months post randomization while the youth are in programming–or if the effects persist once the program has ended.

One important note about our outcome data is that we rely on administrative CPD data. If a youth is not matched to the CPD data they are assumed to have "zero" arrests, victimizations, or stops. However, if youth move out of Chicago, this could undercount the true number of arrests. This would be particularly concerning if this measurement attrition was differential across treatment condition. To assess whether study youth moved out of Chicago at differential rates, we use school enrollment data given that we have baseline CPS data for 99.8% of all study youth. CPS records all youth who transfer, and in particular note if they have transferred outside of Chicago in enrollment data. Using this indicator, we find that the rates of moving out of the city are small even up to 36 months post randomization and no evidence of a differential rate of departure for those assigned to treatment or control groups (see Table A.XXVII in the Appendix).[20]

4.2 Education data

To understand the effect of the C2C program on young people's education outcomes, we connected our study sample to administrative data collected by Chicago Public Schools (CPS). For most of the sample, we received a unique student ID identifier for each referral that allowed us to directly match each study member to all CPS administrative records. For a small number of study members, we

[19] In a companion paper, we develop an alternative method of classifying arrests that takes into account police discretion to better understand the types of arrests youth have more control over in changing (Abdul-Razzak et al., 2025).

[20] 42 months is the longest time outcome period we have CPS enrollment data for all study youth. Specifically, we find that in the 36 months after randomization, study controls spend an average of 74 days not in Chicago (around 7% possible days). Given that youth can move out of Chicago temporarily and then move back into CPS, we believe looking at the total number of days they spend not enrolled in Chicago is a more accurate measure of this censoring issue. Looking at just any transfers out of Chicago, we find that about 13% of the control group is moving out of Chicago at some point in the 36 months post-randomization.

used a probabilistic matching algorithm based on first name, last name, and date of birth to link them to a valid student ID number.[21]

Our primary, pre-registered education outcome is school engagement as measured by an index which was intended to include attendance, a CPS indicator of whether a students was "on-track" to graduate, grade point average (GPA), course-taking, disciplinary incidents, test scores, graduation, school switching, and enrollment in schools within juvenile justice facilities. During the study period, CPS paused and then adjusted the way that "on-track" to graduate was measured. As a result, we do not have a consistent measure of this variable that could be incorporated in the index. Further, given the variation in assessment schedules and in the age and grades of students in our sample, we found that we were missing standardized assessment data for nearly 75% of our sample within the first year of randomization alone. We therefore also excluded this variable from the index. As a result, our final school academic index included: number of days attended, an indicator of whether a student was enrolled in coursework or had graduated, number of misconduct events, GPA, whether the student was enrolled in a school within a juvenile justice facility, and the number of schools attended. Each index component was standardized using the control group mean and standard deviation, and reverse coded if appropriate so that a positive score is always associated with greater school engagement. The scores were then averaged to create an overall standardized index. We examine the effect of the program on the index at semester intervals for up to six semesters post-randomization and explore the effect of the program on individual index components over the same interval.

Because many of the youth in the sample were tenuously connected to school at baseline and in the subsequent semesters (see Figure A.IX), it was important to carefully address missing data in the variables that were incorporated in the school index. Data could be missing because a student moved out of the district, graduated, or dropped out of school and the reason for missingness could plausibly be associated with the effect of the program. To take this into account, we implemented three approaches to impute missing values: (1) Our *basic* approach includes minimal imputation, instead calculating the index for each individual based on the non-missing components. (2) The *mean* imputation approach imputes missing attendance, GPA, and misconduct semesters with

[21] As we mentioned above, we were able to locate 99.7% of our study sample in the data collected by the school district. A total of 9 study members who could not be matched to CPS administrative data or who had already graduated from a CPS school prior to randomization were excluded from all education analyses.

cohort, treatment status, and semester means. (3) The *mixed* imputation method, our preferred approach (and shown in the main table results), attempts to make a best guess based on the available information. Under this approach, if a student is identified as having dropped out in a given semester, we impute zero for attendance and GPA and use the treatment status means for misconduct. If the student is missing data for a reason other than having dropped out (for example - transferring to a non-CPS school where we are no longer able to track outcomes), we mean impute attendance and GPA, and zero impute misconduct. Findings were generally consistent across imputation methods (see Appendix Table A.IX).

Unlike with criminal justice outcomes, we defined educational outcomes based on the number of school semesters relative to randomization given the structure of the education data. We define each youth's first outcome semester as the first full semester after being randomized. The amount of missing data increases with time since randomization as a greater share of the sample has graduated, transferred, or dropped out (See Appendix Figure A.IX). For this reason, in the body of the paper, we focus on education outcomes in the two semesters post randomization where the majority of the sample remains actively engaged in school to minimize our reliance on imputation.

Our final intended pre-specified outcome was employment. Unfortunately, CPS stopped collecting social security numbers from students, which are required to link to employment data in Illinois. As a result, we were not able to examine the effect of the program on employment.

4.3 C2C program take-up and dosage

Table II provides a summary of program take-up and dosage. Of those assigned to treatment, 62% enrolled in the program.[22] We define take-up as participating in at least one trauma-informed CBT session (SPARCS) or receiving at least five hours of mentoring/advocacy services. We find C2C's take-up rate to be marginally higher than other similar programs in the Chicago area despite serving a relatively disadvantaged population.[23]. In fact, the C2C program was able to engage young people who were experiencing risk factors for criminal justice system engagement and/or

[22]There are four youth assigned to the control group who engaged in C2C services, for a total of 657 youth in the program during this time period.

[23]Comparing C2C to the Becoming a Man (BAM) program, a school-based CBT program in Chicago, they saw a take-up rate of 50% (defined as attending at least one CBT session) (Heller et al., 2017). Another program, READI, reaching a higher-risk population of men at risk of engaging in gun violence in Chicago, saw a take-up rate of 55% (Bhatt et al., 2023)

school disengagement at relatively high rates. Of the 206 youth assigned to treatment who had a gap in school enrollment prior to randomization, 45% took up the program. Similarly, 51% of youth who had a baseline arrest enrolled in the program. Dropping out of school, or having a gap in enrollment is often considered the last step in disengaging from school. We also consider an intermediate measure of school disengagement that takes into account low attendance, failing grades, and whether a study member was too old for their grade. Using this school disengagement measure, we find that about 30% of the study population is disengaged from school at baseline, and C2C is still able to engage about 47% of these youth in the program.[24]

Youth who enrolled in the C2C program on average spent *186 hours* or eight hours per week receiving services and almost all of the young people who were served by the program received both mentoring support and attended CBT sessions (82% of those who enrolled in the program engaged in both program components). On average, participants spent nine hours in CBT sessions and 177 hours with their mentor (either one-on-one or in a group setting) over the course of the program. C2C programming is significantly more intense than many existing programs for youth. For example, the school-based program, Becoming a Man (BAM), averaged 13 hours of service for participants in the first *year* of the study and 17 in the second year (Heller et al., 2017). Youth enrolled in the national U.S. mentoring program *Big Brothers, Big Sister*, on average receive an hour a week of services for up to a year. By contrast, even those in the 25th percentile of C2C engagement attended four CBT sessions and received *126* hours of mentoring support. The distribution of dosage in CBT does indicate that youth with more risk factors on average engaged less in CBT, reinforcing the need for the wraparound support services to engage these young people in the program.[25]

[24] We measure school disengagement by youth who missed more than 60% of school days and received a grade of "F" in at least 75% of their courses, had a serious learning disability, were at least two years older than expected for the grade they were enrolled in using administrative CPS data, missing completely from CPS data during their baseline semester, or were inactive at the end of any of the 3 semesters immediately prior to randomization. This measure of school disengagement has been used in other school-based programs in Chicago, such as Becoming a Man (BAM), to determine eligibility based on the rationale that a student would have to be more engaged to benefit from an entirely school-based intervention.

[25] Alternatively, we could have defined participation as an individual who receives at least one CBT/SPARCS session *and* five hours of mentoring. This would have resulted in a take-up rate of 51%. Given the program model and how mentors are also reinforcing CBT skills (in the 8 hours a week on average they spend with youth), we prefer to use the more inclusive measure of take-up. Over time, as the program worked out implementation issues, more youth received SPARCS sessions.

While our study design does not enable us to disentangle the value-add of the two primary program components (intensive mentorship and CBT), qualitative data collected from C2C youth and program staff suggest that the intensive mentorship is critical for engaging harder to reach young people in CBT. Drawing on interviews with front-line staff and focus groups with youth participants, we found that many of the young people who were assigned to C2C were initially skeptical both of the program generally and participating in the CBT sessions in particular. The persistent engagement and the social ties built with the program mentors were critical to overcoming this reticence and supporting engagement with the CBT sessions.[26]

5 RCT Analysis Approach

To estimate the intent to treat (ITT) effects, we run the following regression:

$$Y_{ijt} = \alpha + \beta^{ITT} Z_{ij} + \gamma X_i + \phi_j + \epsilon_{ijt} \tag{1}$$

Where Y_{ijt} represents the outcomes of interest during the post-randomization period t for individual i, in block j, Z_{ij} represents the random assignment indicator for each youth within each block (j). X_i are the baseline covariates included to improve precision by accounting for residual variation in the outcome of interest, and, ϕ_j are the set of dummy variables indicating the observation's randomization block (referral source and cohort). β^{ITT} captures the impact of being offered the C2C program, which is often considered the most policy relevant effect.

However, given that not everyone who is offered the program participates, we present the effects of participating in the program as well. To do this, we estimate the effect of the treatment on the treated (TOT) using a two-stage least squares instrumental variables approach that treats random assignment as an instrument for participation, as follows:[27]

$$P_{ij} = \alpha_1 + \beta_1 Z_{ij} + \gamma_1 X_i + \phi_j + \epsilon_{ij1} \tag{2}$$

[26] See (Lubell et al., 2024) for a qualitative exposition on the details used to engage youth during the program.

[27] This analysis requires the typical relevance and exogeneity assumptions of instrumental variables. In order for the random assignment variable to be a valid instrument, it must be correlated with program participation and uncorrelated with observables. It must shift participation in a uniform direction across people (the monotonicity assumption).

$$Y_{ijt} = \alpha_2 + \beta_2 P_{ij} + \gamma_2 X_i + \phi_j + \epsilon_{ijt2} \tag{3}$$

Where P_{ij} indicates participation in the program after randomization and β_2 captures the TOT. [28] To benchmark the TOT, we calculate the control complier mean (CCM), or the outcome mean for those who would have taken up treatment had they been offered it. This is calculated by taking the mean of the outcome for those that participated in the program and subtracting the TOT. (Katz et al., 2001).[29]

6 RCT Results

6.1 Impact of C2C on arrests

Panel A of Figure Ia shows the Intent-to-Treat (ITT) cumulative arrest estimates for both the extensive and intensive margin over time, up to five years post-randomization. The extensive margin effects appear to peak at 24-months, where we find that being offered C2C substantially reduces the likelihood a young person has any arrest by 6.3 percentage points or 18 percent of the control mean (p-value < 0.01). TOT estimates (highlighted in Table IV) show a 10 percentage point decrease or a 31 percent reduction in the probability of arrest compared to the control complier mean over the same period. The ITT treatment effect is significant at the 1 percent level as early as 12 months post randomization, peaks in terms of magnitude at 24 months post randomization, but maintains significance, with large effects even at 48 months post randomization, with an impact of 5.0 percentage points, or a decrease of 12% of the control mean (see Tables III through V). The intensive margin effect of C2C appears to be noisier and less robust. While we observe consistently negative coefficients for program effects on number of arrests, these estimates have wide confidence intervals during most outcome periods and are not statistically significant at conventional levels.

[28] Given the minor control crossover – 4 youth – this is technically the local average treatment effect (LATE) but given the very low rate of crossover this should be very close to the TOT.

[29] Most of the outcomes we seek to learn about involve variables that take only a limited number of values (e.g: binary variable for being arrested at all or number of arrests within a time frame). Some may argue that nonlinear models such as probit and tobit are preferred in these cases when the outcome of interest is not continuous. All of our main results will use ordinary least squares given that OLS gives us the average causal effect without additional distributional or functional form assumptions. Likewise, we know that OLS will also always give us the minimum mean squared error linear approximation to the conditional expectation function, and IV will capture the local average treatment effect even in the cases where we have dependent variables that take limited values (Angrist and Pischke, 2008).

When we break these arrest impacts down by the charge associated with each arrest, we see most of the extensive margin effect is driven by large reductions in arrests for violent crime offenses, especially leading up to 36 months post randomization. Figure Ib highlights the ITT results for violent-offense arrests over time. Twenty-four months after randomization, we find that being offered C2C significantly reduces the probability of being arrested for a violent offense by 3.7 percentage points (p-value < 0.01, or p-value $< .05$ after adjusting for multiple hypotheses), or 23 percent of the control mean (16% of control youth have an arrest for a violent offense two years after randomization). TOT estimates (presented in Tables III and IV) show a 39 percent reduction in the probability of having any violent arrest. Importantly, we see persistence in these impacts over time as youth in the study age into adulthood. Three years post randomization, we find that being offered C2C significantly reduces the probability of being arrested for a violent offense by 3.1 percentage points or 16 percent of the control mean. Along the intensive margin, we see effect estimates that are in the same direction and magnitude, although slightly noisier over time. During the course of programming, being offered C2C reduces the number of arrests for a violent offense by roughly 36% of the control mean, resulting in roughly 2.3 fewer violent arrests per 100 youth in six months (p-value < 0.05 and $< .10$ after adjusting for multiple hypotheses).

Generally, many of our estimates for overall arrests and violent arrests remain statistically significant at conventional levels after correcting for multiple comparisons using the family-wise error rate (FWER) or the false discovery rate (FDR) up to 24 to 36 months post randomization (Westfall and Young, 1993; Benjamini and Hochberg, 1995). FWER is defined as the chance that at least one of our outcomes in the "family" of outcomes is significant when the null hypothesis of no effect is true. We consider our criminal justice family of outcomes to include all arrest types in a given time frame (6 months, 12 months, 24 months, etc.) by margin (extensive or intensive).[30] We report the q-value from the FDR procedure, that captures the expected proportion of "false positives". Tables III through V also highlight the impact of C2C on other charge categories of arrests: property, drug, and other (anything that does not fall into violence, property or drug). We find some evidence that C2C impacts "other" types of arrests at 36 months and 48 months

[30] FWER uses a bootstrap resampling technique that simulates data under the null hypothesis. Within each permutation, we randomly reassign the treatment indicator with replacement and estimate program impacts on all five of our main outcomes (all arrest categories in each time period). By repeating this procedure 5000 times, we create an empirical distribution of t-statistics that allows us to compare the actual set of t-statistics we find to what we would have found by chance under the null.

post randomization. Although these effects do not remain significant after correction for multiple hypotheses, this does suggest in the initial years, violent related arrests drive the reduction in the extensive margin of arrests, but this begins to shift around three years post-randomization with reductions on other types of arrests. These "other" types of arrests are concentrated in warrants, unlawful use of weapons, criminal trespassing and reckless conduct. Lastly, we do see some evidence that youth who are offered a spot in the C2C program are slightly more likely to be arrested for a drug offense while they are in the program (6 months after randomization) but this finding does not remain significant after correcting for multiple hypotheses.

Further disaggregating the reduction in violent offenses, we see that these largest reductions are driven mainly by a decline in arrests related to interpersonal conflict: aggravated and simple assaults and batteries. Despite large control means in robberies (in 24 months, we see 7 robberies per 100 youth), C2C does not seem to impact these types of violent incidents (see Figure A.VIII in Appendix).[31] We also find some evidence that C2C reduces the number of serious violent victimizations by 5.4 incidents per 100 youth by 36 months post-randomization, or 25% of the control mean (p-value $< .05$, see Table A.XXVIII in Appendix).[32] We did not pre-specify this outcome given victimizations are self-reported to CPD and involvement in C2C could lead to differential reporting rates between the treatment and control groups that could drive differences in outcomes. However, we believe these reductions are promising given these types of serious violent incidences are less likely to be subject to differences in reporting to CPD.[33] Specifically, these results, combined with the reduction in violent arrests, give a strong indication that C2C is effective at reducing violence involvement.

Persistence in Impact To further understand the persistence in these effects, we present 6-month interval non-cumulative estimates (see A.V in the Appendix) and find that the largest reductions occur in the 6 to 12 months after randomization (21% statistically significant reduction in number of arrests). We also observe large decreases in number arrests 24-30 months after randomization, although these are not statistically significant (see Table A.III in Appendix). By 54-

[31]We see about 10 simple assault and batteries per 100 youth and about 4 aggravated assaults and battery per 100 youth in 24 months post randomization among the control youth.

[32]Serious violent victimizations include part-one violent incidents which are homicide, shootings, sexual assault, robbery, aggravated assault, and aggravated battery.

[33]We also focus on serious violent victimizations instead of any violent victimization as the more serious ones are more likely to be reported and less subject to reporting concerns.

60 months post randomization, most impacts are null or close to zero. It is also worth highlighting the steady decrease in control means in arrests over time, showing how the risk of engaging in the criminal justice is decreasing over time as participants age. Taken together, the cumulative and non-cumulative effects suggest that most of the sustained difference between the treatment and control groups over time is driven by large initial drops in youth arrests, though there is some evidence of behavior change further out. Because youth enter the study between the ages of 12 and 20 (with an average age of 16.5), many study youth are entering the program when they are at greatest risk for offending.[34] The persistence in effects, especially in reducing the likelihood of any arrest up to four years post randomization, along with the risk profile of the youth in the study (see Figure A.III in Appendix) suggests that C2C was able to engage youth at a moment of need and help them through this critical period as they transition to adulthood by reducing the likelihood that they had contact with the criminal justice system.[35]

Intensive vs Extensive Margin One puzzling pattern in our arrest findings is the consistent and large impact on the extensive margin, with less or no impact on the intensive margin. We attempt to understand this pattern in a few ways. First, at baseline we observe a large right tail in the number of prior arrests (see Figure A.VI in Appendix). Roughly 11 percent of the study population had five or more baseline arrests prior to randomization. Considering the age of the study participants, this highlights that a small percentage of youth in the study had very intensive prior criminal justice engagement. Nearly two-thirds of the study population had zero arrests prior to randomization, highlighting the wide range of risk levels that C2C was serving. Post-randomization, we find some evidence that outliers may be influencing the estimated effect on number of arrests. Figure A.VII in the Appendix shows the distribution of arrests in the 24 months after randomization among youth in the treatment and control groups. We see clear treatment-control differences for the zero or one arrest margins in 24 months. However this difference fades as we look at three or more arrests. There are also a handful of youth in the treatment group that accumulate more than a dozen arrests in two years. If we top-code our arrest counts at the

[34] The age-crime curve is a phenomenon known widely in Criminology that shows crime peaking in late teenage years and early 20s across the U.S.(Farrington, 1986). In Chicago, using CPD data, we see arrests beginning to peak around age 17 through 24.

[35] Within this five year outcome window, roughly 98% of the study population reaches the age of 18 and youth are 18 or older during 64% of person-days during the period.

99 or 95th percentile in each outcome window prior to running the regressions, we see statistically significant effects at most outcome time periods on the intensive margin (see Table A.VI in the Appendix). While we will explore sub-group effects in the following section in a more consistent manner, we also observe large and statistically significant reductions in the number of arrests at almost every post period among the majority of the sample that had less than five baseline arrests (N=1,852) (see Table A.XII in the Appendix). These findings suggest three takeaways. First, it appears that the program is more effective for youth on the margin of engaging in any crime for the overall sample. Second, we have some outliers in our sample that have accumulated a lot of arrests (both at baseline and in the post period) that increase the variance of our point estimates of the intensive margin of arrests. Lastly, there is some heterogeneity in our treatment effects by baseline engagement in the criminal justice that we explore more systematically in section 6.3.

All of our arrest results are also robust to not including baseline covariates in our estimates (see Table A.V for more details). Given the slight chance imbalance on a few baseline characteristics, we do see some intensive margin estimates are sometimes larger and more statistically significant when not including baseline covariates.[36]

6.1.1 Program effects on police stops

To this point, we have explored the impact of C2C on police contact that actually results in an arrest. To get a better understanding of the type of contact that youth have with the criminal justice system we take a step back and try to understand the first stage of contact that many youth have with police officers: a stop. We find a high level of baseline stops in the study sample; 57% of study youth have been stopped by police officers in either a street or traffic stop at some point prior to randomization.[37] Table A.XXIX examines the effect of the C2C program on the likelihood of being stopped by the police, regardless of whether the stop culminates in an arrest, up to 36 months post randomization (the longest time period we have stops data for the full sample).

Given the proactive policing strategies that are common to Chicago, observing if C2C had an effect on police stops can also help us understand if the program is changing where youth are

[36]As discussed above, we believe there are some outliers in the sample impacting our estimates for the intensive margin. When we winsorize the outcomes at the 95 percentile, the difference between the intensive margin estimates with and without covariates becomes smaller. Regardless, all of our main estimates includes baseline covariates, as we pre-specified in our analysis plan.

[37]See Appendix Section A.1.8 for a detailed discussion on the stops data we use.

hanging out or if youth become more skilled at avoiding police detection. It is notable that there is a high level of interaction between the police and study youth following randomization. Within just six months of randomization, over a fifth of the youth have been stopped by a police officer. Within a year after randomization, almost a third have been stopped. However, the program appears to have no discernible impact on the likelihood that a young person experiences a stop or the number of stops they experience. Twenty four months post randomization, when almost half of the control group have been stopped, we do find a small adverse impact with those in the treatment group slightly more likely to be stopped by a police officer (ITT effect of 3.6 percentage points, with p-value < 0.10, representing an 8% increase in likelihood of being stopped). This could be a result of C2C youth spending more time in new neighborhoods of Chicago and generally participating in more activities outside the house, something we know the program encourages. It is also possible that due to the program's effects, C2C creates an incapacitation effect with control youth being incarcerated at higher rates and therefore less likely to be stopped. Looking at incarceration as an outcome, we do find a negative point estimate. However, these effects are not statistically significant at conventional levels (see Table A.XXVII and section A.1.9 in Appendix). We do not observe an adverse stop impact in other time periods, among the intensive stop margin, or when we break this out by stop type.[38]

Combining our stops findings (null or potentially adverse impact) with the arrest results (significant decrease) suggests that youth are not avoiding police detection and once youth are stopped by a police officer, youth in C2C are either less likely to have demonstrated behavior worthy of an arrest, navigate the interaction more skillfully so they are less likely to result in an arrest, or both. We do find some suggestive evidence that conditional on a stop, C2C reduces the likelihood a stop results in an arrest in 24 to 36 months post randomization (see Table A.XXX in the Appendix).

6.2 Impact of C2C on academic outcomes

Table VI presents the estimated effects of the program on academic outcomes. We see a modest, but statistically significant increase in the overall academic index in the two semesters following

[38]We are able to rule out with the extensive margin estimates that C2C leads to a reduction in the likelihood of being stopped. Here, looking at the 95% confidence interval of any stop within 36 months of randomization, the longest run impact estimate for which we have for full sample, we are able to rule out ITT program reduction effects that are as small as 5 percent of the control mean.

random assignment. The magnitude of these effects range from 0.033 to 0.048 standard deviations (ITT effects) and 0.053 to 0.078 standard deviations for young people served by the program (TOT effects). These results were generally robust to the approach used to impute missing data (see Table A.IX in the Appendix).[39] While we are less confident in the education data further out from randomization because the majority of the sample are no longer actively engaged in school (more youth have either graduated, dropped out, or transferred out of the CPS system), we do show results up to six semesters post-randomization (see Table A.X and A.XI). In these longer-term results, we see null impacts in semester 3 and 4, but more positive and statistically significant results in semester 5 and 6. Given we are relying on more imputation assumptions as we move further out in time post-randomization, we remain cautious in interpreting these persisting effects.

With the main academic index including measures related both to engagement in the academic aspects of school (attendance, enrollment or graduation, and GPA) as well as students' behavior (misconducts incidents, being enrolled in school within a juvenile justice facility, and the number of schools a student is enrolled in), we further broke this overall index into two separate indices. We find that the impact of the program is concentrated in increasing behavioral engagement in school. The program leads to a moderate and statistically significant increase in the behavioral academic index (ITT effects of 0.055 in semester 1 and 0.076 in semester 2; TOT effects of 0.089 in semester 1 and 0.123 in semester two), while program participants see more modest and not consistently significant increases in the academic engagement index. When examining the effect of the program on the index sub-components, the most consistent effect observed is a reduction in the number of misconduct incidents (with estimates ranging from 0.159 to 0.177 fewer incidences in a given semester for TOT estimates, with significance remaining after correcting for multiple hypotheses). We find the program leads to some increase in days attended (2 to 3 more days attended in a given semester – TOT effects) as well as reduced school mobility, although these effects do not remain statistically significant after multiple hypothesis corrections. This pattern of results is consistent with the program's focus on building participants' social and emotional coping, rather than academic skills.

[39]Our results are also robust to not including any baseline covariates (only randomization blocks) or including more baseline education data as covariates in our regressions (see Table A.VIII and A.VII in the Appendix).

Lastly, we also look at the impact of C2C on graduation (see Table VII for details). We examine the impact of any graduation within the six semesters of outcome data we have for the entire sample. For this main outcome ("any graduation"), we do not take into account projected graduation based on youth age or grade at randomization. We also look at a measure of graduation that subsets to the sample of youth for which we should be able to observe their projected graduation year (called "graduated on track"), and include versions of this metric that allow for one or two additional years. First, it is worth highlighting the relatively low graduation rates among the control mean. In the 2023 school year, the average four-year graduation rate for all of Chicago Public Schools was 85%.[40] In contrast, we see the control group achieving a 46% four-year graduation rate, underscoring how academically behind the study population is. We find the program had no effect on participants' likelihood of graduating from high school on-time or with one or two extra years.[41] These findings are consistent with a growing body of literature that finds that interventions that increase engagement in school for youth that are substantially behind academically are less likely to increases in degree attainment if they do not provide personalized academic support (Guryan et al., 2021, 2023).

6.3 Heterogeneity analysis

The C2C program aims to serve a different group of youth than those who are typically reached by school-based programs, specifically targeting young people with previous engagement with the criminal justice system and young people who are beginning to or have disengaged from school. In addition, the program serves both girls and boys in contrast to several well-studied programs that target young men. Given the range of risk levels among the study youth, this raises questions about how treatment effects differ for these groups of students and whether the program's effectiveness could be increased through better targeting of services. To answer these questions we conducted several exploratory subgroup analyses to explore treatment effect heterogeneity. Specifically, we

[40]This is calculated by looking at within the first-time 9th graders in the SY 2019-2020, what percent of them graduate as expected in four years in the Spring of 2023. This rate is up from a 62% CPS-wide graduation rate in 2008.

[41]We also present a version of the graduation results that account for transfers given we do not observe graduation for youth who leave CPS, and find no impact on this measure of graduation either.

examine whether treatment effects vary by having an arrest at baseline, gender, and engagement in school at baseline.[42] [43]

Overall, we find suggestive evidence that all of these groups of youth are benefiting from the program, but that the program may be moving different outcomes for different groups of youth in ways that may be related to baseline propensity to experience a given outcome (See Tables A.XIV, A.XV, A.XVII). These results should be interpreted with caution given we were not powered to detect subgroup effects and there are a large number of hypothesis tests involved. We find that for young people who had not been arrested at baseline, the program led to large and statistically significant decreases in the number of overall arrests. For example, 24 months after randomization, young people assigned to the treatment group had 0.157 fewer arrests, a 42% reduction relative to the control group ($p = 0.002$). These differences remain large and statistically significant up to four years after randomization. In contrast, for youth with an arrest at baseline, we find non statistically significant reductions in the number of arrests and in some time periods the point estimates are even positive.

However, this pattern reverses when we look at program effects on violent offenses. Here, we see that those at higher risk for criminal justice involvement at baseline (those with a prior arrest, males, and those who are not engaged in school at baseline), the program reduces the likelihood of being arrested for a violent offense. We also observe a large reduction in the number of arrests for violent offenses for males and young people who had an arrest at baseline.

Similarly, how connected a young person was to school at baseline was associated with different patterns of effects on the education outcomes we examined in the study. For young people who were engaged in school at baseline, the program led to increased engagement in school in terms of both behavioral and academic engagement (see Appendix Table A.XXI). Specifically, in addition to reducing behavioral misconducts as we saw in the full sample, the program led to increased attendance and GPA for this group of students. Most notably, youth who were engaged at baseline

[42] Drawing on the criteria the BAM program used to determine whether a student was sufficiently engaged in school to enroll in the program, we identify a student as not being engaged if they were inactive at any point in the three semesters prior to randomization, were two or more years too old for their grade (an indicator of having been held back), had low attendance and had failed several courses, or had a severe disability that would preclude them from participating in a group therapy session.

[43] We also pre-specified looking at the type of school enrollment/referral at baseline. These results are provided in Appendix Table A.XVIII, but given it breaks the sample into 3 groups, we are underpowered to detect significant effects.

saw a statistically significant 5.9 percentage point in high school graduation (see Appendix Table A.XXIII). In contrast, for youth who were already disengaging from school at baseline, positive program effects were limited to improved behavioral engagement (predominantly via reduced behavioral incidences) and the program actually significantly decreased high school graduation. This pattern of results reinforced findings from past research that behaviorally focused interventions that do not include a focus on academic remediation may be insufficient to improve academic outcomes for students who have fallen substantially behind their classmates.

Taken as a whole, these results suggest the program may be working differently for different groups of youth in ways related to baseline propensity to experience an arrest or engage in school. To test this explicitly, we stratified our sample into four quartiles of predicted arrest risk following the procedure on endogenous stratification outlined in Abadie et al. (2018) (see Appendix A.1.10 for more detail). Our prediction model ranks participants by their expected number of arrests in the first 24 months after randomization; ITT effects are then estimated within each quartile using both leave-one-out (LOO) and repeated split sample (RSS) estimators. The results using this method echo the patterns found earlier looking at binary splits in the data. For example, for ever being arrested, we see sizable and statistically meaningful reductions appear in the two mid-risk quartiles for the overall case, while reductions in ever arrested for a violent-offense are concentrated in the highest-risk quartile (see Figure A.XII and Figure A.XIII). A similar pattern can be observed for the number of arrest outcomes, albeit less statistically precise given the noise we have documented in this measure (Figure A.X and Figure A.XI). Conversely, improvements on our combined academic index emerge primarily in the upper-mid and highest quartiles, potentially driven by fewer school incidents in the top quartile and improved attendance in the upper-mid risk quartile (see Figure A.XIV). Even for outcomes where we don't see quartile estimates consistently crossing the significance level, the difference in point estimates across groups can be quite informative. For number of violent offense arrests, point estimates suggest that the highest-risk group has a treatment effect at least 5 times as large as the lowest risk group, when comparing them across the same outcome periods.[44] For number of arrests, mid-risk groups 2 and 3 have a treatment effects

[44]This comparison is based solely on the preferred RSS estimator. In a few outcome periods, we observe slightly positive yet not significant RSS estimates for the lower risk group. For consistency, ratios were computed only when both estimates were negative.

at least 3 times as large as effects estimated for the lowest risk group.[45] These results should be interpreted with caution, and are only suggestive. With these (smaller) subgroups and the comparatively large standard errors, it is difficult to pin down the shape of the relationship between risk group and response (Heller, 2022). That said, we believe the alignment between risk-specific needs and observed impacts reinforces the view that treatment effect heterogeneity may be related to baseline risk of a particular outcome.

6.4 Cost-benefit analysis

While C2C does not produce a statistically significant reduction in the total number of arrests (for the full sample), it does meaningfully reduce the likelihood that a young person is arrested at all. Avoiding even a single arrest can have widespread implications for a young person's long term outcomes by reducing their exposure to the criminal justice system, limiting potential disruptions to education or employment, and decreasing the risk of future involvement with the system (Bacher-Hicks and de la Campa, 2020; Geller, 2019; Lochner, 2004). Additionally, C2C leads to a reduction in the number of arrests for violent offenses from 6 to 24 months post-randomization, although these effects are sometimes only marginally significant. Both of these outcomes can have substantial costs for individuals and society (Gobbo, 2023; Chalfin, 2015; Council et al., 2011). As such, reducing either can yield considerable social benefits.

Our cost-benefit estimates are rough approximations and focus solely on benefits gained from reduced involvement with the criminal justice system, as captured through arrest and victimization data. We do not attempt to monetize potential gains from improved educational outcomes such as increased attendance, fewer misconduct incidents, or greater school stability given the challenges in reliably assigning dollar values to these measures. However, since we observe statistically significant improvements on our academic index, the estimates presented here may represent a conservative lower bound on the full benefits of the program.

Because different types of crime carry varying costs on society, we assign a dollar value to each C2C arrest and victimization based on the top charge associated with the event. For each outcome period, we calculate the total cost incurred by summing all arrest and victimization related costs for each study member. We then estimate C2C's benefit in terms of reduced criminal justice costs

[45] ibid.

by using this total cost as the dependent variable. Finally, we compare the estimated cost savings from crime reduction to the program's cost per participant. On average, C2C cost \$4,500 per participant in cohorts 1–5 and \$7,300 in cohorts 6–8, averaging just over \$5,000 per participant across the full study period.[46]

Given the uncertainty in assigning a monetary value to the cost of crime, we bound C2C's benefit by estimating a lower-end and higher-end cost of crime. Our lower end estimates assign the cost of each arrest and victimization based on the "bottom-up" cost estimates from (Cohen and Piquero, 2009), which aim to assign a victim cost, legal system cost, and offender productivity cost to each event. Our higher-end estimates use the Willingness-to-Pay (WTP) estimates from (Cohen and Piquero, 2009), which uses a "top-down" approach aiming to estimate the amount the public would be willing to pay to reduce each crime type. Appendix Table A.XXIV shows the bottom-up and willingess-to-pay cost to society for each crime type.

One challenge in this estimation is how to handle homicide events, which are orders of magnitude more costly than any other crime type and are exceedingly rare in our sample. Specifically, we observe only six homicide arrests and four homicide victimizations over the 24-months post-randomization. As shown in Appendix Figure A.VIII, C2C does not appear to have any effect on homicides. Including these few events introduces a disproportionate amount of noise into these estimates, as one additional homicide can swing our more conservative estimates by tens of thousands of dollars. For this reason, we recode each homicide event as an aggravated assault - assigning it the same cost as a non-fatal shooting. Appendix Table A.XXVI shows how the estimates would change if homicide events were left untouched.

Table IX shows that in the absence of C2C, society would incur between \$14,400 and \$40,600 per control complier after 24-months. We estimate that C2C reduces these costs between \$5,200 and \$12,800 per participant after 24-months (p = .020 and p = .089, respectively), with the lower-end estimate still being enough to cover the cost of programming for each participant after 24-months. When compared to the control complier means, this represents savings between 30-36% per participant. Savings to society are maximized 3 years post-randomization, with savings up to \$6,800 per participant using our lower-end estimates (p = .019) and over \$20,200 using our higher-end estimates (p = .034). This represents a benefit-cost ratio between 1.4:1 and 4:1. A

[46] All costs and benefits have been converted to 2015 dollars - the year the study began.

key limitation of our approach is that we are only able to assign costs to arrest and victimization events that we directly observe in CPD data. This may result in underestimation, particularly if C2C affects underlying behaviors that are not captured in administrative data. We aim to create an upper bound for our estimates in Appendix Table A.XXV by adjusting for crime reporting and clearance rates. When adjusting for reporting and clearance rates, we see immediate savings of \$82,000 per participant after just 6-months ($p = .047$), with large savings persisting overtime, although these estimates are imprecise.[47]

While these figures provide a estimation of C2C's benefits through reduced crime involvement, they understate the substantial external costs of violence. Crime and violence imposes large negative externalities on communities, including contributing to population decline and reduced property values (Cullen and Levitt, 1999; Shapiro et al., 2012). The estimated reductions in violent crime arrests and serious violent crime victimization (see Appendix Table A.XXVIII) suggest that the social returns to the communities that C2C serves may be considerably larger.

7 Discussion

Qualitative Data and Potential Mechanisms To dive into potential mechanisms and explore how C2C achieved positive impacts on behavior, we utilize data from our extensive qualitative work where we conducted over 12 focus groups with 69 C2C participants and 19 interviews with C2C program staff. Some of the key themes that came out of these data suggest that the program leads to increased future orientation, increased ability to process trauma, an increase in self-esteem and self-worth, an increase in social and emotional skills, with new skills in managing stressful situations. The CBT component of the program was critical to developing these new social-emotional skills. As discussed earlier, some exercises in the program are intended to teach youth how to cope with stressful situations in the present that might be triggered by chronic dysregulation of impulses, or reminders of earlier trauma. Research has shown that trauma and stress can deplete mental bandwidth, leading to over-reliance on the automatic, more impulse driven part of our brain (Kahneman, 2011; Mullainathan and Shafir, 2013; Ludwig, 2025). Given the significant past

[47] Comparing the cost-benefit ratios of C2C to other similar programs can be difficult given the variation in assumptions used by researchers. Future work would benefit from using a unified approach such as the Marginal Value of Public Funds (MVPF) to compare policies (Hendren and Sprung-Keyser, 2020).

trauma exposure of youth in the program, and the high-poverty and low-resourced neighborhoods they come from, challenging situations or environments is a frequent encounter for study youth.[48] Youth in the program learn that after extreme stress, they might carry around "emotional left-overs" and even minor challenging situations (e.g. a teacher reprimanding a student for talking too much in class or an argument with a friend) can quickly turn into more serious escalations with negative consequences. The program's group sessions help youth learn to recognize and manage these symptoms of stress on the body, and teach various coping skills during stressful moments, including "distracting and soothing" tools if they can't solve the situation immediately or think through safer or appropriate choices when upset. The arrest results are consistent with the theory that C2C is changing how youth engage in conflict given the violent-arrest findings are concentrated in arrests related to interpersonal conflict such as assaults and batteries (and not robberies that might be financially motivated). For example, one youth said "I learned how to walk away and not act on everything and make a permanent decision on a temporary situation. Furthermore, one C2C therapist highlighted, the shift in perspective from "what is wrong with you?" to asking "what happened to you?" can be a beneficial perspective and reframe existing challenges in a trauma-informed lens that puts healing at the forefront. Given the high levels of trauma and adverse experience exposure among youth in the program, our findings highlight the importance of trauma-informed programming in supporting youth's ability to safely navigate their environment.

Focus group respondents also stressed the importance of the strength-based mentoring from the adult advocate. For example, one youth said, "He'd been there, done that, he'd been my age before. ... actually, listening and taking advice from someone who knows[s] and not just trying to put off their opinion on you, someone who was actually in this situation and overcame it." The consistent, positive adult support and the new experiences youth have in the program can also serve as pathways to restore physical safety. Before any CBT services are offered, the advocate spends time developing a concrete personal relationship with the youth, which is fundamental to generating trust. Intuitively, this kind of human connection – or what the sociology literature calls "social capital" – is at the core of most transformation stories we hear or experience in our own lives (Coleman, 1988). Particular success has been found when positive and relatable adult figures are

[48] On average, youth participants had experienced 7 traumatic or adverse experiences prior to the program starting, and almost all had experienced at least one.

able to use this social capital to go beyond engaging youth directly and increase their connections to pro-social institutions, such as school or after-school activities (Tolan et al., 2013; DuBois et al., 2011; DeLuca et al., 2016). Research has also highlighted that social support is one of the most powerful protections against becoming overwhelmed with stress and trauma (Van der Kolk, 2015). The positive experiences and extensive time with the mentor/advocate served as opportunities for new coping skills to be modeled and practiced. As discussed earlier, one of the unique aspects of the C2C program was the focus on practicing the skills learned in the group CBT sessions with the mentors/advocates in the eight hours they spend together weekly. The existing psychology literature has highlighted that the more frequently a person performs a behavior or skill, the more habitual and automatic it becomes ("rewiring the brain"), which can help youth respond differently in moments of stress (Van der Kolk, 2015; Wilson, 2004; Beck and Beck, 2011; Blattman et al., 2023). The program model allows for many opportunities to practice the new social-emotional skills youth learn in C2C, suggesting this is one potential mechanism for why we see the sustained program impacts.

Targeting and Scaling While the program was initially privately funded, the documented positive effects of this program led to significant public dollar investment from Chicago Public Schools and a scaling of program services to engage more youth throughout Chicago after randomization ended in the last five years. As this program continues to operate, thinking critically about targeting and who this programs works best for will be important to sustain the program, especially during times of austerity.

We did not set up the study to explicitly answer this question, however we do have some suggestive evidence that while all groups of youth are benefiting from the program, participants' baseline propensity to experience a given outcome appear to be related to which outcomes are improved. Those who enter the program with substantial exposure to the criminal justice system have the largest decreases in being arrested for a violent offense, while those who are connected to school at baseline are the only group of students for whom we see increased high school graduation. This pattern of results is in many ways consistent with the program model which is designed to meet young people where they are. This treatment effect heterogeneity could suggest that if policy makers are most interested in reducing violent crime (because of it's high social cost), they may want to target the program more narrowly to those at greatest risk for criminal justice system engagement at baseline. However, we caution that this could unintentionally change the overall effectiveness of the program by changing the composition of participants. Crime and violence often have a social component and almost all C2C activities are group-based. Therefore, changing program targeting or the composition of youth in the program could also change the program experience and in turn program effectiveness.[49] Based on this and the fact that many of the subgroup analyses are underpowered and exploratory, we are cautious in making any clear targeting recommendations for program purposes based on these existing findings and would recommend this as an area for future work. Understanding treatment effect heterogeneity more completely and how CBT-based programs can replicate to other contexts and scale will be critical to guiding public investment decisions moving forward.[50]

Areas for Future Work Lastly, our results suggest that C2C program supports may potentially pair well with other interventions. Regarding the education impacts, it is important to note that we do not observe if youth take the General Educational Development (GED) exam and obtain a GED instead of a traditional high school diploma. From conversations with program providers, some program staff did connect youth who were very behind on credits in school to job opportunities and suggested a GED as an alternative to high school graduation. While this might seem like a logical choice for young people who are a substantial distance from graduation, there is generally a consensus in the literature that GEDs are not equivalent to conventional high school diplomas, and do not offer the same value with respect to post-secondary outcomes and labor market opportunities (Heckman et al., 2011; Cameron and Heckman, 1993; Murnane et al., 2000). The lack of impacts on GPA or graduation overall suggests that C2C might pair well with effective, targeted academic interventions such as high dosage tutoring, especially for youth who need more intensive remediation (Guryan et al., 2023; Bhatt et al., 2024).[51]

[49] A separate but related question is if the program had spillovers on siblings or peers given the substantial treatment effects and the peer based nature of crime among youth. In separate unpublished exploratory work, we do not find any evidence of spillovers using co-offending baseline data and administrative school data to identify a network of social ties. More information is available upon request from the authors on this analysis.

[50] In related research on summer youth programs, research has found summer youth programs impacts on arrests grow linearly with the risk of socially costly behavior each person faces (Heller, 2022).

[51] Anecdotally, we heard from some of the program staff that some youth in the program were quite behind academically, and although they could help them show up more to school, that has limited impact when they are sitting in 10th grade classes with a 4th or 5th grade reading level, for example.

The null effects we see on stops also serves as a useful reminder that individual programs like C2C did not change where youth spend time, where they live or the policing tactics used in their neighborhoods. Our paper can not speak to the tradeoffs or the benefits or costs of various police strategies that lead to high contact with police.[52] However, our paper does highlight a potential collateral effect of these policies by demonstrating a situation where youth behavior on more serious dimensions changes (including a reduction in violence), and yet this did not impact the propensity of being stopped by police. There is ample research documenting the harm even low levels of contact, like stops, can have on youth by negatively impacting health, mental well-being, educational outcomes, and political engagement (Weitzer et al., 2008; Futterman et al., 2016; Geller et al., 2014; Bell, 2017; Weaver and Geller, 2019; Del Toro et al., 2019; Pickett et al., 2021; Bacher-Hicks and de la Campa, 2020). In separate, but related work, we find that arrest-charge level evaluations can obscure aspects of police behavior or police strategy that may influence criminal justice system contact.[53] Given the growing policy focus on improving public safety while decreasing the toll of the criminal justice system, especially in communities of color, further research can better understand how changing policing strategies may moderate or complement individual-focused supports like C2C.

8 Conclusion

This paper evaluates the impact of a program that combines intensive mentoring and trauma-informed CBT in Chicago. The program's main effect on reducing the likelihood of being arrested is encouraging as any contact with the criminal justice system can be harmful to youth. Furthermore,

[52]Some research has shown that hot spot policing is intended to act as a deterrent effect and may have modest effect sizes in reducing crime (Sherman and Weisburd, 1995; Braga et al., 2014; Weisburd, 2021). Yet, existing research is unclear on the benefits of stop and frisk policies above and beyond increased police presence in certain areas (MacDonald et al., 2016; Bacher-Hicks and de la Campa, 2020; Weisburd et al., 2015). Specifically, MacDonald et al. (2016) find that in New York City, the only stops that had a detectable impact on crime were stops based on probable cause, and these kinds of stops were very rare. Bacher-Hicks and de la Campa (2020) find that New York City's Stop and Frisk program did not deter more serious crime, and only had a small effect on misdemeanor crime.

[53]It is well established that arrests incorporate not just civilian behavior but policing behavior as well, with a long history in Criminology and other fields that documents the discretion present in policing (Goldstein, 1963; Linn, 2009; Kelling, 1999; Wu and Lum, 2017; Cho et al., 2021). Our companion paper uses arrest narratives to identify police discretion leading to an arrest, allowing researchers to tease out criminal justice contact that people have more control over versus what police might have more control. We develop a text classification machine learning model that uses narrative data to accurately predict which arrests are discretionary, highlighting variation within arrest charges (Abdul-Razzak et al., 2025).

the large initial reduction in violence engagement is particularly promising given the high cost of violence for both individuals and society as a whole. Our null findings with police stops highlight that the program is not helping youth avoid all contact with police (or avoid detection), suggesting that C2C is changing the underlying behavior worthy of an arrest and/or helping youth navigate police interactions more skillfully. While the school engagement positive effects are modest, they do highlight a consistent narrative in a change in behavior both in school and outside of school that could lead to later downstream positive effects even if graduation is not directly impacted by the program. Importantly, our study follows all study youth into adulthood, highlighting a cost-effective program that can have longer-term impacts and engage youth at critical moments in adolescence.

In the context of other successful CBT-based programs, Choose to Change offers the possibility of scaling CBT with a harder-to-reach population of youth outside of school. Our research, along with the growing CBT literature, highlights a potential program "recipe" where some form of engagement support (e.g. mentoring) is beneficial to encourage participation and practice of CBT-based skills. Continuing to scale and test versions of this program in different contexts is ever more important and relevant given the potential payoffs of reducing violence and overall criminal justice engagement for individuals and communities.

9 References

Abadie, Alberto A, Matthew M Chingos, and Martin R West (2018) "Endogenous stratification in randomized experiments," *The review of economics and statistics*, C (4).

Abdul-Razzak, Nour, Brandon Domash, Cristobal Poehls Pinto, and Kelly Hallberg (2025) "Measuring Discretion in Arrests: Evidence from Chicago,"Technical report, Working Paper.

Aizer, Anna and Joseph J Doyle (2015) "Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges," *The Quarterly Journal of Economics*, 130 (2), 759–803.

Ang, Desmond (2021) "The effects of police violence on inner-city students," *The Quarterly Journal of Economics*, 136 (1), 115–168.

Ang, Desmond and Jonathan Tebes (2020) "Civic Responses to Police Violence," *Harvard Kennedy School, John F. Kennedy School of Government.*

Angrist, Joshua D and Jörn-Steffen Pischke (2008) *Mostly harmless econometrics*: Princeton university press.

Arbour, William (2021) *Can Recidivism be Prevented from Behind Bars?: Evidence from a Behavioral Program*: University of Toronto, Department of Economics4060 1 Online-Resource.

Bacher-Hicks, Andrew and Elijah de la Campa (2020) "Social Costs of Proactive Policing: The Impact of NYC's Stop and Frisk Program on Educational Attainment,"Technical report, Working paper.

Bandes, Susan A, Marie Pryor, Erin M Kerrison, and Phillip Atiba Goff (2019) "The mismeasure of Terry stops: Assessing the psychological and emotional harms of stop and frisk to individuals and communities," *Behavioral sciences & the law*, 37 (2), 176–194.

Barnes, Geoffrey C, Jordan M Hyatt, and Lawrence W Sherman (2017) "Even a little bit helps: An implementation and experimental evaluation of cognitive-behavioral therapy for high-risk probationers," *Criminal Justice and Behavior*, 44 (4), 611–630.

Batistich, Mary Kate, William N Evans, Tyler Giles, and Rebecca Margolit-Chan (2024) "Therapy to Reduce Violence and Improve Institutional Safety During Incarceration,"Technical report, National Bureau of Economic Research.

Beck, JS and AT Beck (2011) "Cognitive behavior therapy: basics and beyond, vol. 2."

Becker, Gary S (1968) “Crime and punishment: An economic approach,” *Journal of political economy*, 76 (2), 169–217.

Bell, Monica C (2017) “Police reform and the dismantling of legal estrangement,” *The Yale Law Journal*, 2054–2150.

Benjamini, Yoav and Yosef Hochberg (1995) “Controlling the false discovery rate: a practical and powerful approach to multiple testing,” *Journal of the Royal statistical society: series B (Methodological)*, 57 (1), 289–300.

Bhatt, Monica P, Jonathan Guryan, Salman A Khan, Michael LaForest-Tucker, and Bhavya Mishra (2024) “Can technology facilitate scale? Evidence from a randomized evaluation of high dosage tutoring,”Technical report, National Bureau of Economic Research.

Bhatt, Monica P, Sara B Heller, Max Kapustin, Marianne Bertrand, and Christopher Blattman (2023) “Predicting and Preventing Gun Violence: An Experimental Evaluation of READI Chicago*,” *The Quarterly Journal of Economics*, qjad031, 10.1093/qje/qjad031.

Blattman, Christopher, Sebastian Chaskel, Julian C Jamison, and Margaret Sheridan (2023) “Cognitive Behavioral Therapy Reduces Crime and Violence over Ten Years: Experimental Evidence,” *American Economic Review: Insights*, 5 (4), 527–545.

Blattman, Christopher, Julian C Jamison, and Margaret Sheridan (2017) “Reducing crime and violence: Experimental evidence from cognitive behavioral therapy in Liberia,” *American Economic Review*, 107 (4), 1165–1206.

Braga, Anthony A, Andrew V Papachristos, and David M Hureau (2014) “The effects of hot spots policing on crime: An updated systematic review and meta-analysis,” *Justice quarterly*, 31 (4), 633–663.

Butler, Paul (2014) “Stop and frisk and torture-lite: police terror of minority communities,” *Ohio St. J. Crim. L.*, 12, 57.

Cameron, Stephen V and James J Heckman (1993) “The nonequivalence of high school equivalents,” *Journal of labor economics*, 11 (1, Part 1), 1–47.

Cattan, Sarah, Daniel A Kamhöfer, Martin Karlsson, and Therese Nilsson (2023) “The long-term effects of student absence: Evidence from Sweden,” *The Economic Journal*, 133 (650), 888–903.

Chalfin, Aaron (2015) “Economic costs of crime,” *The encyclopedia of crime and punishment*, 1–12.

Cho, Sungwoo, Felipe Gonçalves, and Emily Weisburst (2021) "Do Police Make Too Many Arrests?".

Cloitre, Marylene, Bradley C Stolbach, Judith L Herman, Bessel van der Kolk, Robert Pynoos, Jing Wang, and Eva Petkova (2009) "A developmental approach to complex PTSD: Childhood and adult cumulative trauma as predictors of symptom complexity," *Journal of traumatic stress*, 22 (5), 399–408.

Cohen, Mark A and Alex R Piquero (2009) "New Evidence on the Monetary Value of Saving a High Risk Youth," *Journal of Quantitative Criminology*, 25 (1), 25–49.

Coleman, James S (1988) "Social capital in the creation of human capital," *American journal of sociology*, 94, S95–S120.

Council, National Research et al. (2011) "Social and Economic Costs of Violence: Workshop Summary."

Cullen, Julie Berry and Steven D. Levitt (1999) "Crime, Urban Flight, and the Consequences for Cities," *The Review of Economics and Statistics*, 81 (2), 159–169.

Davis, Jonathan MV and Sara B Heller (2020) "Rethinking the benefits of youth employment programs: The heterogeneous effects of summer jobs," *Review of economics and statistics*, 102 (4), 664–677.

De Rosa, R, D Pelcovitz, J Rathus et al. (2004) "SPARCS (Structured Psychotherapy for Adolescents Responding to Chronic Stress)."

Dee, Thomas S (2024) "Higher chronic absenteeism threatens academic recovery from the COVID-19 pandemic," *Proceedings of the National Academy of Sciences*, 121 (3), e2312249121.

Del Toro, Juan, Tracey Lloyd, Kim S Buchanan et al. (2019) "The criminogenic and psychological effects of police stops on adolescent black and Latino boys," *Proceedings of the National Academy of Sciences*, 116 (17), 8261–8268.

DeLuca, Stefanie, Susan Clampet-Lundquist, and Kathryn Edin (2016) *Coming of age in the other America*: Russell Sage Foundation.

Desmond, Matthew, Andrew V Papachristos, and David S Kirk (2016) "Police violence and citizen crime reporting in the black community," *American sociological review*, 81 (5), 857–876.

Dinarte-Diaz, Lelys and Pablo Egana-delSol (2024) “Preventing violence in the most violent contexts: Behavioral and neurophysiological evidence from el salvador,” *Journal of the European Economic Association*, 22 (3), 1367–1406.

Doleac, Jennifer L (2019) “Wrap-around services don’t improve prisoner reentry outcomes,” *Journal of policy analysis and management*, 38 (2), 508–514.

DuBois, David L, Nelson Portillo, Jean E Rhodes, Naida Silverthorn, and Jeffrey C Valentine (2011) “How effective are mentoring programs for youth? A systematic assessment of the evidence,” *Psychological science in the public interest*, 12 (2), 57–91.

Farrington, David P (1986) “Age and crime,” *Crime and justice*, 7, 189–250.

Flannery, Daniel J, Kelly L Wester, and Mark I Singer (2004) “Impact of exposure to violence in school on child and adolescent mental health and behavior,” *Journal of community psychology*, 32 (5), 559–573.

Ford, Julian D, J Kirk Hartman, Josephine Hawke, and John F Chapman (2008) “Traumatic victimization, posttraumatic stress disorder, suicidal ideation, and substance abuse risk among juvenile justice-involved youth,” *Journal of Child & Adolescent Trauma*, 1 (1), 75–92.

Futterman, Craig B, Chaclyn Hunt, and Jamie Kalven (2016) “Youth/police encounters on Chicago’s south side: Acknowledging the realities,” *U. Chi. Legal F.*, 125.

Geller, Amanda (2019) “Policing America’s children: Police contact among urban teens,” *Unpublished Manuscript. Fragile Families Working Paper WP18-02-FF.*

Geller, Amanda, Jeffrey Fagan, Tom Tyler, and Bruce G Link (2014) “Aggressive policing and the mental health of young urban men,” *American journal of public health*, 104 (12), 2321–2327.

Gelman, Andrew, Jeffrey Fagan, and Alex Kiss (2007) “An analysis of the New York City police department’s “stop-and-frisk” policy in the context of claims of racial bias,” *Journal of the American statistical association*, 102 (479), 813–823.

Gobbo, Andre (2023) “The economic costs of gun violence in the United States.”

Goldstein, Herman (1963) “Police discretion: The ideal versus the real,” *Public Administration Review*, 140–148.

Guryan, Jonathan, Sandra Christenson, Ashley Cureton, Ijun Lai, Jens Ludwig, Catherine Schwarz, Emma Shirey, and Mary Clair Turner (2021) "The effect of mentoring on school attendance and academic outcomes: A randomized evaluation of the Check & Connect Program," *Journal of Policy Analysis and Management*, 40 (3), 841–882.

Guryan, Jonathan, Jens Ludwig, Monica P Bhatt et al. (2023) "Not too late: Improving academic outcomes among adolescents," *American Economic Review*, 113 (3), 738–765.

Hausman, David and Dorothy Kronick (2021) "When Police Sabotage Reform by Switching Tactics," *Available at SSRN 3192908*.

Heckman, James J, John Eric Humphries, and Nicholas S Mader (2011) "The GED," *Handbook of the Economics of Education*, 3, 423–483.

Heckman, James J and Tim Kautz (2013) "Fostering and measuring skills: Interventions that improve character and cognition."

Heller, Sara B (2014) "Summer jobs reduce violence among disadvantaged youth," *Science*, 346 (6214), 1219–1223.

——— (2022) "When scale and replication work: Learning from summer youth employment experiments," *Journal of Public Economics*, 209, 104617.

Heller, Sara B, Anuj K Shah, Jonathan Guryan, Jens Ludwig, Sendhil Mullainathan, and Harold A Pollack (2017) "Thinking, fast and slow? Some field experiments to reduce crime and dropout in Chicago," *The Quarterly Journal of Economics*, 132 (1), 1–54.

Hendren, Nathaniel and Ben Sprung-Keyser (2020) "A unified welfare analysis of government policies," *The Quarterly journal of economics*, 135 (3), 1209–1318.

Hjalmarsson, Randi (2008) "Criminal justice involvement and high school completion," *Journal of Urban Economics*, 63 (2), 613–630.

Kahneman, Daniel (2011) *Thinking, fast and slow*: macmillan.

Katz, Lawrence F, Jeffrey R Kling, and Jeffrey B Liebman (2001) "Moving to opportunity in Boston: Early results of a randomized mobility experiment," *The Quarterly Journal of Economics*, 116 (2), 607–654.

Kelling, George L (1999) *Broken windows and police discretion*: US Department of Justice, Office of Justice Programs, National Institute of Justice.

Van der Kolk, Bessel A (2015) *The body keeps the score: Brain, mind, and body in the healing of trauma*: Penguin Books.

Lab, Education (2021) "Data Insights from Chicago's Options Schools: Seizing the Opportunity to Advance Education Equity,"Technical report, University of Chicago.

Landenberger, Nana A and Mark W Lipsey (2005) "The positive effects of cognitive–behavioral programs for offenders: A meta-analysis of factors associated with effective treatment," *Journal of experimental criminology*, 1 (4), 451–476.

Legewie, Joscha and Jeffrey Fagan (2019) "Aggressive policing and the educational performance of minority youth," *American Sociological Review*, 84 (2), 220–247.

Lerman, Amy E and Vesla M Weaver (2014) *Arresting citizenship*: University of Chicago Press.

Linn, Edith (2009) *Arrest decisions: What works for the officer?* (5): Peter Lang.

Liu, Jing, Monica Lee, and Seth Gershenson (2021) "The short-and long-run impacts of secondary school absences," *Journal of Public Economics*, 199, 104441.

Lochner, Lance (2004) "Education, work, and crime: A human capital approach," *International Economic Review*, 45 (3), 811–843.

——— (2011) "Non-production benefits of education: Crime, health, and good citizenship."

——— (2020) "Education and crime," in *The economics of education*, 109–117: Elsevier.

Lochner, Lance and Enrico Moretti (2004) "The effect of education on crime: Evidence from prison inmates, arrests, and self-reports," *American economic review*, 94 (1), 155–189.

Lubell, Max, Nour Abdul-Razzak, and Kelly Hallberg (2024) "How Community Violence Program Staff Develop Social Ties in Heavily Policed Neighborhoods,"Technical report, Working Paper.

Ludwig, Jens (2025) "Unforgiving Places: The Unexpected Origins of American Gun Violence," in *Unforgiving Places*: University of Chicago Press.

MacDonald, John, Jeffrey Fagan, and Amanda Geller (2016) "The effects of local police surges on crime and arrests in New York City," *PLoS one*, 11 (6), e0157223.

Margolin, Gayla and Elana B Gordis (2000) "The effects of family and community violence on children," *Annual review of psychology*, 51 (1), 445–479.

Mullainathan, Sendhil and Eldar Shafir (2013) *Scarcity: Why having too little means so much*: Macmillan.

Murnane, Richard J, John B Willett, and John H Tyler (2000) "Who benefits from obtaining a GED? Evidence from high school and beyond," *Review of Economics and Statistics*, 82 (1), 23–37.

OCO (2025) "Chronic Absenteeism: Supporting student attendance and combatting chronic absenteeism in our nation's schools,"Technical report, U.S. Department of Education, `https://www.ed.gov/teaching-and-administration/supporting-students/chronic-absenteeism`.

Owens, Emily G (2017) "Testing the school-to-prison pipeline," *Journal of Policy Analysis and Management*, 36 (1), 11–37.

Pickett, Justin, Amanda Graham, and Frank Cullen (2021) "The American Racial Divide in Fear of the Police."

Ross, Lee and Richard E Nisbett (2011) *The person and the situation: Perspectives of social psychology*: Pinter & Martin Publishers.

Santiago, Catherine DeCarlo, Tali Raviv, and Lisa H Jaycox (2018) *Creating healing school communities: School-based interventions for students exposed to trauma.*: American Psychological Association.

National Academies of Sciences, Engineering, Medicine et al. (2018) *Proactive policing: Effects on crime and communities*: National Academies Press.

Shah, Anuj K and Jens Ludwig (2016) "Option Awareness: The psychology of what we consider," *American Economic Review*, 106 (5), 425–429.

Shapiro, Robert J, Kevin A Hassett et al. (2012) "The economic benefits of reducing violent crime: A case study of 8 American cities."

Sherman, Lawrence W and David Weisburd (1995) "General deterrent effects of police patrol in crime "hot spots": A randomized, controlled trial," *Justice quarterly*, 12 (4), 625–648.

Stoudt, Brett G, Michelle Fine, and Madeline Fox (2011) "Growing up policed in the age of aggressive policing policies," *NYL Sch. L. Rev.*, 56, 1331.

Tolan, Patrick, David Henry, Michael Schoeny, Arin Bass, Peter Lovegrove, and Emily Nichols (2013) "Mentoring interventions to affect juvenile delinquency and associated problems: A systematic review," *Campbell Systematic Reviews*, 9 (1), 1–158.

Van Dijk, Sheri (2013) *DBT made simple: A step-by-step guide to dialectical behavior therapy*: New Harbinger Publications.

Weaver, Vesla M and Amanda Geller (2019) "De-policing America's youth: Disrupting criminal justice policy feedbacks that distort power and derail prospects," *The ANNALS of the American Academy of Political and Social Science*, 685 (1), 190–226.

Weisburd, David, Michael Davis, and Charlotte Gill (2015) "Increasing collective efficacy and social capital at crime hot spots: New crime control tools for police," *Policing: A Journal of Policy and Practice*, 9 (3), 265–274.

Weisburd, Sarit (2021) "Police presence, rapid response rates, and crime prevention," *Review of Economics and Statistics*, 103 (2), 280–293.

Weitzer, Ronald, Steven A Tuch, and Wesley G Skogan (2008) "Police–community relations in a majority-Black city," *Journal of Research in Crime and Delinquency*, 45 (4), 398–428.

Western, Bruce (2006) *Punishment and inequality in America*: Russell Sage Foundation.

Westfall, Peter H and S Stanley Young (1993) *Resampling-based multiple testing: Examples and methods for p-value adjustment*, 279: John Wiley & Sons.

Wilson, Timothy D (2004) *Strangers to ourselves*: Harvard University Press.

Wu, Xiaoyun and Cynthia Lum (2017) "Measuring the spatial and temporal patterns of police proactivity," *Journal of quantitative criminology*, 33 (4), 915–934.

10 Tables and Figures

Table I: C2C study youth baseline characteristics

Variable	Mean T	Mean C	Difference
Demographics			
Age	16.29	16.26	0.03
% Black	0.94	0.96	-0.01
% Hispanic	0.05	0.04	0.01
% Female	0.43	0.41	0.02
Number of arrests			
Number of prior arrests	1.41	1.57	-0.16
Number of prior violent arrests	0.38	0.45	-0.07*
Number of prior property arrests	0.36	0.36	0.00
Number of prior drug arrest	0.10	0.10	0.00
Number of prior other arrests	0.57	0.67	-0.09
Any arrest			
Any prior arrest	0.35	0.36	-0.01
Any prior violent arrest	0.22	0.23	-0.01
Any prior property arrest	0.17	0.18	-0.02
Any prior drug arrest	0.06	0.05	0.01
Any prior other arrest	0.21	0.23	-0.03*
Victimizations			
Number of prior victimizations	0.90	0.94	-0.04
Any prior victimization	0.49	0.48	0.00
Education			
Attendance rate	0.74	0.72	0.02*
GPA	1.97	1.92	0.05
Misconduct incidents	0.89	0.99	-0.10
Had free/reduce lunch status	0.96	0.95	0.01
School grade at baseline	9.78	9.72	0.06
Had an enrollment gap	0.20	0.20	-0.01
Observations	1052	1022	
P-value on Join F Test	0.55		

Notes: Mean differences for the treatment vs. control groups was estimated using a linear regression with referral level and cohort fixed effects (randomization was conducted within referral source and cohort). Standard errors are robust. Stars indicator the following p-values: *** p<0.01, ** p<0.05, * p<0.1. School grade at baseline is only available for 1,996 youth. Attendance rate, GPA, and misconduct incidents only available for 2,065 youth. Arrests are categorized using FBI code and statutes. Violent arrests include arrests for murder, sexual assault, robbery, aggravated assault or battery, simple assault and battery, sexual misconduct, and some miscellaneous violent statutes. Variables included in the F-test are limited to the variables for which we have data for all study participants – demographics, indicators for every category of prior arrest, indicator for any prior victimization, indicator for prior free/reduced lunch status, and prior disconnection from school (overall any arrest is not including given it's a linear combination of the subcategories). F-test is not sensitive to whether we include extensive measures of arrest or intensive measure of arrests.

Table II: Program take-up, trauma-informed CBT sessions & mentoring hours, all cohorts and select subgroups

			SPARCS/CBT Hours				Mentoring Hours			
	N Treat (Take-up Rate)	Avg Total Hours	25th %tile	50th %tile	75th %tile	Mean	25th %tile	50th %tile	75th %tile	Mean
All Participants	1052 (62%)	186	4	9	14	9	126	177	223	177
Gender										
Female	455 (67%)	173	4	10	14	9	113	164	207	163
Male	597 (58%)	198	3	9	14	10	144	191	233	189
Any Baseline Arrest										
Has prior arrest	363 (51%)	192	3	8	14	9	124	178	224	183
No prior arrest	689 (68%)	184	5	10	14	10	126	177	223	174
School Disengagement Indicator										
Engaged in school at baseline	747 (69%)	189	4	10	14	10	127	184	226	179
Not engaged in school at baseline	305 (46%)	175	2	8	13	8	118	166	209	167

Notes: This table produces the average number of CBT (SPARCS) sessions attended and the average number of YAP/mentoring received during the course of programming for participants in the program. We include the number of assigned treated youth by baseline characteristic, the percent who take-up by baseline characteristic, and the distribution of SPARCS and mentorship engagement among the participants. We use the following criteria for not engaged in school at baseline: any youth who missed more than 60% of school days and received a grade of "F" in at least 75% of their courses in the semester prior to randomization, had a serious learning disability, were at least two years older than expected for the grade they were enrolled in at baseline, or were listed inactive at the end of any of the three semesters immediately prior to randomization

Table III: C2C estimate effects on arrest outcomes, 6 and 12 cumulative Months post-randomization

	Estimates				P-values		
Outcome	CM	ITT	CCM	TOT	Observed ITT	FWER	FDR
Intensive Margin 6 months							
Number of arrests	0.248	-0.013 (0.025)	0.170	-0.021 (0.040)	0.610	0.918	1.000
Number of violent arrests	0.064	-0.023 (0.010)	0.079	-0.037 (0.016)	0.017**	0.078*	0.085*
Number of property arrests	0.057	0.006 (0.012)	0.025	0.009 (0.019)	0.631	0.918	1.000
Number of drug arrests	0.017	0.013 (0.008)	0.000	0.022 (0.012)	0.076*	0.271	0.302
Number of other arrests	0.111	-0.009 (0.015)	0.073	-0.014 (0.025)	0.566	0.918	1.000
Extensive Margin 6 months							
Any arrest	0.161	-0.018 (0.014)	0.140	-0.029 (0.022)	0.199	0.484	0.598
Any violent arrest	0.060	-0.021 (0.009)	0.073	-0.035 (0.014)	0.017**	0.084*	0.087*
Any property arrest	0.049	0.000 (0.009)	0.028	0.000 (0.014)	0.987	0.988	0.987
Any drug arrest	0.012	0.012 (0.005)	0.000	0.019 (0.009)	0.032**	0.122	0.128
Any other arrest	0.085	-0.008 (0.011)	0.065	-0.013 (0.017)	0.452	0.709	0.904
Intensive Margin 12 months							
Number of arrests	0.490	-0.063 (0.040)	0.370	-0.103 (0.064)	0.115	0.397	0.460
Number of violent arrests	0.116	-0.026 (0.014)	0.115	-0.043 (0.023)	0.070*	0.307	0.349
Number of property arrests	0.133	-0.019 (0.019)	0.099	-0.031 (0.030)	0.302	0.659	0.905
Number of drug arrests	0.035	0.002 (0.010)	0.016	0.003 (0.016)	0.861	0.860	0.905
Number of other arrests	0.205	-0.020 (0.023)	0.140	-0.032 (0.036)	0.384	0.659	0.905
Extensive Margin 12 months							
Any arrest	0.244	-0.041 (0.016)	0.223	-0.067 (0.025)	0.008***	0.041**	0.039**
Any violent arrest	0.098	-0.025 (0.011)	0.098	-0.041 (0.018)	0.028**	0.114	0.114
Any property arrest	0.095	-0.012 (0.012)	0.074	-0.020 (0.019)	0.291	0.505	0.582
Any drug arrest	0.025	0.002 (0.007)	0.009	0.003 (0.010)	0.755	0.762	0.755
Any other arrest	0.133	-0.018 (0.012)	0.113	-0.029 (0.020)	0.155	0.410	0.466

Notes: CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. FWER is the family wise error rate p-value, calculated using the outcomes during the same outcome period and margin as the family. FDR is the false discovery rate p-value. Standard errors are robust and shown in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table IV: C2C estimate effects on arrest outcomes, 24 and 36 cumulative months post-randomization

Outcome	Estimates				P-values		
	CM	ITT	CCM	TOT	Observed ITT	FWER	FDR
Intensive Margin 24 months							
Number of arrests	0.871	-0.054 (0.064)	0.635	-0.088 (0.103)	0.398	0.874	1.000
Number of violent arrests	0.226	-0.041 (0.023)	0.202	-0.066 (0.037)	0.076*	0.331	0.378
Number of property arrests	0.218	-0.009 (0.027)	0.160	-0.015 (0.043)	0.730	0.983	1.000
Number of drug arrests	0.062	-0.001 (0.013)	0.034	-0.001 (0.021)	0.946	0.995	1.000
Number of other arrests	0.365	-0.003 (0.036)	0.240	-0.005 (0.057)	0.925	0.995	1.000
Extensive Margin 24 months							
Any arrest	0.345	-0.063 (0.017)	0.328	-0.103 (0.027)	0.000***	0.001***	0.001***
Any violent arrest	0.164	-0.037 (0.014)	0.152	-0.060 (0.022)	0.009***	0.030**	0.034**
Any property arrest	0.137	-0.007 (0.014)	0.106	-0.011 (0.022)	0.634	0.865	1.000
Any drug arrest	0.044	0.003 (0.008)	0.020	0.005 (0.014)	0.717	0.865	1.000
Any other arrest	0.197	-0.017 (0.015)	0.164	-0.027 (0.023)	0.250	0.573	0.750
Intensive Margin 36 months							
Number of arrests	1.175	-0.084 (0.078)	0.915	-0.136 (0.125)	0.285	0.818	1.000
Number of violent arrests	0.289	-0.029 (0.028)	0.245	-0.047 (0.045)	0.299	0.818	1.000
Number of property arrests	0.272	-0.009 (0.032)	0.210	-0.014 (0.051)	0.788	0.955	1.000
Number of drug arrests	0.090	-0.004 (0.016)	0.056	-0.006 (0.026)	0.812	0.955	1.000
Number of other arrests	0.524	-0.042 (0.043)	0.404	-0.069 (0.069)	0.327	0.818	1.000
Extensive Margin 36 months							
Any arrest	0.386	-0.052 (0.018)	0.363	-0.084 (0.028)	0.004***	0.018**	0.018**
Any violent arrest	0.196	-0.031 (0.015)	0.179	-0.051 (0.024)	0.040**	0.134	0.148
Any property arrest	0.159	-0.007 (0.015)	0.128	-0.011 (0.024)	0.641	0.872	1.000
Any drug arrest	0.059	0.004 (0.010)	0.032	0.006 (0.015)	0.694	0.872	1.000
Any other arrest	0.250	-0.033 (0.016)	0.229	-0.053 (0.025)	0.037**	0.134	0.148

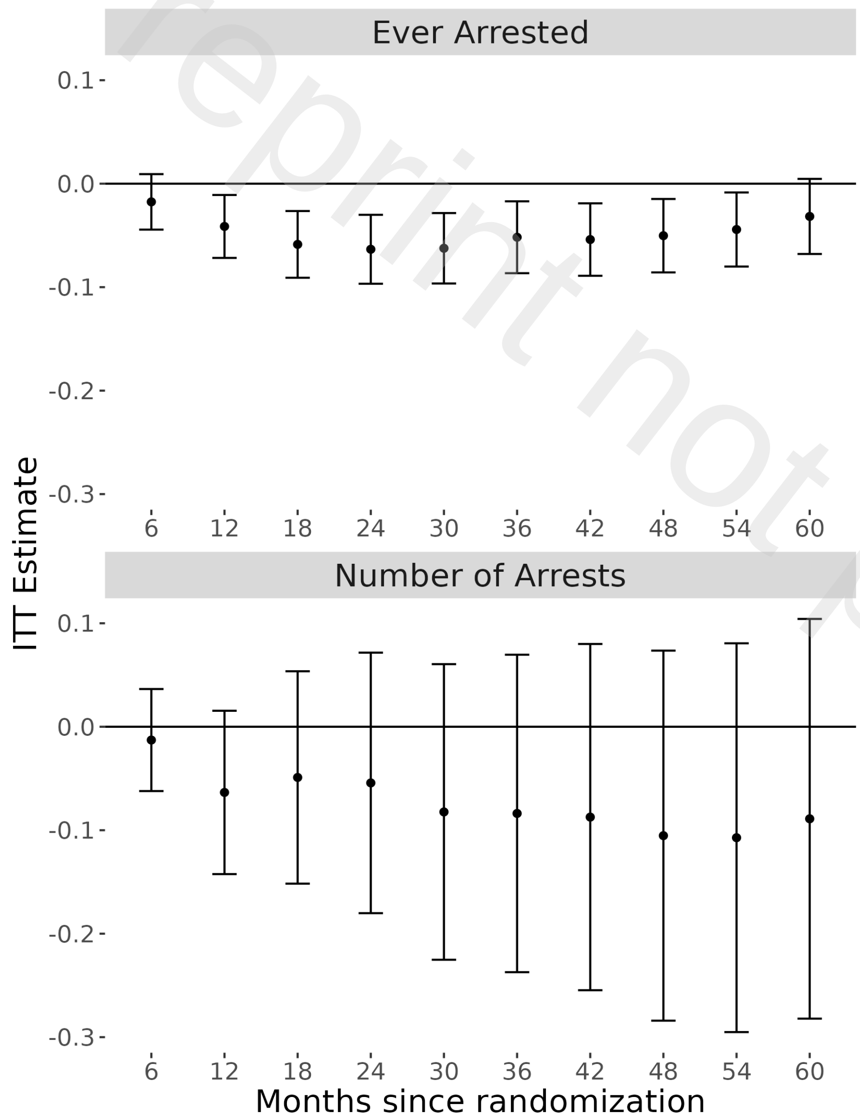
Notes: CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. FWER is the family wise error rate p-value, calculated using the outcomes during the same outcome period and margin as the family. FDR is the false discovery rate p-value. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table V: C2C estimate effects on arrest outcomes, 48 and 60 cumulative months post-randomization

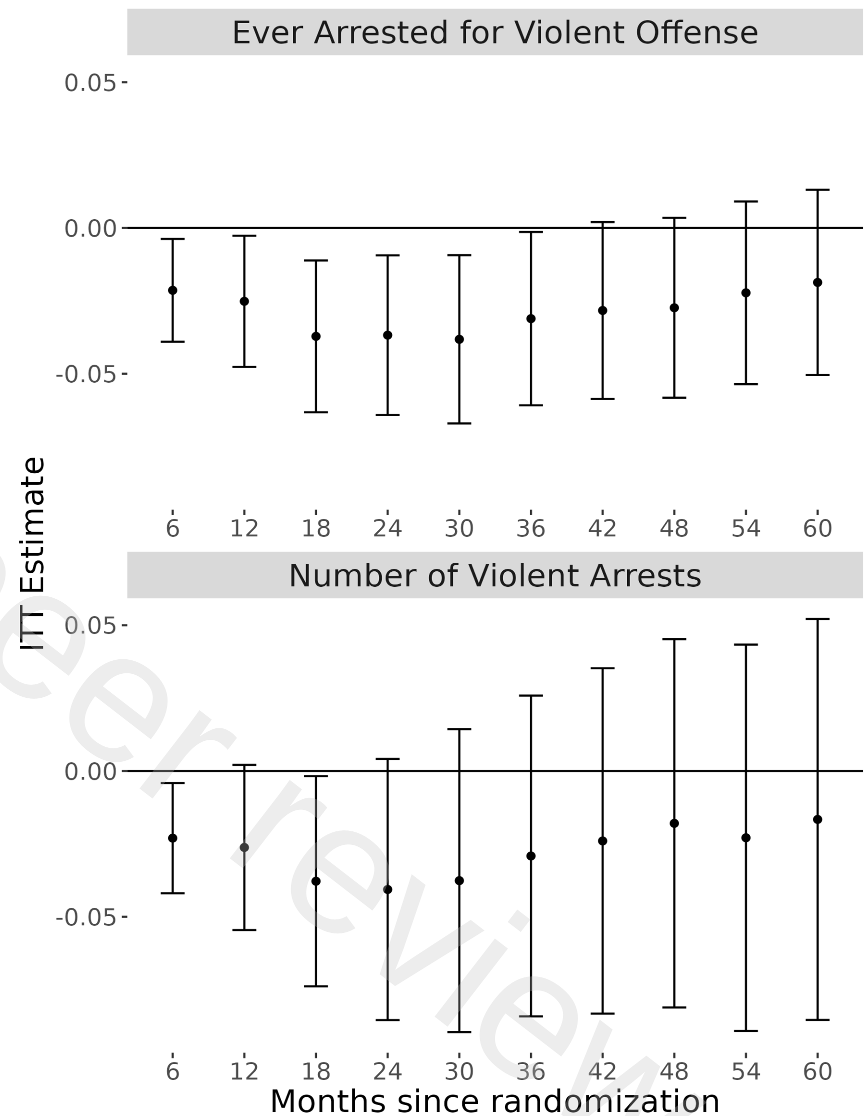
	Estimates				P-values		
Outcome	CM	ITT	CCM	TOT	Observed ITT	FWER	FDR
Intensive Margin 48 months							
Number of arrests	1.425	-0.105 (0.091)	1.124	-0.171 (0.146)	0.249	0.682	0.995
Number of violent arrests	0.336	-0.018 (0.032)	0.268	-0.029 (0.052)	0.577	0.926	1.000
Number of property arrests	0.318	-0.014 (0.036)	0.255	-0.022 (0.058)	0.707	0.926	1.000
Number of drug arrests	0.103	-0.002 (0.018)	0.064	-0.003 (0.029)	0.908	0.926	1.000
Number of other arrests	0.668	-0.072 (0.051)	0.537	-0.117 (0.081)	0.158	0.581	0.789
Extensive Margin 48 months							
Any arrest	0.421	-0.050 (0.018)	0.395	-0.082 (0.029)	0.006***	0.026**	0.028**
Any violent arrest	0.215	-0.027 (0.016)	0.194	-0.045 (0.025)	0.082*	0.231	0.246
Any property arrest	0.180	-0.015 (0.015)	0.159	-0.025 (0.025)	0.326	0.546	0.651
Any drug arrest	0.068	0.002 (0.010)	0.043	0.003 (0.016)	0.839	0.832	0.839
Any other arrest	0.289	-0.036 (0.016)	0.261	-0.058 (0.026)	0.030**	0.117	0.120
Intensive Margin 60 months							
Number of arrests	1.586	-0.089 (0.099)	1.262	-0.145 (0.158)	0.367	0.839	1.000
Number of violent arrests	0.377	-0.017 (0.035)	0.306	-0.027 (0.056)	0.636	0.952	1.000
Number of property arrests	0.341	-0.010 (0.038)	0.269	-0.017 (0.060)	0.785	0.952	1.000
Number of drug arrests	0.115	0.000 (0.020)	0.075	0.000 (0.033)	0.994	0.992	1.000
Number of other arrests	0.753	-0.062 (0.055)	0.611	-0.101 (0.088)	0.257	0.781	1.000
Extensive Margin 60 months							
Any arrest	0.438	-0.032 (0.019)	0.407	-0.051 (0.030)	0.088*	0.374	0.438
Any violent arrest	0.230	-0.019 (0.016)	0.205	-0.030 (0.026)	0.249	0.586	0.748
Any property arrest	0.189	-0.013 (0.016)	0.161	-0.022 (0.025)	0.389	0.625	0.778
Any drug arrest	0.073	0.001 (0.011)	0.052	0.002 (0.017)	0.922	0.918	0.922
Any other arrest	0.311	-0.027 (0.017)	0.282	-0.044 (0.027)	0.108	0.374	0.438

Notes: CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. FWER is the family wise error rate p-value, calculated using the outcomes during the same outcome period and margin as the family. FDR is the false discovery rate p-value. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure I: C2C's estimated effects on arrest outcomes over time (ITT estimates, cumulative)



(a) C2C program effects on overall arrests, ITT



(b) C2C program effects on arrests for violent offenses, ITT

Notes: Figures show the ITT estimates for different categories of arrests over time. The whisker lines represent unadjusted 95% confidence intervals. Estimates and confidence intervals are generated from regression with robust standard errors that include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations.

Table VI: C2C estimated effects on educational outcomes

	Estimates				P-values		
Outcome	CM	ITT	CCM	TOT	Observed ITT	FWER	FDR
Semester 1							
Academic Index - Combined	0.000	0.033	0.088	0.053	0.029**		
		(0.015)		(0.024)			
Academic Index - Engagement	0.000	0.028	0.196	0.045	0.267		
		(0.025)		(0.040)			
Attendance Days	54.420	1.610	59.489	2.600	0.089*	0.252	0.266
		(0.945)		(1.494)			
Ever Enrolled or Graduated	0.848	-0.018	0.934	-0.029	0.247	0.436	0.495
		(0.015)		(0.024)			
GPA	1.754	0.041	1.997	0.067	0.289	0.436	0.495
		(0.039)		(0.062)			
Academic Index - Behavioral	0.000	0.055	-0.048	0.089	0.016**		
		(0.023)		(0.037)			
Misconduct Incidents	0.597	-0.098	0.677	-0.159	0.017**	0.050**	0.051*
		(0.041)		(0.065)			
Ever Enrolled in Juv. Justice	0.035	0.002	0.013	0.003	0.817	0.811	0.817
		(0.007)		(0.011)			
No. of CPS Schools Enrolled In	0.927	-0.039	1.018	-0.063	0.054*	0.101	0.107
		(0.020)		(0.032)			
Semester 2							
Academic Index - Combined	0.000	0.048	0.042	0.078	0.001***		
		(0.014)		(0.023)			
Academic Index - Engagement	0.000	0.041	0.128	0.066	0.086*		
		(0.024)		(0.038)			
Attendance Days	49.247	1.987	52.961	3.210	0.041**	0.123	0.122
		(0.970)		(1.537)			
Ever Enrolled or Graduated	0.793	-0.028	0.878	-0.045	0.109	0.213	0.218
		(0.017)		(0.028)			
GPA	1.721	0.058	1.883	0.095	0.119	0.213	0.218
		(0.037)		(0.060)			
Academic Index - Behavioral	0.000	0.076	-0.076	0.123	0.001***		
		(0.022)		(0.036)			
Misconduct Incidents	0.477	-0.109	0.563	-0.177	0.005***	0.017**	0.016**
		(0.039)		(0.063)			
Ever Enrolled in Juv. Justice	0.033	-0.010	0.029	-0.017	0.124	0.173	0.174
		(0.007)		(0.011)			
No. of CPS Schools Enrolled In	0.858	-0.037	0.947	-0.060	0.087*	0.173	0.174
		(0.022)		(0.035)			

Notes: Our index is equal to an unweighted average of the different components, after all have been normalized to Z-score form using the control group's distribution. Index components used "mixed" imputation methodology for missing values, see Appendix Table A.IX for outcome sensitivity to imputation methodology. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, number of prior victimizations, and a dummy indicator for whether the outcome has been imputed. Index outcomes include multiple dummies for whether each component was imputed. FWER is the family wise error rate p-values, where the three components of each separate index during each semester constitute each family. FDR is the false discovery rate. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table VII: C2C estimated effects on graduation

		ITT			TOT		
	Model N	Control Mean	Estimate	P-Value	CCM	Estimate	P-Value
Graduated Only							
Any graduation	2065	0.550	0.003	0.872	0.657	0.005	0.869
			(0.019)			(0.031)	
Graduated on-track	2054	0.459	0.019	0.333	0.564	0.031	0.323
			(0.019)			(0.031)	
Graduated on-track + 1	1998	0.497	0.000	0.998	0.614	0.000	0.998
			(0.019)			(0.031)	
Graduated on-track + 2	1913	0.492	-0.004	0.853	0.630	-0.006	0.851
			(0.020)			(0.032)	
Graduated or Transferred							
Any graduation	2065	0.645	-0.003	0.851	0.743	-0.006	0.849
			(0.018)			(0.029)	
Graduated on-track	2054	0.550	0.009	0.651	0.647	0.014	0.645
			(0.019)			(0.031)	
Graduated on-track + 1	1998	0.588	-0.014	0.470	0.700	-0.023	0.463
			(0.019)			(0.031)	
Graduated on-track + 2	1913	0.574	-0.013	0.509	0.710	-0.021	0.502
			(0.019)			(0.031)	

Notes: For on-track graduation outcomes, a projected graduation year is calculated for each study member based on their grade level at the time of graduation. For those with missing grade levels at randomization, we calculate a projected graduation year using the highest grade level listed in any CPS masterfile for each student prior to randomization. Graduated on-track + 1 allows for one extra school year beyond each person's projected graduation semester; Graduated on-track + 2 allows for two extra school years. Graduated on-track outcomes subset to the sample to those for which we can observe their projected graduation year. Any graduation outcomes does not take projected graduation year into account, and considers the entire sample. Graduated or Transferred outcomes include study members who have an out-of-CPS transfer as their last leave code prior to their projected graduation year, providing an upper-bound on graduation rates. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. *** p<0.01, ** p<0.05, * p<0.1

Table IX: C2C's estimated effect on the social cost of crime over time

	Bottom-up Estimates				Willingness-to-pay Estimates			
	CM	ITT	CCM	TOT	CM	ITT	CCM	TOT
6 months	$4,494	-$1,333**	$4,743	-$2,168**	$12,569	-$4,008**	$13,064	-$6,516**
		($544)		($872)		($1,738)		($2,783)
12 months	$8,242	-$1,538*	$7,527	-$2,501*	$23,985	-$3,864	$21,983	-$6,283
		($818)		($1,310)		($2,948)		($4,718)
24 months	$16,237	-$3,208**	$14,371	-$5,217**	$47,727	-$7,839*	$40,590	-$12,745*
		($1,396)		($2,232)		($4,684)		($7,487)
36 months	$22,444	-$4,203**	$20,133	-$6,834**	$67,034	-$12,424**	$60,306	-$20,201**
		($1,788)		($2,860)		($5,770)		($9,230)
48 months	$27,027	-$3,776*	$23,808	-$6,140*	$80,908	-$11,393*	$71,967	-$18,525*
		($2,031)		($3,248)		($6,512)		($10,414)
60 months	$31,659	-$3,856*	$28,490	-$6,269*	$94,915	-$12,241*	$85,943	-$19,903*
		($2,262)		($3,618)		($7,300)		($11,675)

Notes: All estimates inflated to 2015 dollars. Program costs are $5,070 per C2C participant, or $3,147 per youth randomized to C2C treatment, in 2015 dollars. Bottom-up estimates assign the cost of each arrest and victimization based on the "bottom-up" cost estimates from Cohen and Piquero (2009), which aim to sum up costs to victims, legal system cost, and offender productivity cost for each event. Willingness-to-pay etimates use estimates from Cohen and Piquero (2009), which uses a "top-down" approach aiming to estimate the amount the public would be willing to pay to reduce each crime type. Cost inputs for each crime type can be found in Table A.XXIV. Homicides are recoded as aggravated assaults to match the cost of non-fatal shootings, given how rare these events are within our sample and how sensitive these estimates are each additional homicide. For each outcome period, we calculate the total cost incurred by each person by summing their arrest and victimization costs, which is used as the dependent variable in our regressions. Bottom-up and willingness-to-pay costs are estimated based on observed arrests and victimizations only. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix

A.1.1 Details on the C2C Program

Table A.I: C2C Program Elements, Wraparound Services

Wraparound Services	
Component	**Purpose**
Around-the-clock support and crisis intervention	YAP advocates are available 24/7 to provide immediate support when crises arise and help youth learn new ways of responding to challenges.
Creating a sense of security	The program helps youth and their family meet basic needs, as it includes consistent contact, care, wraparound services, and engagement with families to create a social safety net that youth can rely on even after their program participation ends.
New Experiences & group activities	New opportunities with peers can help youth process trauma by experiencing safety and relaxation in new environments. It also allows youth the opportunity to practice new ways of coping with peers that are working to make the same changes. C2C also exposes young people to new career opportunities and employment opportunities both inside and outside of their communities.
Individualized Support	YAP advocates formally identify young people's strengths, needs, preferences and goals across a spectrum of life domains to develop individualized service plans that become the basis of support.

Table A.II: C2C Program Elements, SPARCS†

SPARCS	
Topic	**Purpose**
Mindfulness exercises	Students are taught to be present, pay attention in a particular way on purpose and without judgement. They recognition and labeling of the link between emotion and the body, linking trauma to somatic symptoms. The goal is to affect regulation and impulsivity. An example of mindfulness exercises includes controlled breathing and weekly SOS (slow down, orient, self-check) practice.
Skill building: coping in the moment	Youth are given the tools to handle situations they can't immediately fix or change or in which action may worsen the situation. These tools involve distress tolerance skills such as distract and self-soothe techniques, identifying MUPS (coping strategies tat "Mess you UP") and defining ways these strategies may actually exacerbate or form new problems.
Skill building: Problem solving and creating meaning	Youth are taught skills to address past traumatic experiences or situations they do have control over. Youth are encouraged to construct a sense of purpose and meaning in their lives despite trauma. Utilizing LET'M GO (losing it, emotions, thoughts, meaning, goals and options) practice, they map out elements of reactions to an emotion situation and teach problem solving skills to address those emotions mindfully.
Skill building: Collaboration and communication	To address youth problems with alienation and trust youth learn and regularly practice communication skills through collaborative group work. The aim is also to assist youth in identifying and strengthening sources of social support. The MAKE A LINK technique provides a step-by-step guide so youth can more effectively manage their interpersonal interactions in any environment.
Immersive and experiential	SPARCS lessons are carried out in ways that tie in personal life experiences to better connect youth with the materials they learn about on paper. Lessons like, Portrait of my Life, helps participants learn about two key concepts, triggers and regulation of emotions, anger as well as the varying levels of intensity in which these feelings can occur.

†Structured Psychotherapy for Adolescents Responding to Chronic Stress

Figure A.I: Cohort 1-8 Communities Served

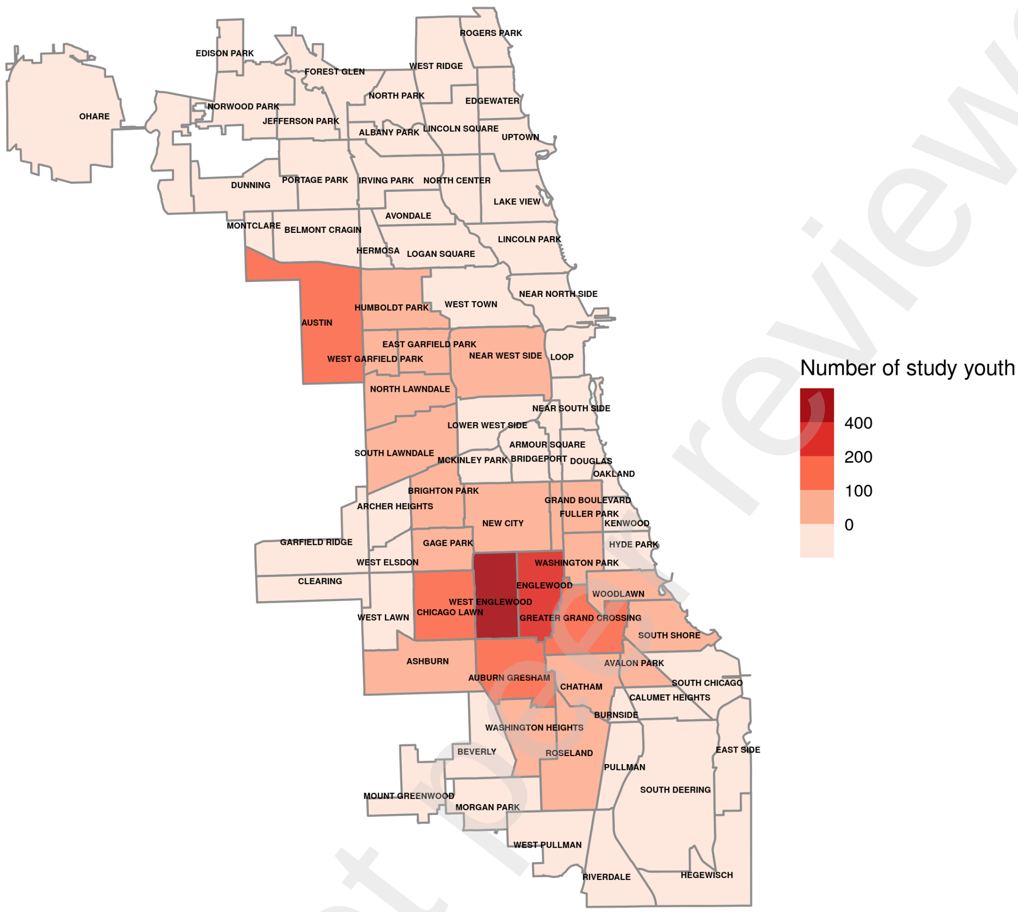
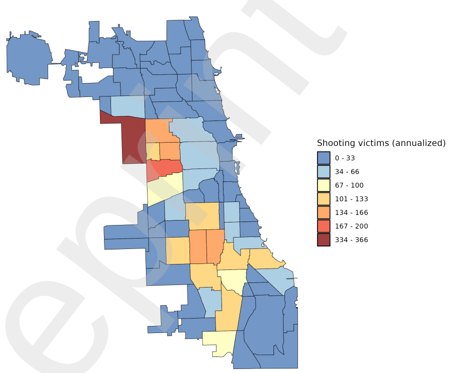
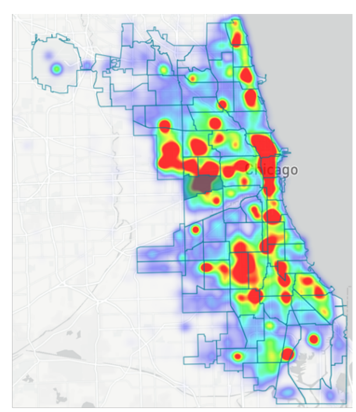


Figure A.II: Exposure to interpersonal violence and police



(a) Annual average number of shootings by neighborhood, Source: University of Chicago Crime Lab



(b) CPD Complaints, 32 year total, Source: Invisible Institute

A.1.2 Details on C2C Arrest Effects

Figure A.III: C2C control means by 6-month period relative to randomization, number of arrests

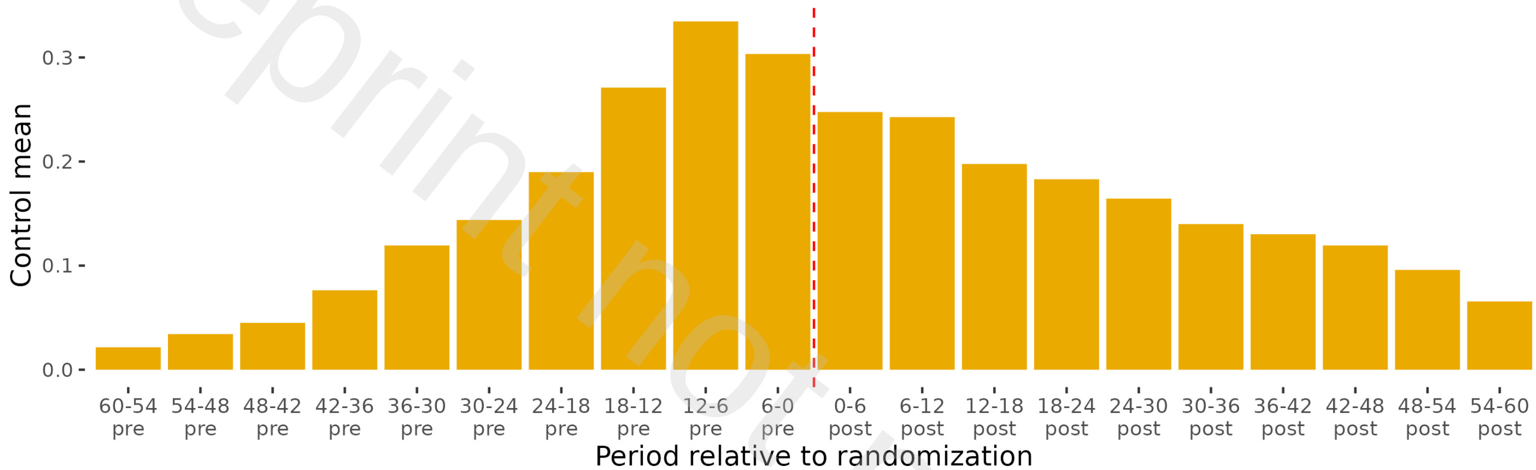
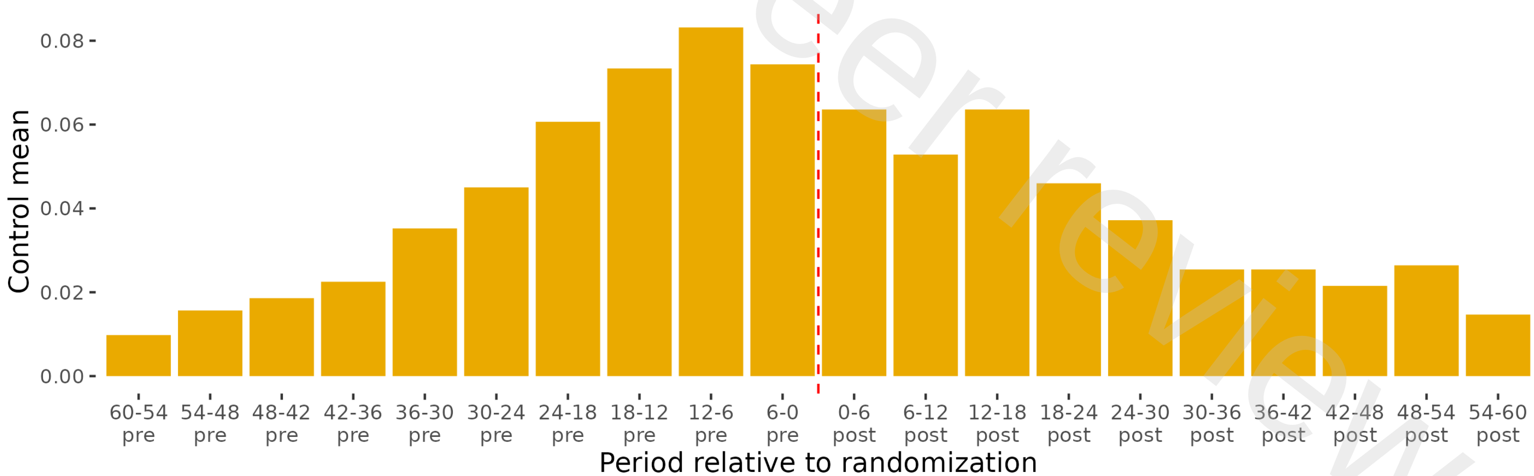
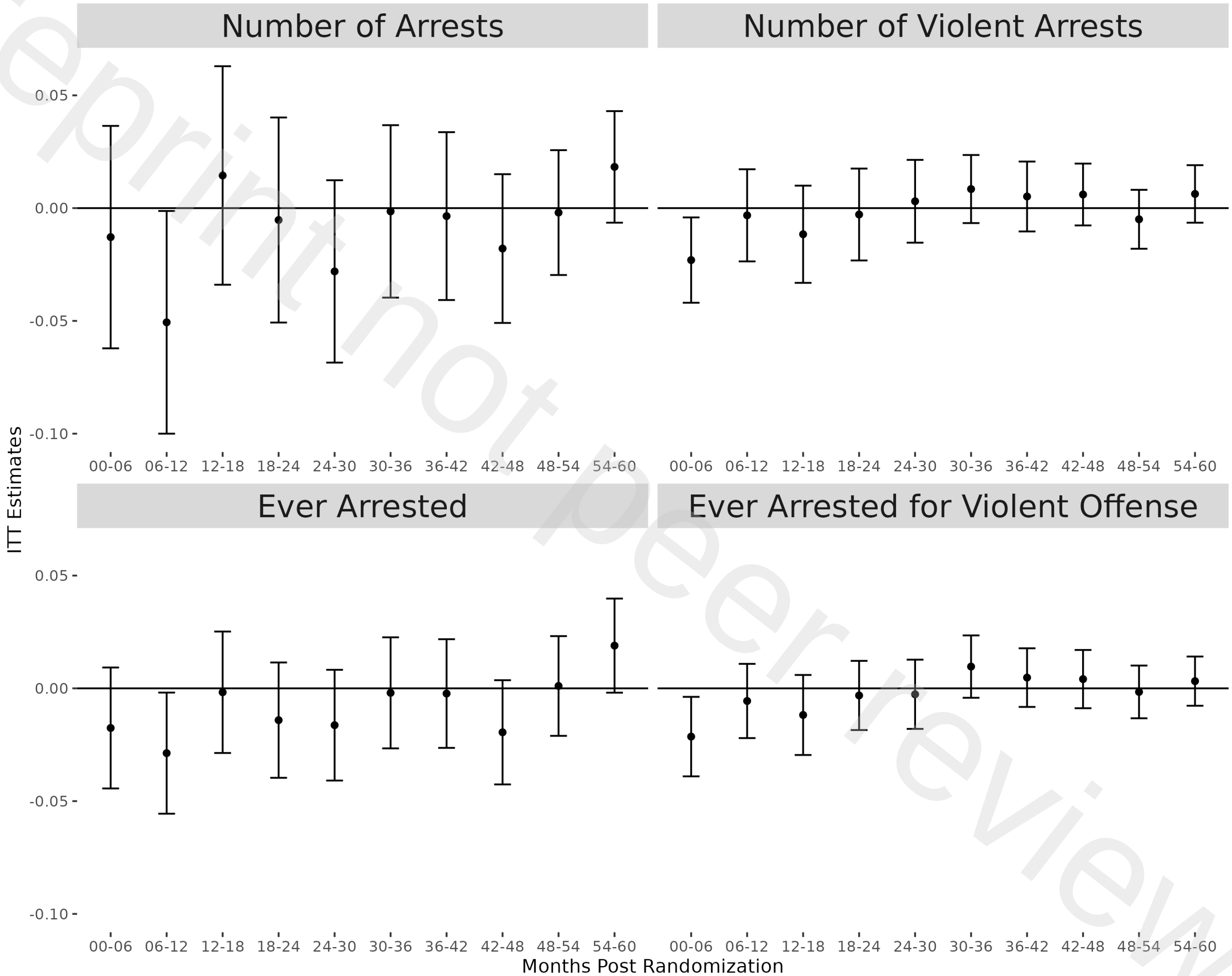


Figure A.IV: C2C control means by 6-month period relative to randomization, number of arrests for a violent offenses



Note: Pre bars represent the number of months prior to each study member's randomization date, post bars represent the number of months post-randomization. Figures shows control means for total number of arrests and number of arrests for violent offenses, for each 6-month interval leading up to randomization and post-randomization.

Figure A.V: C2C non-cumulative effects, ITT



Note: Figures show the ITT estimates for different categories of arrests over non-cumulative 6-month periods. The whisker lines represent unadjusted 95% confidence intervals. Estimates and confidence intervals are generated from regression with robust standard errors that include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations

Table A.III: C2C non-cumulative effects, all arrests

		ITT		TOT		
	Control Mean	Estimate	P-Value	CCM	Estimate	P-Value
Number of Arrests						
0-6 months	0.248	-0.013	0.610	0.170	-0.021	0.603
		(0.025)			(0.040)	
6-12 months	0.243	-0.051	0.045**	0.199	-0.082	0.041**
		(0.025)			(0.040)	
12-18 months	0.198	0.014	0.558	0.126	0.024	0.552
		(0.025)			(0.040)	
18-24 months	0.183	-0.005	0.820	0.140	-0.009	0.817
		(0.023)			(0.037)	
24-30 months	0.164	-0.028	0.174	0.164	-0.046	0.167
		(0.021)			(0.033)	
30-36 months	0.140	-0.001	0.941	0.115	-0.002	0.940
		(0.020)			(0.031)	
36-42 months	0.130	-0.004	0.852	0.095	-0.006	0.850
		(0.019)			(0.030)	
42-48 months	0.119	-0.018	0.287	0.114	-0.029	0.279
		(0.017)			(0.027)	
48-54 months	0.096	-0.002	0.889	0.082	-0.003	0.887
		(0.014)			(0.023)	
54-60 months	0.066	0.018	0.148	0.057	0.030	0.140
		(0.013)			(0.020)	
Ever Arrested						
0-6 months	0.161	-0.018	0.199	0.140	-0.029	0.192
		(0.014)			(0.022)	
6-12 months	0.159	-0.029	0.036**	0.135	-0.047	0.033**
		(0.014)			(0.022)	
12-18 months	0.135	-0.002	0.899	0.095	-0.003	0.898
		(0.014)			(0.022)	
18-24 months	0.126	-0.014	0.280	0.109	-0.023	0.272
		(0.013)			(0.021)	
24-30 months	0.115	-0.016	0.192	0.110	-0.027	0.186
		(0.013)			(0.020)	
30-36 months	0.104	-0.002	0.874	0.091	-0.003	0.872
		(0.013)			(0.020)	
36-42 months	0.097	-0.002	0.851	0.072	-0.004	0.849
		(0.012)			(0.020)	
42-48 months	0.099	-0.019	0.098*	0.099	-0.032	0.093*
		(0.012)			(0.019)	
48-54 months	0.080	0.001	0.926	0.066	0.002	0.925
		(0.011)			(0.018)	
54-60 months	0.058	0.019	0.075*	0.048	0.031	0.070*
		(0.011)			(0.017)	

Notes: Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. FWER is the family wise error rate p-value, calculated using the outcomes during the same outcome period and margin as the family. FDR is the false discovery rate p-value. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.IV: C2C non-cumulative effects, arrests for violent offenses

		ITT		TOT		
	Control Mean	Estimate	P-Value	CCM	Estimate	P-Value
Number of Violent Arrests						
0-6 months	0.064	-0.023	0.017**	0.079	-0.037	0.016**
		(0.010)			(0.016)	
6-12 months	0.053	-0.003	0.759	0.036	-0.005	0.755
		(0.010)			(0.017)	
12-18 months	0.064	-0.012	0.292	0.050	-0.019	0.284
		(0.011)			(0.018)	
18-24 months	0.046	-0.003	0.785	0.037	-0.005	0.782
		(0.010)			(0.017)	
24-30 months	0.037	0.003	0.746	0.027	0.005	0.742
		(0.009)			(0.015)	
30-36 months	0.025	0.008	0.272	0.016	0.014	0.264
		(0.008)			(0.012)	
36-42 months	0.025	0.005	0.514	0.012	0.008	0.507
		(0.008)			(0.013)	
42-48 months	0.022	0.006	0.388	0.012	0.010	0.380
		(0.007)			(0.011)	
48-54 months	0.026	-0.005	0.458	0.023	-0.008	0.450
		(0.007)			(0.011)	
54-60 months	0.015	0.006	0.334	0.014	0.010	0.325
		(0.007)			(0.010)	
Ever Arrested for Violent Offense						
0-6 months	0.060	-0.021	0.017**	0.073	-0.035	0.016**
		(0.009)			(0.014)	
6-12 months	0.045	-0.006	0.505	0.035	-0.009	0.498
		(0.008)			(0.013)	
12-18 months	0.056	-0.012	0.192	0.045	-0.019	0.185
		(0.009)			(0.014)	
18-24 months	0.038	-0.003	0.687	0.031	-0.005	0.682
		(0.008)			(0.013)	
24-30 months	0.035	-0.003	0.737	0.032	-0.004	0.733
		(0.008)			(0.013)	
30-36 months	0.023	0.010	0.172	0.012	0.016	0.165
		(0.007)			(0.011)	
36-42 months	0.023	0.005	0.473	0.011	0.008	0.465
		(0.007)			(0.011)	
42-48 months	0.022	0.004	0.534	0.013	0.007	0.527
		(0.007)			(0.011)	
48-54 months	0.023	-0.002	0.791	0.018	-0.003	0.787
		(0.006)			(0.010)	
54-60 months	0.015	0.003	0.566	0.015	0.005	0.559
		(0.006)			(0.009)	

Notes: Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. FWER is the family wise error rate p-value, calculated using the outcomes during the same outcome period and margin as the family. FDR is the false discovery rate p-value. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.V: C2C effects without covariates, using only randomization blocks

	Estimates				P-values
	CM	ITT	CCM	TOT	Observed ITT
6 months					
Number of arrests	0.248	-0.037 (0.029)	0.209	-0.059 (0.047)	0.212
Number of violent arrests	0.064	-0.027 (0.010)	0.085	-0.044 (0.016)	0.008***
Any arrest	0.161	-0.029 (0.015)	0.159	-0.048 (0.024)	0.049**
Any violent arrest	0.060	-0.025 (0.009)	0.079	-0.040 (0.015)	0.007***
12 months					
Number of arrests	0.490	-0.100 (0.045)	0.430	-0.163 (0.073)	0.026**
Number of violent arrests	0.116	-0.034 (0.015)	0.128	-0.055 (0.025)	0.027**
Any arrest	0.244	-0.054 (0.017)	0.243	-0.087 (0.027)	0.002***
Any violent arrest	0.098	-0.030 (0.012)	0.105	-0.048 (0.019)	0.012**
24 months					
Number of arrests	0.871	-0.116 (0.070)	0.736	-0.189 (0.113)	0.098*
Number of violent arrests	0.226	-0.058 (0.024)	0.231	-0.095 (0.039)	0.015**
Any arrest	0.345	-0.079 (0.019)	0.353	-0.128 (0.030)	0.000***
Any violent arrest	0.164	-0.047 (0.015)	0.169	-0.077 (0.024)	0.001***
36 months					
Number of arrests	1.175	-0.163 (0.088)	1.044	-0.266 (0.142)	0.063*
Number of violent arrests	0.289	-0.052 (0.029)	0.281	-0.084 (0.047)	0.078*
Any arrest	0.386	-0.068 (0.020)	0.389	-0.110 (0.032)	0.001***
Any violent arrest	0.196	-0.042 (0.016)	0.197	-0.069 (0.026)	0.008***
48 months					
Number of arrests	1.425	-0.195 (0.103)	1.269	-0.317 (0.166)	0.058*
Number of violent arrests	0.336	-0.040 (0.034)	0.304	-0.065 (0.054)	0.232
Any arrest	0.421	-0.064 (0.020)	0.417	-0.104 (0.032)	0.001***
Any violent arrest	0.215	-0.039 (0.017)	0.213	-0.063 (0.027)	0.020**
60 months					
Number of arrests	1.586	-0.192 (0.112)	1.430	-0.313 (0.180)	0.086*
Number of violent arrests	0.377	-0.042 (0.037)	0.347	-0.068 (0.059)	0.251
Any arrest	0.438	-0.047 (0.020)	0.433	-0.077 (0.033)	0.020**
Any violent arrest	0.230	-0.032 (0.017)	0.226	-0.052 (0.028)	0.067*

Notes: Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. *** p<0.01, ** p<0.05, * p<0.1

Table A.VI: Winsorized effects for number of arrests outcomes

	Control Mean	ITT Estimate	ITT P-Value	TOT CCM	TOT Estimate	TOT P-Value
6-months						
Number of arrests	0.248	-0.013 (0.025)	0.610	0.170	-0.021 (0.040)	0.603
Number of arrests - top 1% winsorized	0.233	-0.018 (0.022)	0.416	0.177	-0.029 (0.035)	0.408
Number of arrests - top 5% winsorized	0.161	-0.018 (0.014)	0.199	0.140	-0.029 (0.022)	0.192
12-months						
Number of arrests	0.490	-0.063 (0.040)	0.115	0.370	-0.103 (0.064)	0.109
Number of arrests - top 1% winsorized	0.468	-0.064 (0.036)	0.074*	0.367	-0.104 (0.057)	0.069*
Number of arrests - top 5% winsorized	0.424	-0.063 (0.029)	0.032**	0.349	-0.103 (0.047)	0.030**
24-months						
Number of arrests	0.871	-0.054 (0.064)	0.398	0.635	-0.088 (0.103)	0.390
Number of arrests - top 1% winsorized	0.854	-0.074 (0.058)	0.200	0.652	-0.120 (0.092)	0.193
Number of arrests - top 5% winsorized	0.727	-0.070 (0.042)	0.097*	0.585	-0.114 (0.067)	0.092*
36-months						
Number of arrests	1.175	-0.084 (0.078)	0.285	0.915	-0.136 (0.125)	0.277
Number of arrests - top 1% winsorized	1.153	-0.115 (0.070)	0.098*	0.930	-0.188 (0.111)	0.092*
Number of arrests - top 5% winsorized	1.053	-0.097 (0.058)	0.094*	0.858	-0.159 (0.093)	0.089*
48-months						
Number of arrests	1.425	-0.105 (0.091)	0.249	1.124	-0.171 (0.146)	0.241
Number of arrests - top 1% winsorized	1.404	-0.133 (0.084)	0.112	1.141	-0.217 (0.134)	0.107
Number of arrests - top 5% winsorized	1.279	-0.122 (0.068)	0.072*	1.056	-0.198 (0.108)	0.068*
60-months						
Number of arrests	1.603	-0.096 (0.100)	0.337	1.266	-0.156 (0.159)	0.329
Number of arrests - top 1% winsorized	1.580	-0.120 (0.093)	0.196	1.281	-0.196 (0.149)	0.189
Number of arrests - top 5% winsorized	1.404	-0.103 (0.072)	0.153	1.160	-0.168 (0.116)	0.147

Notes: Top 1% winsorized outcomes top-code values above the 99th percentile; top 5% winsorized outcomes top-code values above the 95th percentile. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. FWER is the family wise error rate p-value, calculated using the outcomes during the same outcome period and margin as the family. FDR is the false discovery rate p-value. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure A.VI: Baseline number of arrests distribution

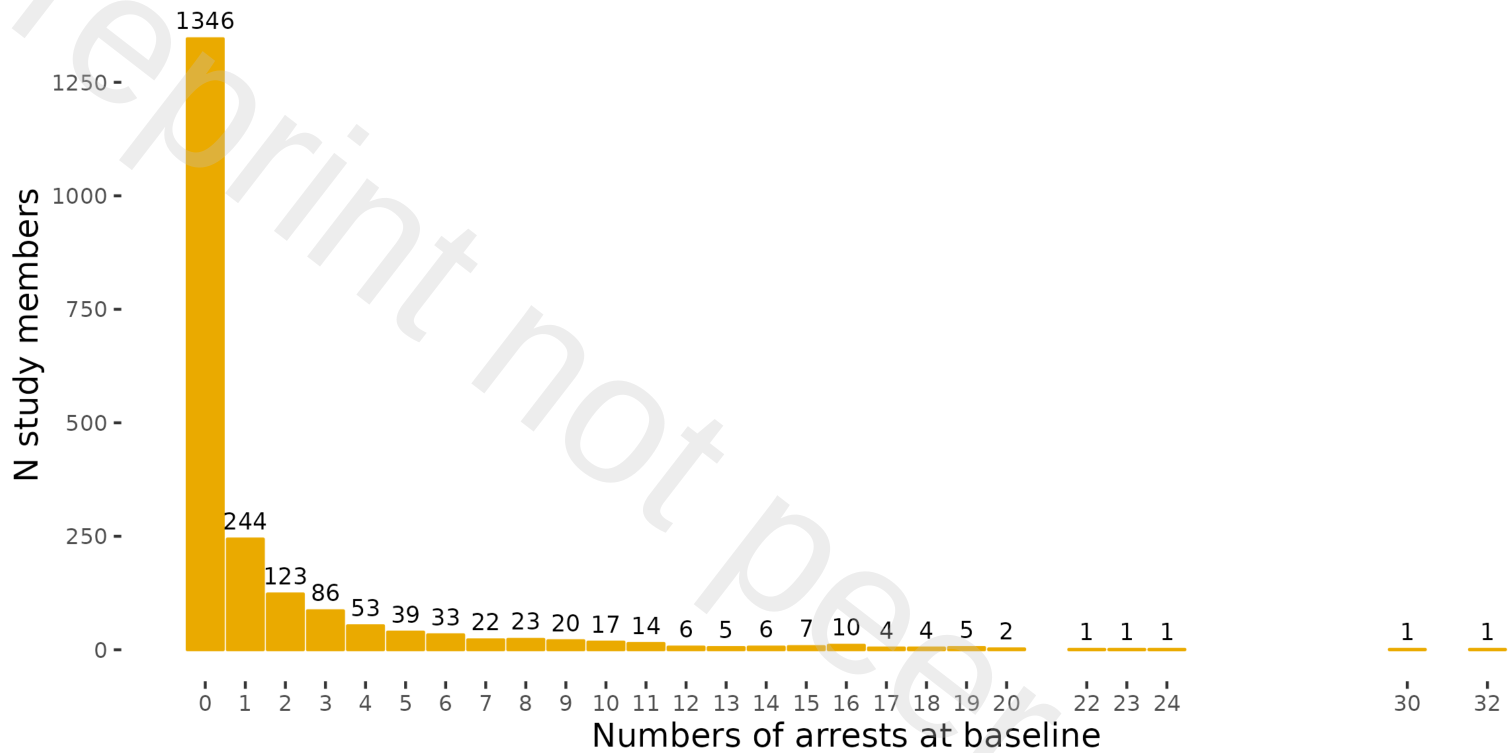


Figure A.VII: Distribution of arrests, 24-months post randomization

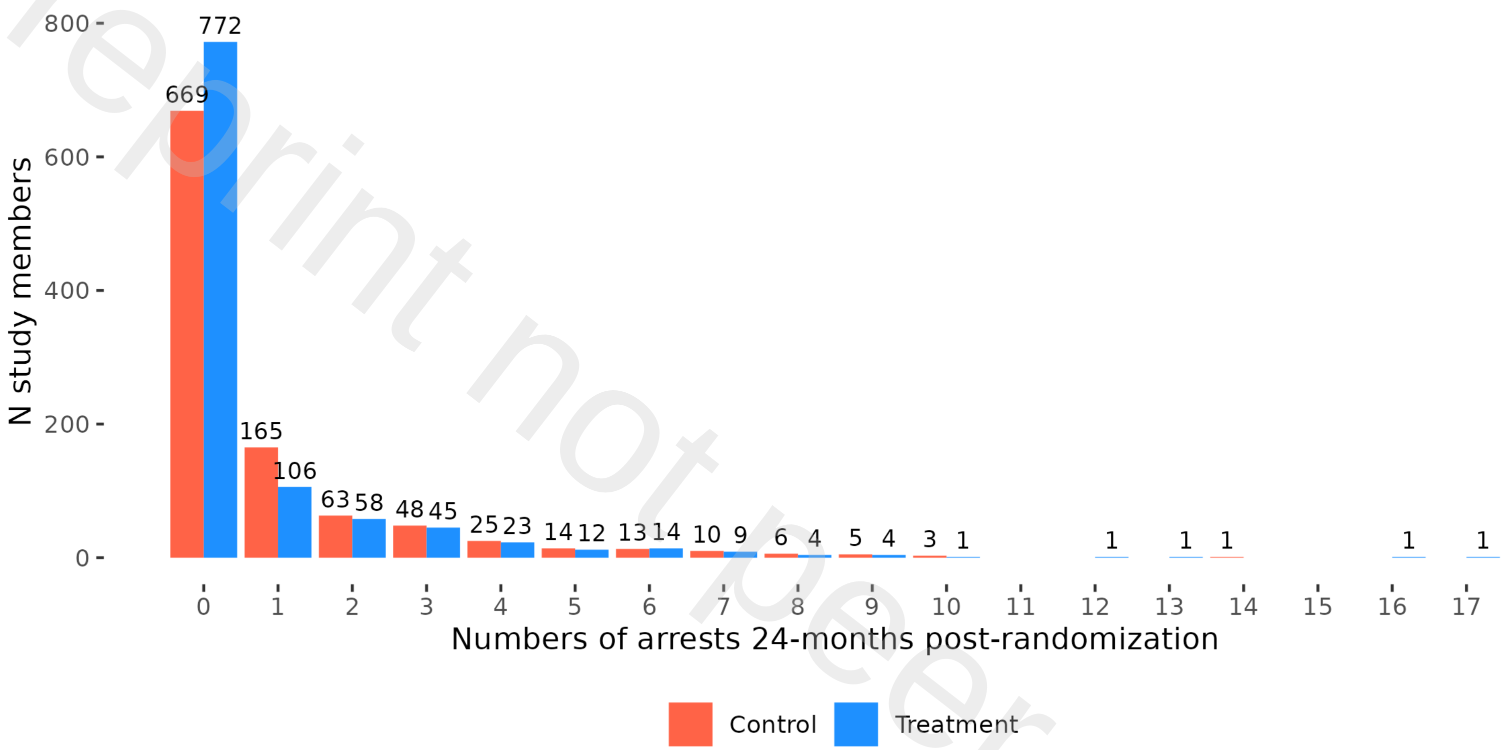
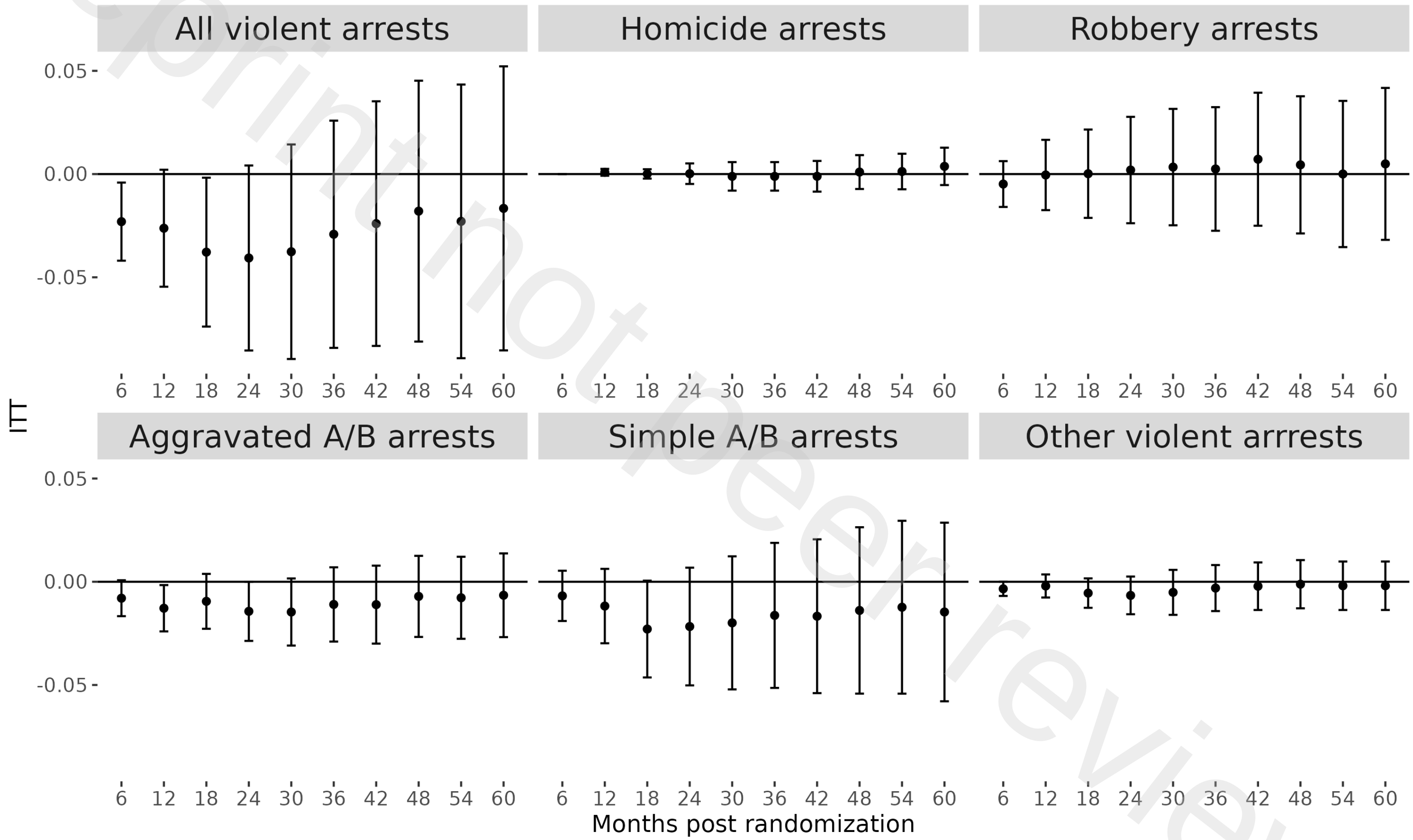


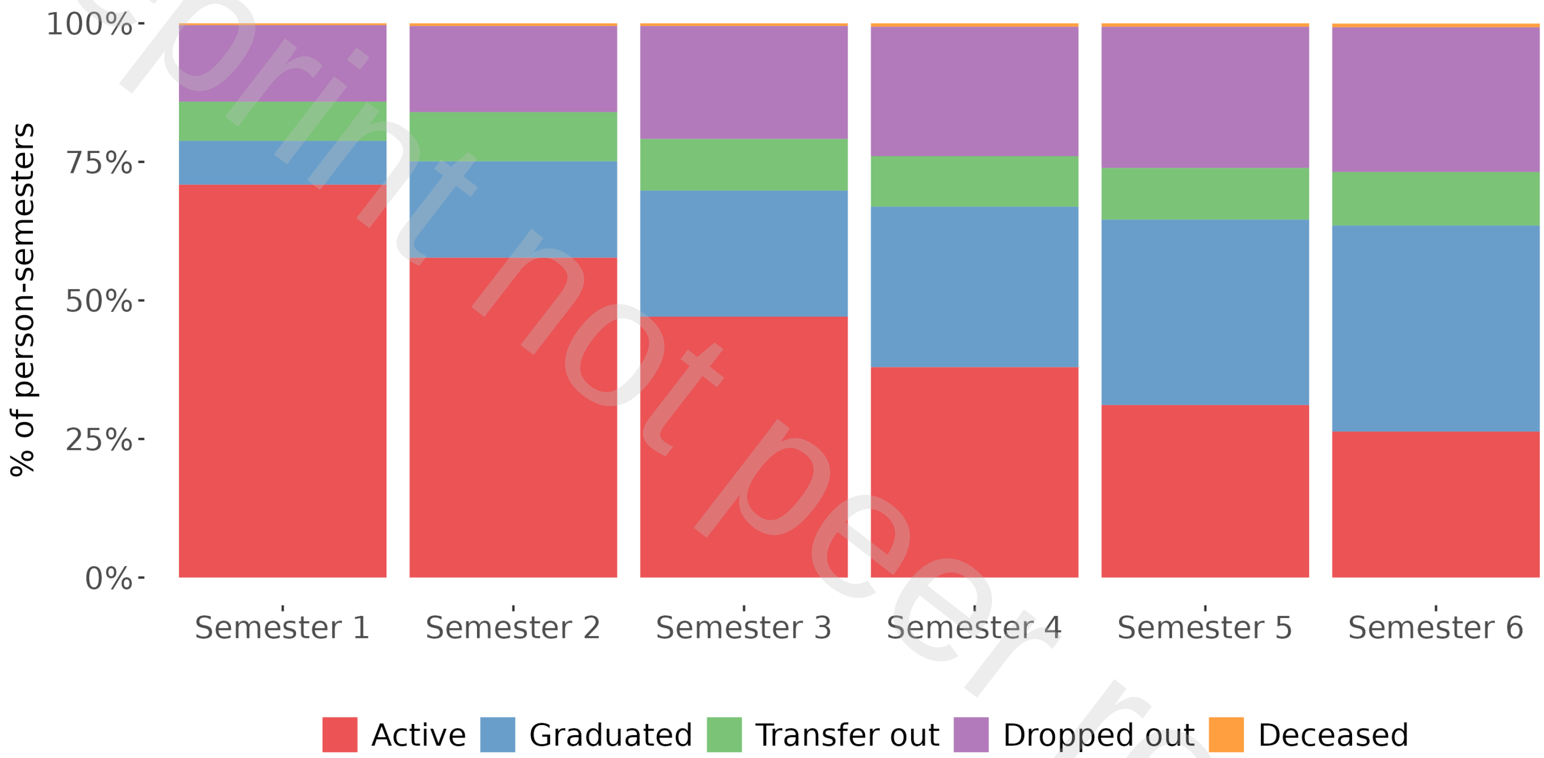
Figure A.VIII: Arrests for violent offenses by subtypes



Note: Arrests classified using arrest FBI codes and statute descriptions. Violent arrests are the sum of the five component arrest types shown in the other panels. Simple A/B arrests combines simple assaults and simple batteries. Aggravated A/B arrests combines aggravated assaults and aggravated batteries. Other violent arrests includes arrests for sexual assault and misconduct (N=3 after 60-months), violent disorderly conduct statutes (N=8 after 60-months), and miscellaneous violent offenses (N=24 after 60-months)

A.1.3 More details on education impacts

Figure A.IX: Enrollment status by outcome semester



Note: Y-axis shows percent of study members who fall into each enrollment status at the end of each semester for the entire study sample, treatment and control. Enrollment status categories are mutually exclusive, and are classified using leave codes from CPS masterfiles at the end of each person-semester. Transfer out status includes those who transferred to a Chicago non-public school, those who transferred to a school outside of Chicago, and those who transferred to a school outside of Chicago. Dropped-out status includes youth with a leave code corresponding to dropping-out, those who never arriving to school for the entirety of a semester, those with an unverified transfer within CPS listed at the end of a semester, or those with a leave code for being legally committed to a correctional institution

Table A.VII: C2C estimated effects on educational outcomes, using additional education baseline covariates

Outcome	Estimates				P-values		
	CM	ITT	CCM	TOT	Observed ITT	FWER	FDR
Semester 1							
Academic Index - Combined	0.000	0.027	0.097	0.044	0.061*		
		(0.015)		(0.023)			
Academic Index - Engagement	0.000	0.016	0.215	0.026	0.501		
		(0.024)		(0.037)			
Attendance Days	54.420	1.029	60.423	1.666	0.247	0.429	0.507
		(0.888)		(1.409)			
Ever Enrolled or Graduated	0.848	-0.021	0.940	-0.034	0.169	0.423	0.507
		(0.015)		(0.024)			
GPA	1.754	0.028	2.018	0.046	0.449	0.453	0.507
		(0.037)		(0.059)			
Academic Index - Behavioral	0.000	0.055	-0.048	0.089	0.016**		
		(0.023)		(0.037)			
Misconduct Incidents	0.597	-0.095	0.673	-0.155	0.019**	0.057*	0.057*
		(0.041)		(0.065)			
Ever Enrolled in Juv. Justice	0.035	0.002	0.012	0.003	0.769	0.772	0.769
		(0.007)		(0.011)			
No. of CPS Schools Enrolled In	0.927	-0.041	1.023	-0.067	0.040**	0.079*	0.080*
		(0.020)		(0.032)			
Semester 2							
Academic Index - Combined	0.000	0.044	0.049	0.071	0.002***		
		(0.014)		(0.022)			
Academic Index - Engagement	0.000	0.031	0.143	0.050	0.171		
		(0.023)		(0.036)			
Attendance Days	49.247	1.455	53.813	2.358	0.117	0.220	0.238
		(0.929)		(1.477)			
Ever Enrolled or Graduated	0.793	-0.030	0.882	-0.049	0.079*	0.220	0.238
		(0.017)		(0.028)			
GPA	1.721	0.049	1.898	0.080	0.175	0.220	0.238
		(0.036)		(0.058)			
Academic Index - Behavioral	0.000	0.076	-0.078	0.124	0.001***		
		(0.022)		(0.036)			
Misconduct Incidents	0.477	-0.111	0.565	-0.180	0.006***	0.017**	0.018**
		(0.040)		(0.064)			
Ever Enrolled in Juv. Justice	0.033	-0.010	0.028	-0.016	0.151	0.158	0.151
		(0.007)		(0.011)			
No. of CPS Schools Enrolled In	0.858	-0.039	0.950	-0.063	0.074*	0.140	0.149
		(0.022)		(0.035)			

Notes: Our index is equal to an unweighted average of the different components, after all have been normalized to Z-score form using the control group's distribution. Index components used "mixed" imputation methodology for missing values, see Appendix Table A.IX for outcome sensitivity to imputation methodology. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, educational attainment in the semester prior to randomization (attendance days/GPA/misconduct incidents), prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, number of prior victimizations, and a dummy indicator for whether the outcome has been imputed. Index outcomes include multiple dummies for whether each component was imputed. FWER is the family wise error rate p-values, where the three components of each separate index during each semester constitute each family. FDR is the false discovery rate. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.VIII: C2C estimated effects on educational outcomes, using only no baseline covariates

Outcome	Control Mean	ITT Estimate	ITT P-Value	TOT CCM	TOT Estimate	TOT P-Value
Semester 1						
Academic Index - Combined	0.000	0.044 (0.016)	0.007***	0.070	0.071 (0.026)	0.006***
Academic Index - Engagement	0.000	0.043 (0.026)	0.106	0.172	0.069 (0.042)	0.101
Attendance Days	54.420	1.933 (0.984)	0.050**	58.972	3.117 (1.565)	0.047**
Ever Enrolled or Graduated	0.848	-0.013 (0.016)	0.402	0.927	-0.021 (0.025)	0.399
GPA	1.754	0.064 (0.040)	0.108	1.960	0.104 (0.064)	0.104
Academic Index - Behavioral	0.000	0.065 (0.024)	0.006***	-0.064	0.106 (0.039)	0.006***
Misconduct Incidents	0.597	-0.116 (0.042)	0.006***	0.707	-0.188 (0.069)	0.006***
Ever Enrolled in Juv. Justice	0.035	-0.003 (0.008)	0.722	0.020	-0.004 (0.012)	0.720
No. of CPS Schools Enrolled In	0.927	-0.035 (0.020)	0.084*	1.012	-0.057 (0.033)	0.083*
Semester 2						
Academic Index - Combined	0.000	0.058 (0.015)	0.000***	0.027	0.093 (0.024)	0.000***
Academic Index - Engagement	0.000	0.052 (0.025)	0.039**	0.109	0.084 (0.040)	0.037**
Attendance Days	49.247	2.461 (1.014)	0.015**	52.200	3.971 (1.617)	0.014**
Ever Enrolled or Graduated	0.793	-0.026 (0.017)	0.142	0.875	-0.042 (0.028)	0.140
GPA	1.721	0.071 (0.038)	0.067*	1.864	0.114 (0.062)	0.064*
Academic Index - Behavioral	0.000	0.087 (0.023)	0.000***	-0.095	0.142 (0.037)	0.000***
Misconduct Incidents	0.477	-0.126 (0.040)	0.002***	0.589	-0.204 (0.064)	0.002***
Ever Enrolled in Juv. Justice	0.033	-0.013 (0.007)	0.060*	0.033	-0.021 (0.011)	0.058*
No. of CPS Schools Enrolled In	0.858	-0.040 (0.022)	0.068*	0.952	-0.064 (0.035)	0.067*

Notes: Our index is equal to an unweighted average of the different components, after all have been normalized to Z-score form using the control group's distribution. Index components used "mixed" imputation methodology for missing values, see Appendix Table A.IX for outcome sensitivity to imputation methodology. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. Index outcomes include multiple dummies for whether each component was imputed. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.IX: C2C education outcomes sensitivity to imputation, semesters 1 and 2

	Imputation A (Basic)			Imputation B (Mean)			Main Results		
	N	CM	ITT	N	CM	ITT	N	CM	ITT
Semester 1									
C2C Academic Index - Combined	2065	-0.009	0.023	2065	0.000	0.043***	2065	0.000	0.033**
			(0.015)			(0.015)			(0.015)
C2C Academic Index - Engagement	2065	-0.143	0.028	2065	0.000	0.052**	2065	0.000	0.028
			(0.030)			(0.024)			(0.025)
Attendance Days	1729	57.639	2.003**	2065	58.097	2.146**	2065	54.420	1.610*
			(1.012)			(0.868)			(0.945)
Ever Enrolled or Graduated	2065	0.848	-0.018	2065	0.848	-0.018	2065	0.848	-0.018
			(0.015)			(0.015)			(0.015)
GPA	1338	2.107	0.046	2065	2.095	0.074**	2065	1.754	0.041
			(0.050)			(0.033)			(0.039)
C2C Academic Index - Behavioral	2065	-0.002	0.044*	2065	0.000	0.050**	2065	0.000	0.055**
			(0.025)			(0.023)			(0.023)
Misconduct Incidents	1674	0.501	-0.065	2065	0.503	-0.083**	2065	0.597	-0.098**
			(0.050)			(0.041)			(0.041)
Ever Enrolled in Juvenile Justice	2065	0.035	0.002	2065	0.035	0.002	2065	0.035	0.002
			(0.007)			(0.007)			(0.007)
Number of CPS Schools Enrolled In	2065	0.927	-0.039*	2065	0.927	-0.039*	2065	0.927	-0.039*
			(0.020)			(0.020)			(0.020)
Semester 2									
C2C Academic Index - Combined	2065	-0.010	0.036**	2065	0.000	0.058***	2065	0.000	0.048***
			(0.015)			(0.014)			(0.014)
C2C Academic Index - Engagement	2065	-0.201	0.045	2065	0.000	0.069***	2065	0.000	0.041*
			(0.029)			(0.023)			(0.024)
Attendance Days	1566	55.018	2.386**	2065	55.398	2.602***	2065	49.247	1.987**
			(1.073)			(0.862)			(0.970)
Ever Enrolled or Graduated	2065	0.793	-0.028	2065	0.793	-0.028	2065	0.793	-0.028
			(0.017)			(0.017)			(0.017)
GPA	1127	2.190	0.063	2065	2.189	0.084***	2065	1.721	0.058
			(0.053)			(0.031)			(0.037)
C2C Academic Index - Behavioral	2065	0.000	0.062**	2065	0.000	0.068***	2065	0.000	0.076***
			(0.025)			(0.023)			(0.022)
Misconduct Incidents	1557	0.371	-0.058	2065	0.382	-0.082**	2065	0.477	-0.109***
			(0.050)			(0.040)			(0.039)
Ever Enrolled in Juvenile Justice	2065	0.033	-0.010	2065	0.033	-0.010	2065	0.033	-0.010
			(0.007)			(0.007)			(0.007)
Number of CPS Schools Enrolled In	2065	0.858	-0.037*	2065	0.858	-0.037*	2065	0.858	-0.037*
			(0.022)			(0.022)			(0.022)

Notes: Semester 1 is the first full semester post-randomization. Imputation A uses minimal imputation, only imputing post-graduate semesters with a students most recent attendance and GPA, and imputing 0 for misconducts. Missing values after the fact are left missing, and are dropped from component regressions. Indices are calculated using the average of all non-missing components, where each component is standardized such that the control mean is 0. Imputation B imputes all values that were still missing under Imputation A using cohort-treatment status-outcome semester means. Main results impute all values that were still missing under Imputation A using the following logic: for GPA and attendance, impute with 0 if study member has a drop-out leave code at the end of the semester, otherwise impute with cohort-treatment status-outcome semester means. For misconduct, impute with 0 if study member is listed as active in CPS masterfiles at the end of semester, otherwise impute with cohort-treatment status-outcome semester means. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.X: C2C education outcomes sensitivity to imputation, semesters 3 and 4

	Imputation A (Basic)			Imputation B (Mean)			Main Results		
	N	CM	ITT	N	CM	ITT	N	CM	ITT
Semester 3									
C2C Academic Index - Combined	2065	-0.005	0.019	2065	0.000	0.036**	2065	0.000	0.013
			(0.015)			(0.015)			(0.015)
C2C Academic Index - Engagement	2065	-0.218	0.047*	2065	0.000	0.065***	2065	0.000	0.013
			(0.028)			(0.024)			(0.023)
Attendance Days	1469	56.013	1.113	2065	56.149	0.964	2065	47.662	0.236
			(1.165)			(0.868)			(0.977)
Ever Enrolled or Graduated	2065	0.735	-0.008	2065	0.735	-0.008	2065	0.735	-0.008
			(0.018)			(0.018)			(0.018)
GPA	1001	2.208	0.052	2065	2.206	0.086***	2065	1.609	-0.016
			(0.060)			(0.030)			(0.038)
C2C Academic Index - Behavioral	2065	0.009	0.028	2065	0.000	0.029	2065	0.000	0.034
			(0.025)			(0.023)			(0.023)
Misconduct Incidents	1557	0.242	0.009	2065	0.246	-0.004	2065	0.339	-0.016
			(0.042)			(0.032)			(0.031)
Ever Enrolled in Juvenile Justice	2065	0.029	-0.005	2065	0.029	-0.005	2065	0.029	-0.005
			(0.007)			(0.007)			(0.007)
Number of CPS Schools Enrolled In	2065	0.784	-0.026	2065	0.784	-0.026	2065	0.784	-0.026
			(0.021)			(0.021)			(0.021)
Semester 4									
C2C Academic Index - Combined	2065	-0.001	0.007	2065	0.000	0.005	2065	0.000	-0.004
			(0.015)			(0.016)			(0.015)
C2C Academic Index - Engagement	2065	-0.229	0.027	2065	0.000	0.017	2065	0.000	0.012
			(0.028)			(0.024)			(0.023)
Attendance Days	1401	55.781	1.355	2065	55.656	1.475*	2065	44.452	0.878
			(1.212)			(0.858)			(0.980)
Ever Enrolled or Graduated	2065	0.691	0.004	2065	0.691	0.004	2065	0.691	0.004
			(0.019)			(0.019)			(0.019)
GPA	918	2.273	-0.049	2065	2.263	-0.020	2065	1.491	-0.022
			(0.063)			(0.028)			(0.037)
C2C Academic Index - Behavioral	2065	0.008	0.008	2065	0.000	0.001	2065	0.000	-0.011
			(0.026)			(0.026)			(0.025)
Misconduct Incidents	1637	0.163	0.035	2065	0.156	0.036	2065	0.231	0.062*
			(0.044)			(0.034)			(0.034)
Ever Enrolled in Juvenile Justice	2065	0.037	-0.007	2065	0.037	-0.007	2065	0.037	-0.007
			(0.008)			(0.008)			(0.008)
Number of CPS Schools Enrolled In	2065	0.726	-0.009	2065	0.726	-0.009	2065	0.726	-0.009
			(0.021)			(0.021)			(0.021)

Notes: Imputation A uses minimal imputation, only imputing post-graduate semesters with a students most recent attendance and GPA, and imputing 0 for misconducts. Missing values after the fact are left missing, and are dropped from component regressions. Indices are calculated using the average of all non-missing components, where each component is standardized such that the control mean is 0. Imputation B imputes all values that were still missing under Imputation A using cohort-treatment status-outcome semester means. Main results impute all values that were still missing under Imputation A using the following logic: for GPA and attendance, impute with 0 if study member has a drop-out leave code at the end of the semester, otherwise impute with cohort-treatment status-outcome semester means. For misconduct, impute with 0 if study member is listed as active in CPS masterfiles at the end of semester, otherwise impute with cohort-treatment status-outcome semester means. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. Standard errors are robust and shown in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A.XI: C2C education outcomes sensitivity to imputation, semesters 5 and 6

	Imputation A (Basic)			Imputation B (Mean)			Main Results		
	N	CM	ITT	N	CM	ITT	N	CM	ITT
Semester 5									
C2C Academic Index - Combined	2065	-0.004	0.029**	2065	0.000	0.048***	2065	0.000	0.029**
			(0.014)			(0.015)			(0.014)
C2C Academic Index - Engagement	2065	-0.239	0.029	2065	0.000	0.068***	2065	0.000	0.025
			(0.027)			(0.024)			(0.022)
Attendance Days	1339	58.844	2.083*	2065	58.728	2.562***	2065	45.681	1.009
			(1.236)			(0.824)			(0.963)
Ever Enrolled or Graduated	2065	0.675	-0.012	2065	0.675	-0.012	2065	0.675	-0.012
			(0.019)			(0.019)			(0.019)
GPA	894	2.254	0.015	2065	2.233	0.056*	2065	1.403	0.028
			(0.064)			(0.029)			(0.037)
C2C Academic Index - Behavioral	2065	0.008	0.045*	2065	0.000	0.043*	2065	0.000	0.048**
			(0.024)			(0.022)			(0.022)
Misconduct Incidents	1693	0.182	-0.016	2065	0.183	-0.023	2065	0.264	-0.035
			(0.037)			(0.030)			(0.029)
Ever Enrolled in Juvenile Justice	2065	0.037	-0.010	2065	0.037	-0.010	2065	0.037	-0.010
			(0.007)			(0.007)			(0.007)
Number of CPS Schools Enrolled In	2065	0.706	-0.022	2065	0.706	-0.022	2065	0.706	-0.022
			(0.021)			(0.021)			(0.021)
Semester 6									
C2C Academic Index - Combined	2065	0.000	0.030**	2065	0.000	0.048***	2065	0.000	0.029**
			(0.013)			(0.014)			(0.013)
C2C Academic Index - Engagement	2065	-0.249	0.042*	2065	0.000	0.071***	2065	0.000	0.024
			(0.025)			(0.022)			(0.021)
Attendance Days	1288	53.957	1.868	2065	53.407	2.223***	2065	40.181	0.524
			(1.196)			(0.764)			(0.912)
Ever Enrolled or Graduated	2065	0.652	0.000	2065	0.652	0.000	2065	0.652	0.000
			(0.019)			(0.019)			(0.019)
GPA	853	2.312	0.040	2065	2.307	0.059**	2065	1.403	0.016
			(0.062)			(0.027)			(0.035)
C2C Academic Index - Behavioral	2065	0.007	0.035	2065	0.000	0.033	2065	0.000	0.041*
			(0.023)			(0.022)			(0.021)
Misconduct Incidents	1737	0.132	-0.026	2065	0.129	-0.026	2065	0.193	-0.041*
			(0.028)			(0.023)			(0.022)
Ever Enrolled in Juvenile Justice	2065	0.029	-0.010	2065	0.029	-0.010	2065	0.029	-0.010
			(0.006)			(0.006)			(0.006)
Number of CPS Schools Enrolled In	2065	0.668	-0.001	2065	0.668	-0.001	2065	0.668	-0.001
			(0.020)			(0.020)			(0.020)

Notes: Imputation A uses minimal imputation, only imputing post-graduate semesters with a students most recent attendance and GPA, and imputing 0 for misconducts. Missing values after the fact are left missing, and are dropped from component regressions. Indices are calculated using the average of all non-missing components, where each component is standardized such that the control mean is 0. Imputation B imputes all values that were still missing under Imputation A using cohort-treatment status-outcome semester means. Main results impute all values that were still missing under Imputation A using the following logic: for GPA and attendance, impute with 0 if study member has a drop-out leave code at the end of the semester, otherwise impute with cohort-treatment status-outcome semester means. For misconduct, impute with 0 if study member is listed as active in CPS masterfiles at the end of semester, otherwise impute with cohort-treatment status-outcome semester means. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. Standard errors are robust and shown in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

A.1.4 Heterogeneous Treatment Effects

Table A.XII: C2C estimated effects on arrests, by number of prior arrest at baseline

	0-4 arrests at baseline (N=1852)				5+ arrests at baseline (N=222)			
	CM	ITT	% Change	P-Value	CM	ITT	% Change	P-Value
Number of Arrests								
6 months	0.161	-0.031	-19%	0.135	0.885	0.134	15%	0.435
		(0.021)				(0.171)		
12 months	0.336	-0.092	-27%	0.010***	1.631	0.140	9%	0.565
		(0.035)				(0.242)		
24 months	0.639	-0.121	-19%	0.037**	2.582	0.459	18%	0.193
		(0.058)				(0.353)		
36 months	0.850	-0.136	-16%	0.058*	3.574	0.331	9%	0.433
		(0.072)				(0.422)		
48 months	1.054	-0.156	-15%	0.069*	4.156	0.278	7%	0.549
		(0.086)				(0.464)		
60 months	1.177	-0.122	-10%	0.193	4.607	0.150	3%	0.758
		(0.093)				(0.485)		
Number of Violent Arrests								
6 months	0.049	-0.017	-34%	0.064*	0.172	-0.082	-48%	0.111
		(0.009)				(0.052)		
12 months	0.091	-0.026	-29%	0.053*	0.303	-0.045	-15%	0.549
		(0.014)				(0.075)		
24 months	0.186	-0.042	-23%	0.051*	0.525	-0.063	-12%	0.579
		(0.022)				(0.114)		
36 months	0.229	-0.031	-13%	0.235	0.730	-0.051	-7%	0.720
		(0.026)				(0.142)		
48 months	0.271	-0.029	-11%	0.340	0.811	0.040	5%	0.791
		(0.031)				(0.151)		
60 months	0.302	-0.020	-7%	0.562	0.926	-0.029	-3%	0.853
		(0.034)				(0.155)		
Ever Arrested								
6 months	0.109	-0.012	-11%	0.365	0.549	-0.070	-13%	0.287
		(0.013)				(0.066)		
12 months	0.180	-0.038	-21%	0.016**	0.713	-0.072	-10%	0.238
		(0.016)				(0.061)		
24 months	0.278	-0.063	-23%	0.000***	0.844	-0.067	-8%	0.200
		(0.018)				(0.052)		
36 months	0.320	-0.052	-16%	0.006***	0.877	-0.056	-6%	0.245
		(0.019)				(0.048)		
48 months	0.357	-0.053	-15%	0.007***	0.893	-0.039	-4%	0.389
		(0.019)				(0.046)		
60 months	0.376	-0.033	-9%	0.100	0.902	-0.033	-4%	0.461
		(0.020)				(0.045)		
Ever Arrested for Violent Offense								
6 months	0.046	-0.014	-30%	0.104	0.164	-0.091	-56%	0.041**
		(0.008)				(0.045)		
12 months	0.074	-0.019	-26%	0.083*	0.270	-0.086	-32%	0.126
		(0.011)				(0.056)		
24 months	0.134	-0.032	-24%	0.019**	0.385	-0.085	-22%	0.171
		(0.014)				(0.062)		
36 months	0.160	-0.028	-17%	0.067*	0.459	-0.073	-16%	0.259
		(0.015)				(0.064)		
48 months	0.177	-0.026	-15%	0.106	0.500	-0.057	-11%	0.376
		(0.016)				(0.065)		
60 months	0.187	-0.013	-7%	0.438	0.549	-0.084	-15%	0.196
		(0.016)				(0.065)		

Notes: Estimates for each outcome are from a single regression that interacts subgroup indicators with treatment. 5 arrests at baseline represents the 90th percentile of baseline arrests in the C2C sample. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.XIII: C2C estimated effects on arrests, by age group

	Age under median (N=1037)				Age over median (N=1037)			
	CM	ITT	% Change	P-Value	CM	ITT	% Change	P-Value
Number of Arrests								
6 months	0.221	0.045 (0.035)	21%	0.200	0.274	-0.071 (0.036)	-26%	0.050*
12 months	0.468	-0.009 (0.055)	-2%	0.865	0.513	-0.117 (0.059)	-23%	0.049**
24 months	0.877	0.037 (0.089)	4%	0.678	0.865	-0.145 (0.092)	-17%	0.117
36 months	1.211	0.017 (0.113)	1%	0.883	1.139	-0.183 (0.108)	-16%	0.090*
48 months	1.515	0.002 (0.139)	0%	0.991	1.335	-0.210 (0.117)	-16%	0.073*
60 months	1.695	0.021 (0.151)	1%	0.888	1.477	-0.197 (0.126)	-13%	0.118
Number of Violent Arrests								
6 months	0.057	-0.009 (0.013)	-15%	0.498	0.070	-0.037 (0.015)	-53%	0.011**
12 months	0.115	-0.015 (0.020)	-13%	0.464	0.117	-0.038 (0.021)	-32%	0.076*
24 months	0.231	-0.021 (0.033)	-9%	0.524	0.221	-0.061 (0.031)	-27%	0.054*
36 months	0.307	-0.006 (0.041)	-2%	0.890	0.270	-0.053 (0.038)	-20%	0.162
48 months	0.358	0.011 (0.048)	3%	0.818	0.313	-0.047 (0.042)	-15%	0.260
60 months	0.411	0.011 (0.054)	3%	0.840	0.342	-0.044 (0.044)	-13%	0.314
Ever Arrested								
6 months	0.153	-0.005 (0.018)	-3%	0.784	0.170	-0.030 (0.020)	-18%	0.139
12 months	0.233	-0.026 (0.021)	-11%	0.215	0.254	-0.057 (0.023)	-22%	0.013**
24 months	0.339	-0.052 (0.023)	-15%	0.027**	0.352	-0.076 (0.025)	-21%	0.002***
36 months	0.382	-0.040 (0.025)	-11%	0.101	0.391	-0.064 (0.026)	-16%	0.013**
48 months	0.403	-0.030 (0.025)	-7%	0.241	0.438	-0.071 (0.026)	-16%	0.006***
60 months	0.427	-0.011 (0.026)	-3%	0.673	0.450	-0.052 (0.026)	-12%	0.048**
Ever Arrested for Violent Offense								
6 months	0.053	-0.009 (0.012)	-17%	0.445	0.067	-0.034 (0.014)	-51%	0.014**
12 months	0.094	-0.011 (0.016)	-12%	0.475	0.102	-0.039 (0.017)	-38%	0.021**
24 months	0.157	-0.021 (0.019)	-14%	0.270	0.172	-0.053 (0.020)	-31%	0.010***
36 months	0.196	-0.014 (0.021)	-7%	0.518	0.196	-0.049 (0.022)	-25%	0.024**
48 months	0.213	-0.013 (0.022)	-6%	0.561	0.217	-0.043 (0.023)	-20%	0.062*
60 months	0.227	0.005 (0.023)	2%	0.834	0.233	-0.042 (0.023)	-18%	0.068*

Notes: Estimates for each outcome are from a single regression that interacts subgroup indicators with treatment. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.XIV: C2C estimated effects on arrests, for those who had a prior arrest vs those with no prior arrests at baseline

	No arrests at baseline (N=1346)				Ever arrested at baseline (N=728)			
	CM	ITT	% Change	P-Value	CM	ITT	% Change	P-Value
Number of Arrests								
6 months	0.067	-0.015 (0.017)	-22%	0.379	0.573	-0.009 (0.069)	-2%	0.897
12 months	0.177	-0.080 (0.031)	-45%	0.011**	1.055	-0.037 (0.102)	-4%	0.717
24 months	0.370	-0.157 (0.051)	-42%	0.002***	1.773	0.129 (0.160)	7%	0.422
36 months	0.505	-0.154 (0.067)	-31%	0.021**	2.381	0.040 (0.189)	2%	0.835
48 months	0.654	-0.175 (0.084)	-27%	0.036**	2.811	0.015 (0.214)	1%	0.944
60 months	0.746	-0.149 (0.093)	-20%	0.110	3.099	0.016 (0.226)	1%	0.945
Number of Violent Arrests								
6 months	0.021	-0.003 (0.008)	-15%	0.686	0.140	-0.059 (0.023)	-42%	0.011**
12 months	0.047	-0.010 (0.012)	-21%	0.420	0.241	-0.057 (0.035)	-24%	0.101
24 months	0.108	-0.032 (0.020)	-30%	0.110	0.438	-0.058 (0.054)	-13%	0.283
36 months	0.139	-0.024 (0.024)	-17%	0.332	0.559	-0.041 (0.067)	-7%	0.536
48 months	0.174	-0.032 (0.030)	-18%	0.294	0.627	0.006 (0.074)	1%	0.931
60 months	0.202	-0.029 (0.035)	-14%	0.406	0.690	0.007 (0.077)	1%	0.932
Ever Arrested								
6 months	0.055	-0.012 (0.012)	-22%	0.309	0.353	-0.028 (0.032)	-8%	0.390
12 months	0.108	-0.036 (0.016)	-34%	0.021**	0.488	-0.053 (0.034)	-11%	0.118
24 months	0.190	-0.068 (0.019)	-36%	0.000***	0.625	-0.058 (0.033)	-9%	0.078*
36 months	0.221	-0.045 (0.021)	-20%	0.031**	0.685	-0.069 (0.032)	-10%	0.034**
48 months	0.245	-0.038 (0.022)	-16%	0.082*	0.737	-0.079 (0.032)	-11%	0.013**
60 months	0.262	-0.016 (0.023)	-6%	0.496	0.756	-0.068 (0.031)	-9%	0.029**
Ever Arrested for Violent Offense								
6 months	0.020	-0.002 (0.008)	-11%	0.783	0.132	-0.056 (0.022)	-43%	0.009***
12 months	0.041	-0.009 (0.010)	-22%	0.384	0.200	-0.055 (0.027)	-28%	0.039**
24 months	0.082	-0.026 (0.014)	-32%	0.053*	0.312	-0.058 (0.031)	-18%	0.064*
36 months	0.100	-0.017 (0.016)	-17%	0.269	0.367	-0.059 (0.033)	-16%	0.072*
48 months	0.111	-0.017 (0.016)	-15%	0.305	0.403	-0.049 (0.034)	-12%	0.144
60 months	0.119	-0.007 (0.017)	-6%	0.686	0.430	-0.042 (0.034)	-10%	0.216

Notes: Estimates for each outcome are from a single regression that interacts subgroup indicators with treatment. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.XV: C2C estimated effects on arrests, by gender

	Male (N=1199)				Female (N=875)			
	CM	ITT	% Change	P-Value	CM	ITT	% Change	P-Value
Number of Arrests								
6 months	0.350	-0.005 (0.040)	-1%	0.897	0.100	-0.023 (0.025)	-23%	0.345
12 months	0.684	-0.072 (0.063)	-11%	0.251	0.212	-0.052 (0.043)	-24%	0.234
24 months	1.239	-0.030 (0.104)	-2%	0.776	0.343	-0.088 (0.062)	-26%	0.154
36 months	1.708	-0.081 (0.127)	-5%	0.524	0.412	-0.087 (0.073)	-21%	0.230
48 months	2.085	-0.122 (0.150)	-6%	0.414	0.479	-0.082 (0.082)	-17%	0.315
60 months	2.327	-0.106 (0.162)	-5%	0.512	0.524	-0.066 (0.088)	-13%	0.455
Number of Violent Arrests								
6 months	0.088	-0.040 (0.015)	-45%	0.008***	0.029	0.000 (0.011)	-1%	0.973
12 months	0.164	-0.049 (0.023)	-30%	0.031**	0.048	0.005 (0.015)	10%	0.762
24 months	0.327	-0.079 (0.037)	-24%	0.031**	0.081	0.012 (0.022)	15%	0.601
36 months	0.422	-0.065 (0.046)	-15%	0.158	0.098	0.019 (0.026)	20%	0.463
48 months	0.488	-0.049 (0.052)	-10%	0.353	0.117	0.024 (0.030)	20%	0.420
60 months	0.545	-0.041 (0.057)	-8%	0.474	0.136	0.016 (0.032)	12%	0.606
Ever Arrested								
6 months	0.223	-0.021 (0.021)	-10%	0.302	0.074	-0.012 (0.016)	-17%	0.441
12 months	0.329	-0.058 (0.022)	-18%	0.010***	0.121	-0.019 (0.021)	-16%	0.361
24 months	0.458	-0.080 (0.024)	-17%	0.001***	0.183	-0.041 (0.024)	-22%	0.082*
36 months	0.508	-0.061 (0.025)	-12%	0.013**	0.212	-0.039 (0.025)	-18%	0.128
48 months	0.548	-0.058 (0.025)	-11%	0.021**	0.238	-0.040 (0.026)	-17%	0.129
60 months	0.568	-0.037 (0.025)	-6%	0.152	0.252	-0.025 (0.027)	-10%	0.356
Ever Arrested for Violent Offense								
6 months	0.081	-0.037 (0.014)	-45%	0.008***	0.029	-0.001 (0.011)	-2%	0.954
12 months	0.135	-0.044 (0.017)	-33%	0.012**	0.045	0.001 (0.014)	1%	0.970
24 months	0.231	-0.067 (0.021)	-29%	0.001***	0.069	0.004 (0.017)	6%	0.803
36 months	0.274	-0.061 (0.022)	-22%	0.007***	0.083	0.009 (0.019)	11%	0.634
48 months	0.296	-0.049 (0.023)	-17%	0.033**	0.100	0.003 (0.020)	3%	0.894
60 months	0.316	-0.038 (0.024)	-12%	0.112	0.107	0.007 (0.021)	7%	0.720

Notes: Estimates for each outcome are from a single regression that interacts subgroup indicators with treatment. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.XVI: C2C baseline differences, by gender

	All	Male	Female	P-value, Test of subgroup difference
N	2074	1199	875	
Take-up Rate	0.62	0.58	0.67	
Demographics				
Age	16.28	16.25	16.32	0.271
% Black	0.95	0.94	0.96	0.071*
% Hispanic	0.04	0.05	0.04	0.291
% Female	0.42	0.00	1.00	
Number of arrests				
Number of prior arrests	1.49	2.20	0.52	0.000***
Number of prior violent arrests	0.41	0.56	0.21	0.000***
Number of prior property arrests	0.36	0.50	0.16	0.000***
Number of prior drug arrest	0.10	0.17	0.01	0.000***
Number of prior other arrests	0.62	0.97	0.15	0.000***
Any arrest				
Any prior arrest	0.35	0.45	0.22	0.000***
Any prior violent arrest	0.22	0.29	0.13	0.000***
Any prior property arrest	0.18	0.23	0.10	0.000***
Any prior drug arrest	0.06	0.09	0.01	0.000***
Any prior other arrest	0.22	0.31	0.09	0.000***
Victimizations and stops				
Number of prior victimizations	0.92	0.89	0.97	0.229
Any prior victimization	0.49	0.49	0.48	0.694
Number of prior stops	4.20	6.71	0.76	0.000***
Education				
Attendance rate	0.73	0.72	0.75	0.012**
GPA	1.95	1.76	2.21	0.000***
Misconduct incidents	0.94	1.08	0.75	0.000***
Had free/reduce lunch status	0.96	0.96	0.95	0.067*
School grade at baseline	9.74	9.65	9.86	0.002***
Had an enrollment gap	0.20	0.21	0.19	0.185

Notes: Standard errors are robust. Stars indicator the following p-values: *** p<0.01, ** p<0.05, * p<0.1. School grade at baseline is only available for 1,996 youth. Attendance rate, GPA, and misconduct incidents only available for 2,065 youth. Arrests are categorized using FBI code and statutes. Violent arrests include arrests for murder, sexual assault, robbery, aggravated assault or battery, simple assault and battery, sexual misconduct, and some miscellaneous violent statutes. *** p<0.01, ** p<0.05, * p<0.1

Table A.XVII: C2C estimated effects on arrests, by school engagement at baseline

	Engaged in school at baseline (N=1454)				Not engaged in school at baseline (N=620)			
	CM	ITT	% Change	P-Value	CM	ITT	% Change	P-Value
Number of Arrests								
6 months	0.144	0.007 (0.024)	5%	0.777	0.479	-0.058 (0.068)	-12%	0.396
12 months	0.331	-0.049 (0.038)	-15%	0.202	0.848	-0.096 (0.103)	-11%	0.352
24 months	0.644	-0.051 (0.065)	-8%	0.426	1.381	-0.059 (0.149)	-4%	0.693
36 months	0.876	-0.040 (0.081)	-5%	0.620	1.848	-0.184 (0.179)	-10%	0.306
48 months	1.109	-0.078 (0.097)	-7%	0.420	2.133	-0.168 (0.209)	-8%	0.420
60 months	1.267	-0.087 (0.107)	-7%	0.416	2.302	-0.093 (0.221)	-4%	0.674
Number of Violent Arrests								
6 months	0.040	-0.009 (0.010)	-22%	0.381	0.117	-0.056 (0.024)	-48%	0.019**
12 months	0.083	-0.015 (0.015)	-18%	0.339	0.190	-0.052 (0.034)	-27%	0.129
24 months	0.170	-0.031 (0.023)	-18%	0.183	0.352	-0.062 (0.054)	-18%	0.251
36 months	0.218	-0.018 (0.029)	-8%	0.534	0.448	-0.054 (0.065)	-12%	0.405
48 months	0.262	-0.022 (0.033)	-8%	0.515	0.502	-0.008 (0.074)	-2%	0.917
60 months	0.301	-0.024 (0.038)	-8%	0.526	0.546	0.002 (0.078)	0%	0.981
Ever Arrested								
6 months	0.115	-0.018 (0.015)	-15%	0.226	0.267	-0.017 (0.031)	-6%	0.575
12 months	0.187	-0.036 (0.017)	-19%	0.040**	0.371	-0.054 (0.032)	-15%	0.095*
24 months	0.296	-0.075 (0.020)	-25%	0.000***	0.457	-0.035 (0.033)	-8%	0.277
36 months	0.338	-0.057 (0.021)	-17%	0.006***	0.495	-0.039 (0.033)	-8%	0.235
48 months	0.375	-0.063 (0.022)	-17%	0.004***	0.524	-0.021 (0.033)	-4%	0.534
60 months	0.395	-0.044 (0.022)	-11%	0.048**	0.537	-0.003 (0.033)	-1%	0.924
Ever Arrested for Violent Offense								
6 months	0.040	-0.009 (0.010)	-22%	0.361	0.105	-0.050 (0.020)	-48%	0.014**
12 months	0.072	-0.014 (0.013)	-20%	0.255	0.156	-0.050 (0.025)	-32%	0.047**
24 months	0.132	-0.030 (0.016)	-23%	0.053*	0.238	-0.051 (0.029)	-22%	0.080*
36 months	0.154	-0.017 (0.017)	-11%	0.320	0.289	-0.064 (0.031)	-22%	0.041**
48 months	0.175	-0.024 (0.018)	-13%	0.188	0.305	-0.036 (0.032)	-12%	0.264
60 months	0.190	-0.021 (0.018)	-11%	0.260	0.321	-0.014 (0.033)	-4%	0.680

Notes: We use the following criteria for not engaged in school at baseline: any youth who missed more than 60% of school days and received a grade of "F" in at least 75% of their courses in the semester prior to randomization, had a serious learning disability, were at least two years older than expected for the grade they were enrolled in at baseline, or were listed inactive at the end of any of the three semesters immediately prior to randomization. Estimates for each outcome are from a single regression that interacts subgroup indicators with treatment. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.XVIII: C2C estimated effects on arrests, by referral type

	Referred from traditional schools (N=1027)				Referred from non-schools (N=673)				Referred from option schools (N=374)			
	CM	ITT	% Change	P-Value	CM	ITT	% Change	P-Value	CM	ITT	% Change	P-Value
Number of Arrests												
6 months	0.129	-0.015	-12%	0.565	0.405	-0.005	-1%	0.938	0.283	-0.020	-7%	0.769
		(0.027)				(0.063)				(0.067)		
12 months	0.265	-0.047	-18%	0.268	0.810	-0.092	-11%	0.329	0.522	-0.058	-11%	0.576
		(0.042)				(0.094)				(0.105)		
24 months	0.550	-0.039	-7%	0.584	1.357	-0.069	-5%	0.638	0.859	-0.069	-8%	0.634
		(0.071)				(0.147)				(0.145)		
36 months	0.737	0.010	1%	0.914	1.845	-0.218	-12%	0.205	1.147	-0.099	-9%	0.574
		(0.092)				(0.172)				(0.175)		
48 months	0.950	-0.009	-1%	0.939	2.182	-0.269	-12%	0.172	1.337	-0.075	-6%	0.710
		(0.111)				(0.197)				(0.202)		
60 months	1.088	0.040	4%	0.745	2.399	-0.301	-13%	0.147	1.462	-0.063	-4%	0.772
		(0.125)				(0.207)				(0.217)		
Number of Violent Arrests												
6 months	0.044	-0.017	-39%	0.131	0.095	-0.033	-35%	0.124	0.060	-0.022	-36%	0.397
		(0.011)				(0.022)				(0.026)		
12 months	0.076	-0.022	-30%	0.153	0.190	-0.046	-24%	0.184	0.092	-0.002	-2%	0.956
		(0.016)				(0.034)				(0.035)		
24 months	0.163	-0.040	-24%	0.138	0.327	-0.046	-14%	0.356	0.212	-0.032	-15%	0.552
		(0.027)				(0.050)				(0.055)		
36 months	0.207	-0.033	-16%	0.312	0.420	-0.041	-10%	0.498	0.272	0.003	1%	0.959
		(0.033)				(0.061)				(0.066)		
48 months	0.241	-0.025	-10%	0.518	0.497	-0.042	-8%	0.553	0.299	0.042	14%	0.568
		(0.038)				(0.070)				(0.074)		
60 months	0.283	-0.025	-9%	0.585	0.557	-0.058	-10%	0.430	0.304	0.079	26%	0.285
		(0.045)				(0.073)				(0.073)		
Ever Arrested												
6 months	0.096	-0.010	-10%	0.551	0.247	-0.029	-12%	0.290	0.185	-0.018	-10%	0.608
		(0.016)				(0.027)				(0.035)		
12 months	0.163	-0.034	-21%	0.086*	0.363	-0.075	-21%	0.013**	0.245	-0.002	-1%	0.955
		(0.020)				(0.030)				(0.039)		
24 months	0.255	-0.052	-20%	0.023**	0.479	-0.100	-21%	0.001***	0.348	-0.028	-8%	0.496
		(0.023)				(0.031)				(0.041)		
36 months	0.287	-0.029	-10%	0.238	0.533	-0.104	-20%	0.001***	0.391	-0.021	-5%	0.630
		(0.025)				(0.032)				(0.043)		
48 months	0.321	-0.030	-9%	0.241	0.565	-0.091	-16%	0.005***	0.429	-0.034	-8%	0.444
		(0.025)				(0.032)				(0.044)		
60 months	0.341	-0.005	-1%	0.862	0.583	-0.082	-14%	0.011**	0.440	-0.015	-3%	0.739
		(0.026)				(0.032)				(0.045)		
Ever Arrested for Violent Offense												
6 months	0.042	-0.017	-40%	0.118	0.089	-0.034	-38%	0.078*	0.054	-0.012	-23%	0.587
		(0.011)				(0.019)				(0.023)		
12 months	0.072	-0.027	-38%	0.051*	0.146	-0.037	-25%	0.128	0.082	0.000	0%	0.998
		(0.014)				(0.024)				(0.029)		
24 months	0.124	-0.035	-29%	0.046**	0.229	-0.042	-18%	0.142	0.158	-0.032	-20%	0.348
		(0.018)				(0.028)				(0.034)		
36 months	0.143	-0.023	-16%	0.242	0.277	-0.042	-15%	0.163	0.190	-0.034	-18%	0.351
		(0.020)				(0.030)				(0.036)		
48 months	0.151	-0.006	-4%	0.774	0.315	-0.062	-20%	0.043**	0.207	-0.023	-11%	0.536
		(0.021)				(0.031)				(0.037)		
60 months	0.163	-0.001	-1%	0.958	0.339	-0.059	-17%	0.066*	0.212	0.005	2%	0.903
		(0.021)				(0.032)				(0.038)		

Notes: Estimates for each outcome are from a single regression that interacts subgroup indicators with treatment. Non-schools includes N=416 youth referred from Student Outreach and Re-Engagement Centers (SOAR), N=124 from the CPS office of safety and security, N=59 from Cook County Juvenile Probation, N=43 youth who were being detained and soon-to-be released from CPS Juvenile Justice centers, N=21 from the Mayor's Mentoring Initiative, and N=10 from Comprehensive Community-Based Youth Services. Option schools include six alternative CPS schools throughout Chicago, including Excel (Englewood and Southwest), TEAM Englewood, Ombudmans, Youth Connection Charter School (YCCS), and Community Youth Development Institute (CYDI). CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.XIX: C2C estimated effects on academic outcomes, for those who had a prior arrest vs those with no prior arrests at baseline

	No arrests at baseline (N=1343)				Ever arrested at baseline (N=722)			
	CM	ITT	% Change	P-Value	CM	ITT	% Change	P-Value
Semester 1								
Academic Index - Combined	0.124	0.031		0.068*	-0.226	0.038		0.204
		(0.017)				(0.030)		
Academic Index - Engagement	0.191	0.016		0.589	-0.346	0.053		0.262
		(0.029)				(0.047)		
Attendance Days	60.507	1.655	3%	0.130	43.374	1.718	4%	0.336
		(1.093)				(1.784)		
Ever Enrolled or Graduated	0.883	-0.023	-3%	0.202	0.785	-0.009	-1%	0.750
		(0.018)				(0.029)		
GPA	2.032	0.007	0%	0.882	1.249	0.108	9%	0.128
		(0.046)				(0.071)		
Academic Index - Behavioral	0.058	0.066		0.007***	-0.106	0.035		0.464
		(0.024)				(0.047)		
Misconduct Incidents	0.533	-0.107	-20%	0.037**	0.714	-0.089	-13%	0.195
		(0.051)				(0.069)		
Ever Enrolled in Juv. Justice	0.003	0.001	32%	0.796	0.094	0.004	5%	0.827
		(0.004)				(0.019)		
No. of CPS Schools Enrolled In	0.956	-0.049	-5%	0.038**	0.876	-0.021	-2%	0.585
		(0.024)				(0.038)		
Semester 2								
Academic Index - Combined	0.126	0.027		0.076*	-0.228	0.090		0.003***
		(0.015)				(0.030)		
Academic Index - Engagement	0.204	0.029		0.282	-0.371	0.065		0.154
		(0.027)				(0.046)		
Attendance Days	55.617	1.912	3%	0.084*	37.686	2.330	6%	0.210
		(1.108)				(1.857)		
Ever Enrolled or Graduated	0.845	-0.043	-5%	0.031**	0.699	0.000	0%	0.988
		(0.020)				(0.033)		
GPA	2.003	0.024	1%	0.591	1.209	0.126	10%	0.064*
		(0.045)				(0.068)		
Academic Index - Behavioral	0.047	0.062		0.005***	-0.085	0.101		0.035**
		(0.022)				(0.048)		
Misconduct Incidents	0.393	-0.068	-17%	0.126	0.630	-0.194	-31%	0.008***
		(0.044)				(0.073)		
Ever Enrolled in Juv. Justice	0.006	-0.004	-61%	0.278	0.083	-0.020	-24%	0.262
		(0.003)				(0.018)		
No. of CPS Schools Enrolled In	0.901	-0.055	-6%	0.026**	0.779	-0.004	-1%	0.913
		(0.025)				(0.041)		

Notes: Estimates for each outcome are from a single regression that interacts subgroup indicators with treatment. CM is the control mean. Our index is equal to an unweighted average of the different components, after all have been normalized to Z-score form using the control group's distribution. Index components used "mixed" imputation methodology for missing values. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.XX: C2C estimated effects on academic outcomes, by gender

	Male (N=1193)				Female (N=872)			
	CM	ITT	% Change	P-Value	CM	ITT	% Change	P-Value
Semester 1								
Academic Index - Combined	-0.056	0.016		0.446	0.081	0.056		0.009***
		(0.021)				(0.021)		
Academic Index - Engagement	-0.068	-0.009		0.797	0.097	0.078		0.041**
		(0.033)				(0.038)		
Attendance Days	54.024	0.277	1%	0.830	54.988	3.472	6%	0.012**
		(1.294)				(1.374)		
Ever Enrolled or Graduated	0.832	-0.025	-3%	0.226	0.871	-0.008	-1%	0.735
		(0.021)				(0.022)		
GPA	1.595	0.001	0%	0.990	1.982	0.097	5%	0.116
		(0.050)				(0.062)		
Academic Index - Behavioral	-0.045	0.055		0.098*	0.065	0.054		0.074*
		(0.033)				(0.030)		
Misconduct Incidents	0.650	-0.091	-14%	0.097*	0.523	-0.106	-20%	0.095*
		(0.055)				(0.063)		
Ever Enrolled in Juv. Justice	0.057	0.002	3%	0.886	0.005	0.002	36%	0.730
		(0.012)				(0.005)		
No. of CPS Schools Enrolled In	0.913	-0.043	-5%	0.119	0.947	-0.034	-4%	0.251
		(0.028)				(0.029)		
Semester 2								
Academic Index - Combined	-0.060	0.062		0.003***	0.086	0.030		0.112
		(0.021)				(0.019)		
Academic Index - Engagement	-0.069	0.034		0.279	0.099	0.051		0.162
		(0.032)				(0.036)		
Attendance Days	48.010	1.994	4%	0.122	51.019	2.021	4%	0.171
		(1.290)				(1.477)		
Ever Enrolled or Graduated	0.783	-0.049	-6%	0.036**	0.807	0.002	0%	0.945
		(0.023)				(0.026)		
GPA	1.569	0.074	5%	0.136	1.939	0.039	2%	0.501
		(0.050)				(0.058)		
Academic Index - Behavioral	-0.050	0.112		0.001***	0.072	0.027		0.280
		(0.034)				(0.025)		
Misconduct Incidents	0.560	-0.148	-26%	0.016**	0.357	-0.057	-16%	0.164
		(0.061)				(0.041)		
Ever Enrolled in Juv. Justice	0.052	-0.019	-36%	0.086*	0.007	0.001	17%	0.809
		(0.011)				(0.005)		
No. of CPS Schools Enrolled In	0.845	-0.049	-6%	0.097*	0.876	-0.021	-2%	0.513
		(0.029)				(0.032)		

Notes: Estimates for each outcome are from a single regression that interacts subgroup indicators with treatment. Our index is equal to an unweighted average of the different components, after all have been normalized to Z-score form using the control group's distribution. Index components used "mixed" imputation methodology for missing values. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.XXI: C2C estimated effects on academic outcomes, by school engagement at baseline

	Engaged in school at baseline (N=1454)				Not engaged in school at baseline (N=611)			
	CM	ITT	% Change	P-Value	CM	ITT	% Change	P-Value
Semester 1								
Academic Index - Combined	0.083	0.034		0.041**	-0.187	0.029		0.362
		(0.017)				(0.032)		
Academic Index - Engagement	0.153	0.057		0.043**	-0.347	-0.042		0.429
		(0.028)				(0.053)		
Attendance Days	59.325	2.957	5%	0.005***	43.306	-1.630	-4%	0.424
		(1.041)				(2.039)		
Ever Enrolled or Graduated	0.887	-0.003	0%	0.856	0.760	-0.053	-7%	0.113
		(0.017)				(0.034)		
GPA	1.944	0.056	3%	0.207	1.324	0.007	1%	0.933
		(0.044)				(0.079)		
Academic Index - Behavioral	0.012	0.025		0.311	-0.028	0.125		0.012**
		(0.025)				(0.049)		
Misconduct Incidents	0.596	-0.097	-16%	0.063*	0.601	-0.099	-17%	0.114
		(0.052)				(0.063)		
Ever Enrolled in Juv. Justice	0.013	0.010	78%	0.105	0.087	-0.018	-20%	0.362
		(0.006)				(0.019)		
No. of CPS Schools Enrolled In	0.969	-0.019	-2%	0.419	0.833	-0.088	-11%	0.030**
		(0.023)				(0.041)		
Semester 2								
Academic Index - Combined	0.076	0.050		0.001***	-0.172	0.044		0.174
		(0.015)				(0.033)		
Academic Index - Engagement	0.141	0.071		0.007***	-0.320	-0.031		0.546
		(0.026)				(0.051)		
Attendance Days	53.505	3.224	6%	0.003***	39.599	-0.956	-2%	0.643
		(1.069)				(2.060)		
Ever Enrolled or Graduated	0.835	-0.004	0%	0.827	0.699	-0.084	-12%	0.024**
		(0.019)				(0.037)		
GPA	1.907	0.092	5%	0.032**	1.300	-0.021	-2%	0.782
		(0.043)				(0.075)		
Academic Index - Behavioral	0.010	0.040		0.077*	-0.024	0.161		0.002***
		(0.023)				(0.052)		
Misconduct Incidents	0.427	-0.057	-13%	0.162	0.589	-0.234	-40%	0.010**
		(0.041)				(0.091)		
Ever Enrolled in Juv. Justice	0.020	-0.008	-40%	0.210	0.064	-0.016	-24%	0.375
		(0.006)				(0.017)		
No. of CPS Schools Enrolled In	0.902	-0.014	-2%	0.562	0.756	-0.092	-12%	0.038**
		(0.025)				(0.044)		

Notes: We use the following criteria for not engaged in school at baseline: any youth who missed more than 60% of school days and received a grade of “F” in at least 75% of their courses in the semester prior to randomization, had a serious learning disability, were at least two years older than expected for the grade they were enrolled in at baseline, or were listed inactive at the end of any of the three semesters immediately prior to randomization. Estimates for each outcome are from a single regression that interacts subgroup indicators with treatment. Our index is equal to an unweighted average of the different components, after all have been normalized to Z-score form using the control group’s distribution. Index components used “mixed” imputation methodology for missing values. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.XXII: C2C estimated effects on academic outcomes, by age group

	Age under median (N=1033)				Age over median (N=1032)			
	CM	ITT	% Change	P-Value	CM	ITT	% Change	P-Value
Semester 1								
Academic Index - Combined	0.039	0.012		0.590	-0.039	0.054		0.009***
		(0.022)				(0.021)		
Academic Index - Engagement	0.125	-0.002		0.949	-0.125	0.058		0.128
		(0.032)				(0.038)		
Attendance Days	60.122	0.748	1%	0.540	48.730	2.510	5%	0.083*
		(1.221)				(1.449)		
Ever Enrolled or Graduated	0.857	-0.032	-4%	0.147	0.839	-0.004	0%	0.850
		(0.022)				(0.021)		
GPA	1.912	0.013	1%	0.802	1.596	0.070	4%	0.216
		(0.054)				(0.056)		
Academic Index - Behavioral	-0.047	0.045		0.224	0.047	0.065		0.019**
		(0.037)				(0.028)		
Misconduct Incidents	0.704	-0.085	-12%	0.232	0.491	-0.110	-22%	0.007***
		(0.071)				(0.041)		
Ever Enrolled in Juv. Justice	0.037	0.008	21%	0.443	0.033	-0.004	-13%	0.650
		(0.010)				(0.010)		
No. of CPS Schools Enrolled In	0.941	-0.047	-5%	0.116	0.914	-0.031	-3%	0.248
		(0.030)				(0.027)		
Semester 2								
Academic Index - Combined	0.033	0.051		0.015**	-0.033	0.046		0.016**
		(0.021)				(0.019)		
Academic Index - Engagement	0.115	0.038		0.222	-0.115	0.044		0.215
		(0.031)				(0.035)		
Attendance Days	53.393	2.090	4%	0.108	45.110	1.922	4%	0.180
		(1.300)				(1.434)		
Ever Enrolled or Graduated	0.809	-0.050	-6%	0.041**	0.776	-0.005	-1%	0.825
		(0.024)				(0.024)		
GPA	1.892	0.065	3%	0.229	1.550	0.053	3%	0.307
		(0.054)				(0.052)		
Academic Index - Behavioral	-0.049	0.094		0.008***	0.048	0.058		0.030**
		(0.035)				(0.027)		
Misconduct Incidents	0.577	-0.147	-26%	0.038**	0.377	-0.072	-19%	0.037**
		(0.071)				(0.034)		
Ever Enrolled in Juv. Justice	0.035	-0.007	-20%	0.454	0.031	-0.013	-43%	0.147
		(0.009)				(0.009)		
No. of CPS Schools Enrolled In	0.880	-0.055	-6%	0.081*	0.835	-0.019	-2%	0.521
		(0.031)				(0.030)		

Notes: Estimates for each outcome are from a single regression that interacts subgroup indicators with treatment. Our index is equal to an unweighted average of the different components, after all have been normalized to Z-score form using the control group's distribution. Index components used "mixed" imputation methodology for missing values. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.XXIII: C2C estimated effects on graduation, by school engagement at baseline

	Engaged in school at baseline (N=1454)					Not engaged in school at baseline (N=611)				
	N	CM	ITT	% Change	P-Value	N	CM	ITT	% Change	P-Value
Graduated Only										
Any graduation	1454	0.622	0.044	7%	0.059*	611	0.385	-0.094	-24%	0.009***
			(0.023)					(0.036)		
Graduated on-track	1447	0.536	0.047	9%	0.047**	607	0.285	-0.050	-18%	0.142
			(0.024)					(0.034)		
Graduated on-track + 1	1399	0.578	0.026	4%	0.273	599	0.319	-0.062	-19%	0.076*
			(0.023)					(0.035)		
Graduated on-track + 2	1330	0.568	0.019	3%	0.426	583	0.329	-0.056	-17%	0.114
			(0.024)					(0.036)		
Graduated or Transferred										
Any graduation	1454	0.717	0.028	4%	0.186	611	0.481	-0.079	-16%	0.034**
			(0.021)					(0.037)		
Graduated on-track	1447	0.626	0.031	5%	0.180	607	0.379	-0.044	-12%	0.245
			(0.023)					(0.038)		
Graduated on-track + 1	1399	0.669	0.004	1%	0.863	599	0.410	-0.056	-14%	0.143
			(0.022)					(0.038)		
Graduated on-track + 2	1330	0.649	0.003	0%	0.889	583	0.412	-0.049	-12%	0.202
			(0.022)					(0.038)		

Notes: For on-track graduation outcomes, a projected graduation year is calculated for each study member based on their grade level at the time of graduation. For those with missing grade levels at randomization, we calculate a projected graduation year using the highest grade level listed in any CPS masterfile for each student prior to randomization. Graduated on-track + 1 allows for one extra school year beyond each person's projected graduation semester; Graduated on-track + 2 allows for two extra school years. Graduated on-track outcomes subset to the sample to those for which we can observe their projected graduation year. Any graduation outcomes does not take projected graduation year into account, and considers the entire sample. Graduated or Transferred outcomes include study members who have an out-of-CPS transfer as their last leave code prior to their projected graduation year, providing an upper-bound on graduation rates. We use the following criteria for not engaged in school at baseline: any youth who missed more than 60% of school days and received a grade of "F" in at least 75% of their courses in the semester prior to randomization, had a serious learning disability, were at least two years older than expected for the grade they were enrolled in at baseline, or were listed inactive at the end of any of the three semesters immediately prior to randomization. Estimates for each outcome are from a single regression that interacts subgroup indicators with treatment. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

A.1.5 Additional Cost-Benefit Tables

Table A.XXIV: Inputs to social cost of crime estimates, 2015 dollars

	Bottom-Up Components						
Crime Type	Victim Costs	Legal System Costs	Offender Productivity Costs	Total Bottom-Up Costs	Willingness-to-Pay Costs	Estimated Clearance Rate	Estimated Reporting Rate
Murder	5,258,346	342,936	160,037	5,715,594	13,488,801	0.37	1.00
Rape	154,321	9,488	5,144	171,468	331,504	0.28	0.32
Armed Robbery	33,150	16,804	9,145	57,156	320,073	0.09	0.52
Robbery	13,717	8,459	4,572	26,292	44,582	0.09	0.52
Aggravated Assaults	42,295	15,432	7,316	62,872	97,165	0.15	0.56
Simple Assaults	5,144	5,716	1,486	12,574	21,719	0.11	0.37
Burglary	2,286	2,629	1,143	5,716	40,009	0.09	0.48
Motor Vehicle Theft	6,287	3,315	1,143	10,288	19,433	0.12	0.78
Larceny	514	1,943	800	3,201	4,572	0.11	0.28
Drunk Driving Crash	32,007	1,943	800	34,294	68,587	0.09	1.00
Arson	65,158	1,943	800	68,587	131,459	0.09	1.00
Vandalism	423	720	0	1,143	2,286	0.09	0.28
Fraud	1,257	1,943	800	4,001	6,287	0.09	0.28
Other	0	572	0	572	1,143	0.09	0.28

Notes: Bottom-up components and willingness-to-pay cost estimates taken from Cohen and Piquero (2009) and inflated to 2015 dollars. Estimated clearance rates and reporting rates from Bhatt et al. (2023)

Table A.XXV: C2C estimated effect on the social cost of crime, accounting for clearance and reporting rates

	Bottom-up Estimates		Willingness-to-pay Estimates		Upper Bound Estimates	
	CCM	TOT	CCM	TOT	CCM	TOT
6 months	$4,743	-$2,168**	$13,064	-$6,516**	$136,535	-$82,144**
		($872)		($2,783)		($41,302)
12 months	$7,527	-$2,501*	$21,983	-$6,283	$239,228	-$82,564
		($1,310)		($4,718)		($70,896)
24 months	$14,371	-$5,217**	$40,590	-$12,745*	$379,726	-$79,198
		($2,232)		($7,487)		($114,658)
36 months	$20,133	-$6,834**	$60,306	-$20,201**	$500,695	-$91,264
		($2,860)		($9,230)		($132,652)
48 months	$23,808	-$6,140*	$71,967	-$18,525*	$554,255	-$44,848
		($3,248)		($10,414)		($149,320)
60 months	$28,490	-$6,269*	$85,943	-$19,903*	$657,038	-$70,988
		($3,618)		($11,675)		($167,268)

Notes: Program costs are $5,070 per C2C participant, or $3,147 per youth randomized to C2C treatment, in 2015 dollars. All estimates are inflated to 2015 dollars. Bottom-up estimates assign costs of crime by summing up victim cost, legal system costs, and offender productivity costs from Cohen and Piquero (2009). Willingness-to-pay (WTP) costs from Cohen and Piquero (2009). Bottom-up and willingness-to-pay estimates only assign costs to observed arrests and victimization events. Upper bound estimates scale willingness-to-pay costs by reporting rate and clearance rate following the procedure outlined in Bhatt et al. (2023). More specifically, for each arrest, the portion of the WTP cost not assigned to legal system or offender productivity costs is scaled by 1/(reporting rate x clearance rate). For each victimization, the portion of the WTP cost not assigned to legal system or offender productivity costs is scaled by 1/(reporting rate), using the clearance rates and reporting rates in Appendix Table A.XXIV. All cost inputs, reporting rates, clearance rates can be found in Appendix Table A.XXIV. Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A.XXVI: C2C estimated effect on the social cost of crime, without recoding homicides

	Bottom-up Estimates				Willingness-to-pay Estimates			
	CM	ITT	CCM	TOT	CM	ITT	CCM	TOT
6 months	$4,494	-$1,333**	$4,743	-$2,168**	$12,569	-$4,008**	$13,064	-$6,516**
		($544)		($872)		($1,738)		($2,783)
12 months	$8,242	$8,815	$0	$14,332	$23,985	$21,671	$0	$35,237
		($7,358)		($11,782)		($18,270)		($29,249)
24 months	$38,062	$9,057	$0	$14,726	$100,140	$23,427	$0	$38,092
		($18,293)		($29,270)		($44,295)		($70,880)
36 months	$80,935	-$8,440	$51,138	-$13,723	$211,171	-$21,928	$137,674	-$35,655
		($26,713)		($42,697)		($65,098)		($104,049)
48 months	$122,184	-$2,804	$71,939	-$4,560	$316,769	-$10,099	$193,690	-$16,420
		($33,273)		($53,190)		($81,592)		($130,427)
60 months	$142,597	$17,644	$76,137	$28,689	$370,085	$38,362	$210,043	$62,376
		($36,808)		($58,871)		($90,412)		($144,595)

Notes: All estimates inflated to 2015 dollars. Program costs are $5,070 per C2C participant, or $3,147 per youth randomized to C2C treatment, in 2015 dollars. Bottom-up estimates assign the cost of each arrest and victimization based on the "bottom-up" cost estimates from Cohen and Piquero (2009), which aim to sum up costs to victims, legal system cost, and offender productivity cost for each event. Willingness-to-pay etimates use estimates from Cohen and Piquero (2009), which uses a "top-down" approach aiming to estimate the amount the public would be willing to pay to reduce each crime type. Cost inputs for each crime type can be found in Table A.XXIV. For each outcome period, we calculate the total cost incurred by each person by summing their arrest and victimization costs, which is used as the dependant variable in our regressions. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

A.1.6 Matching to CPD Data

To match our study participants to Chicago Police Department arrest data, we designed an algorithm using name, date of birth and address to match to Individual Record (IR) Numbers, a unique person-level identifier used by CPD. Specifically, if name and DOB are an exact match, we kept the match. However, given that CPD data is not accurate in every arrest record and that typos in name and DOB exist, we created several other pathways to ensure we would not miss accurate matches and designed the algorithm in a way to minimize false-positive arrests and manual review of particular cases. To generate potential matches, we match all CPD records to the C2C roster using DOB components (month/year, day/month, day/year) and compute a name similarity score for each match by calculating the Jaro-Winkler distance between each full name in the CPD data and C2C roster. Similarity scores are 0 if the names do not match at all, and 1 if they are an exact match. We consider any resulting match with a similarity score above 0.85 as a potential match. We allow for highly similar names (similarity score > 0.95) with a highly similar date of births (differing on regular time intervals that may be attributed to a typo such as one year/month/day) with a matching address. Matches were kept automatically if they have the same date of birth and same name or highly similar names and same date of birth. If the date of birth is the same, but the name match was not highly similar (similarity score < 0.95), the case was manually reviewed, taking into account address information. Matches that went to manual reviews were only kept if at least two independent reviewers (out of three) consider the match good. We considered these type of "low name matches" with exact date of birth because the algorithm only looks at string distance and this does not account for phonetically similar names that are spelled differently.

If the date of birth was highly similar, we went through a different set of decision nodes before any matches were kept. In this context, we accounted for name matches that were similar but also "common" names found in the CPD data base (specifically among the 100 most common first names or the 100 most common last names found in CPD arrest files for people born during or after 1995). If date of birth was highly similar yet the name is a common CPD name, the match went to manual review (or dropped if name was not highly similar by string distance function and the address does not match). If date of birth was highly similar but the name is not a common CPD name, then the match was kept if the name is highly similar and the address matches, or sent to manual review otherwise. Only low name matches and non-address matches were dropped. Lastly, matches where the date of birth was not highly similar were either dropped outright (given how closely the name matches) or sent to manual review based on address.

A.1.7 Transfers out of Chicago Public Schools, Impact of C2C on Incapacitation, and Data Censoring Overall

Table A.XXVII: C2C Impact on Transfers out of Chicago and Incarceration

			ITT		TOT		
Outcome	N	Control Mean	Estimate	P-Value	CCM	Estimate	P-Value
Days Incarcerated							
6 months	2061	4.998	0.091	0.926	1.028	0.147	0.925
			(0.982)			(1.560)	
12 months	2061	11.032	-0.863	0.648	3.895	-1.395	0.643
			(1.893)			(3.006)	
24 months	2061	26.866	-3.049	0.455	13.230	-4.926	0.448
			(4.083)			(6.485)	
36 months	2061	52.054	-9.057	0.171	34.190	-14.635	0.164
			(6.610)			(10.507)	
Days Transferred out of Chicago							
6 months	2061	6.578	-0.604	0.627	3.551	-0.975	0.621
			(1.244)			(1.973)	
12 months	2061	18.507	-1.234	0.662	13.167	-1.994	0.657
			(2.826)			(4.486)	
24 months	2061	46.220	-4.324	0.489	38.615	-6.987	0.481
			(6.248)			(9.916)	
36 months	2061	74.238	-7.723	0.428	66.356	-12.478	0.42
			(9.745)			(15.468)	
Days Censored							
6 months	2061	11.576	-0.513	0.741	4.579	-0.829	0.736
			(1.549)			(2.459)	
12 months	2061	29.540	-2.097	0.529	17.062	-3.389	0.521
			(3.330)			(5.283)	
24 months	2061	73.085	-7.373	0.313	51.845	-11.913	0.305
			(7.313)			(11.604)	
36 months	2061	126.292	-16.780	0.144	100.546	-27.113	0.137
			(11.475)			(18.214)	

Notes: Data for these outcomes is currently limited to 42 months post randomization, and therefore we present results up to 36 months post randomization to match the time periods shared in all other results. Given that youth can move out of Chicago temporarily and then move back into Chicago Public Schools (CPS), we believe looking at the total number of total days they spend not enrolled in Chicago is a more accurate measure of this censoring issue. Looking at just any transfers out of Chicago, we find that about 13% of the control group is moving out of Chicago at some point in the 36 months post-randomization with C2C not having a statistically significant effect. We track incarceration data through Illinois Department of Corrections entry and exit data which allows us to see all prison spells, as well as those who have a leave code of "Legally committed to a correctional institution" in CPS data. We were not able to obtain data from Cook County Sheriff's office to track jail spells, so this measure might be under counted. Days transferred out of Chicago uses CPS leave codes to determine those who have transferred to a school outside of Chicago. CM is the control mean. Days censored is the combination of days incarcerated and days transferred out of Chicago. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.XXVIII: C2C Victimization Outcomes, Ever Victimized and Number of Victimizations

Outcome	Control Mean	ITT Estimate	ITT P-Value	TOT CCM	TOT Estimate	TOT P-Value
Number of Violent Victimizations						
6 months	0.085	-0.011 (0.013)	0.398	0.090	-0.018 (0.021)	0.390
12 months	0.159	-0.013 (0.019)	0.487	0.155	-0.021 (0.030)	0.480
24 months	0.319	-0.047 (0.032)	0.140	0.308	-0.076 (0.050)	0.133
36 months	0.472	-0.059 (0.041)	0.150	0.458	-0.095 (0.065)	0.143
48 months	0.618	-0.072 (0.051)	0.160	0.612	-0.116 (0.081)	0.153
60 months	0.758	-0.076 (0.061)	0.212	0.767	-0.124 (0.098)	0.204
Number of Serious Violent Victimizations						
6 months	0.037	-0.006 (0.008)	0.460	0.036	-0.010 (0.013)	0.453
12 months	0.069	-0.001 (0.012)	0.907	0.055	-0.002 (0.020)	0.906
24 months	0.144	-0.031 (0.018)	0.086*	0.131	-0.051 (0.029)	0.080*
36 months	0.214	-0.054 (0.022)	0.014**	0.203	-0.088 (0.035)	0.012**
48 months	0.264	-0.054 (0.025)	0.033**	0.251	-0.087 (0.040)	0.030**
60 months	0.321	-0.052 (0.029)	0.074*	0.315	-0.084 (0.046)	0.070*
Any Violent Victimization						
6 months	0.074	-0.008 (0.011)	0.481	0.081	-0.013 (0.018)	0.474
12 months	0.128	-0.006 (0.014)	0.673	0.124	-0.010 (0.023)	0.667
24 months	0.211	-0.015 (0.017)	0.399	0.204	-0.024 (0.028)	0.391
36 months	0.293	-0.019 (0.019)	0.318	0.290	-0.032 (0.031)	0.310
48 months	0.344	-0.020 (0.021)	0.329	0.353	-0.033 (0.033)	0.322
60 months	0.383	-0.015 (0.021)	0.478	0.391	-0.024 (0.034)	0.471
Any Serious Violent Victimization						
6 months	0.036	-0.009 (0.008)	0.233	0.038	-0.015 (0.012)	0.225
12 months	0.060	0.002 (0.010)	0.876	0.042	0.003 (0.016)	0.874
24 months	0.115	-0.018 (0.013)	0.170	0.095	-0.029 (0.021)	0.162
36 months	0.169	-0.036 (0.015)	0.017**	0.155	-0.059 (0.024)	0.015**
48 months	0.204	-0.035 (0.017)	0.036**	0.194	-0.057 (0.027)	0.033**
60 months	0.229	-0.013 (0.018)	0.458	0.214	-0.022 (0.029)	0.450

Notes: Victimizations classified using FBI codes and statute descriptions. Serious violent victimizations include homicide, sexual assault, robbery, and aggravated assault and battery victims. Violent victimiztion includes all serious violent victimizations plus simple assault and batteries, sexual abuse, and miscellaneous violence victims. CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

A.1.8 More details on CPD Stops

The last few decades have seen a steep rise in the use of proactive police strategies, which intentionally raise the frequency of police contact through street stops and field interrogations. These strategies have also been heavily criticized for targeting minority and young populations in specific neighborhoods (Gelman et al., 2007; Stoudt et al., 2011; National Academies of Sciences et al., 2018). [1a] To our knowledge, we are unaware of any program evaluation for individuals that look at street stops as a criminal justice outcome, despite street stops being the most common interaction individuals have with law enforcement and their documented negative impact on health, mental well-being, educational outcomes, fear of police, and political engagement, particularly among communities of color (Weitzer et al., 2008; Butler, 2014; Bandes et al., 2019; Futterman et al., 2016; Geller, 2019; Bell, 2017; Weaver and Geller, 2019; Del Toro et al., 2019; Pickett et al., 2021; Bacher-Hicks and de la Campa, 2020).

The legal rules governing stop and frisk permit officers to have a wide degree of discretion in who they can stop and under what circumstances. The standard set from the 1968 Supreme Court Terry v. Ohio (1968) involved a pedestrian stop that defined the parameters around "reasonable suspicion." The court ruled that it is not a violation of the Fourth Amendment's prohibition on unreasonable searches and seizures when a police officer stops a suspect on the street and questions them without probable cause to arrest, so long as the police officer has a reasonable suspicion that the person has committed, is committing, or is about to commit a crime.[2a]

In Chicago, police officers are supposed to fill out a stop record (called investigatory stop report (ISR) at CPD) anytime a contact has officially become a stop: when someone is being detained and they can't walk away voluntary. This is different from a field investigation where police officers can ask people questions, but the civilian is free to go at anytime. Documentation from CPD states that police officers are supposed to make the person aware when an encounter is a field

[1a]In Chicago, the ACLU came to an agreement with CPD to institute new stop regulations in 2016 after they found that CPD engaged in a pattern of unconstitutional street stops. The newly instituted street stop practices and procedures led to a substantial reduction in street stops coupled with a corresponding increase in vehicle stops (Hausman and Kronick, 2021). Given this behavioral change, we will investigate C2C's impact on both pedestrian and vehicle stops given the common use of pretext traffic stops as a policing strategy in Chicago. Only a small portion of our outcome period for this study happens prior to 2016 (November 2015-December 2015), where 18% of our study sample had been randomized. The vast majority of our study sample and study period for stops will cover the time period after the stops policy change.

[2a]Several subsequent court decisions after Terry v. Ohio, the concept of "reasonable suspicion" was expanded to include location as well as behavior. For instance, in the Supreme Court case of Illinois v. Wardlow (2000), a person's presence in a "high-crime area" can be relevant in determining whether a person's individual behavior is sufficiently suspicious. This court case is key to the detainment of youth who flee from officers in high crime areas, a very common scenario on the streets of Chicago. The courts have also determined within the police discretion to make a warrantless custodial arrest for a very minor offense, such as a seat belt violation, that is punishable only by a fine (Atwater v. Lago Vista, 2001).

investigation (a person doesn't have to answer and they are free to go) vs when something is an official investigatory stop (the police officer suspects a crime is happening, about to happen, or has happened). [3a]. Officer-initiated pedestrian and vehicle stops are often sensitive to officer effort and discretion as officers choose whether to investigate and/or intervene. Given that police stops are a measure of justice contact with a high degree of discretion, we believe this is an important juvenile justice outcome to understand as it is often the first documented point of contact between police and civilians.

A.1.8.1 Matching to CPD Stop Data

To match our study participants to Chicago Police Department (CPD) stops data, we followed a similar but more nuanced protocol outlined in the Matching to CPD arrest data section. Additional criteria was required to determine a good match to CPD for many cases, as a large proportion of traffic and pre-2016 stops had incomplete date of birth fields (some traffic stops only include the birth-year of the person being stopped, and many pre-2016 have no information on birth date). Because there are no unique identifiers in the stops data that link one person to multiple stops (such as IR number in the CPD arrest data), we created a pseudo-identifier that combined an individual's first name, last name, and birth date to match to study participants. Similar to the arrest matching protocol, if name and DOB or name and address are an exact match, we kept the match (these account for nearly 70% of our matches). For cases in the stops data with complete DOB fields that did not match perfectly to a C2C participant, we used the same probabilistic matching criteria outlined in the arrest matching section, while also considering the potential match's race and gender to reduce the number of manual reviews the team would consider. These probabilistic matches account for approximately 16% of the C2C matches to the stops data.

For cases which a stop record only had birth-year, we required a higher standard to keep a potential match. Given the lack of available information to match on, we only considered potential matches that had the same birth-year. Matches with a high name-similarity score, a high address-similarly score, matching gender, and which did not have a common name in the CPD data are kept. Other matches with a high name-similarity score were manually reviewed. Matches with birth-year only account for approximately 12% of our matches. For cases with completely missing DOB information, we had to rely on the age field recorded in the stops data and compare it to the implied age of a given study participant at the date of the stop. Because the age field in the

[3a] For more details on the CPD documentation required for both street and traffic stops, please see section A.1.8.2

stops data might be a guess by the CPD officer, we considered all matches with a recorded age being within one year of a given study participant's real age on the date of the stop. We only keep matches with a missing DOB field if the name matches perfectly, the address matches perfectly with a high name-similarity score, or a very high name-similarly score with a name that is unique to the CPD data (1 or 0 unique IR numbers in the arrest data). These cases make up only 2% of our matches.

A.1.8.2 What Constitutes a Stop

CPD officially defines an investigative stop as non-voluntary contact. Specifically, guidance for officers states, that "a voluntary contact is a consensual encounter between an officer and a person during which the person must feel free to leave the officer's presence. An officer may approach any person at any time for any reason on any basis. However, absent reasonable suspicion or probable cause, that person must be free to walk away at any time. An officer's ability to articulate that no factors existed that would make a reasonable person perceive they were not free to leave is important. The following are some factors the court may consider to determine whether or not a consensual encounter has elevated to an Investigatory Stop or an arrest: Threatening presence of several officers, Display of a weapon by an officer; Use of language or tone of voice indicating that compliance with the officer's request might be compelled; Officer blocks a person's path; or Choice to end the encounter is not available to the person." Officers are required to complete ISR reports, which are quite lengthy and require detailed characteristics about the person and the reason or factors that led to the stop intended to demonstrate reasonable, articulable suspicion. They also detail whether or not a protective pat down was conducted, the reasons for the pat down, and if and why a subsequent search was then done. A narrative field also details the factors and situation of the stop. Our stops data includes every stop (traffic or street) that was formally recorded by CPD, even if that stop resulted in an arrest. In cases where an arrest came from a stop, these incidents are recorded in both the stops and the arrest data.

For traffic stops, police officers in Chicago are required to fill out traffic stop data sheet that specifies the driver's name, race, the reason for the stop, whether or not a search was conducted, whether or not contraband was found during the search, and the action resulted.[4a]

[4a]Since 2004, given concern around the racial disparities in stops, the Illinois Traffic and Pedestrian Stop Statistical Study Act became law and required all Illinois law enforcement to document the race of the driver and the reason for the stop, and report traffic stops to the Illinois Department of Transportation. Illinois, and subsequently Chicago, has some of the most rigorous data collection requirements around traffic stops in the nation.

Table A.XXIX: C2C Stop Outcomes, Ever Stopped and Number of Stops

		ITT		TOT		
Outcome	Control Mean	Estimate	P-Value	CCM	Estimate	P-Value
Number of Stops						
6 months	0.437	-0.020	0.634	0.315	-0.032	0.628
		(0.041)			(0.066)	
12 months	0.850	-0.025	0.732	0.579	-0.040	0.728
		(0.072)			(0.115)	
24 months	1.696	-0.017	0.894	1.333	-0.028	0.892
		(0.129)			(0.206)	
36 months	2.642	-0.062	0.741	2.209	-0.101	0.737
		(0.188)			(0.300)	
Number of Street Stops						
6 months	0.393	-0.019	0.616	0.280	-0.031	0.610
		(0.038)			(0.060)	
12 months	0.753	-0.023	0.719	0.511	-0.038	0.714
		(0.064)			(0.103)	
24 months	1.363	-0.011	0.917	1.021	-0.018	0.916
		(0.106)			(0.170)	
36 months	1.945	-0.069	0.624	1.530	-0.112	0.618
		(0.141)			(0.225)	
Number of Traffic Stops						
6 months	0.044	-0.001	0.963	0.035	-0.001	0.963
		(0.014)			(0.022)	
12 months	0.097	-0.002	0.952	0.069	-0.003	0.951
		(0.026)			(0.041)	
24 months	0.333	-0.006	0.914	0.312	-0.010	0.913
		(0.058)			(0.093)	
36 months	0.697	0.007	0.943	0.679	0.012	0.942
		(0.100)			(0.160)	
Ever Stopped						
6 months	0.218	0.009	0.593	0.158	0.014	0.587
		(0.016)			(0.026)	
12 months	0.307	0.027	0.127	0.236	0.044	0.121
		(0.018)			(0.029)	
24 months	0.446	0.036	0.060*	0.391	0.059	0.056*
		(0.019)			(0.031)	
36 months	0.546	0.010	0.607	0.512	0.016	0.601
		(0.020)			(0.031)	
Ever Stopped (Street)						
6 months	0.208	-0.003	0.869	0.163	-0.004	0.867
		(0.016)			(0.025)	
12 months	0.287	0.024	0.161	0.219	0.039	0.154
		(0.017)			(0.028)	
24 months	0.398	0.021	0.272	0.350	0.033	0.264
		(0.019)			(0.030)	
36 months	0.471	0.008	0.683	0.433	0.013	0.678
		(0.019)			(0.030)	
Ever Stopped (Traffic)						
6 months	0.027	0.008	0.304	0.010	0.013	0.297
		(0.008)			(0.012)	
12 months	0.052	0.003	0.788	0.036	0.004	0.784
		(0.009)			(0.015)	
24 months	0.124	0.014	0.339	0.092	0.022	0.331
		(0.014)			(0.023)	
36 months	0.212	-0.008	0.662	0.197	-0.012	0.656
		(0.017)			(0.027)	

Notes: CM is the control mean. Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors. CCM is the control complier mean. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.XXX: C2C Effects on Arrests, Conditional on a Stop

				ITT		TOT		
Outcome	N(stop)	N(ind)	Control Mean	Estimate	P-Value	CCM	Estimate	P-Value
Stop to Arrest Rate								
6 months	858	451	0.127	-0.001	0.976	0.094	-0.002	0.975
				(0.023)			(0.050)	
12 months	1680	649	0.122	-0.014	0.409	0.121	-0.033	0.401
				(0.017)			(0.039)	
24 months	3401	950	0.116	-0.021*	0.079	0.127	-0.041*	0.078
				(0.012)			(0.023)	
36 months	5296	1129	0.111	-0.022**	0.026	0.123	-0.041**	0.026
				(0.010)			(0.019)	

Notes: Intent-to-treat (ITT) and Treatment on the treated (TOT) estimates were calculated using randomization block fixed effects and robust standard errors, clustered at the individual level. CCM is the control complier mean. N(stop) shows the number of stop records corresponding to a C2C participant within each follow-up window; N(ind) shows the number of C2C participants with a stop record for that follow-up window. We include the following baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. In addition, for this analysis we also include a few extra stop-informed covariates: ever stopped for a street stop, ever stopped for a traffic stop, number of street stops, and number of traffic stops. Standard errors are robust and shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

A.1.9 Incarceration

If youth are accumulating violent arrests as well as other significant charges, they may be more likely to be incarcerated (and for long periods of times) and therefore unable to be stopped or arrested by police officers. Our findings may be difficult to interpret if there are substantially more control youth incapacitated compared to treatment youth (and therefore more present in the neighborhood, etc). First, we want to highlight that incarceration rates among youth have fallen substantially in the last 5-10 years in Chicago and Illinois more broadly due to concerted policy efforts. For example, in July 2019 there were 264 youth detained in the Illinois Department of Juvenile Justice (IDJJ) centers across the state. [5a] In Chicago, the Juvenile Temporary Detention Center (JTDC) of Cook County on average houses roughly 100 youth at any given point in time, with stays averaging around a week or two.

Despite the general downward trends, we also attempt to confirm incarceration rates among our study youth. Although we don't have juvenile incarceration data available, we can use Chicago Public Schools (CPS) data as a proxy for incarceration. CPS records leave reasons for youth, including the reason of being legally committed to a correctional institution. CPS also has indicators for whether or not youth are attending the alternative schools available to detained students.[6a] We use the combination of these indicators, as well as publicly available adult incarceration data from the Illinois Department of Corrections (IDOC) as a combined measure of being incarcerated post randomization. In Table A.XXVII we present the RCT effects looking at incarceration as an outcome up to 36 months post randomization. Although the point estimates are negative, we find that the base rates are small (about 52 days incarcerated in 36 months), and there is no statistically significant differential impact between treatment and control youth in any outcome period. This suggests that incapacitation effects due to incarceration will not change the interpretation of our findings. We also combine this incarceration measure with transfer out of Chicago data to create an indicator for any form of data censoring, and find no differential data censoring in our study sample (see Table A.XXVII).

[5a] https://www2.illinois.gov/idjj/Pages/Data-and-Reports.aspx

[6a] There are two CPS schools that serve students that are detained. Nancy B Jefferson Alternative High School serves youth in the JTDC. While York Alternative High School works with students 17 and older who are detained in the Cook County Jail.

A.1.10 Program effects by predicted arrest risk quartiles

Aiming to further understand the different effects the C2C program has in the diverse population it serves, we conduct an analysis in which we divide the population into quartiles defined by the predicted risk of contact with the criminal justice system, following the procedure outlined in Abadie et al. (2018). Specifically, we train a predictive model using as target variable the observed number of arrests 24 months after randomization for those in the control group, conditional on pre-specified baseline characteristics. ITT effects are then estimated within each one of the predicted subgroups using both leave-one-out (LOO) and repeated split sample (RSS) estimators.[7a] We present results using both estimators for full transparency, with RSS being our primary estimator given that in simulations studies it has shown lower bias and better coverage properties than LOO (see Abadie et al. (2018)). Standard errors are obtained using 500 bootstrap iterations, for each estimator and outcome-period. Within each risk quartile, we study the following outcomes: number of arrests, number of violent offense arrests, ever arrested, ever arrest for a violent offense arrest, combined academic index.

Results obtained suggest that the C2C program works differently for youth across different risk groups. The most robust result we see is for the ever arrested category: we observe significant (or borderline significant) reductions concentrated in mid-risk quartiles 2 and 3 for the probability of ever being arrested (see Fig. A.XII); and significant (or borderline significant) reductions in the probability of ever being arrested for a **violent offense** arrest solely concentrated in the highest risk quartile (4), as shown in Fig. A.XIII. A similar trend can be observed when looking at the number of arrest outcomes, with negative point estimates concentrated in mid-risk quartiles 2-3 when looking at reductions for overall number of arrests (see Fig. A.X), and most meaningful reductions again concentrated in the highest risk quartile when looking at the number of **violent offense** arrests (see Fig. A.XI), significant up to 12 months after randomization. Finally, when looking at academic outcomes as represented by our combined index, we see that improvements appear to be concentrated in mid-to-high risk groups 3 and 4.

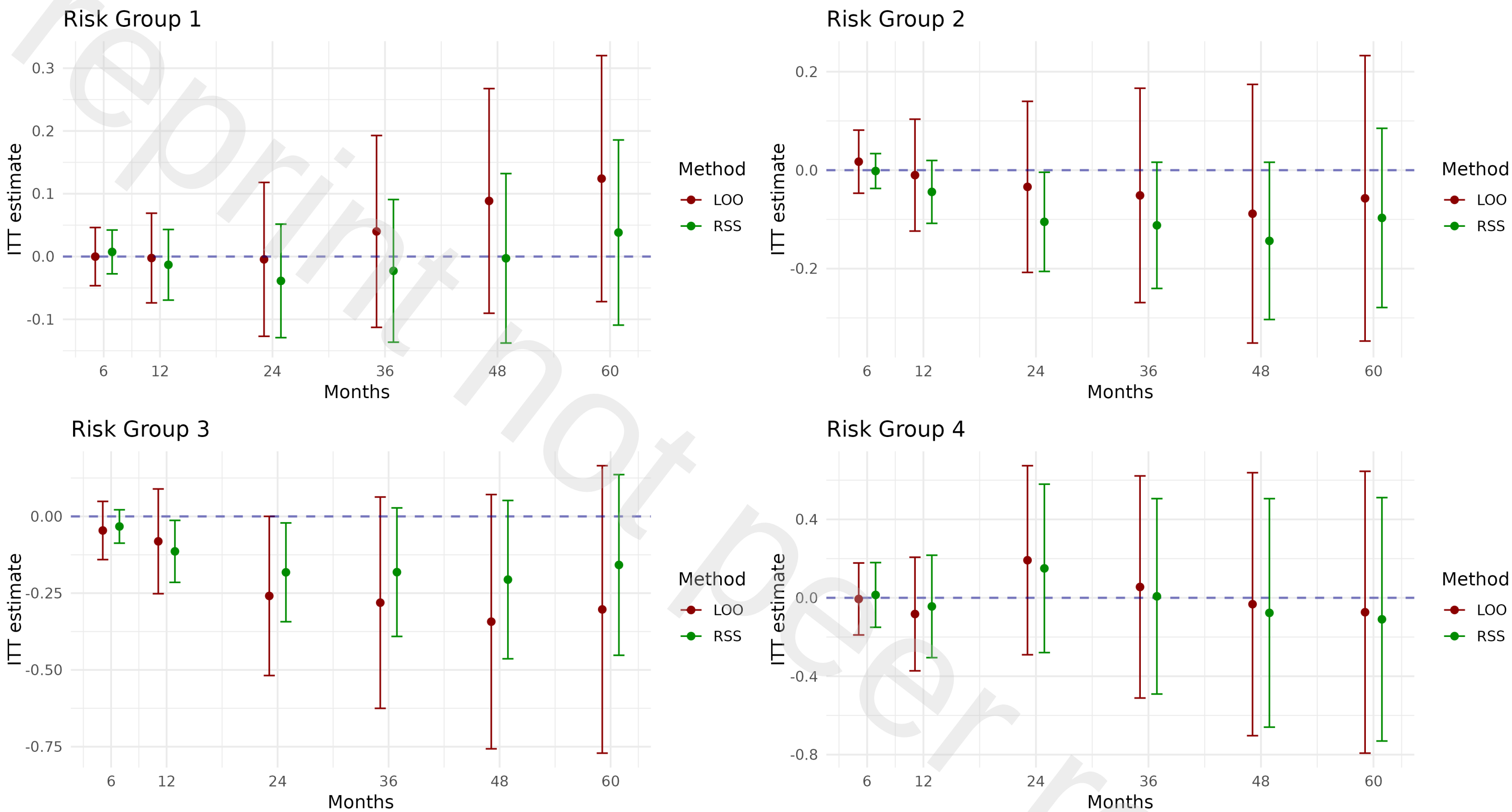
These findings are interesting and meaningful. C2C's unique combination of wraparound mentorship and CBT sessions allow it to be a program that "meets youth where they are." As such, it should not surprise us to see reductions in violent offense arrests concentrated in the highest risk quartile, as it is expected that both mentors and therapists would have emphasized these kind of

[7a] In the case of RSS, estimates where obtained using 100 repetitions.

behavioral changes with youth at the highest risk for committing violent acts, based on their direct interactions with them. Likewise, reductions in less serious arrests being concentrated in mid-risk quartiles make sense given that for these groups, mentors and therapist would have reinforced behavioral changes adequate to the challenges faced by them. Academic results concentrated in upper-mid and highest quartiles seem a bit more puzzling, but could potentially be explained by the fact that the academic movement we see is mainly due to behavioral changes, in particular reductions in school incidents (see Table VI); and perhaps to a lesser extent to engagement changes such as attendance. It seem plausible that reductions in school incidents are coming primarily from youth in the top risk quartile. And, at the same time, improvements in the engagement dimension, such as increased attendance, are concentrated in the upper-mid quartile. This interpretation is in line with what we see from Table A.XXI, where academic results are shown for youth engaged in school at baseline vs not. If this interpretation is correct, C2C mentors and therapists would effectively be aiding these different risk groups in the specific school challenges they face.

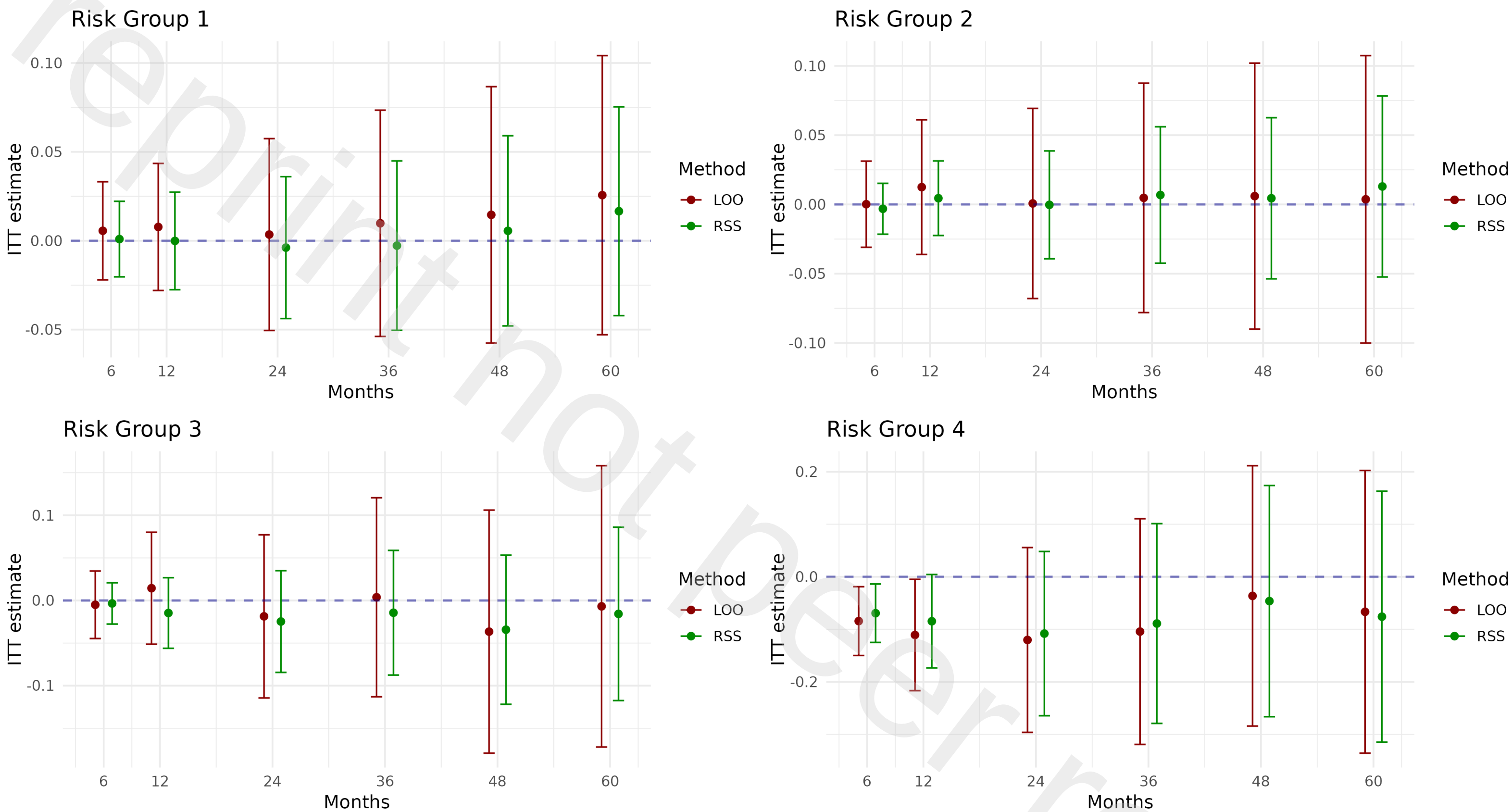
The results presented in this section should be interpreted with caution, and are only suggestive. In most trends discussed above, only some of the estimates are significant at the 95% level. This is especially true in the case of our number of arrest outcomes (Figs. A.X and A.XI) and academic outcomes (Fig A.XIV), where **most** estimates don't cross the 95% significance line. We believe this to be related to the decreased statistical power when dividing our sample into quartiles. That said, we still think these results are informative and meaningful enough to be included here.

Figure A.X: ITT effects by arrest risk, number of arrests



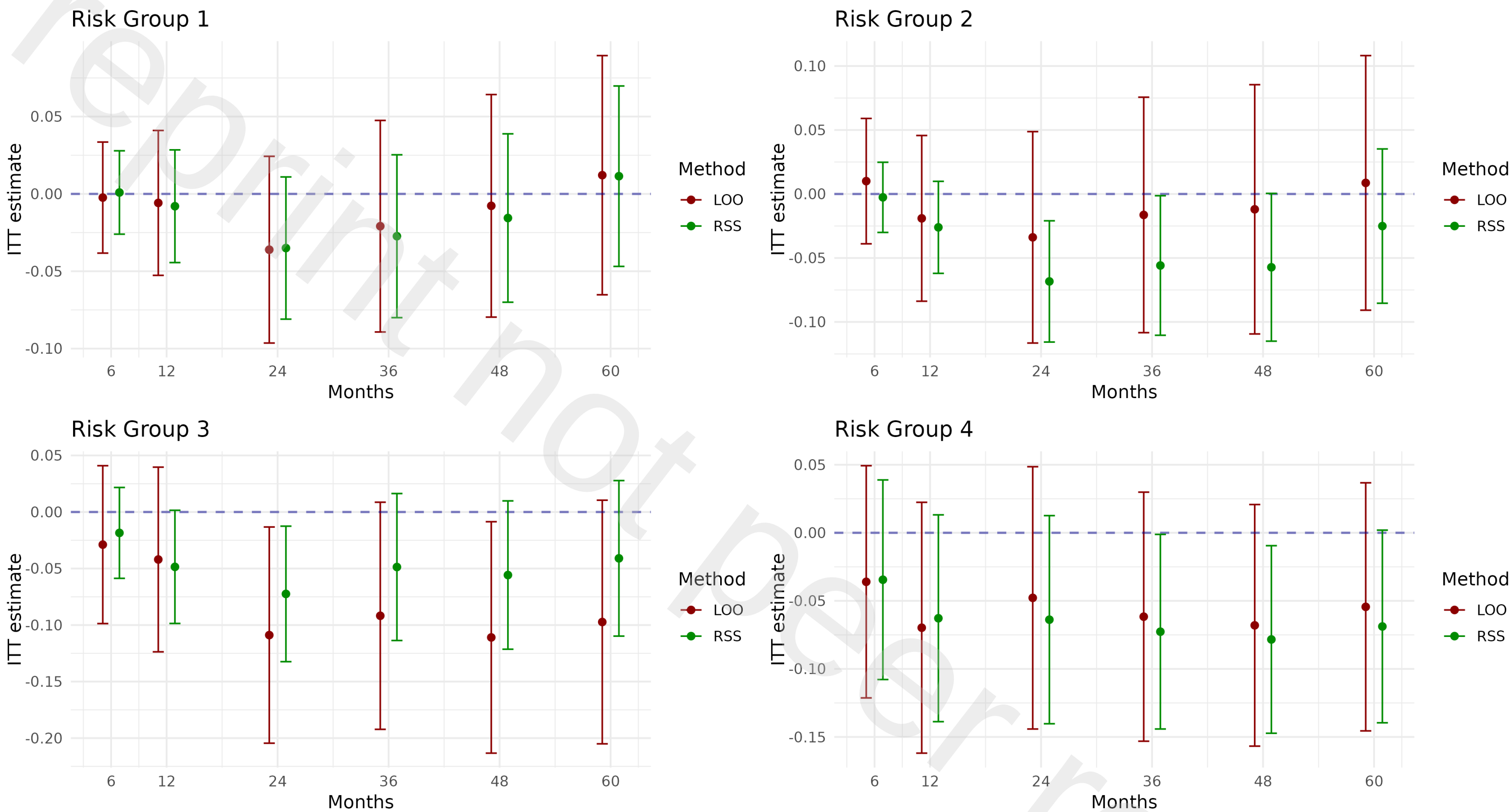
Notes: We report point estimates with 95% confidence intervals from both the repeated split-sample (RSS) and leave-one-out (LOO) estimators. Prediction and estimation models control for baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. Standard errors are robust.

Figure A.XI: ITT effects by arrest risk, number of violent offense arrests



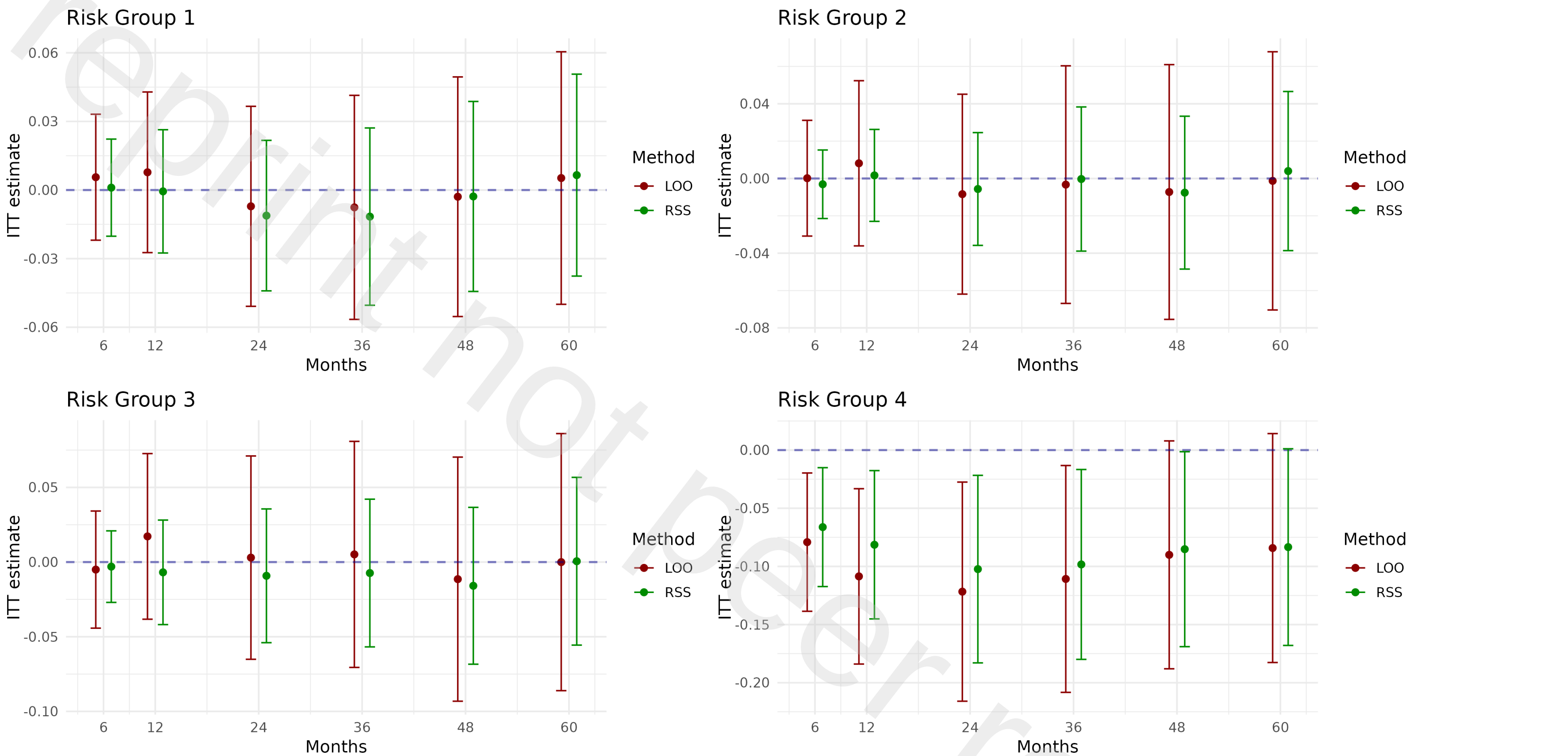
Notes: We report point estimates with 95% confidence intervals from both the repeated split-sample (RSS) and leave-one-out (LOO) estimators. Prediction and estimation models control for baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. Standard errors are robust.

Figure A.XII: ITT effects by arrest risk, ever arrested



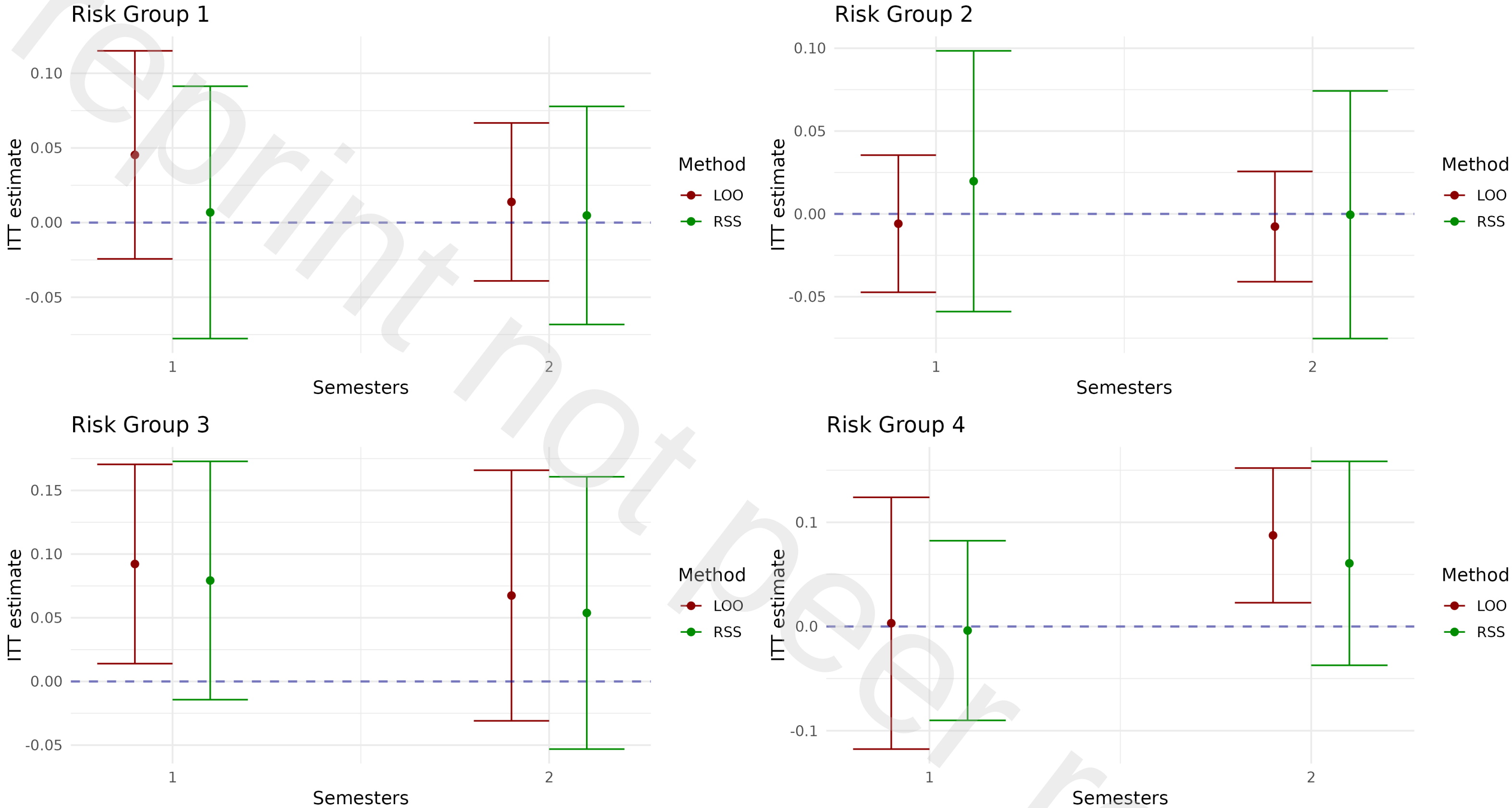
Notes: We report point estimates with 95% confidence intervals from both the repeated split-sample (RSS) and leave-one-out (LOO) estimators. Prediction and estimation models control for baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. Standard errors are robust.

Figure A.XIII: ITT effects by arrest risk, ever arrested for a violent offense



Notes: We report point estimates with 95% confidence intervals from both the repeated split-sample (RSS) and leave-one-out (LOO) estimators. Prediction and estimation models control for baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. Standard errors are robust.

Figure A.XIV: ITT effects by arrest risk, ed combined index



Notes: We report point estimates with 95% confidence intervals from both the repeated split-sample (RSS) and leave-one-out (LOO) estimators. Prediction and estimation models control for baseline characteristics: demographic covariates (age/race/gender dummies), school grade at randomization indicators, prior enrollment in free/reduced lunch benefits, prior gap in enrollment indicator, prior arrest records by type (numbers and indicators), indicator for any prior victimization, and number of prior victimizations. Standard errors are robust.