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

| | |
|---|---|
| Document Title: | The Effects of Summer Jobs on Youth Violence |
| Author(s): | Sara Heller, Harold Pollack, Johnathan M.V. Davis |
| Document Number: | 251101 |
| Date Received: | August 2017 |
| Award Number: | 2012-MU-FX-0002 |

# Final Technical Report for Grant 2012-MU-FX-0002
# The Effects of Summer Jobs on Youth Violence

PIs: Sara Heller (University of Pennsylvania) and Harold Pollack (University of Chicago)*
With Jonathan M.V. Davis, University of Chicago

## Abstract

For decades, policymakers have attempted to use employment programs to improve job prospects and reduce crime among disadvantaged youth. But most empirical evidence suggests that changing youths' behavior with these programs is difficult and costly. This report presents somewhat more optimistic findings from a randomized controlled trial of an intervention that has been largely absent from the rigorous evaluation literature: summer jobs. In 2012, we randomly assigned 1,634 disadvantaged youth applicants from 13 Chicago public high schools to a program called One Summer Chicago Plus (OSC+) or to a control group. The program offered an 8-week summer job at minimum wage, an adult job mentor, and for some youth, a cognitive behavioral therapy-based curriculum. We track youth in administrative data sources and find that the main effect of the program was to dramatically reduce violence. In the first year, violent-crime arrests dropped by 45 percent (4.5 fewer arrests per 100 participants). The decline does not continue in the second year, although the cumulative effect provides suggestive evidence that the program could have a long-term impact on violence. There are no significant changes in other types of crime. The mechanism at work does not appear to be incapacitation during summer work hours (the drop occurs mostly after the end of the program), nor increased time or effort in high school the following year. We also find no effects on formal-sector employment or college enrollment, although those results have a variety of limitations. One possibility is that the program improves how youth handle or avoid conflict, which might affect violence without changing other outcomes. Although more research is needed to determine why the program works, for whom, and in what contexts, the study results highlight the utility of rethinking what youth employment programs can do. Even without improving employment or changing the total number of arrests, summer jobs programs can reduce a hugely socially costly outcome at a relatively low cost; we estimate that social benefits are likely to justify program costs, and may outweigh them by as much as 11 to 1.

Acknowledgements: This project was supported by Award B139634411 from the U.S. Department of Labor and Grant 2012-MU-FX-0002 from the Office of Juvenile Justice and Delinquency Prevention, Office of Justice Programs, U.S. Department of Justice. We acknowledge the City of Chicago and Mayor Rahm Emanuel, the work of the Chicago Department of Family Support Services, especially Evelyn Diaz, Jennifer Axelrod, Andrew Fernandez, and Jennifer Welch, and the organizations that implemented the program: Phalanx Family Services, Sinai Community Institute, Saint Sabina Employment Resource Center, SGA Youth and Family Services, and Youth Guidance. Thanks to data providers at the Chicago Public Schools, the Chicago Police Department, and the Illinois Department of Employment Security, as well as to Roseanna Ander, Stephen Coussens, Gretchen Cusick, Meg Egan, Richard Harris, Nathan Hess, Addie Métivier, and Janey Rountree for research assistance or project support. Some of the data was provided by and belongs to the Chicago Police Department, Chicago Public Schools, Chicago Department of Family and Support Services, and Illinois Department of Employment Security. Any further use of this data must be approved by those agencies. All content is the responsibility of the authors and does not represent the official position or policies of any of the funders or data providers.

* We note that Dr. Pollack's role has been advisory. He was initially the sole PI because Dr. Heller was a graduate student at the beginning of the project, and students are not allowed to serve as PIs at the University of Chicago. The research started as part of Dr. Heller's dissertation, and the initial publication was sole-authored. The additional analyses reported here are with Jonathan Davis.

## I. Introduction

Minority and low-income youth experience strikingly disparate socio-economic outcomes. Forty percent of African-American young adults were jobless year-round in 2011, compared with only 24 percent of whites (Sum, et al. 2014). One in three black men will spend time in prison during their lifetimes, but only 1 in 17 white men (Bonczar 2003). And homicide kills more young African-American males than the 9 other leading causes of death *combined* (while killing not even one-tenth the number of young white males who die from the single leading cause) (Center for Disease Control and Prevention 2014).

These racial and socio-economic inequalities in crime and employment are a pressing social problem. They not only generate enormous human costs, but also drain hundreds of billions of dollars from state budgets and the national economy (Council of Economic Advisors 2015). The causes of these disparities are complex, but seminal social science theory by economist Gary Becker (1968) and sociologist William Julius Wilson (1996) has been influential in identifying one candidate lever for intervention: employment. These theorists argue that poor employment prospects drive crime and violence by lowering the cost of punishment and crippling inner-city neighborhoods. If so, it might seem logical that policies to enrich job prospects among disadvantaged youth should reduce their involvement in crime and violence.

Empirically, however, employment programs have had mixed success, especially among young people. Only very intensive and expensive interventions appear to improve employment among disadvantaged youth (Kemple, Willner & MDRC 2008; Millenky, et al. 2011; Roder & Elliott 2011; Schochet, Burghardt & McConnell 2008), and even fewer reduce crime.[1] The apparent difficulty of improving youth outcomes via jobs programs has often led to the conclusion that such programs require too much investment to be worth their costs (e.g., Heckman 2006).[2]

Yet summer jobs programs have been noticeably absent from this discussion, perhaps because of how little rigorous research has evaluated them (Fernandes-Alcantara 2011; LaLonde 2003). This report discusses what was, to the author's knowledge, the first experimental study to estimate the effects of summer jobs on crime – a study of the One Summer Chicago Plus (OSC+) program.[3] Initial results from the large-scale randomized controlled trial have been published as Heller (2014). This technical report, written as the final product for OJJDP Grant 2012-MU-FX-0002, briefly summarizes those results and adds analysis of additional crime, employment, and college data.

Section II gives a very brief overview of the summer jobs literature. Section III describes the program. Sections IV and V describe the data and methods respectively. Section VI summarizes

---

[1] Of the large, well-evaluated youth employment programs, only Job Corps and JobSTART reduce crime (Cave, et al. 1993; Schochet, Burghardt & McConnell 2008) (whereas the Job Training Partnership Act may actually increase crime among male youth) (Bloom, et al. 1997). The crime reductions, however, fade out quickly after the end of the programs, raising the possibility that intensive programs reduce crime because incapacitate youth for enough time during the program itself to reduce offending. The National Supported Work Demonstration also appears to have reduced crime among older participants, but not among youth (Uggen 2000).

[2] Although see, for example, (Heinrich & Holzer 2011) for a more optimistic read of the literature.

[3] Note that prior work has referred to the program as One Summer Plus (OSP). The City of Chicago has since updated its acronym, so this report uses the current OSC+ abbreviation.

the findings previously reported in Heller (2014). Section VII presents additional results, and Section VIII concludes.

## II. Related Literature

Despite over 50 years of federal funding and widespread implementation in most large U.S. cities, summer jobs have historically received almost no rigorous research attention (LaLonde 2003). A few studies on programs carried out from the 1960s to 1980s, which included both summer jobs and a range of non-summer services, find some positive effects on schooling or earnings, especially for black males (Farkas, Smith & Stromsdorfer 1983; Grossman & Sipe 1992; Somers & Stromsdorfer 1972). Yet all but one use non-experimental designs susceptible to selection bias. The single experiment, which tests the STEP program (Grossman & Sipe 1992; Walker & Vilella-Velez 1992), provides *both* treatment and control groups with training and employment. It identifies only the effect of an additional life skills and sex education curriculum which was not offered to the control group, making the study more about an education intervention than about summer jobs. A more recent study of a Philadelphia-area program claims to be a randomized controlled trial, but appears to report analyses based only on observational analyses (McClanahan, Sipe & Smith 2004).[4]

There are promising indications in non-experimental settings that summer jobs can reduce delinquency (e.g., Sum, Trubskyy & McHugh 2013). But the risk of selection bias, as well as a reliance on self-reported outcome data, means that any observed behavioral differences could be because participants are more hesitant than the comparison group to report wrong-doing, or from pre-existing differences between participants and non-participants.[5]

The first study to estimate the effects of summer jobs using a random-assignment mechanism (lotteries that allocate program slots in New York City) is Leos-Urbel (2014) and a follow-up paper (Schwartz, Leos-Urbel & Wiswall 2015). These papers show very small increases in school attendance and test-taking among the subpopulation that attends school. Heller (2014), discussed in detail below, was the first to experimentally test the effects of a summer jobs program on crime. A second study of New York City's program found results that seem consistent with the Chicago results reported here: a decline in incarceration for adult offenses and a 20 percent decline in mortality, likely driven by reduced homicide, but no improvements in employment or college-going (Gelber, Isen & Kessler 2014).

## III. One Summer Chicago Plus

OSC+, like many summer jobs programs around the country, provides a supported summer job. The program was designed by the government agency that administers the program, Chicago's

---

[4] The methodology appendix explains that the control youth who actually received treatment were dropped from the analysis, and the program effects were estimated by including an indicator variable for participation – not random assignment – in a regression analysis. This effectively compares actual participants to those who did not participate (both controls and treatment non-participants), thus re-introducing selection bias concerns.

[5] The treatment and comparison groups were quite different on observable characteristics in this study. For example, 48 percent of the treatment group was female versus 37 percent in the comparison group, and half of treatment youth were African-American, compared to 35 percent of the comparison group.

Department of Family and Support Services (DFSS), primarily as a violence-reduction intervention. The program model described here is for summer 2012, when DFSS first adapted the City's more general summer programming into a specific model for youth at elevated risk of violence involvement. It is worth noting that DFSS considers this an experimental program with which they can learn how best to improve youth outcomes. As such, there is not a single program manual, nor specific tests of implementation fidelity. Instead, the City has continued to experiment over time with variations to the program model in order to maximize positive impacts and test candidate mechanisms. Future work will report the results of this continued program testing.

To implement OSC+, DFSS contracts with community-based non-profit agencies, which are responsible for recruiting and serving youth participants. In 2012, those agencies were Sinai Community Institute, St. Sabina Employment Resource Center, and Phalanx Family Services. Program providers were responsible for implementing all aspects of the program, including finding jobs for youth. Because of a limitation imposed by one of the funders of the program, the 2012 program included only government and non-profit jobs, not private sector employment (this restriction was relaxed in later years). Youth served as summer camp counselors, assisted in aldermen's offices, cleared vacant lots to plant and maintain community gardens, and engaged in a variety of other jobs.

Providers recruited applicants at 13 high-violence, high-poverty Chicago public high schools, which were selected for their large numbers of youth at risk of violence involvement. The school-based recruiting system, which focused largely on the south and west sides of the city, successfully targeted youth living in high-violence neighborhoods (see Figure 1, which shows applicants overlaid with community area violent crime rates).

Participants in OSC+ were offered 8 weeks of programming for 5 hours per day, 5 days per week at Illinois' minimum wage ($8.25 per hour in 2012). In the study year, there were two versions of the program: one where youth worked at worksites for all 5 daily hours, and another where they worked fewer hours (3 per day) and spent the additional 2 hours per day participating in a social-emotional learning curriculum based on cognitive behavioral therapy (CBT) principles (described below).[6]

In both versions of the program, youth were assigned to job mentors – adults whose job was to teach youth to be successful employees and help them deal with barriers to employment – at a ratio of about 10 to 1. Characteristics of mentors varied: Some were staff at the program providers, some were college students home for the summer, and some were individuals who applied for the mentor jobs directly. Mentors participated in a one-day training (which has been revised and extended in later years of the program) and were paid a salary.

---

[6] OSC+ was originally designed to run over 7 summer weeks, but additional funding allowed for an optional week-long extension of the jobs component. Eight weeks of programming were offered but not required, and in the 8th week there was no CBT programming. Anecdotally, program providers reported that 2 hours per day was too much time for the CBT curriculum. Later iterations of the program spent less time on non-job activities. One service provider also offered access to additional, optional programming outside of OSC+ (like drama, graphic design, and fitness activities), but these activities were not funded by the program. Program impacts were not limited to this provider, so these activities seem unlikely to be the key driver of the results.

One hypothesis for why prior youth employment programs require intensive intervention to improve outcomes is that disadvantaged adolescents may lack the "soft skills" to benefit from lower-intensity programming. The addition of the CBT-based curriculum, given the general umbrella term "social-emotional learning" or SEL since it was not solely cognitive behavioral therapy, was designed to test whether targeting some of these soft skills could improve the impact of the program. The motivating idea for the SEL programming was to help youth learn to understand and manage the aspects of their emotions and behavior that might interfere with successful participation and employment (e.g., the inclination, not uncommon among adolescents, to snap defensively at a someone offering constructive criticism).

SEL sessions, delivered by the two non-profit agencies SGA Youth and Family Services and Youth Guidance, focused on emotional and conflict management, social information processing, and goal setting. The curriculum differed somewhat across the two providers, but both were based on a manualized curriculum guided by cognitive behavioral therapy principles, which focus on helping youth to track how their thoughts and beliefs lead to actions, and how to better control that process. Prior research has shown that similar programming can reduce violent crime and create lasting improvements in school engagement on its own (Heller, et al. 2015); its inclusion in OSC+ was to test whether, in combination with employment, it could increase program participation and improve outcomes more than the jobs alone.

## IV. Data

To keep costs low, the study relies exclusively on existing administrative data sources. Applicant information and participation data is from DFSS program records. Using name, date of birth, and gender, applicants were matched with probabilistic matching algorithms to individual-level Chicago Police Department arrest records and Chicago Public Schools student records. Data in this report cover the school year after the program (academic year 2012-13) and arrest records through two post-lottery years. Further discussion of the details of these data is in the web appendix to Heller (2014).[7]

Employment records are from the Unemployment Insurance databases maintained by the Illinois Department of Employment Security (IDES). For each employer at which a youth worked, these data report the total earnings, employer name, and industry by quarter. As with all Unemployment Insurance data, the records only include employment eligible for UI withholding. That excludes many agricultural and domestic positions, family employment, and any employment in the informal sector. Our current data includes complete records from quarter 1 of 2005 through quarter 1 of 2013 (one year after random assignment, or 2 quarters after the end of the program).

In order to match youth to UI data, IDES requires youths' social security numbers (SSNs). Since DFSS only collected SSNs for treatment youth who participated in the program, we took advantage of the fact that the school district has historically asked for SSNs during the enrollment process (they no longer do so, but most of the study youth enrolled at a time when

[7] The accepted version of Heller (2014) and appendix (prior to layout or copyediting) are included with this report. They are included by permission of the AAAS, for personal use, not for redistribution. The definitive version was published in Science (Vol 346, 5 December 2014), doi: 10.1126/science.1257809

they did).[8] The school district provided the numbers directly to IDES without researcher involvement, and removed them before we received the data. The district also removed 7% of the matches because of "significant" conflicts between the names in the two files (i.e., they removed apparent false positives). Although having a SSN is a baseline characteristic and should be balanced across treatment groups, treatment youth are slightly more likely to have an SSN available for matching (80 versus 77 percent, $p = 0.10$). Our analysis treats anyone without an SSN or who was removed as a bad match as missing.[9] This approach assumes that cases are missing completely at random. Although this is a strong assumption, the observable baseline characteristics we have available are still balanced across treatment and control youth with non-missing SSNs ($F(19, 1235 = 0.36$, $p = 0.995$, $n = 1{,}280$).[10] For youth with "valid" SSNs,[11] we assume anyone not matched to an Unemployment Insurance wage record never worked in the formal sector and assign zeros for employment and earnings.

To measure college outcomes for those old enough to have enrolled in post-secondary education, we use the National Student Clearinghouse data. Although reporting to the Clearinghouse is voluntary, the data include post-secondary enrollment information for over 3,600 colleges and universities covering 96% of students (National Student Clearinghouse 2015). The school district performs its own match for all students in the district to this data, which is linked to student identification numbers. We accessed the college data using the student identification numbers from our match to school district data. The data cover college enrollment through 2014, two years after the program. We limit the analysis to youth who were in 10th grade or older during the pre-program year (2011-12), since if they continued their grade progression with no delay, they are the youngest group that could have reached post-secondary education by fall 2014.

## V. Methods

### *A. Experimental Design and Study Population*

The full experimental design is described in the web appendix to Heller (2014), included as an attachment to this report. What follows is a basic summary of the key design elements.

Prior to the program, DFSS selected 13 Chicago public high schools to participate. Because OSC+ was designed to prevent violence, the schools chosen had the highest number of youth at risk of violence involvement in the city, as identified by a separate research partner. Program providers encouraged youth at these schools to apply to the program, marketing it as a summer jobs program with more work hours (and so more opportunity for income) than Chicago's standard summer programming.

---

8 Prior to May 2011, CPS asked parents and guardians to include SSNs in students' enrollment information. So any program applicant who was enrolled before that date had the chance to provide SSNs, although the school district did not validate them, nor require their submission. Since the decision to provide an SSN (or a valid SSN) is a pre-program characteristic, missing data should in theory be balanced across treatment and control groups.

9 Treating the matches that CPS removed as zeros rather than missing does not appreciably change the results.

10 When we obtain more than 2 post-program quarters of data in future work, we will assess the sensitivity of the results to different ways of treating the missing data that rely on weaker assumptions.

11 We call an SSN "valid" if it was both a) submitted to IDES and b) not removed by CPS because of a name mismatch.

A total of 1,634 youth in the study schools chose to apply for the 700 available program slots. The research team blocked youth on school and gender (the former to match youth to the closest program provider and the latter to over-select males, who are disproportionately involved in violence). We then randomly selected 350 youth for the jobs-only treatment arm and 350 for the jobs + social-emotional learning treatment arm. The remaining applicants were randomly ordered within blocks and treatment groups to form a waitlist. When 30 treatment youth declined to participate, the first 30 control youth (in the same block and treatment group as the decliners) were offered the program, for a total treatment group of 730. Control youth were completely embargoed from OSC+ but were free to pursue other summer opportunities, including other City programs. Among control youth with employment data available, only 12 percent were hired to a UI-covered job during the program quarters. The program did not crowd out much of this employment; 9 percent of the treatment group also worked in a UI-covered job during the program quarters.

Table 1 describes the study population in further detail, and shows tests of treatment/control balance. Only one of the pre-program differences is statistically significant at the 10 percent level, and the differences across all available baseline characteristics are not jointly significant ($F(20, 1588) = 0.61$, $p = 0.907$). In other words, randomization successfully balanced baseline covariates.

On average, study youth were just over 16 years old and in 10$^{th}$ grade. The applicants were almost entirely African-American (96 percent) and almost entirely from poor households (92 percent were eligible for free and reduced price lunch). They missed an average of about 6 weeks of school in the year before the program, and 19 percent had an arrest record. Among youth with available employment data, only 8 percent had worked in the year before the program (which is fairly consistent with statewide employment statistics for African-American teens).

### *B. Analysis Plan*

The analysis plan is as follows: Let $Y_{ibt}$ denote some post-program outcome for individual *i* in block *b* during post-randomization period *t.* This outcome, $Y_{ibt}$, will be a function of treatment group assignment, denoted by $Z_{ib}$, and observed variables from administrative records measured at or before baseline, $X_{ib(t-1)}$, as in equation (1) below.[12] We control for the blocking variable with block fixed effects, $\xi_b$. The "Intent-To-Treat effect" (ITT) captures the effect of being offered the chance to participate in the program, and is given by the estimate of coefficient $\delta_1$ in

[12] Baseline covariates include controls for demographic characteristics and neighborhood characteristics, as well as for pre-program criminal involvement, academics, and formal employment. Demographic controls include indicators for age at the start of the program and for being male, Black, or Hispanic. Neighborhood controls include the census tract's median income, proportion of those over 25 with a high school diploma or equivalent, and home ownership rate. Crime controls include separate indicators for having been arrested for 1 or 2 or more violent, property, drug, or other crimes. Academic controls include days absent and indicators for the student's free lunch status, special education status, enrollment status in the year prior to the program (determined by June 2012 CPS enrollment status and 2012 attendance), and grade level, as well as the number of As, Bs, Cs, Ds, and Fs received. Finally, employment controls include indicators for having a valid SSN and for having any formal employment in the year before the program. We impute zeros for missing data and include indicator variables that equal one if a variable was missing.

equation (1). Although baseline characteristics are not necessary for identification, we include them in the regression to improve the precision of estimates by accounting for residual variation in the outcomes.

$$Y_{ibt} = Z_{ib}\delta_1 + X_{ib,t-1}\delta_2 + \xi_b + \mu_{ibt} \quad (1)$$

The ITT framework fully exploits the strength of the randomized experimental design. Moreover, the coefficient $\delta_1$ in equation (1) may be especially useful for policy, as it directly addresses the impact of offering services on the outcome Y. But because not all youth offered the treatment participate, the ITT estimates will understate the effects of actually participating in the program on those youth who participate. Under the typical relevance and exogeneity assumptions for instrumental variables,[13] this latter set of effects can be recovered from the experimental data. We perform this estimation through a two-stage least squares strategy, in which random assignment, $Z_{ib}$, is an instrument for program participation, $P_{ibt}$, which is an indicator variable for starting the program, and $\hat{P}_{ibt}$ is the predicted probability of participation from equation (2):

$$P_{ibt} = Z_{ib}\,\pi_1 + X_{ib,t-1}\,\pi_2 + \gamma_b + \nu_{ibt} \ (2)$$
$$Y_{ibt} = \hat{P}_{ibt}\beta_1 + X_{ib,t-1}\beta_2 + \alpha_b + \varepsilon_{ibt} \ (3)$$

If all youth respond the same way to participating in the program (that is, if treatment effects are constant, or homogenous, across youth), then the coefficient $\beta_1$ in the system of equations (2) and (3) is interpretable as the average treatment effect (ATE) across this population of disadvantaged youth, which will also equal the effects of treatment on the treated (TOT). If treatment effects are heterogeneous across youth, then $\beta_1$ represents the local average treatment effect (LATE), or the effect of treatment on youth who complied with random assignment (though in our case, with no control crossover and so no always-takers, the LATE equals the TOT). To help judge the magnitude of the LATE estimates, we will also estimate the average outcomes of those youth in the control group who would have complied with treatment had they been assigned to treatment – or the "control complier mean" (CCM) (see Katz, Kling & Liebman 2001).

While ordinary least squares provides the best linear unbiased estimate of the treatment effect under the Gauss-Markov assumptions, we also explore the robustness of the results to non-linear specifications when appropriate (e.g., count and binary outcomes). For binary outcomes, we will focus our main analysis on a linear probability model to test the treatment-control difference of means, in part because the linear model simplifies the IV estimation of the TOT. But in order to ensure that results are not sensitive to the functional form assumptions underlying the LPM, we also test the robustness of results using logistic regression. These results are shown in Appendix B, Table B1. To analyze treatment-control differences in the number of arrests – a count variable – we use a Poisson quasi-maximum likelihood estimator with Huber-White robust standard errors to allow for over-dispersion, relaxing the Poisson distributional constraint that the mean equals the variance. These results are shown in Appendix B, Table B2. We exclude baseline

[13] In order for the random assignment variable, $Z_{ib}$, to be a valid instrument, it must be correlated with program participation, $P_{ibt}$, and uncorrelated with $\varepsilon_{ibt}$. Moreover, if treatment effects are heterogenous, it must shift participation in a uniform way across people. For example, we must assume there are no youth who would participate if assigned to the control group but not if assigned to the treatment group.

covariates in order to ensure convergence. The results do not substantively differ from the main results (intent-to-treat estimates excluding controls are also shown in the table for comparison). Again, the substantive conclusions do not differ from the main results.

In any experiment testing program effects on multiple outcomes, one might worry that the probability of Type I error increases with the number of tests conducted. To address this concern, we present both unadjusted p-values and p-values which are adjusted using a free-step down permutation method. The step-down method controls the family-wise error rate (FWER), or the probability that at least one of the true null hypotheses in a family of hypothesis tests is rejected (Anderson 2008; Westfall & Young 1993).[14] The FWER approach is useful for controlling the probability of making any Type I error, but it trades off power for this control. An alternative is to control the probability that a null rejection is a Type I error (the false discovery rate, or FDR), increasing the power of individual hypothesis tests in exchange for allowing some specified proportion of rejections to be false (Benjamini & Hochberg 1995; Benjamini, Krieger & Yekutieli 2006). We define our families of outcomes as: 1) the four types of crime (violence, property, drug, and other, excluding total arrests since it is a linear combination of the rest), 2) the three main schooling outcomes across the subset of the sample that would still be in school (re-enrollment, days present, and GPA), and 3) employment and earnings over the whole employment sample. We note that the violence effect was the primary pre-specified outcome of interest.

## VI. Prior Results

Heller (2014) reports the effects of the program on 16 months of crime data and the following academic year of schooling data. The main result is that the program decreased arrests for violent crimes by 43 percent (3.95 fewer arrests per 100 youth offered the program). There were no statistically significant changes in other types of crime (property, drug, or other).

Although the use of administrative data limits the ways in which potential mechanisms can be measured, the paper rules out a number of candidate explanations for the decline in violence. The violence drop was not simply from a mechanical "incapacitation" effect, where youth were physically kept busier over the summer. Instead, the violence decline continues to accrue after the program ends (and is still statistically significant in only the post-summer period, excluding the program months). The time path of the behavioral change suggests that the program changes something about youths' future behavior, not just their summer time use.

---

[14] We estimate the distribution of our test statistics accounting for all of the tests within a particular family by randomly sampling permutations of treatment status within blocks and recording all of the test statistics for each permutation. Under the null hypothesis of no treatment effect, each permutation should be identically distributed. Therefore, we are able to approximate the joint distribution of our test statistics with the distribution of the test statistics across permutations. For a particular hypothesis, we are able to estimate a level α critical value, $c(\alpha)$, with the $(1-\alpha)^{\text{th}}$ percentile of the estimated test statistic distribution. For a family of hypothesis tests, we determine the critical values using the step-down procedure outlined in Anderson (2008). Specifically, we sort the test statistics within a family of hypothesis tests from largest to smallest. Then we determine the adjusted critical value for the test with the largest test statistic using the distribution of the maximum test statistic within the family across permutations. We then drop the test with the highest test statistic and repeat the procedure for the test with the second highest test statistic. This continues until the last test in the family. We estimate the test statistic distributions using 5000 permutations.

The data are also not consistent with a different kind of change in future time use – increased time at school. The point estimates for days present during the following school year are small, negative, and not statistically significant. The top of the confidence interval rules out more than a two-day increase in attendance, making increased school attendance an implausible explanation for the 43 percent decline in violent-crime arrests.[15] The data also rule out relatively small changes in re-enrollment and GPA.

Lastly, the initial published results suggest that the program effect is not driven solely by the CBT-based curriculum. The effects of spending the entire program period in a job are statistically indistinguishable (and numerically very similar) to the effects of substituting 2 work hours per day for a curriculum targeting "soft" skills. The fact that the two types of activities appear substitutable suggests one of two explanations: either the employer and job mentor – which both treatment arms received – provide similar enough instruction that a formal curriculum is not needed to teach the relevant skills, or the mechanism at work is something else that both treatment arms share (changes in income, peer groups, other types of time use, or attitudes and beliefs). Anecdotally, the employers and job mentors provide much of the same types of instruction as the curriculum, especially in terms of helping youth manage and mediate conflicts with supervisors.

The paper suggests two reasons that such a low-intensity program could have such dramatic behavioral effects, which is surprising given the existing literature. First, most of the youth employment studies that look at crime as an outcome do not separate violence from other types of crime. Given that violence seems to respond differentially (which is also the case in numerous other studies of social interventions), the past literature may have aggregated away any violence impacts. Second, one of the major differences between summer jobs programs and the rest of the youth employment literature is that the former generally targets youth who are still in school and have not yet faced the full-time labor market, while the latter targets almost exclusively disconnected (out-of-school, out-of-work) youth. If prevention is easier than remediation, then providing a supported introduction to the labor force to in-school youth may require less intensive intervention than what is needed to change behavior for youth who have already been struggling on their own.

## VII. Additional Results

Since the Heller (2014) publication, we have collected additional data on employment, college-going, and arrests. Table 2 reports the employment and college results. Column 1 shows the program impact on whether youth appear in the Unemployment Insurance records at all during the year after the lottery, excluding the two quarters when the program occurred. Column 2 shows total earnings (the records do not report hours worked, so earnings can differ either from differences in time spent working or from differences in pay rates). Both estimates show no detectable effects on formal-sector employment, but they are imprecise zeros: Standard errors are as big or bigger than the point estimates. Employment is fairly low overall (only 13 percent of control compliers work in the formal sector), perhaps in part because most of the study youth

[15] A two-day increase in attendance is very close to the positive schooling effects in Leos-Urbel (2014). The Chicago evidence is therefore quite consistent with the prior evidence from New York City, although the latter looks at attendance conditional on going to school at all, while this analysis includes 0 days present.

(82%) return to school in the fall, and the data cover only two post-summer quarters (through March of the post-program year). Longer-term employment records may be a more appropriate measure of whether the program changes labor market outcomes.

As noted in the data section, the employment records are imperfect in a number of ways. About 20 percent of the sample did not have SSNs available for matching. Despite the fact that SSN availability was determined prior to the lottery (based on whether the school district had an SSN on record), treatment youth were 3 percentage points more likely to have one. UI records also only cover jobs that pay into the Unemployment Insurance system, which excludes self-employment and informal sector jobs. Field work by sociologists and ethnographers suggests that the informal economy may be a non-trivial source of income for youth living in low-income neighborhoods (e.g., Goffman 2015; Venkatesh 2006). These limitations, combined with the lack of statistical power for these short-term effects, means that the strength of our conclusions about employment outcomes is limited.

The college enrollment data, from the National Student Clearinghouse, suggest a similar story. There is no statistically significant change in college-going (column 3 of Table 2), although the confidence interval is relatively large (from a 8 percentage point decline to a 4 percentage point increase, relative to a base rate among control compliers of 40 percent[16]). Only two-thirds of the sample is old enough to be included in this analysis (we include only youth in 10th grade or higher at baseline, since 9th graders have not yet had time to reach college), so longer-term data may add some precision.

It is worth noting that the New York City summer jobs evaluation found a very precise null effect on employment and college-going, and a very small negative effect (about $85 per year) on wages over 4 post-program years (Gelber, Isen & Kessler 2014). The two sets of results may not be exactly comparable; New York has a different program model, population, and labor market, and the New York study covered program years 2005 – 2008, when economic conditions were different than in 2012. But to the extent that the results there might generalize to Chicago, or at least provide some indication of general mechanisms, they suggest that improved employment outcomes and increased post-secondary education may not be a central explanation for the violence decline.

Table 3 shows two years of arrest results broken down by year. For ease of interpretation, the outcome is scaled to the number of arrests per 100 youth. The table shows both the ITT and the IV estimate, though we focus our discussion on the effects of actually participating in the program (the IV estimate). In year 1 (top panel), violent-crime arrests decline by 4.5 per 100 participants (a 45 percent decline relative to the control complier mean).[17] The point estimates on other types of crime are positive but not statistically significant, so that the effect on total arrests

---

[16] Among the control group, 74% of youth graduated from high school and 19% of these graduates enrolled in a 4-year college. The graduation rates are comparable to the broader average across all of Chicago Public Schools, but the college-going rate is about 50% lower (among those who were first-time 9th graders in 2010-11 in CPS, 73% of students graduate and 40% of graduates enroll in a 4-year college) (Healey, Nagaoka & Michelman 2014).

[17] Note that these numbers differ slightly from Heller (2014) both because they cover a slightly different time period (12 months rather than 16 months) and because we have obtained additional baseline covariates. In a finite sample, changing covariates moves the point estimates by a small amount. Excluding baseline covariates entirely does not change the substantive conclusions, although estimates become somewhat less precise (see Appendix B).

is not significantly different from zero. The year 1 violence result is slightly less precise when adjusting for multiple hypothesis testing within that year, but still marginally significant: across the family of four crime outcomes, the FWER-adjusted p-value on violent-crime arrests is 0.08. If we instead control for the proportion of false null hypotheses rejections we wish to allow to for in the family (this is the q-value, or the FDR's p-value equivalent), we can reject the violent-crime null of no effects at q = 0.10.

There is no change in violent-crime arrests during year 2 (middle panel), suggesting the decline stops accruing after about a year. There continue to be no significant differences between treatment arms. It is worth noting that the control complier mean drops considerably during year 2 (from 10.0 to 4.5 violent-crime arrests per 100 youth). There are several possible explanations, including the fact that crime was declining in Chicago over this period of time.[18] Another interesting possibility is that study youth may be starting to age out of violent crime. There is a great deal of evidence in the criminological literature that violent offending peaks in the late teens and dramatically drops in the early 20s (Farrington 1986; Moffitt 1993). If this is happening in the study population as well, the time pattern of results might suggest that the program was optimally timed to coincide with the peak of violent offending among study youth, but does not continue to change behavior as violent offending slows down. There are no statistically significant changes in other types of crime during year 2, although drug and other crimes show imprecise, proportionally large (31 and 22 percent respectively) declines.

The bottom panel of Table 3 shows the cumulative effect over two years. Although the cumulative violence decline is just past the traditional threshold of statistical significance (p = 0.107) and not robust to adjustments for multiple testing, it is a proportionally large decline that is only a bit smaller than the year 1 effect (4.25 fewer violent crimes per 100 youth, which is a 29 percent decline). This provides suggestive evidence that treatment youth do not catch back up after year 1, but rather that the program may end up reducing lifetime violent-crime arrests.[19]

## VIII. Discussion

The results of this study suggest that the main effect of Chicago's supported summer jobs program for disadvantaged youth is to reduce violent-crime arrests. The violence decline is large (a 45 percent drop) and occurs mainly in the first year after the program. Although the cumulative effect over two years is a bit noisy, it is roughly the same size as the decline after the first year. If this pattern continues, it is possible that the program will generate a permanent decline in lifetime violent-crime arrests. The program has no detectable effect on other types of arrests, employment in the first two post-program quarters, or college-going, although confidence intervals for these outcomes are large.

As in any study that conducts hypothesis tests across a range of outcomes, one might be concerned that a change in a single outcome among many could be a false positive (Type I error). There are a number of reasons this is unlikely to be the case here. First, the number of arrests for violent crime was the primary outcome of interest prior to the start of the study (the study's pre-

---

[18] For example, UCR data suggests that aggravated assaults in Cook County declined by about 36% from 2012 to 2014 (Federal Bureau of Investigation 2012, 2014).
[19] Preliminary data from part of year 3 (not reported) is consistent with this idea and somewhat more precise.

registration listed this outcome first); in fact, DFSS's recruiting strategy and program model were both designed around the goal of violence reduction. This fact should help to assuage any concerns that we may have kept cutting the data until we found significant results.

Second, the violence decline in the first year is relatively robust to two different types of adjustments for multiple hypothesis testing (FWER-adjusted p-value = 0.08, and FDR q-value = 0.10). Some argue that adjusting p-values for multiple hypothesis testing for the primary pre-specified outcome of interest is not necessary, although whether and how to make these adjustments is debated in the literature (Bender & Lange 2001; Gelman, Hill & Yajima 2012). At a minimum, the fact that year-one result are still marginally significant after various adjustments for multiple testing suggests that the violence decline immediately following the program is unlikely to be a by-product of multiple hypothesis tests.

Despite being fairly similar in size to the year 1 violence decline, the cumulative change in violent-crime arrests over the two-year follow-up period is less precise and less robust to adjustments for multiple tests. This makes it less certain (at least until more time goes by) about whether the program reduces lifetime violent-crime arrests or only decreases violence during the first year or so. Adding additional cohorts of experimental data (we conducted new RCTs in summers 2013 and 2015) may add statistical power.

In some policy circles, this pattern of results – a big change in the short term that may not keep accumulating in the longer term – is a sign of program failure. Even OJJDP's own Model Programs Guide uses review criteria that assigns a study an "insufficient evidence" rating if the year 2 results are not the same as the year 1 results (on the basis of inconsistent outcome evidence). But from a policy perspective, this may not be the optimal way of judging a program. Consider violence as an example, which is a much bigger problem among youth (the incidence drops dramatically starting in the early 20s), and which carries a huge social cost for every occurrence. In that case, big short-term changes in the late teens might matter much more in terms of societal impact than small permanent changes. One way to capture that possibility is not to judge a program based on the duration of its impacts, but instead measure whether its benefits outweigh its costs. In the case of OSC+, this approach is particularly appealing; in addition to capturing the extraordinarily high social cost of violence, a benefit-cost comparison also aggregates the findings over time and across crime types into a single outcome measure, thus addressing multiple hypothesis testing concerns as well.

Assigning social costs to each type of crime involves a great deal of uncertainty: how to assign dollar figures to victim suffering, which of the costs (to victims, offenders, and offenders' families) to include, and even whether to limit the analysis to statistically significant changes in crime or to incorporate what may just be sampling error into the estimates. Rather than take a stand on what is inherently an uncertain exercise, our strategy is to take a range of different approaches from the literature and be transparent about how the different assumptions affect the conclusions. Details of our different approaches are described in Appendix A.

It is important to note that we do not aim to conduct a full benefit-cost analysis (which might include the deadweight loss from the part of program costs that were covered by taxes rather than private philanthropy, the opportunity cost of the buildings and staff, projections of future behavior, etc.). Instead, we perform the simpler exercise of estimating how the social benefits of reduced crime compare to the program's direct costs during the time we observe program impacts. We also note that although the standard errors from our regressions will reflect the huge variation in the social costs across different types of crime, they will understate the amount of conceptual uncertainty involved in assigning costs to crimes.

The administrative costs of the program itself were about $3,000 per youth in 2012. This figure includes wages paid to youth, which benefit-cost analyses of employment programs tend to treat as a transfer rather than an actual cost to society (LaLonde 1995; McConnell & Glazerman 2001). Since youth averaged about $1,400 in wages, the net administrative cost of the program is about $1,600. As this is the cost per participant, not per youth assigned to treatment, we report IV estimates for the social savings per participant as well.

Table 4 shows a variety of estimates for the social savings of reduced crime. The panels differ by which costs are included, as indicated in the bottom rows of the table. Within panel, the columns differ by whether costs are estimated with a "bottom-up" approach (constructing estimates of social harm from jury awards, medical costs to victims, and other direct costs) or a "top-down" approach (asking people how much they value avoiding crime in contingent valuation surveys).

When we use the time period during which crime is significantly declining (year 1), the program benefits generally outweigh its costs, with the most extreme benefit-cost ratio as high as 11 to 1. Results vary a bit more when we add the second year of data. The biggest difference across panels stems from whether we adjust for the fact that we measure arrests, which capture only a fraction of actual crime committed. If we inflate our arrest impacts by crime- and age-specific arrest-to-incidence ratios (which is fairly standard in these kinds of exercises[20]), the cumulative present value of benefits almost always exceed costs. If we rely only on observed arrests, the substantive conclusion is more sensitive to which cost estimates are used.

We consider Table 4 at least suggestive evidence that the program benefits outweigh its costs in the short-term, although longer-term data are needed before strong conclusions are merited. Replication studies, especially those in different locations with varying study populations and program models, are also needed to establish how general these results are. In that respect, we find the results from New York quite heartening; even with a different program, population, time period, and set of outcome measures, there is still a (cost-effective) decline in incarceration and violence.

An obvious question arises from these results: why do these kinds of summer jobs programs reduce violence without changing other outcomes? Most of the mechanisms we usually think about in terms of youth employment programs do not seem particularly consistent with a violence-specific effect. Increased income seems most likely to reduce property crime, as there would be less need to steal.[21] Changes in perceptions about the returns to schooling should

[20] For example, both prominent Perry Preschool benefit-cost analyses use estimated incidence-to-arrest ratios to estimate the benefits of reduced crimes rather than just reduced arrests (Belfield, et al. 2006; Heckman, et al. 2010). Using only arrests, as in panels 3 and 4, is probably somewhat conservative.

[21] It is true that robbery is considered a violent crime and involves acquiring more money; however, the program effects are driven mainly by assault, not robbery.

increase time or effort spent in school, which is not consistent with the education data. Increased opportunity cost from better employment prospects would likely increase the costs of all types of crime, not just violence.[22] Similarly, general improvements in pro-social beliefs, self-efficacy, or other attitudes and beliefs affected either by the mentoring or the work experience would seem likely to affect all types of crimes (and potentially school or work outcomes).

One possible explanation is that the program affects how youth handle (or avoid) conflict. By definition, violent crime differs from other types of crime in that it involves conflicts with other people. Perhaps the process of learning to work for the first time in a supported environment teaches better conflict management and self-regulation. If so, the fact that OSC+ serves in-school youth may be part of the explanation for its success: Youth generally start to age out of violent crime in their early 20s (Farrington 1986; Moffitt 1993), and violent crime may be concentrated when youth are at school (Jacob & Lefgren 2003). Teaching in-school youth to better manage conflict during the peak of their violence involvement may be a particularly well-timed and well-targeted lesson.

Anecdotally, both mentors and employers identify youths' tendencies to be reactionary and defensive in response to perceived slights as a key obstacle to successful program completion. Both report helping youth learn how to better navigate potential conflict situations as part of the program.[23] Of course, anecdotal evidence does not prove that this is the key mechanism at work; we continue to partner with Chicago on additional evaluations to isolate program mechanisms and learn exactly what works for whom.

In the meantime, this study provides some early evidence on the promise of a supported summer jobs program to reduce violence among disadvantaged youth. Considering how few programs have been rigorously shown to improve outcomes among this population, these findings are a promising beginning. More research will help to determine how widely the impacts generalize, and to sort out why such a low-intensity program can change behavior. But at a minimum, the current results suggest the utility of rethinking the conventional wisdom about what youth employment programs can (or can't) do.

---

[22] It is not impossible to tell a story where youth know that the probability of being caught for a violent crime is higher than other types of crime, so the increased opportunity cost of punishment affects violence differentially. But if better employment prospects are increasing the opportunity cost of crime, we would probably expect employment to increase in the long-run. We do not have enough data to assess long-run employment. But to the extent that the New York evidence generalizes to our setting, it does not seem that summer jobs programming increases employment, or at least reported income, over the following 4 years.
[23] This may be part of the reason that the two treatment arms had similar effects; mentors and employers may have taught many of the same lessons in the curriculum.

**Figure 1: Map of OSC+ Applicants by Community Area Violent Crime Rate**

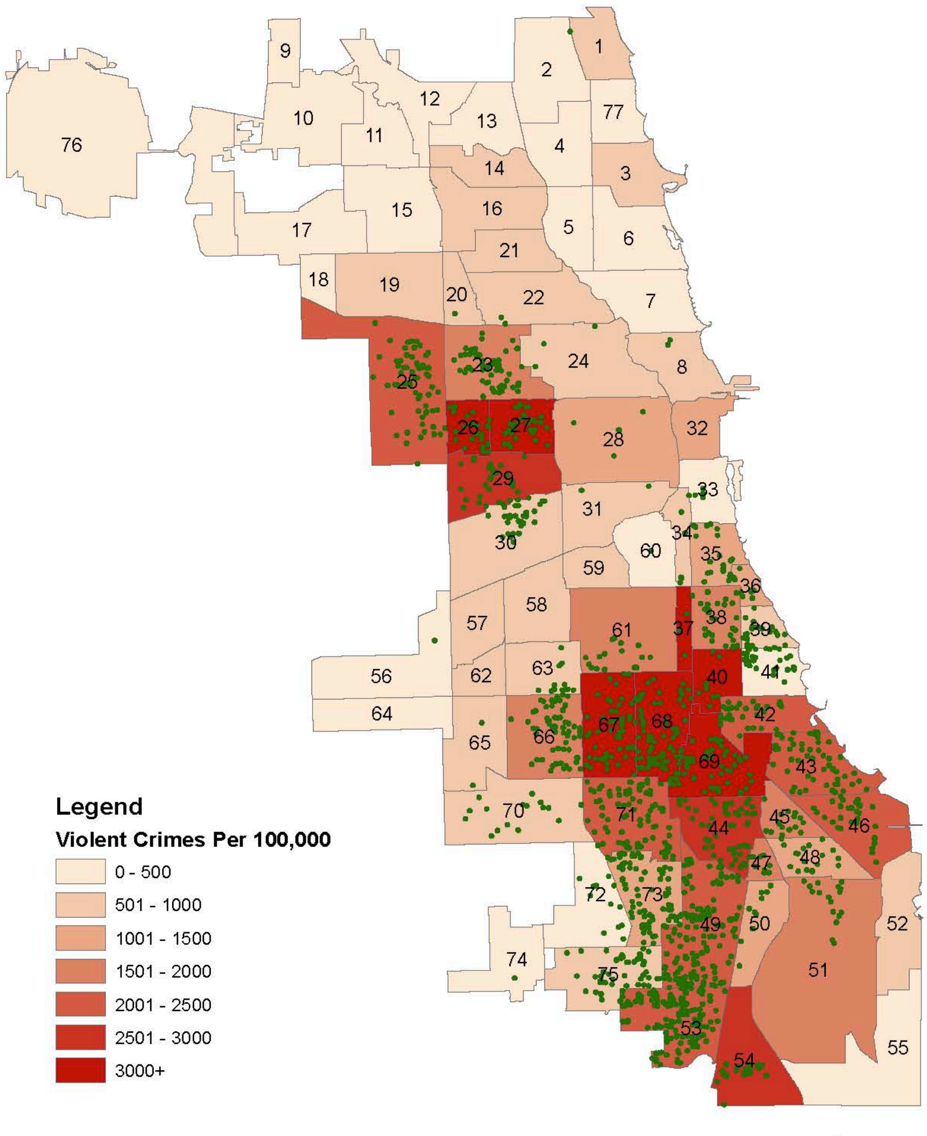

**Table 1. Descriptive Statistics**

| | Control Mean | Coefficient on Treatment | Standard Error |
|---|---|---|---|
| **Demographics** | | | |
| Age | 16.30 | -0.05 | (0.07) |
| Black | 0.96 | 0.00 | (0.01) |
| Hispanic | 0.03 | 0.00 | (0.01) |
| **Arrests** | | | |
| Ever Arrested | 0.19 | 0.02 | (0.02) |
| Number of Arrests for: | | | |
| Violent Crime | 0.13 | 0.04 | (0.03) |
| Property Crime | 0.09 | -0.01 | (0.02) |
| Drug Crime | 0.05 | 0.01 | (0.02) |
| Other Crime | 0.15 | 0.02 | (0.03) |
| **Schooling** | | | |
| Grade | 10.15 | -0.04 | (0.06) |
| Days Attended[a] | 136.92 | 0.70 | (1.40) |
| Free/Reduced Price Lunch | 0.92 | 0.00 | (0.01) |
| **Employment and Earnings[b]** | | | |
| Had Valid Social Security Number | 0.77 | 0.03 | (0.02) |
| Any Earnings in Year Before Program | 0.08 | -0.02 | (0.01) |
| Total Earnings in Year Before Program | 265.41 | -154.70 | (123.59) |
| **Census Block Group Characteristics[c]** | | | |
| Unemployment Rate | 19.07 | -0.03 | (0.42) |
| Share Below Poverty Line | 33.26 | 0.28 | (0.73) |
| Share of Households with High School Credential | 72.91 | -0.85 | (0.79) |
| Share of Population: Black | 87.18 | 0.22 | (0.92) |

Notes: ***p<0.01, ** p<0.05, *p<0.1. The coefficient on treatment status is from a regression of the specified baseline covariate on an indicator for treatment controlling for block fixed effects with heteroskedasticity-robust standard errors. Demographic data is from the program application, CPS, and CPD. Academic data is from CPS. Arrest data is from CPD.

a. The pre-program school year was about 170 days.

b. The first measure in this panel shows the share of the full dataset with a valid Social Security Number. The rest of the measures in this panel are conditional on having a valid SSN (n = 1,280). An SSN is "valid" if it was submitted to IDES and deemed to be associated with the correct person by CPS.

c. Census tract characteristics are from 2009-2013 ACS Census Block Group characteristics.

**Table 2. Employment and College Enrollment Program Effects**

| Outcome: | Any Formal Employment | Average Quarterly Earnings | Any College Enrollment |
|---|---|---|---|
| ITT | -0.02 | -29.08 | -0.02 |
| | (0.02) | (36.75) | (0.02) |
| CM | 0.13 | 150.91 | 0.40 |
| | | | |
| LATE | -0.02 | -39.08 | -0.02 |
| | (0.02) | (48.15) | (0.03) |
| CCM | 0.13 | 177.59 | 0.40 |
| N | 1280 | 1280 | 1083 |

Notes: ***p<0.01, ** p<0.05, *p<0.1. Employment data from Illinois Unemployment Insurance quarterly wage records. Only youth who had a social security number available in the Chicago Public Schools records are included (see text). Employment outcomes for one post-program year exclude the quarters of the program, so include 2012Q4 and 2013Q1. College enrollment from National Student Clearinghouse data. The sample is restricted to youth who were at least in 10th grade in the 2011-12 school year. Heteroskedasticity-robust standard errors in parentheses. CM is control mean; CCM is control complier mean.

**Table 3. Program Effect on Number of Arrests per 100 Youth**

| Crime: | Total | Violent | Property | Drugs | Other |
|---|---|---|---|---|---|
| | | | Year 1 | | |
| ITT | -1.698 | -3.334** | 0.735 | 0.352 | 0.548 |
| | (3.540) | (1.513) | (0.943) | (1.523) | (1.876) |
| CM | 21.792 | 7.19 | 2.655 | 3.65 | 8.296 |
| LATE | -2.285 | -4.487** | 0.989 | 0.474 | 0.738 |
| | (4.671) | (2.000) | (1.243) | (2.009) | (2.475) |
| CCM | 28.94 | 10.002 | 3.422 | 4.305 | 11.211 |
| | | | Year 2 | | |
| ITT | -2.805 | 0.173 | 0.575 | -1.733 | -1.82 |
| | (3.111) | (1.197) | (0.956) | (1.228) | (1.906) |
| CM | 19.801 | 3.761 | 2.544 | 4.978 | 8.518 |
| LATE | -3.775 | 0.233 | 0.774 | -2.332 | -2.45 |
| | (4.102) | (1.580) | (1.262) | (1.621) | (2.513) |
| CCM | 25.282 | 4.546 | 2.535 | 7.296 | 10.906 |
| | | | Cumulative (24 months) | | |
| ITT | -4.503 | -3.161 | 1.31 | -1.38 | -1.272 |
| | (5.581) | (2.001) | (1.406) | (2.231) | (3.063) |
| CM | 41.593 | 10.951 | 5.199 | 8.628 | 16.814 |
| LATE | -6.06 | -4.254 | 1.764 | -1.858 | -1.712 |
| | (7.360) | (2.641) | (1.856) | (2.943) | (4.040) |
| CCM | 54.222 | 14.548 | 5.957 | 11.601 | 22.117 |

Notes: n = 1,634, ***p<0.01, ** p<0.05, *p<0.1. Outcomes are number of arrests per 100 youth as recorded in Chicago Police Department arrest records. Top panels show year-by-year effects; bottom panel shows cumulative effect across both years. Heteroskedasticity-robust standard errors in parentheses. CM is control mean: CCM is control complier mean.

**Table 4. Estimated Social Savings per Participant from Crime Reduction**

| | (1) | | (2) | | (3) | | (4) | |
|---|---|---|---|---|---|---|---|---|
| Year One | -17146** | -6815* | -16818** | -6692* | -3426 | -1342 | -3616* | -1532 |
| | (8536) | (3858) | (8371) | (3786) | (2112) | (993) | (2105) | (982) |
| Year Two | 4715 | 1940 | 4631 | 1911 | 1085 | 348 | 1263 | 526 |
| | (6319) | (2638) | (6188) | (2587) | (1329) | (753) | (1307) | (716) |
| Cumulative | -12431 | -4875 | -12187 | -4781 | -2340 | -994 | -2352 | -1006 |
| | (10624) | (4778) | (10413) | (4687) | (2513) | (1281) | (2494) | (1247) |
| Source of cost estimates | CV | Direct | CV | Direct | CV | Direct | CV | Direct |
| Includes: | | | | | | | | |
| Collateral costs of incarceration | X | | | | | | | |
| Adjustment for crimes per arrest | X | | X | | | | | |
| Social cost of drug use | X | | X | | X | | | |

Notes: n = 1,634, ***p<0.01, ** p<0.05, *p<0.1. Heteroskedasticity-robust standard errors in parentheses. All estimates in 2012 dollars. Columns use different social costs of crime. Homicide trimmed to cost of aggravated assault. Columns using "direct" costs based on estimates from Cohen & Piquero (2009); columns using CV based on contingent valuation estimates from Cohen (2001). See text and Appendix A for details. Columns without social cost of drug use only count direct cost of drug arrests to the criminal justice system. Benefits are discounted monthly at a 5% annual rate beginning in mid-September following the program.

## Works Cited

Aizer, Anna, and Joseph J Doyle Jr, "Juvenile incarceration, human capital and future crime: Evidence from randomly-assigned judges," (National Bureau of Economic Research, 2013).

Anderson, ML, "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," Journal of the American Statistical Association, 103 (2008), 1481-1495.

Becker, Gary S, "Crime and Punishment: An Economic Approach," The Journal of Political Economy, 76 (1968), 169-217.

Belfield, Clive R, Milagros Nores, Steve Barnett, and Lawrence Schweinhart, "The High/Scope Perry Preschool Program Cost-Benefit Analysis Using Data from the Age-40 Followup," Journal of Human Resources, 41 (2006), 162-190.

Bender, Ralf, and Stefan Lange, "Adjusting for multiple testing—when and how?," Journal of Clinical Epidemiology, 54 (2001), 343-349.

Benjamini, Y., and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," Journal of the Royal Statistical Society. Series B (Methodological), (1995), 289-300.

Benjamini, Y., A.M. Krieger, and D. Yekutieli, "Adaptive linear step-up procedures that control the false discovery rate," Biometrika, 93 (2006), 491-507.

Bloom, H.S., L.L. Orr, S.H. Bell, G. Cave, F. Doolittle, W. Lin, and J.M. Bos, "The benefits and costs of JTPA Title II-A programs: Key findings from the National Job Training Partnership Act study," Journal of Human Resources, (1997), 549-576.

Bonczar, Thomas P., "Prevalence of Imprisonment in the U.S. Population, 1974-2001," (U.S. Department of Justice, Bureau of Justice Statistics, 2003).

Cave, George, Hans Bos, Fred Doolittle, and Cyril Toussaint "JOBSTART. Final Report on a Program for School Dropouts," (MDRC, 1993).

Center for Disease Control and Prevention, "Web-based Injury Statistics Query and Reporting System (WISQARS)," (2014).

Chalfin, Aaron, "The Economic Cost of Crime," (University of Cincinnati, 2013).

Cohen, Mark A, and Alex R Piquero, "New evidence on the monetary value of saving a high risk youth," Journal of Quantitative Criminology, 25 (2009), 25-49.

Cohen, Mark, Roland Rust, Sara Steen, and Simon Tidd, "Willingness to pay for crime control programs.," Criminology, 42 (2004), 86-106.

Council of Economic Advisors, "Economic Costs of Youth Disadvantage and High-Return Opportunities for Change," (Washington, DC, 2015).

Farkas, George, D. Alton Smith, and Ernst W. Stromsdorfer, "The youth entitlement demonstration: Subsidized employment with a schooling requirement," Journal of Human Resources, 18 (1983), 557-573.

Farrington, David P., "Age and crime," Crime and Justice, 7 (1986), 189-250.

Federal Bureau of Investigation, "Crime in the United States 2012," in *Uniform Crime Report*, (Washington DC: U.S. Department of Justice, 2012).

---, "Crime in the United States 2014," in *Uniform Crime Report*, (Washington DC: U.S. Department of Justice, 2014).

Fernandes-Alcantara, Adrienne L., "Vulnerable Youth: Federal Funding for Summer Job Training and Employment," (Congressional Research Service, 2011).

Gelber, Alexander, Adam Isen, and Judd Kessler, "The Effect of Youth Employment on Future Earnings: Evidence from Summer Youth Employment Program Lotteries," NBER Working Paper 20810, (2014).

Gelman, Andrew, Jennifer Hill, and Masanao Yajima, "Why we (usually) don't have to worry about multiple comparisons," Journal of Research on Educational Effectiveness, 5 (2012), 189-211.

Goffman, Alice, *On the run: Fugitive life in an American city* (Macmillan, 2015).

Grossman, Jean Baldwin, and Cynthia L. Sipe, "Summer Training and Education Program (STEP): Report on long-term impacts," (Philadelphia PA: Public/Private Ventures, 1992).

Healey, Kaleen, Jenny Nagaoka, and Valerie Michelman, "The Educational Attainment of Chicago Public Schools Students," Consortium on Chicago School Research, ed. (2014).

Heckman, James J, Seong Hyeok Moon, Rodrigo Pinto, Peter A Savelyev, and Adam Yavitz, "The rate of return to the HighScope Perry Preschool Program," Journal of Public Economics, 94 (2010), 114-128.

Heckman, James J., "Skill Formation and the Economics of Investing in Disadvantaged Children," Science, 312 (2006), 3p.

Heinrich, Carolyn J., and Harry J. Holzer, "Improving education and employment for disadvantaged young men: Proven and promising strategies," Annals of the American Academy of Political and Social Science, 635 (2011), 163.

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

Heller, Sara B, Anuj K Shah, Jonathan Guryan, Jens Ludwig, Sendhil Mullainathan, and Harold A Pollack, "Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago," (National Bureau of Economic Research, 2015).

Jacob, Brian A., and Lars Lefgren, "Are Idle Hands the Devil's Workshop? Incapacitation, Concentration, and Juvenile Crime," American Economic Review, 93 (2003), 1560-1577.

Katz, L.F., J.R. Kling, and J.B. Liebman, "Moving to opportunity in Boston: Early results of a randomized mobility experiment," Quarterly Journal of Economics, 116 (2001), 607-654.

Kemple, J.J., C.J. Willner, and MDRC, *Career academies: Long-term impacts on labor market outcomes, educational attainment, and transitions to adulthood* (MDRC New York, 2008).

Kling, Jeffrey R., Jens Ludwig, and Lawrence F. Katz, "Neighborhood effects on crime for female and male Youth: Evidence from a randomized housing mobility experiment," Quarterly Journal of Economics, 120 (2005), 87-130.

LaLonde, Robert J., "The promise of public sector-sponsored training programs," Journal of Economic Perspectives, 9 (1995), 149-168.

---, "Employment and training programs," in *Means-tested transfer programs in the United States*, Robert A. Moffitt, ed. (Chicago, IL: University of Chicago Press, 2003).

Lee, David S, and Justin McCrary, "Crime, punishment, and myopia," (National Bureau of Economic Research Working Paper 11491, 2005).

Leos-Urbel, Jacob, "What Is a Summer Job Worth? The Impact of Summer Youth Employment on Academic Outcomes," Journal of Policy Analysis and Management, 33 (2014), 891-911.

Leos‐Urbel, Jacob, "What is a Summer Job Worth? The Impact of Summer Youth Employment on Academic Outcomes," Journal of Policy Analysis and Management, 33 (2014), 891-911.

Levitt, Steven D, "The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation," The Quarterly Journal of Economics, (1996), 319-351.

McClanahan, Wendy S., Cynthia L. Sipe, and Thomas J. Smith, "Enriching Summer Work: An Evaluation of the Summer Career Exploration Program," (Philadelphia, PA: Public/Private Ventures, 2004).

McConnell, Sheena, and Steven Glazerman, "National Job Corps Study: The Benefits and Costs of Job Corps," (Washington, DC: US Department of Labor, Employment and Training Administration, 2001).

Millenky, M., D. Bloom, S. Muller-Ravett, and J. Broadus, "Staying on Course: Three-Year Results of the National Guard Youth ChalleNGe Evaluation ", (MDRC, 2011).

Miller, Ted R., Mark A. Cohen, and Brian Wiersema, "Victim costs and consequences: A new look," in *National Institute of Justice Research Report*, (U.S. Deptartment of Justice, National Institute of Justice, 1996).

Moffitt, Terrie E., "Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy," Psychological Review, 100 (1993), 674-701.

Mueller-Smith, Michael, "The Criminal and Labor Market Impacts of Incarceration," (2014).

National Student Clearinghouse, "NSC Fact Sheet," (2015).

Roder, Anne, and Mark Elliott, "A Promising Start: Year Up's Initial Impacts on Low-Income Young Adults' Careers," (New York: Economic Mobility Corporation, 2011).

Schochet, PZ, J Burghardt, and S McConnell, "Does Job Corps work? Impact findings from the National Job Corps Study," American Economic Review, 98 (2008), 1864-1886.

Schwartz, Amy Ellen, Jacob Leos-Urbel, and Matthew Wiswall, "Making Summer Matter: The Impact of Youth Employment on Academic Performance," (National Bureau of Economic Research w21470, 2015).

Somers, Gerald G., and Ernst W. Stromsdorfer, "A cost-effectiveness analysis of in-school and summer Neighborhood Youth Corps: a nationwide evaluation," Journal of Human Resources, 7 (1972), 446-459.

Sum, Andrew, Ishwar Khatiwada, Mykhaylo Trubskyy, and Sheila Palma, "The plummeting labor market fortunes of teens and young adults," Washington, DC: The Brookings Institution, (2014).

Sum, Andrew, Mykhaylo Trubskyy, and Walter McHugh, "The Summer Employment Experiences and the Personal/Social Behaviors of Youth Violence Prevention Employment Program Participants and those of a Comparison Group," (Boston, MA: Center for Labor Market Studies, Northeastern University, 2013).

Uggen, Christopher, "Work as a turning point in the life course of criminals: A duration model of age, employment, and recidivism," American sociological review, (2000), 529-546.

Venkatesh, Sudhir Alladi, *Off the books* (Harvard University Press, 2006).

Walker, Gary, and Frances Vilella-Velez, "Anatomy of a Demonstration: The Summer Training and Education Program (STEP) from Pilot through Replication and Postprogram Impacts," (Philadelphia, PA: Public/Private Ventures, 1992).

Westfall, P.H., and S.S. Young, *Resampling-based multiple testing: Examples and methods for p-value adjustment* (Wiley-Interscience, 1993).

Wilson, William J, "When work disappears: The world of the urban poor," NY: Alfred Knopf, (1996).

## Appendix A
## Benefit-Cost Comparison Details

Table 5 reports a variety of estimates for the social savings from reduced crime. As discussed in the main text, assigning costs to crime is an inherently uncertain and difficult exercise. This appendix outlines our approach to dealing with various sources of uncertainty and explains how we form our social cost estimates.

*Source of Social Cost Estimates*

The cost of crime to society comes in many parts – harm to victims (which include direct costs like lost property or medical costs as well as indirect costs like harm and suffering, or fear and behavioral changes to avoid crime), costs to the criminal justice system (police, courts, and incarceration), and costs to the offender (lost productivity and any collateral costs of arrests and incarceration on earnings, future crime, and family).

There are two basic approaches in the literature to estimating these costs: "bottom up" and "top down." The bottom up approach focuses mostly on direct costs, combining evidence from jury awards, the costs of medical care, lost wages, and other relatively observable costs of being a victim of crime. The most widely cited estimates using this approach come from Miller, Cohen, and Wiersema (1996); we use an updated and slightly expanded version of these estimates from Cohen and Piquero (2009). The Cohen and Piquero update includes costs to the criminal justice system and approximates lost offender productivity for the small proportion of crimes that end in incarceration.

The "top down" approach includes more indirect costs like fear and behavioral changes by soliciting willingness-to-pay (wtp) for crime avoidance using contingent valuation. Conceptually, this approach may capture more of the relevant costs, but it also suffers from the typical problems of obtaining true wtp measures through survey questions. Since both top down and bottom up approaches have strengths and weaknesses, we show estimates using both versions. Our top down estimates come from Cohen et al. (2004). We transform all dollar values into 2012 dollars using the Bureau of Labor Statistics' Consumer Price Index, and we discount the costs based on the time of arrest relative to the end of the program (using a monthly discounting that translates into a 5% annual discount rate).

One challenge that all cost-of-crime techniques face is in assigning a statistical value of a life to fatal crimes (homicide). In practice, these costs are so large as to swamp all other crimes. This is a particular problem in finite data sets where homicide is rare, as in our data. We simply do not have the power to identify a program effect on homicide. As such, if we assigned the statistical value of a life to these incidents, we would be capitalizing on what is effectively chance in our cost estimates (whether treatment or control youth happen to have one or two more of a hugely costly outcome). To avoid this problem, we assign homicide charges the cost of an aggravated assault. This may not accurately capture the true social cost of a homicide, but it prevents our cost estimates from being dramatically swayed by an extremely rare outcome for which we lack power to estimate program effects.

*Arrests versus Crimes*

We measure arrests, but it is well established that only a fraction of crimes committed result in arrest (e.g., Federal Bureau of Investigation 2012). If what we care about is the social cost of crime, we want to assign costs to all crimes, not just arrests. The common approach in the literature is to assume that crime changes in proportion to observed arrests, and multiply each arrest by an estimate of crimes-per-arrest. For example, both oft-cited Perry Preschool benefit-cost analyses take this approach (Belfield, et al. 2006; Heckman, et al. 2010), as do other economics of crime and cost of crime papers such as Levitt (1996) and Cohen and Piquero (2009). We use the incidence-to-arrest ratios from Cohen and Piquero (column 1 of Table 1 for arrests while under 18 and the more conservative version for adults, column 3 of Table 1).

For the bottom-up estimates, we only multiply the victim costs by these scaling factors, since the costs to the criminal justice system and to offenders are only incurred when someone is actually arrested. The scaling is a little trickier for the top-down estimates, since wtp does not separate criminal justice costs from victim costs. For simplicity, we assume that people's willingness-to-pay for the criminal justice and lost offender productivity costs are the same as in the bottom-up estimates. We then take the remaining difference between top-down and bottom-up cost estimates as the victim costs and multiply that difference by the scaling factor. If people value criminal justice costs or the opportunity cost of offender time more than is reflected in the bottom-up estimates, this approach may slightly overstate the victim costs.

*Which Costs are Included*

Most cost-of-crime estimates focus on the costs to victims and the criminal justice system. In theory, there are also collateral costs to the offender: being arrested may make it harder to find a job in the future, or have a causal impact on families and future crime (we usually imagine those effects would be negative, though it is an empirical question). Yet estimating these costs is made difficult by the scarcity of causal evidence on how crime and incarceration affects these other collateral outcomes.

We attempt to incorporate a few of these costs, though we recognize that there is mixed evidence on how juvenile crime and arrests affect these outcomes. As such, we consider this a suggestive exercise rather than a compelling argument. We limit ourselves to outcomes where there is some causal evidence relevant to our population. For example, Aizer and Doyle (2013) estimate the effect of juvenile incarceration – at the same detention center in Cook County that our study youth would be detained in – on the probability of adult incarceration by age 25. We use their estimates to approximate how much an arrest during the study period is likely to increase future incarceration.[24] We then use Mueller-Smith's (2014) estimates of how adult incarceration affects future wages to project how much this increased adult incarceration decreases future earnings.

To estimate these collateral costs of incarceration, we first need to estimate how much an arrest for a particular type of crime (which we observe) changes the probability of incarceration for future offenses (which we do not observe). Aizer and Doyle (2013) report two pieces of the

[24] They also estimate the effect of juvenile detention on high school graduation. But since we will be able to measure that outcome directly, we do not incorporate those estimates.

puzzle: what proportion of youth who are *charged* in juvenile court with various crime types end up in juvenile detention[25], and how much juvenile detention increases the probability of adult incarceration by age 25. By combining Aizer and Doyle's detention rates by charge with the fact that about 15 percent of juvenile arrests result in juvenile detention,[26] we estimate that 65% of arrests result in charges. These three proportions (charges per arrest, detentions per charge, and adult incarceration per detention) give us a crime-specific estimate of how much adult incarceration will increase from each arrest.

We then assign costs to future incarceration. Aizer and Doyle's measure is a 0/1 indicator for any adult incarceration by age 25, which does not capture multiple incarcerations, the reason for incarceration, or the timing of the incident (which matters for discounting the costs). We take a fairly conservative approach: we assume the increase in adult incarceration corresponds to a single arrest, for a similar type of crime as the observed juvenile arrest, 10 years in the future. We then use the same social costs of crime that we use for arrests during the study period, discounted at a 5% annual rate (over 10 years). This is a bit of an over-simplification, but one that we believe conservatively leans towards under-estimating costs.

We also calculate the impact of this additional adult incarceration on later earnings using Mueller-Smith (2014). The immediate loss of income during incarceration is already incorporated in the bottom up cost of crime estimates, so we use the Mueller-Smith estimates to approximate post-incarceration earnings loss. He estimates that each year of incarceration reduces future earnings by about $2,000 per year over the following 5 years. Lee and McCrary (2005) show that the average adult duration of incarceration is 13.09 weeks, or about a quarter of a year. So we extrapolate that an average adult incarceration episode will reduce offender earnings by $500 per year for each of five years. Since we are assuming that any adult incarceration occurs 10 years in the future, we also assume this income would have been earned from 10 to 15 years in the future and discount it at a 5% annual rate.

As Table 5 makes clear, adding these collateral costs does not make much of a difference to our cost-of-crime estimates. This is in part because the probability of future incarceration conditional on arrest is quite low. But it is also because we make extremely conservative assumptions in this process and do not include many potential costs outside of incarceration and earnings. In reality, we expect that the collateral costs are much higher than is suggested by our table. Future work will assess how much different assumptions matter.

We also vary how we treat the social costs of drug use. Most of the cost-of-crime literature excludes drug crime altogether, in part because it is often seen as a "victimless" crime (Chalfin 2013). This is not to say that drug use cannot have social costs (e.g., for government service provision to addicts, family stress and lost productivity, etc.). But much of the literature assigns zero cost to these types of offenses. We follow the Moving to Opportunity approach (Kling, Ludwig & Katz 2005) by showing one version that includes estimates of the social costs of drug arrests (from the MTO study) and another that sets the victim costs to zero (in that case, we do still assign criminal justice costs to drug arrests).

---

[25] They do not report these proportions directly, but we calculate them from the summary statistics.
[26] http://www.steansfamilyfoundation.org/pdf/Juvenile_Justice_in_Illinois.pdf

## Appendix B
## Main Results with Alternative Functional Forms

This appendix shows results using alternative functional forms. In order to ensure convergence, we omit baseline covariates other than the randomization block fixed effects. This makes the results slightly less precise, but not substantively different from the results reported in the main text.

Table B1 uses a logit for the regressions with a dichotomous dependent variable, reporting both the logit coefficient and the average marginal effects (which are more directly comparable to the ordinary least squares results). The number of observations is slightly lower than in the main tables because some of the block fixed effects perfectly predict the outcome, so those observations are dropped. Table B2 uses Poisson regression (with robust standard errors to relax the constraint that the mean equals the variance) to estimate the intent-to-treat effect on the counts of arrests. It also shows average marginal effects, as well as the ordinary least squares ITT with no baseline covariates for comparison. All results are very similar across the two estimation techniques.

### Table B1. Logistic Regression Estimates of Employment and College Program Effects

**Logistic Regression Estimates of Program Effects Outside of High School**

| Outcome: | Any Formal Employment | Any College Enrollment |
|---|---|---|
| Estimate | -0.18 | -0.15 |
| | (0.21) | (0.18) |
| AME | -0.02 | -0.02 |
| | 1265 | 1020 |

Notes: ***p<0.01, ** p<0.05, *p<0.1. Employment data from Illinois Unemployment Insurance quarterly wage records. Only youth who had a social security number available in the Chicago Public Schools records are included (see text). Employment outcomes for one post-program year exclude the quarters of the program, so include 2012Q4 and 2013Q1. College enrollment from National Student Clearinghouse data. The sample is restricted to youth who were at least in 10th grade in the 2011-12 school year. Heteroskedasticity-robust standard errors in parentheses.

**Table B.2 Poisson ITT Effects on Number of Arrests per 100 Participants**

| Crime: | Total | Violent | Property | Drugs | Other |
|---|---|---|---|---|---|
| | | Year 1 | | | |
| Poisson Estimate | 0.052 | -0.453* | 0.294 | 0.209 | 0.209 |
| | (0.160) | (0.237) | (0.281) | (0.339) | (0.199) |
| AME | 1.260 | -2.784 | 0.955 | 0.943 | 2.091 |
| ITT w/ no X | 1.300 | -2.833* | 0.977 | 0.983 | 2.172 |
| | (4.035) | (1.538) | (0.958) | (1.679) | (2.119) |
| | | Year 2 | | | |
| Estimate | -0.042 | 0.125 | 0.303 | -0.291 | -0.100 |
| | (0.153) | (0.282) | (0.299) | (0.249) | (0.215) |
| AME | -0.894 | 0.523 | 0.962 | -1.438 | -0.903 |
| ITT w/ no X | -0.926 | 0.531 | 0.989 | -1.507 | -0.940 |
| | (3.399) | (1.202) | (0.989) | (1.314) | (2.032) |
| | | Cumulative (24 months) | | | |
| Estimate | 0.008 | -0.218 | 0.299 | -0.053 | 0.062 |
| | (0.135) | (0.192) | (0.214) | (0.244) | (0.171) |
| AME | 0.362 | -2.264 | 1.917 | -0.500 | 1.186 |
| ITT w/ no X | 0.374 | -2.302 | 1.966 | -0.523 | 1.233 |
| | (6.414) | (2.059) | (1.441) | (2.431) | (3.446) |

Notes: n = 1,634, ***p<0.01, ** p<0.05, *p<0.1. AME shows average marginal effects. Outcomes are arrests per 100 youth per month as recorded in Chicago Police Department arrest records. Top panels show year-by-year effects; bottom panel shows cumulative effect across both years. Unlike other results, these results do not include baseline covariates in order to ensure convergence. ITT estimates without additional controls are shown for comparison (but all regressions include randomization block fixed effects). Heteroskedasticity-robust standard errors in parentheses.