YOUTH POLICY LAB
UNIVERSITY OF MICHIGAN

JANUARY 2020

GROW DETROIT'S YOUNG TALENT

AUTHORS:

Brian Jacob

Jeremy Guardiola

Cameron Conrad

KEY FINDINGS:

1. In the years following participation, GDYT youth from the 2015, 2016, and 2017 cohorts are somewhat more likely to be enrolled in school, take the SAT, and graduate from high school compared to non-participating applicants; they are also less likely to be chronically absent. Though modest in size, these differences are statistically significant.

2. Three-quarters of 2017 GDYT youth completed their work commitments by working at least 102 hours. The remaining 25% of youth exhibit higher rates of baseline chronic absenteeism and are more likely to be older and male. Note that youth may not complete for positive reasons (e.g. starting college, participating in sports camp).

3. Youth who complete GDYT outperform those who work but do not complete across our academic outcomes. We must interpret these results carefully, however, as we know youth who do and do not complete their work commitments differ in a number of important ways before they ever participate in GDYT (see Finding #2).

INTRODUCTION

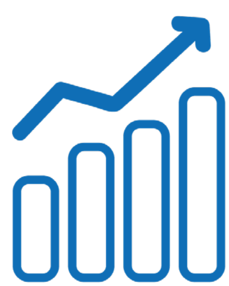
Grow Detroit's Young Talent (GDYT), Detroit's citywide summer youth employment program, serves as an introduction to the workplace for thousands of youth across Detroit every year. By providing youth the opportunity to cultivate professional skills and gain job experience, GDYT aims to both set individual participants on stable career pathways and propel the broader city's continued economic resurgence.

Youth must be between the ages of 14 and 24 to participate. Depending on their age and prior work experience, successful applicants work in jobs that fall within one of the program's three tiers: Career Exploration, Ready for Work, and Career Pathways Internships. Selected applicants may work up to 120 hours over the summer. GDYT considers youth who work at least 85% of the maximum 120 hours (i.e., 102 hours) to have completed their work commitments.

In April 2018, the Youth Policy Lab (YPL) published findings from our initial study of GDYT. We found that GDYT youth from the 2015 cohort modestly outperformed non-participating applicants during the two years after their summer employment: higher school enrollment (1.6%), lower rates of chronic absenteeism (3%), greater probability of taking the SAT (5.2%), and increased likelihood of graduating high school (4.1%).[i]

The results in this brief build upon the work we began with the 2015 cohort. We now have enough post-participation data to study multiple cohorts and data for the 2017 group is much more detailed relative to previous years. Whereas we only know whether or not 2015 and 2016 applicants ultimately worked in GDYT jobs, we can see exactly how many hours 2017 participants actually worked. This allows us to ask more nuanced questions of the data, such as whether or not youth outcomes vary by level of program engagement.

"GDYT aims to both set individual participants on stable career pathways and propel the broader city's continued economic resurgence."



METHODS

In order to assess GDYT participants' academic outcomes relative to what one might expect had they not participated in the program, we compare participants to peers who applied to GDYT but did not ultimately participate. While we cannot distinguish youth who were not selected to participate from those who were but declined the opportunity, by comparing participants to non-participating applicants we ensure that all youth we study expressed an initial interest in the program and completed all application steps. To further ensure we are making "apples-to-apples" comparisons, we use statistical methods to match GDYT youth to comparison youth on the basis of grade, gender, race, and school attended. For example, we compare a Black, female, 9th-grade student who attended High School X the year she applied to GDYT to other Black, female, 9th grade students who attended High School X that same year. Finally, our statistical models account for age, neighborhood characteristics, prior academic achievement, and prior school attendance via multiple regression. See Appendix B for more details.

Despite these efforts to make sound comparisons between similar youth, we cannot say that the differences we observe between GDYT youth and non-participating applicants are *caused by* the program. Because invitations into the program are not randomly assigned, outcome differences may instead be attributable to other characteristics that distinguish GDYT youth from non-participating applicants. In fact, our data show that before they even apply, non-participating applicants are more likely to be chronically absent from school relative to program participants. Moreover, GDYT youth in some cohorts exhibit slightly better levels of prior academic achievement as measured by 8th-grade standardized test scores. Fundamental differences such as these mean we should interpret post-program differences with caution.

FINDINGS

How consistent are post-GDYT academic outcomes?

With limited data available during our initial analysis of GDYT, we focused our attention on the 2015 summer cohort. We found that in the two academic years following their participation in GDYT, youth were somewhat more likely to remain enrolled in school, take the SAT, and graduate high school. Conversely, they were marginally less likely to be chronically absent. Now that we have additional post-participation data for the 2016 and 2017 cohorts, we can assess whether those findings hold true for other cohorts.

When we examine the 2016 and 2017 GDYT cohorts individually, we find results that are similar to 2015. Participants typically outperform non-participating applicants, even after controlling for demographic, neighborhood characteristics, and prior academic achievement.

Since we find that outcomes are broadly similar across groups, we pool data from all three cohorts: three years of post-participation data for the 2015 cohort, two years for 2016, and one year for 2017.[ii] This allows us to estimate an overall program impact that is not specific to any one cohort.

Table 1 shows results from this approach. We see that GDYT youth generally perform slightly better on all four measures of school performance compared to non-participating applicants. Note that we have also included statistics on Detroit youth who did not apply to GDYT but who attended the same high school, were the same grade level, and shared the same race/ethnicity and gender as GDYT applicants for reference. The fact that both GDYT youth and non-participating applicants both outperform comparison youth suggests that youth who choose to apply to the program are not representative of the broader population. This is why we focus our analyses on the set of youth who applied to GDYT.

TABLE 1 – GDYT youth from the 2015, 2016, and 2017 cohorts modestly outperform non-participating applicants across a range of academic outcomes

OUTCOME	COMPARISON DETROIT YOUTH	NON-PARTICIPATING APPLICANTS	GDYT YOUTH	DIFFERENCE (RAW)	DIFFERENCE (PERCENT)
Enrolled in school	91.5%	93.6%	94.4%	0.8*	0.9%*
Chronically Absent	41.9%	42.9%	41.4%	-1.5*	-3.5%*
Took SAT	63.5%	66.9%	70.1%	3.2*	4.8%*
Graduated HS	77.6%	80.8%	83.8%	3.0*	3.7%*

* Indicates statistically significant difference between non-participating applicants and GDYT youth.

To best interpret our findings, it is helpful to consider the raw differences between GDYT youth and non-participating applicants in relation to both groups' total rate in each outcome. For example, the 3.2-percentage point lead GDYT youth exhibit in SAT test-taking translates to a 4.8% increase in probability of taking the exam.[iii] This is the largest difference we observe in both absolute (raw magnitude) and relative (participant rate as a percentage of non-participating applicant rate) terms across all four outcomes. So although we find statistically significant differences between the two groups, their academic outcomes are relatively similar. That said, we must note that even small differences in educational outcomes are encouraging. Every additional youth who remains enrolled in school and graduates brings us one step closer to our goal of setting all Detroit's young people on paths to economic prosperity.

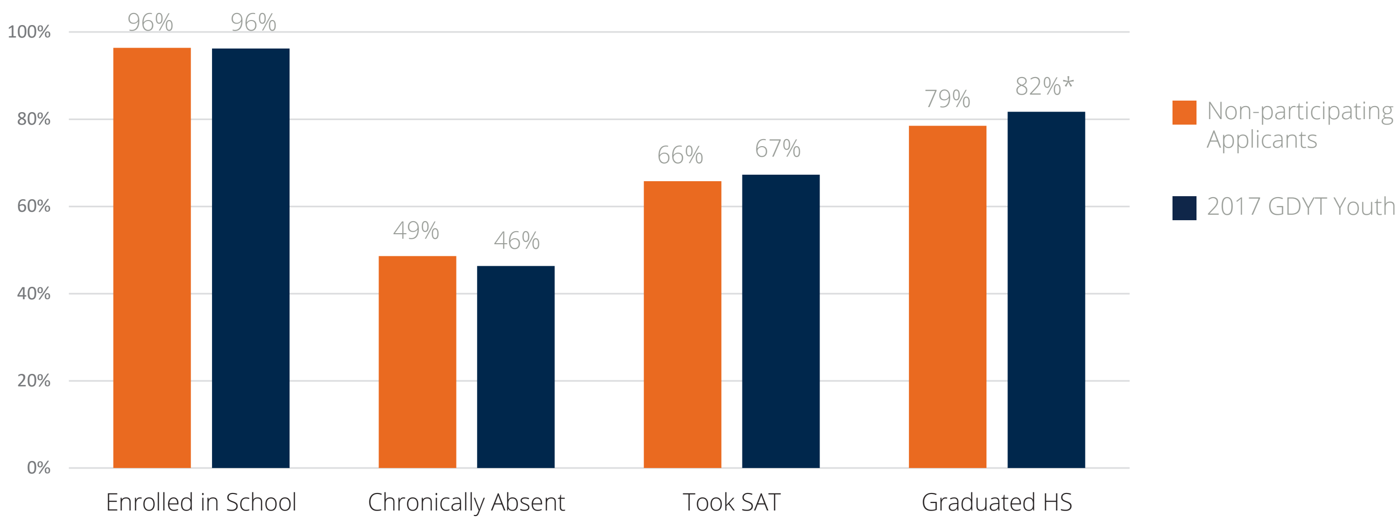
It is also important to acknowledge that our analysis does not capture all the ways GDYT might impact participants. For example, our data do not include measures of work readiness like hard or soft skill development. And although we might hope that summer employment would correlate with improved educational experiences, we might expect outcomes like employment and earnings to be more relevant. We plan to examine some of these workforce outcomes in future work.

Focusing on 2017

The data available for the 2017 cohort are more detailed than those we have for either 2015 or 2016. In addition to the type of employer or provider youth were placed with, we know the number of hours youth worked in their summer jobs. This means that for the 2017 cohort we can employ a more nuanced measure of program participation and analyze outcomes according to the number of hours youth worked. Before we discuss any such subgroup outcomes, however, it is helpful to establish some baseline parameters about the cohort as a whole.

Figure 1 shows the difference between 2017 GDYT youth and non-participating applicants across our four key outcome measures. Overall, we see that these results are largely similar to those we found after combining all three cohorts. GDYT youth from 2017 are slightly more likely to graduate high school and are somewhat less likely to be chronically absent. They exhibit school enrollment and SAT test-taking rates that are comparable to non-participating applicants. With these reference points now established, we can divide the overall cohort into work commitment completers and non-completers to compare these groups.

FIGURE 1 – 2017 GDYT youth exhibit lower rates of chronic absenteeism and higher probability of graduating high school relative to non-participating applicants.



* Indicates statistically significant difference between non-participating applicants and GDYT youth.

What share of GDYT youth complete their work commitments?

For the most part, GDYT youth are able to work a maximum of 120 hours over the summer, although some youth do work additional hours.[iv] Youth who work at least 85% of the 120-hour benchmark (i.e., 102 hours) are considered to have completed their work commitments. Figure 2 shows the distribution of hours worked among 2017 participants. Each bar's height corresponds to the share of GDYT youth whose total hours worked fell within the corresponding range of hours.

Approximately three-quarters (75.3%) of 2017 GDYT youth completed their work commitments. The majority of non-completing youth worked at least 80 hours and less than 10% of all GDYT youth worked less than half of the 120 hours available to them. Note that we do not know why one-quarter of youth did not complete their work commitments and any commentary on the subject at this point would be mere speculation. Moreover, it is possible that youth who fell short of completion did so for positive reasons. For example, some youth may stop working early to participate in pre-season athletics or prepare for college. Either way, we can study pre-GDYT applicant data to better understand how similar these groups are to begin with.

FIGURE 2 – Approximately three-quarters of 2017 GDYT youth completed their work commitments.

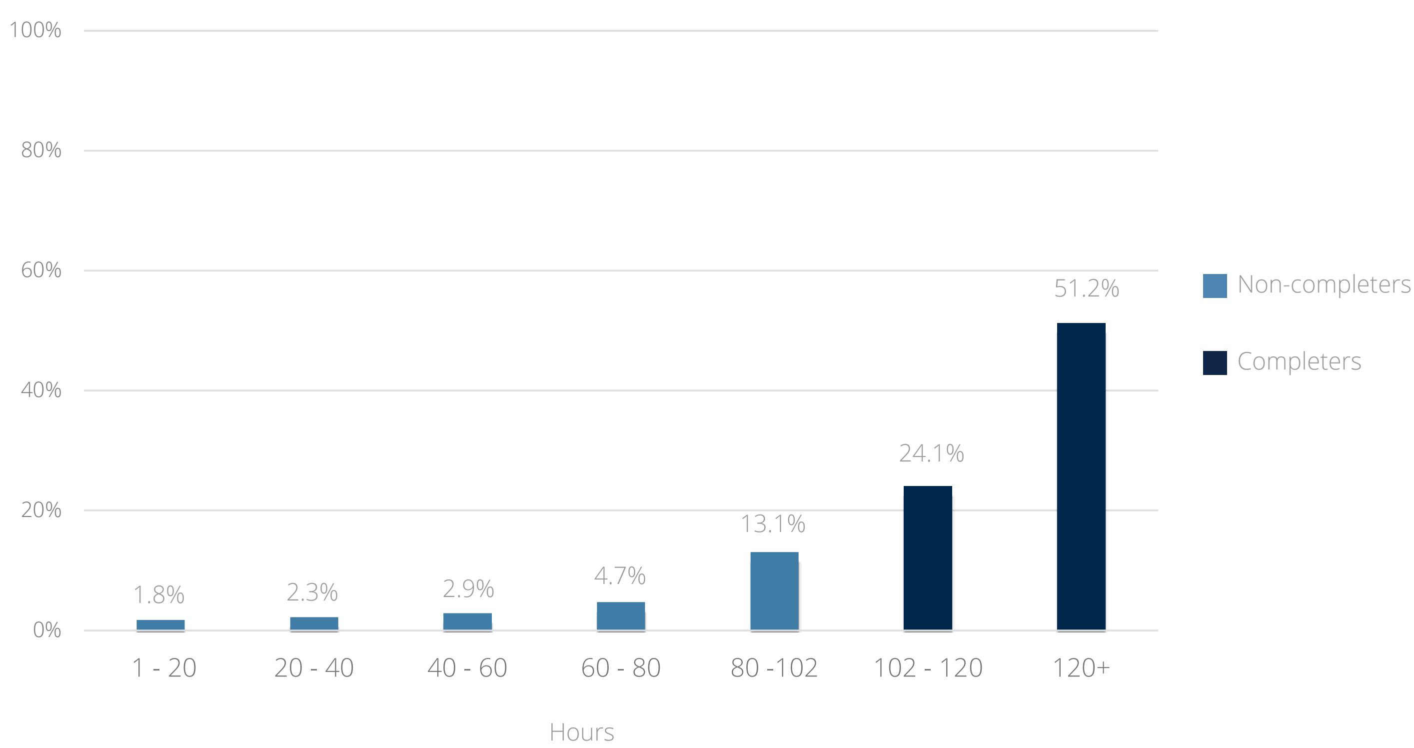


TABLE 2 – 2017 GDYT completers are more likely to be female and have lower rates of pre-program chronic absenteeism.

	COMPARISON DETROIT YOUTH	NON-PARTICIPATING APPLICANTS	NON-COMPLETERS	COMPLETERS
Female	52.4%	59%	52%	56.8%
Prior Chronic Absenteeism	44.1%	42%	43.6%	36.2%
Low-Income	77.6%	81.6%	81.5%	80.3%
8th Grade Reading Proficiency	27.5%	27.1%	28.2%	27%
8th Grade Math Proficiency	7%	6.9%	6.5%	6.8%

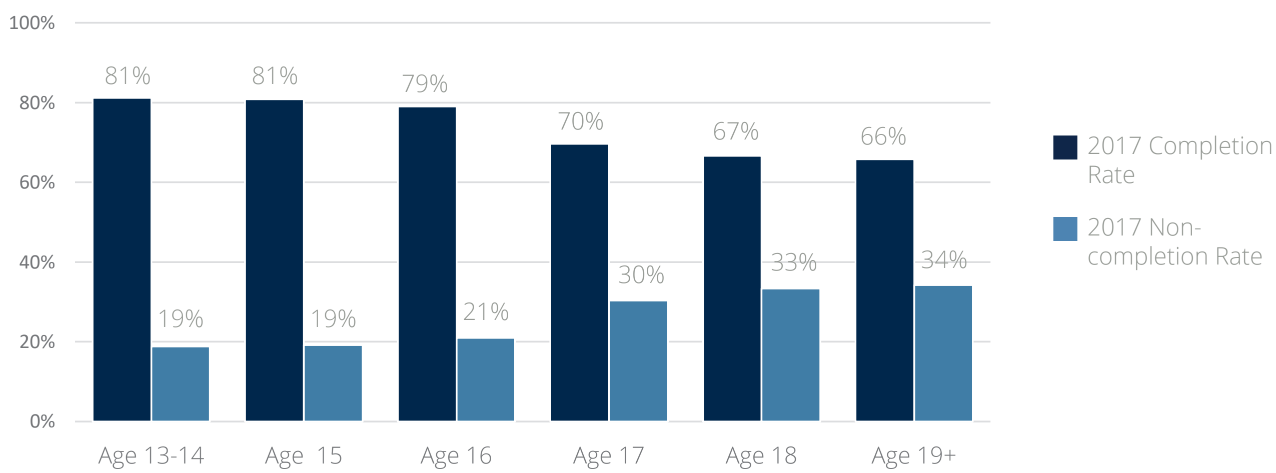
How do completers and non-completers compare before applying to GDYT?

Table 2 displays rates of key baseline (i.e., pre-GDYT) characteristics among non-participating applicants, GDYT work completers, and GDYT work non-completers. We find that although completers and non-completers exhibit comparable socioeconomic status and academic achievement, these groups differ with respect to gender and chronic absenteeism. Boys make up a greater share of non-completers than completers, and the rate of pre-GDYT chronic absenteeism is 20.4% higher among non-completers.[v] The latter statistic should perhaps not surprise us. We might expect that youth who miss a significant number of school days would be likely to miss work as well.

We also find that the share of GDYT youth who complete decreases with age.

Together, these data points bring our understanding of GDYT participants into clearer focus and shed light on factors associated with program completion status. Our results indicate that youth who ultimately complete their work commitments are more likely to be girls, are typically younger, and exhibit higher school attendance prior to participating in the program. These systematic differences warn us against making strong inferences when comparing outcomes between groups.

FIGURE 3 – Work commitment completion rates decline with age.

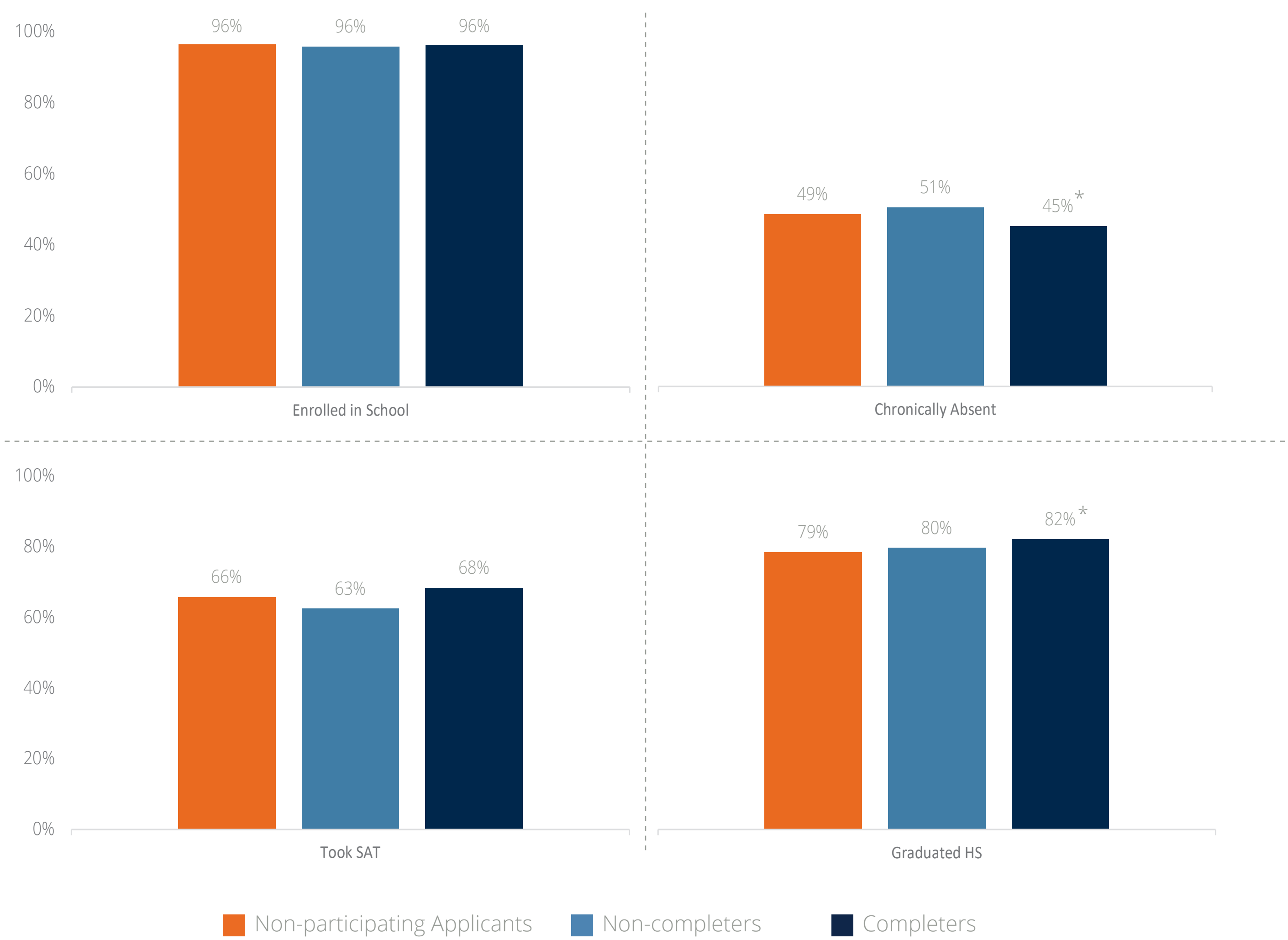


How do outcomes vary by work commitment completion status?

We next deconstruct the overall 2017 cohort outcomes into separate results for completers and non-completers. A couple of caveats are worth noting. First, youth are not randomly assigned a given number of hours to work each summer and we can reasonably assume that many youth who do not complete their work commitments face special circumstances that most youth who complete do not face. Second, we have seen that completers and non-completers differ with respect to a number of key observable characteristics (e.g., gender, age, chronic absenteeism).

As such, we cannot interpret any positive findings among completers as *caused by* working at least 102 hours. The results we present below are merely descriptive and meant to help us better understand how academic outcomes of youth who do and do not complete their work commitments compare.

FIGURE 4 – 2017 cohort results are driven by youth who completed their work commitments.



* Indicates statistically significant difference between non-participating applicants and GDYT youth.

Figure 4 shows academic outcomes by completion status for 2017 GDYT youth one year after participating in the program. The reference category for both groups is still non-participating applicants.[vi] We see that the positive outcomes we observed for the overall cohort are driven by youth who completed their work commitments. These youth are somewhat less likely to be chronically absent and more likely to have graduated high school relative to non-participating applicants. Their enrollment rate is essentially identical to that of non-participating applicants. Conversely, GDYT youth who did not complete their work commitments appear marginally more likely to be chronically absent and less likely to have taken the SAT. Neither of these results for non-completers is statistically significant, however, meaning we cannot say with confidence that they differ from non-participating applicants.

It is worth taking a moment to consider these results in light of the demographic differences we found between completers and non-completers. We should not be surprised that non-completers, a sub-group of youth who miss particularly high amounts of school prior to their summer jobs, continue to do so after the program. Moreover, since students must be in school to take the SAT, this group's lower rate of sitting for the exam is also understandable.

"Youth who completed their work requirements are somewhat less likely to be chronically absent and more likely to have graduated high school relative to non-participating applicants."

CONCLUSION

We find that GDYT youth from three cohorts marginally outperform non-participating applicants across a number of academic outcomes in subsequent years. Our results also indicate that these cohort-level results mask differences between completing and non-completing youth. Youth who complete their work commitments outperform those who do not. That said, neither acceptance into the program nor the number of hours youth work is randomly assigned. Whether the results we observe are caused by youth participating in GDYT therefore remains unknown. The best way to answer this question is to randomly assign GDYT invitations. We will soon report the results from one such pilot study.

YPL partnered with GDYT in 2018 to implement a randomized controlled trial of the Junior Police Cadets (JPC) program, a job for youth aged 14 to 15. In addition to randomizing invitations, we also distributed a follow-up survey to all eligible applicants that asked questions about school engagement, work readiness, and perceptions of the police.

Once data from the 2018-19 school year are available, we will conduct academic outcome analyses and publish those findings along with survey results. These results will represent our first opportunity to make causal claims about one component of GDYT's impact on youth and, importantly, broaden our set of outcome measures.

Beyond the JPC study, we will further expand our understanding of GDYT's relationship with youth outcomes by linking program records with workforce data from the Michigan Unemployment Insurance Agency. While the academic measures we have analyzed thus far are surely important, they are not necessarily the most relevant outcomes we might evaluate. Employment and earnings records, however, bear direct and obvious relevance to summer youth employment. Studying them will allow us to assess how GDYT youth compare to non-participating applicants in the labor market and thus shine further light on an important program that touches thousands of lives each year.

The Youth Policy Lab would like to thank our partners at Connect Detroit and Detroit Employment Solutions Corporation for their support of this work.

About the Authors

Brian A. Jacob is the Walter H. Annenberg Professor of Education Policy and Professor of Economics at the University of Michigan. His current research focuses on urban school reform, youth workforce development and program evaluation.

Jeremy Guardiola is a project manager at the Youth Policy Lab. He focuses on workforce development programs and policies.

Cameron Conrad is a data analyst at the Youth Policy Lab. His work focuses on evaluating education and workforce development programs.

University of Michigan Youth Policy Lab
5201 Institute for Social Research
426 Thompson St
Ann Arbor, MI 48104

734-647-8829
@YouthPolicyLab
youthpolicylab.umich.edu

Support the Youth Policy Lab's efforts to use data for good.

The University of Michigan Youth Policy Lab helps community and government agencies make better decisions by measuring what really works. We're data experts who believe that government can and must do better for the people of Michigan. We're also parents and community members who dream of a brighter future for all of our children. At the Youth Policy Lab, we're working to make that dream a reality by strengthening programs that address some of our most pressing social challenges.

We recognize that the wellbeing of youth is intricately linked to the wellbeing of families and communities, so we engage in work that impacts all age ranges. Using rigorous evaluation design and data analysis, we're working closely with our partners to build a future where public investments are based on strong evidence, so all Michiganders have a pathway to prosperity.

TECHNICAL APPENDIX

Appendix A- Data

This appendix describes our data sources, record linkage, and how we define our sample and create variables for analysis.

A.1 Application and Participation Data

Connect Detroit and Detroit Employment Solutions Corporation (DESC), the agencies that administer GDYT, use the Philadelphia Youth Network's (PYN's) database management system for GDYT application design, submission, and data management. PYN provided application data files to Connect Detroit and DESC which then sent them to YPL for analysis. The key variables we use from application data include application status; personal identifiers such as first and last name, birth date, gender, race/ethnicity, home address; and referral code, which indicates whether youth entered the program with a referral from a particular worksite. PYN maintained the application data for all three cohorts from 2015-17.
We also obtained payroll data on GDYT participants. For the 2015 and 2016 cohorts, payroll data was maintained in decentralized systems, with multiple organizations maintaining payroll records in different data management systems. In these years, we are missing important data about work placements, positions, and hours worked. Due to these inconsistencies, we consider any youth who appeared in any payroll data source to be a worker for these cohorts. In 2017, payroll data were centralized with PYN's payroll data management system, so this is the first cohort that we have more consistent data on job placements, hours, and wages. Youth are classified as participants in 2017 if youth worked any positive number of hours or were one of the small percentage of youth who worked for an affiliate and have missing hours data.

A.2 Education Data

YPL obtained administrative education data from the Michigan Department of Education (MDE) and the Center for Educational Performance and Information (CEPI). These data include information on all students in Michigan (MI) public schools, including public charter schools. Records include information on enrollment, demographics, test scores, and graduation. We also link to college enrollment data from National Student Clearinghouse (NSC) for youth who attended MI public schools at some point during their K-12 education.

A.3 Record Linkage

We clean and link data across sources using the following steps:

1. Clean, link, and de-duplicate application and payroll data: we follow relatively standard data cleaning procedures by removing non-alphabetic characters from and standardizing capitalization in names. We also convert string variables to categorical variables for ease of analysis. Once data have been cleaned, we merge application data to payroll data using a common identifier. Finally, we de-duplicate data using Record Linkage, R's probabilistic de-duplication program. De-duplication is necessary because there are duplicates in both the application and payroll data: some youth apply to GDYT multiple times, while other youth may have multiple payroll records due to changes in job placement.

2. Link program data to state education identifier from CEPI: we match clean program data to records from MDE that are provided by CEPI. CEPI uses a 17-step quasi-probabilistic record-linking algorithm to generate "matching percent" scores indicating the likelihood of a match. The algorithm uses first name, middle initial (optional), last name, suffix (optional), birth date, and gender to create scores. The "matching percent" score for each record leads to one of three results: exact match, non-match, and requires resolution.

3. Review CEPI matches manually: we manually reviewed the set of records that required resolution. In the event of uncertainty, we assume no match in an effort to minimize false positive matches. After all matching and manual review was complete, we matched 94 percent of the applicant sample to state education records (see Table A.3.a). Based on self-reported school enrollment from the application data, we suspect that less than 1 percent of applicants attended private schools. Thus we were unable to match 5 percent of the sample for other reasons such as out-of-state moves.

4. Further de-duplicate data after completing CEPI matching: After the completion of linking to CEPI data, we performed further de-duplication because CEPI's algorithm found additional duplicates that had not previously been identified.

Table A.3.a: Selected Characteristics of Applicants and Participants

	2015		2016		2017	
	Applicants	Participants	Applicants	Participants	Applicants	Participants
Total Number	12,236	2,807	14,098	6,870	15,143	5,261
Matched to Education Data	94%	90%	92%	87%	95%	95%
	% of Applicants	% of Participants	% of Applicants	% of Participants	% of Applicants	% of Participants
Demographics						
Black	95%	90%	94%	92%	94%	91%
Hispanic	3%	7%	4%	5%	4%	6%
White	2%	3%	2%	2%	2%	2%
Asian American	1%	0%	0%	0%	0%	0%
Female	58%	55%	57%	55%	56%	54%
Limited Eng. Prof.	3%	5%	3%	4%	3%	5%
Special Education	13%	15%	14%	14%	15%	15%
Low-Income	84%	85%	84%	83%	85%	84%
Age						
Under 14 Years	3%	3%	4%	4%	3%	2%
14-18 Years	82%	81%	81%	83%	83%	82%
19-21 Years	11%	12%	10%	9%	10%	12%
22-24 Years	4%	4%	4%	3%	4%	5%
Enrollment Status						
Enrolled (HS or College)	89%	90%	88%	91%	89%	88%
Not Enrolled	11%	10%	12%	9%	11%	12%
Neighborhood Characteristics						
BA Degree or Higher	14%	13%	14%	13%	14%	14%
Below Poverty Line	35%	35%	34%	35%	34%	34%
Owner Occupied Housing	43%	43%	43%	42%	43%	42%
Employed (Age 16+)	76%	76%	77%	76%	76%	76%

Some categories will not sum to 100% due to data that is missing or records not matched to education data. Some summary statistics may have changed marginally from our April 2018 policy brief on GDYT due to the receipt of updated program data and further data cleaning.

A.4 Sample Definition and Variable Construction

The following tables describe how we define our sample and create key explanatory, control, and baseline variables.

Table A.4.a: Sample Definition

Variable	Description
Applicant	Indicator variable for if youth completed an application or worked even if she/he did not complete an application. All workers are inherently considered applicants in our analysis.
Worked	2015-16: in payroll data. 2017: worked greater than 0 hours or worked for an affiliate and hours data are missing.

Applicant and Worked are key explanatory variables in our basic analysis of post-participation outcomes and baseline characteristics.

Table A.4.b: 2015-17 De-Duplicated Record Counts and Percentages

	Totals	Percentages
Applicants	41,477	100%
Non-Participating Applicants	26,539	64%
Participants	14,938	36%

Table A.4.c: Work Completion Variables

Variable	Description
Worked Missing Hours	Youth who worked for an affiliate, but hours data are missing in 2017.
Worked Non-Completer	Worked greater than 0, but less than 102 hours in 2017.
Worked Completer	Worked greater than or equal to 102 hours in 2017.

There were 150 youth who worked for an affiliate, but had missing hours data in 2017. These youth accounted for less than 3 percent of all workers in 2017.

Connect Detroit and DESC define work completion as working at least 85 percent of the available program hours. This is at least 102 of the maximum 120 hours available for most youth.

Most control and baseline variables have standard definitions, but the few that require additional explanation are shown in the table on page 14.

Table A.4.d: Control & Baseline Variables

Control Variables	Description
Low-Income	A binary indicator for whether a student is eligible for free or reduced price lunch.
Percent Poverty	The percentage of families in a Census Block Group that live below the poverty level.
Chronic Absenteeism	A binary indicator for whether a student had an attendance rate less than or equal to 90 percent.
Control Flags for Missing Values	Indicator variable adjustments to account for missing values in control variables. Control flags are set to one if that control is missing and the control variable itself is set to zero instead of missing.
Baseline Variables	**Description**
Reading/Math Proficiency	Reading and math proficiency are binary indicators of whether a student scored above the proficiency threshold determined by the State of Michigan on their 8th grade standardized reading and/or math test. If the test was taken before 2012, the proficiency threshold is adjusted to be consistent with the more strict proficiency standards which were implemented in 2012. Defined for students in 8th grade or above in baseline year. Proficiency determined by highest 8th grade score if a student repeated 8th grade.

The control variables are measured in the baseline year and serve as control variables in our post-participation outcome regressions. The baseline variables serve as the dependent variables in our baseline regressions that test for statistically significant differences in characteristics among comparison youth, non-participating applicants, and participants in the baseline year.

The baseline year is the year of GDYT application or the most recent year before application that the youth was enrolled in MI K-12 public schools. For example, if a youth was enrolled in the 2014-15 school year and applied for GDYT in 2015, the youth's baseline characteristics are created based on data from the 2014-15 school year.

APPENDIX B – ANALYSIS METHODOLOGY

B.1 Analysis Goals

The goals of our analysis include:

1. Assess how participation in GDYT influences educational outcomes across three different cohorts
2. Assess whether outcomes vary by work completion status
3. Assess how work completers and non-completers compare in baseline characteristics

In our basic analysis of post-participation outcomes, there are three potential sources of bias to account for:

1. The self-selection of youth into the applicant pool
2. The employer selection among applicants
3. The self-selection of youth chosen by employers in their decision about whether to work

We account for the first source of selection bias by limiting our analysis sample to just the set of youth who applied to GDYT and comparing the difference in outcomes between participants and non-participating applicants. We do not have data on job offers so we cannot distinguish between employer selection and youths' decisions about whether to work.

In our analysis that uses work completion variables as explanatory variables, there is an additional source of bias. Ideally we would compare the difference in outcomes between work completers and non-participating applicants who would have completed work as well as the difference in outcomes between work non-completers and non-participating applicants who would not have completed work.

However, we are unable to determine which non-participating applicants would and would not have completed work. Thus we compare the difference in outcomes between work completers and all non-participating applicants as well as work non-completers and all non-participating applicants. We cannot interpret these estimates as causal because of this bias.

B.2 Analysis Goals

Research Design

We assess how GDYT participation influences four outcomes of interest (see Table B.2.b):

1. Enrollment in MI K-12 Public Schools
2. Chronic Absenteeism
3. Took SAT
4. Graduated High School (HS)

Our method of analysis for post-participation outcomes replicates the matching research design and fixed effects regression model used in our April 2018 GDYT policy brief. The match group for each applicant consists of all eligible non-applicant Detroit youth who are the same with respect to the following five characteristics:

1. Match year: the baseline year
2. Match school: school in the baseline year
3. Match grade: grade in the baseline year
4. Race
5. Gender

For applicants who were not enrolled in a MI K-12 school at the time of application, a match group is created based on the most recent year the youth was enrolled.
In the 2015-17 cohorts, 0.11 percent of applicants were missing data on gender and none were missing data on their race. To construct match groups for youth who were missing gender, we impute gender to be female as majority (57 percent) of applicants were female.

For youth *i* in match group *j* in application year *t*, we use a fixed effects regression model to identify outcome differences among matched comparison youth, applicants, and participants:

$$y_{ijt} = \beta_0 + \beta_1 Applicant_{ijt} + \beta_2 Worked_{ijt} + \beta_3 X_{ijt} + \gamma_{ijt} + \varepsilon_{ijt} \quad (1)$$

1. y_{ijt} is an one of our four outcomes of interest (see Table B.2.b)

2. $Applicant_{ijt}$ and $Worked_{ijt}$ are binary variables for whether youth applied for and worked in GDYT (see Table A.4.a)

3. X_{ijt} is our set of controls variables, including baseline measures of demographic and neighborhood characteristics, 8th grade test scores, and education characteristics

4. γ_{ijt} are the match group fixed effects

5. ε_{ijt} is our error term with standard errors clustered by match school

In specification (1), β_0 is the average outcome measure within each match group for comparison youth, β_1 is the difference in outcomes between applicants and comparison youth, and β_2 is the difference between participants and applicants.

The variables included in X_{ijt} are: indicators for receipt of special education services, limited English proficient status, and low-income (all measured in the match year); baseline school attendance and indicators for whether youth: were chronically absent, had graduated high school, or had been retained in school; linear and quadratic terms of a youth's age in the match year; 8th grade standardized math and reading scores, and an interaction term between the two; and the following census block measures: share of individuals with at least a Bachelor's degree, the share of families living below the poverty level, the percent of owner-occupied housing units, and the share of the civilian population 16 years or older who are in the labor force.

Table B.2.a shows the post-participation data available for each cohort and Table B.2.b defines our outcome variables for each post-participation year.

Table B.2.a: Outcome Data by Cohort

Application Year Cohort	Years of Outcome Data Available
2015	Post Year 1 (2015-16 school year) Post Year 2 (2016-17 school year) Post Year 3 (2017-18 school year)
2016	Post Year 1 (2016-17 school year) Post Year 2 (2017-18 school year)
2017	Post Year 1 (2017-18 school year)

Table B.2.b: Average Outcome Variable Definitions

Outcome Variable	Post Year 1-3 Definitions
Average Enrolled	Post Year 1: 11^{th} grade or below in application year Post Year 2: 10^{th} grade or below in application year Post Year 3: 9^{th} grade or below in application year
Average Chronic Absenteeism	Post Year 1: 11^{th} grade or below in application year and enrolled in Post Year 1 Post Year 2: 10^{th} grade or below in application year and enrolled in Post Year 2 Post Year 3: 9^{th} grade or below in application year and enrolled in Post Year 3
Average Took SAT	Post Year 1: 10th grade in application year Post Year 2: 9^{th} grade in application year OR in 10^{th} grade in application year and did not take SAT in Post Year 1 Post Year 3: 8^{th} grade in application year OR 9^{th} grade in application year and did not take the SAT in Post Year 2 OR 10^{th} grade in application year and did not take SAT in Post Year 1 and did not take SAT in Post Year 2
Average Graduated HS	Post Year 1: 11^{th} grade in application year and did not graduate HS in or before application year Post Year 2: 10^{th} grade in application year and did not graduate HS in or before application year OR 11^{th} grade in application year and did not graduate HS in or before application year and did not graduate HS in Post Year 1 Post Year 3: 9^{th} grade in application year and did not graduate HS in or before application year OR 10^{th} grade in application year and did not graduate HS in or before application year and did not graduate HS in Post Year 2 OR 11^{th} grade in application year and did not graduate HS in or before application year and did not graduate HS in Post Year 1 and did not graduate HS in Post Year 2

Enrolled is defined only for students in 11th grade or below in the application year so that we obtain estimates only for youth who are eligible to be enrolled (12 grade and below) in Post Year 1.

Similar logic applies for our chronic absenteeism and graduated high school outcome variable definitions. Took SAT is defined for students in 10th grade in their application year because all MI students take the SAT as part of the Michigan Merit Examination in 11th grade.

Consistency of Outcomes Across 2015-17 Cohorts

To test the consistency of post-participation outcomes across the three application cohorts, we pool data across the 2015, 2016, and 2017 application year cohorts and generate variables that take the average of Post Year 1, Post Year 2, and Post Year 3 outcome variables. Table B.2.b shows how the outcome variable definitions vary depending on how many years of post-participation data are available. Outcome variable definitions must vary so that outcomes are defined for youth who are eligible for a certain outcome and set to missing for youth who are not eligible for a certain outcome. For example, a youth who was in 10th grade in the application year would have the Took SAT Post Year 1 outcome defined since MI students take the SAT in 11th grade. On the other hand, a youth who was in 12th grade in the application year would have Took SAT Post Year 1 set to missing because we would not expect this youth to take the SAT in Post Year 1 given this youth's grade level. Post Year 2 and Post Year 3 outcomes for Took SAT and Graduated HS are defined to account for youth who may have been retained, dropped out and re-enrolled, and other potential issues.

We use regression equation (1) defined above to obtain estimates using our pooled data.

For the average Enrolled and average Chronic Absenteeism outcomes, the variables can take on a range of values from 0 to 1 for the 2015 and 2016 cohorts depending on how many years in the post-analysis the youth was enrolled or chronically absent. For the 2017 cohort, average Enrolled and average Chronic Absenteeism can only take on a value of 0 or 1 for youth who were in 11th grade and below in the application year.

Average Took SAT is calculated based on whether the student took the SAT within three years of the application year for the 2015 cohort, two years of the application year for the 2016 cohort, and one year of the application year for the 2017 cohort. Similar logic applies for average Graduated HS. Both are binary variables that only take on values of 0 or 1 since these outcome variables measure one-time events.

Analysis of Post-Participation Outcomes by Work Completion Status

Our estimates from Figure 1 are obtained using regression equation (1) with data from the 2017 cohort, which only has Post Year 1 outcome data available. These estimates serve as a comparison point for our analysis of post-participation outcomes by work completion status.

For this analysis, we build on regression equation (1) by separating the Worked explanatory variable into three separate work completion status variables (see Table A.4.c):

$$y_{ijt} = \beta_0 + \beta_1 Applicant_{ijt} + \beta_2 WorkedNoHours_{ijt} + \beta_3 WorkedNonCompleter_{ijt} + \beta_4 WorkedCompleter_{ijt} + \beta_3 X_{ijt} + \gamma_{ijt} + \varepsilon_{ijt} \quad (2)$$

In this specification (2), β_0 reports the average outcome measure within each match group for comparison youth who did not apply to GDYT, β_1 is the difference in outcomes between applicants and comparison youth, β_2 is the difference between workers with missing hours data and applicants, β_3 is the difference between work non-completers and applicants, β_4 is the difference between work completers and applicants.

B.3 Analysis of Baseline Characteristics

For the 2017 cohort, we assess how non-participating applicants, work non-completers, and work completers compare in terms of their baseline characteristics using the same matching research design we used for analysis of post-participation outcomes. When race or gender is the baseline characteristic of interest, a separate match group is used that excludes that characteristic so that there is still variation in that baseline measure within match groups. Youth who are missing a baseline measure are excluded from the analysis for that measure.

We use a similar fixed effects regression model that we used in our analysis of post-participation outcomes by work completion status, except we use baseline characteristics as the dependent variables for this analysis. For youth *i* in match group *j* in application year *t*, we identify differences in baseline characteristics between the matched comparison youth, applicants, workers with missing hours, work non-completers, and work completers:

$$y_{ijt} = \beta_0 + \beta_1 Applicant_{ijt} + \beta_2 WorkedNoHours_{ijt} + \beta_3 WorkedNonCompleter_{ijt} + \beta_4 WorkedCompleter_{ijt} + \gamma_{ijt} + \varepsilon_{ijt} \quad (3)$$

1. y_{ijt} is the baseline measure of interest

2. $Applicant_{ijt}$, $WorkedNoHours_{ijt}$, $WorkedNonCompleter_{ijt}$, and $WorkedCompleter_{ijt}$ are binary variables for whether youth applied for and worked in GDYT (see Table A.4.a and Table A.4.c)

3. γ_{ijt} are the match group fixed effects

4. ε_{ijt} is our error term with standard errors clustered by match school

In this specification (3), β_0 is the average baseline measure within each match group for comparison youth who did not apply to GDYT, β_1 is the difference in baseline measures between applicants and comparison youth, β_2 is the difference between workers with missing hours data and applicants, β_3 is the difference between work non-completers and applicants, β_4 is the difference between work completers and applicants.

DISCLAIMER

This analysis utilizes data obtained through a confidential data application process submitted to the Michigan Education Data Center (MEDC)/Michigan Education Research Institute (MERI). Youth Policy Lab at the University of Michigan requested data access and completed the analysis included in this report. The data are structured and maintained by the MERI-Michigan Education Data Center (MEDC). MEDC data is modified for analysis purposes using rules governed by MEDC and are not identical to those data collected and maintained by the Michigan Department of Education (MDE) and/or Michigan's Center for Educational Performance and Information (CEPI). Results, information and opinions solely represent the analysis, information and opinions of the author(s) and are not endorsed by, or reflect the views or positions of, grantors, MDE and CEPI or any employee thereof.

---

[i] Readers may notice that these numbers are slightly different than what we reported in our April 2018 policy brief. The figures presented here derive from updated regression models that include additional control variables. These new covariates slightly changed the magnitude of the reported coefficients, although the finding that GDYT participation is associated with lower rates of chronic absenteeism among the 2015 cohort is no longer statistically significant.

[ii] Data from the 2018-19 academic year were not available at the time of publication.

[iii] 70.1/66.9 = 1.0478

[iv] Foster youth, for example, are eligible to work up to 240 hours.

[v] 43.6/36.2 = 1.204

[vi] Although comparing GDYT youth to non-participating applicants helps us account for the factors that lead youth to apply to the program in the first place, we face an added hurdle when we divide youth into completers and non-completers. Among workers, we know precisely which youth did and did not complete their work commitments. For non-participating applicants, however, we have no way of determining which youth would have completed their commitments and which would not have. Ideally, we would be able to compare completers to youth who would have completed had they been invited to participate and non-completers to applicants who would not have completed if their applications had been successfully reviewed, but we are unable to do so. As such, these comparisons are imperfect.

[vii] In the vast majority of cases, an applicant's match year corresponds to the year in which she/he applied to GDYT. In a minority of cases, an applicant's match year is not the application year and is instead the year in which she/he was most recently enrolled in a MI K-12 public school prior to the application year. If an applicant's match year was before her/his application year, this match year may represent the year that this applicant switched from public school to private school or moved out of state. In the 2015-17 cohorts, 80 percent of applicants had a match year equal to their application year, 11 percent had a match year within two years of their application, and 9 percent had a match year greater than two years before they applied.