Pre-Analysis Plan
Impact Evaluation of Girls Inc. EmpowerHub

Mary Kate Batistich
University of Notre Dame

Jonathan Tebes
University of Notre Dame

August 2024

I. Overview

This is a pre-analysis plan for the randomized evaluation of the Girls Inc. EmpowerHub program. We begin with the motivation for the program and the randomized evaluation in Section II. Section III lays out the target population, recruitment strategies, and eligibility criteria. Section IV describes our planned data sources, and Section V defines our primary and secondary outcomes. In Section VI, we present and discuss our power calculations relative to observational evidence and related studies. Finally, our empirical strategy is laid out in Section VII.

II. Background

The healthy development of social-emotional skills, including behavioral and emotional regulation, leadership, decision-making, self-esteem, and relationship-building, is important for children's present and future well-being and success. For instance, behavioral and emotional self-regulation allow youth to relate healthily to others in the present, explore their own interests and aspirations, and promote healthy social functioning in adulthood (Pulkkinen et al., 2002). Middle school is a pivotal time in youth development. Youth are living through a mental health crisis, where anxiety, depression, suicide have reached startingly high rates (Office of the Surgeon General, 2021). During early adolescence, youth develop a sense of personal identity through autonomous decision making and goal-setting, and benefit from the exploration of these questions in a safe social environment with positive guidance (Wong et al., 2010). However, many middle school students do not receive adequate emotional support. Suicide rates fall when school is not in session, whether for summer break or during COVID-19 disruptions (Hansen et al., 2024). Meanwhile, chronic absenteeism remains at an all time high (Mervosh and Parish, 2024) and school enrollment has failed to recover to pre-pandemic levels (Burtis et al., 2024). School engagement can be a challenge, particularly for low-income students attending urban public schools.

An in-school counseling program for high school girls in Chicago led to mental health benefits, decreasing PTSD symptoms, depression, and anxiety (Bhatt et al., 2023). Girls Inc.'s EmpowerHub offers a similar potential solution catered toward the specific needs of middle school girls in Indianapolis communities, incorporating elements of mentorship, social-emotional learning SEL), and academic support. Adolescents' later academic and economic success depends on social factors in addition to early educational investments, particularly for under-resourced communities.

A range of youth programs have been implemented across the United States, displaying varying degrees of recorded effectiveness. Prior studies show positive findings on the effectiveness of youth programs, including those incorporating mentorship (Falk et. al, 2020; Resnjanki et al., 2021; Zimmerman et. al, 2018), social-emotional learning (Battistich et. al, 2004; Crean and Johnson, 2013), and trauma-informed therapy (Bhatt et al., 2023). Other studies of programs incorporating mentoring high-risk populations find no consistent program effect for school grades, test scores, on-time graduation (Maxfield et al., 2003; Heppen et al., 2018), school engagement (Guryan et. al, 2021), risky behaviors (Rodriguez-Planas, 2012; 2017), or student well-being and behavior (Berry et. al, 2016). This study seeks to determine the impact of EmpowerHub–which is uniquely in-school, focuses on middle school girls, and draws upon both SEL and mentorship–not only behavioral outcomes, but also school engagement and academic achievement. This study will inform Girls Inc., IPS, and the greater academic community about the effectiveness of girl-centered comprehensive programming for middle school students. If this study shows that EmpowerHub has a positive effect on the identified outcomes for participants, it can be used as evidence to secure further funding or legislation to expand services to benefit more youth.

III. Evaluation Design

A. Research Questions

What is the impact of EmpowerHub on academic performance in terms of grades, school attendance, homework completion, and disciplinary outcomes?

B. Eligibility

Students are eligible to enroll if they (1) attend one of Girls Inc.'s seven IPS partner schools, (2) are in sixth grade, and (3) identify as a girl. Eligibility will be determined by baseline surveys at the seven middle schools in which EmpowerHub will run. A child interest form will be administered to children at informational presentations during lunch hours, and a parent/guardian application will be administered at back-to-school events. Eligible children who provide assent and whose parent/guardian provides parental consent will be enrolled in the study.

C. Randomization

In September 2024, the research team will randomize study participants into two equal groups, treatment and control, by school, contingent on sufficient enrollment. Depending on program capacity constraints, subsequent years may not have equal sized treatment and control groups. The treatment group will be invited to begin EmpowerHub programming soon after randomization, and the control group will continue attending standard in-school programming. EmpowerHub participants will continue attending programming throughout middle school.

D. Intervention

The EmpowerHub program is delivered by trained program coordinators and is comprised of three sections – Strong (health and wellness), Smart (academic readiness), Bold (civic responsibility and leadership). The program is designed to be consistent across each program site. EmpowerHub's strategy encompasses a multifaceted approach that includes reinforcing academic behaviors, fostering perseverance, cultivating positive mindsets, enhancing learning strategies, and developing social skills, all integral to shaping well-rounded individuals and engaging students in their own futures.

This program focuses on non-cognitive factors such as academic behaviors, mindsets, perseverance, learning strategies, and social skills. EmpowerHub's curricular interventions are strategically designed to develop non-cognitive skills shown by research to be critical in enhancing students' cognitive and emotional realms, thereby improving academic outcomes.Programmatic components include mentorship from program coordinators, activities in STEM and arts, support in building academic and life skills through goal setting and leadership training, and other engaging sessions.

EmpowerHub operates within the school setting, taking place once per week during advisory periods, with a focus on addressing middle school girls' disengagement and academic underperformance. The goal is to provide participants at each partnering location with 35-50 hours of high-impact programming throughout the academic year. The EmpowerHub program will be available to cohorts of 25-40 middle school girls at each of the seven participating IPS middle schools. Students who are not in EmpowerHub–the control group and all other students not in the study–will continue engaging in standard advisory period activities like teacher meetings and homework time.

IV. Data Sources

A. Indianapolis Public Schools (IPS)

IPS holds administrative data for all students from the seven public middle schools participating in EmpowerHub. We plan to use these records to verify eligibility prior to randomization and measure the effect of EmpowerHub on a range of educational outcomes. Through a data sharing partnership with IPS, the research team will pull academic outcomes for study participants from the PowerSchool and Schoology platforms, as well as other administrative data collected by the school system. The research team will also collaborate with IPS on their annual school climate questionnaire, Panorama, that surveys students on emotional and social topics, and the responses of study participants will be used for research purposes.

B. Girls Inc. of Greater Indianapolis

Girls Inc. surveys program applicants to collect information about their household structure and parental education and income, which will allow us to document differences in program engagement and outcomes by family background. Program records from Girls Inc. will provide engagement measures including enrollee attendance to in-school programming, various outside events, and program director interactions which will allow for descriptive work of the treatment and treatment on the treated analysis. LEO has a data sharing agreement with Girls Inc.

V. Study Outcomes

EmpowerHub's strategy encompasses a multifaceted approach that includes reinforcing academic behaviors, fostering perseverance, cultivating positive mindsets, enhancing learning strategies, and developing social skills, all integral to shaping well-rounded individuals and re-engaging students in their future. That is why this evaluation will primarily focus on middle school educational outcomes of participants, pulled from administrative data at IPS. We are also interested in examining the effects of EmpowerHub on disciplinary and social outcomes.

A. *Primary Outcomes*

- Standardized GPA: cumulative GPA, normalized by the control group mean and standard deviation at the end of year 1 within a given school
- We will combine outcomes into a standardized summary index or alternative summary aggregated index (see, e.g. Rose et al, 2022), behavioral index:
    - Incident reports: Number of reported school incidents, suspensions, and detentions, and an indicator for whether an incident occurred during the school year
    - Chronic absenteeism: indicator for whether participant has been chronically absent
    - Grade repetition: indicator for whether the student did not progress to the next grade on time
    - Homework completion: cumulative rate of homework completion as reported by school records

All of these outcomes will be measured at the end of a given school year.

B. *Secondary Outcomes*

- Aggregated and disaggregated standardized test scores: standardized test scores in ILearn (math and ELA)
- Chronic absenteeism: indicator for whether participant has been chronically absent
- Grade repetition: indicator for whether the student did not progress to the next grade on time
- Homework completion: cumulative rate of homework completion as reported by school records
- GPA by subject category: standardized GPA broken into subject category
- Participation in other extracurricular and athletic activities: indicator for whether student participates in other extracurricular and athletic activities outside of EmpowerHub
- School attendance rate: fraction of school days attended
- Full attendance: indicator for attending 94% or more of school days
- Disciplinary outcomes: number of school suspensions and detentions, and an indicator for whether the participant had any suspensions or detentions during the school year
- Incident reports: Number of reported school incidents, and an indicator for whether an incident occurred during the school year

- Social outcomes: collected at the end of the academic year, constructed from student responses to a survey administered in schools by IPS with questions on self-efficacy and belonging. We are collaborating with IPS on the survey content, so outcome construction is dependent upon data availability and granularity.

VI. Statistical Power and Sample Size

We plan to enroll a study sample of approximately 1,050 individuals in the study over three years (academic years 2024-25 to 2026-27), with approximately 40% of these assigned to the treatment group (offered a spot in EmpowerHub), and expect a program take-up rate of around 80%. Because full attendance (attending 94% or more of school days) is readily available rather than chronic absenteeism (missing 10% or more of school days) for the schools we are studying, we report full attendance for our power calculations. Based on school district data, 47.3% of students at the seven participating schools attend school for at least 94% of school days. We are powered to detect a 10.42 percentage point change in full attendance, or a 22% increase.

For standardized outcomes in test scores and grades, we are powered to detect a 0.158 standard deviation change between the control group and the treatment group.

VII. Empirical Strategy

A. Main Specification

Our primary specification estimates the impact of the offer to enroll in EmpowerHub or *intention-to-treat* (ITT) effects of the program on outcomes. We will estimate ITT effects by OLS using the following regression:

$$Y_i = \beta_0 + \beta_1 T_i + \alpha_s + \mu_l + \lambda_t + \epsilon_{islt}$$

Y_i is the outcome for enrolled participant *i*. T_i indicates random assignment of person i to the treatment group, α_s are strata fixed effects (i.e., randomization-block indicators). We will also control for school fixed effects (μ_l) and month/year of randomization fixed effects (λ_t) and ϵ_{islt} is the error term. The coefficient of interest, β_1, estimates the average difference in outcomes between treatment and control groups, or the causal effect of the opportunity to enroll in EmpowerHub, conditional on strata fixed effects. β_1 is referred to as the intent-to-treat (ITT) effect. This is our preferred specification provided that observable baseline characteristics are shown to be balanced across the treatment and control groups.

We will also estimate treatment effects conditional on control vector X_i of baseline characteristics to account for any sampling variation in the composition of treatment and control groups:

$$Y_i = \beta_0 + \beta_1 T_i + X_i \delta + \alpha_s + \mu_l + \lambda_t + \epsilon_{islt}$$

where X_i includes a set of individual-level characteristics collected at baseline such as pre-middle school grades and test scores (where available), age, race, ethnicity, and household characteristics. Standard errors will be clustered at the strata-level and we will also report permuted p-values for our main outcomes.[1]

B. Treatment on Treated Specifications

In addition to the reduced-form estimates obtained in the equations above, we are also interested in estimating the causal impact of EmpowerHub program participation, also known as the *treatment-on-treated* (TOT) effect. To this end, we will estimate the TOT by instrumenting for program participation with treatment assignment. Using a two-stage-least-squares approach (2SLS), we will estimate the system:

$$Y_i = \beta_{D_i} D_i + X_{islt}\delta + \eta_s + \omega_l + \gamma_t + \epsilon_i$$
$$D_i = \pi T_i + X_{islt}\rho + \varsigma_s + \phi_l + \sigma_t + v_i$$

Where D_i is one of two different measures of program enrollment and engagement. The first is an indicator for whether the individual received any programming as a part of EmpowerHub. This indicator will be called $enrolled_i$. Second, $engaged_i$ will be a measure of participant *i*'s cumulative treatment exposure or "dosage". This measure is the fraction of the middle school years that individual *i* was "actively engaged" throughout the first three years of middle school. Under reasonable assumptions, β_{D_i} captures the causal impact of enrollment and engagement in the program on outcome Y_i. Note that this parameter equals the intent-to-treat parameter (β_1) divided by the regression-adjusted take-up rate (π).[2]

We plan to enroll 1,050 students into the study over three years, with approximately 426 of these in treatment and around 341 participating in the program. These numbers may vary depending on the take-up rate for EmpowerHub; for example, if a higher than expected share of the treatment group turns down the program, more students will need to be randomized into treatment in order to fill the spots available in EmpowerHub.

In all models, we will use heteroskedasticity-robust standard errors.

[1]Our randomized design falls within the standard framework described in Athey et al. (2023). We will cluster standard errors at the level that is most recommended at the time of our analysis. According to Athey et al. (2023), the "right" level may even be at the individual level.

[2] This approach relies on the assumption that there was no average effect of being offered Thread enrollment on those who did not take up the program and that the control group was not affected by losing the lottery for the opportunity to enroll in the program.

C. Treatment Effect Heterogeneity and Subgroup Analyses

Given that this study will recruit a broad range of students, the EmpowerHub program will likely have different effects within different sub-groups. Understanding whether the program works broadly, for some sub-population of policy interest, or is most beneficial for some surprising sub-sample provides crucial information to governments on how they might scale this program in the event of a positive finding.

The study will estimate the impact of EmpowerHub across several outcome categories and subgroups. The research team is interested in determining whether the intervention is more effective for certain populations relative to others. Our main sub-group analysis will explore heterogeneity of treatment effects by baseline cognitive and non-cognitive measures prior to randomization. We will use grades, standardized test scores, and other academic information prior to randomization to predict baseline cognitive measures. Predicted non-cognitive measures will draw upon available information on disciplinary outcomes, attendance, and grade repetition prior to randomization. If these predictions have a low R^2, we may instead explore heterogeneity by baseline grades and attendance rates. We will split the sample at the median of a baseline indices discussed below, but will also explore graph treatment effects across the distribution of each index (provided sufficient statistical power).

- Baseline cognitive measures index: We will use grades and standardized test scores prior to randomization to predict baseline cognitive/academic performance.
- Baseline non-cognitive measures index: We will use disciplinary outcomes, attendance, homework completion, and grade repetition prior to randomization to predict baseline non-cognitive traits.
- To measure the level of poverty and family resources at baseline, we will create a family resources index: This index will consist of family structure, family income, parent/guardian education level, family homelessness status, all measured at baseline.
- Racial congruence: We will use student and program coordinator race to investigate the impact of same-race program coordinators on participant outcomes.

D. Exploring Heterogeneity using Machine Learning

We may draw on an emerging literature that leverages machine-learning methods to explore heterogeneity of causal effects (Chernozhukov et al., 2018; Athey and Imbens, 2015, 2019; Davis and Heller, 2017). This methodology will enable us to learn as much as possible from our data using a disciplined and data-driven approach. Since the "state of the art" is still evolving, we cannot commit to a particular approach at present. However, we plan to pre-specify our approach prior to running this analysis and will interpret our results as suggestive.

E. Multiple Hypothesis Testing

Testing multiple hypotheses raises the likelihood that any one hypothesis is found to be statistically significant purely by chance. We will supplement our results by reporting summary indices that aggregate multiple outcome variables within a common outcome domain. Aggregation not only improves the statistical power within a given domain but also vastly reduces the number of hypotheses examined. This plan pre-specifies what data will be collected, primary and secondary outcomes, the main specification, and subgroups of interest. By committing to a set of analyses in advance, we avoid concerns about data-mining and specification searching, and credibly commit to a few hypotheses that, together, comprise the central test of EmpowerHub's model. Classic p-values will be reported for all outcomes, which will provide a reader with full information that they can use to make multiple hypothesis testing corrections if they desire. We will also conduct non-parametric permutation tests and report permuted p-values for the main sets of analyses following Chetty et al. (2016).[3]

[3] This approach entails randomly re-assigning treatment status to students in the main sample and running the main specification thousands of times to simulate a counterfactual distribution of T-statistics. Relative to this counterfactual distribution, we can then compute permuted p-values as likelihood of observing our realized T-statistic. The same approach can be applied to sets of hypotheses to calculate the likelihood of observing by chance the magnitudes of treatment effects observed in the study.

References

Battistich, V., Schaps, E. & Wilson, N. Effects of an Elementary School Intervention on Students' "Connectedness" to School and Social Adjustment During Middle School. The Journal of Primary Prevention 24, 243–262 (2004).

Berry, V., Axford, N., Blower, S. et al. The Effectiveness and Micro-costing Analysis of a Universal, School-Based, Social–Emotional Learning Programme in the UK: A Cluster-Randomised Controlled Trial. School Mental Health 8, 238–256 (2016). https://doi.org/10.1007/s12310-015-9160-1

Bhatt, Monica P. Randomized evaluation of a school-based, trauma-informed group intervention for young women in Chicago. ScienceAdvances (2023). DOI:10.1126/sciadv.abq2077

Burtis, E., Sofoklis Goulas, et al. Breaking Down Enrollment Declines in Public Schools. Brookings (2024). www.brookings.edu/articles/breaking-down-enrollment-declines-in-public-schools/.

Chetty, Raj et al. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *The American economic review* vol. 106,4 (2016): 855-902. doi:10.1257/aer.20150572

Crean, H. F., & Johnson, D. B. Promoting Alternative Thinking Strategies (PATHS) and elementary school aged children's aggression: results from a cluster randomized trial. American journal of community psychology, 52(1-2), 56–72 (2013). https://doi.org/10.1007/s10464-013-9576-4

Falk, Armin, Fabian Kosse, and Pia Pinger. Mentoring and School Decisions: Causal Evidence. CESifo Working Papers (2020).

Guryan, Jonathan, Christenson, Sandra, Cureton, Ashley, Lai, Ijun, Ludwig, Jens, Schwarz, Catherine, Shirey, Emma and Turner, Mary Clair, The Effect of Mentoring on School Attendance and Academic Outcomes: A Randomized Evaluation of the Check & Connect Program, Journal of Policy Analysis and Management, 40, issue 3, p. 841-882 (2021), https://EconPapers.repec.org/RePEc:wly:jpamgt:v:40:y:2021:i:3:p:841-882.

Hansen, B., Sabia, J. & Schaller, J. In-Person Schooling and Youth Suicide, Journal of Human Resources, vol 59(S), pages S227-S255 (2024).

Mervosh, Sarah, and Francesca Paris. Why School Absences Have 'Exploded' Almost Everywhere. The New York Times (2024). www.nytimes.com/interactive/2024/03/29/us/chronic-absences.html.

Office of the Surgeon General (OSG). Protecting Youth Mental Health: The U.S. Surgeon General's Advisory. US Department of Health and Human Services (2024).

Pulkkinen, L., Nygren, H., & Kokko, K. Successful development: Childhood antecedents of adaptive psychosocial functioning in adulthood. Journal of Adult Development, 9, 251–265 (2002). https://doi.org/10.1023/A:1020234926608

Resnjanskij, S., Ruhose, J., Wiederhold, S., & Woessmann, L. IZA DP No. 14097: Can Mentoring Alleviate Family Disadvantage in Adolescence? A Field Experiment to Improve Labor-Market Prospects. Journal of Political Economy, 132(3) (2021).

Rodríguez-Planas, Núria. "Longer-Term Impacts of Mentoring, Educational Services, and Learning Incentives: Evidence from a Randomized Trial in the United States." American Economic Journal: Applied Economics,4 (4): 121-39 (2012).

Rodríguez-Planas, Núria. "School, drugs, mentoring, and peers: Evidence from a randomized trial in the US," Journal of Economic Behavior & Organization, Elsevier, vol. 139(C), pages 166-181 (2017).

Rose, Evan K., Jonathan T. Schellenberg, and Yotam Shem-Tov. *The Effects of Teacher Quality on Adult Criminal Justice Contact*. No. w30274. National Bureau of Economic Research (2022).

Wong, N. T., Zimmerman, M. A., & Parker, E. A. A typology of youth participation and empowerment for child and adolescent health promotion. American Journal of Community Psychology, 46, 100-114 (2010). https://doi.org/10.1007/s10464-010-9330-0

Zimmerman MA, Eisman AB, Reischl TM, et al. Youth Empowerment Solutions: Evaluation of an After-School Program to Engage Middle School Students in Community Change. Health Education & Behavior, 45(1):20-31 (2018).