

Supporting Statement B for the Youth Transition Exploration Demonstration (YTED)
OMB No. 0960-NEW

B. Collection of Information Employing Statistical Methods

SSA contracted with Mathematica to conduct Youth Transition Exploration Demonstration (YTED) which will help youth with disabilities transition successfully into the adult labor force and competitive, integrated employment. In addition, Mathematica is partnering with the Pennsylvania Office of Vocational Rehabilitation (OVR) to recruit youth and deliver intervention services and the University of Maryland’s Center for Transition and Career Innovation (UMD) to provide training and technical assistance to OVR. The YTED will provide SSA with empirical evidence on the impact of the intervention on youth in several outcome areas: (1) employment and earnings; (2) SSI and SSDI benefit receipt; and (3) other related outcomes, such as satisfaction and well-being. A rigorous evaluation of YTED is important to help SSA and other interested parties assess promising options to improve employment-related outcomes and decrease benefit receipts.

1. Statistical Methodology

Recruitment materials and baseline survey: Mathematica will implement the YTED to all residents in the city of Philadelphia and the surrounding four counties in Pennsylvania (Bucks, Chester, Delaware, and Montgomery), who meet the study’s basic eligible criteria (described below). The recruitment for YTED will occur over a 24-month period from 2024 to 2026.

Selection of eligible participants: Recruitment will occur on a rolling basis as new demonstration-eligible youth become known to OVR. The general approach to recruitment includes OVR staff working both internally and with partners like SSA and the SDP to identify and enroll demonstration-eligible youth. The enrollment target for the demonstration is 700 youth. Mathematica will randomly assign youth who enroll in YTED to the treatment or control group. Treatment group members will receive the TE intervention, and control group members will receive information about OVR and how to apply for OVR services. Enrollees will have an equal probability of being assigned by Mathematica to either experimental group. Random assignment will allow Mathematica to attribute any variation in outcomes between the treatment group and control group to the TE intervention. Baseline characteristics both observed and unobserved should be balanced across experimental groups in expectation because of random assignment. Some imbalances in baseline characteristics, however, might occur in the study sample by chance even if random assignment is implemented correctly.

Mathematica will use block random assignment to ensure balance across key observable characteristics. Prior to random assignment, Mathematica will group enrollees into blocks based on observable baseline characteristics that are likely correlated with outcomes of interest to the YTED. Mathematica will then randomly assign enrollees within each block equally between the treatment group and control group. This approach

will help ensure balance across experimental groups for all the baseline characteristics used to create the blocks. The baseline characteristics used to create the blocks will likely be related to employment, age, and education status.

Volunteer rate. Across a two year-period, OVR will approach as many study-eligible youth as possible to solicit their participation in the demonstration. The volunteer rate for YTED is difficult to predict and is likely to vary by referral source, so a key feature of the YTED enrollment plan is to recruit youth from as many referral sources as possible and to train all OVR staff who interact with youth to act as recruiters. Mathematica will only randomly assign participants who provide consent and complete the baseline enrollment and who become a participant in the demonstration. By design, the response rate for the baseline survey will be 100 percent among demonstration enrollees. The baseline survey will take 15 minutes to complete and will collect information about youth's education, employment, health, sources of income, and demographic characteristics. Most youth will complete the baseline survey on paper, either on their own or with help from OVR staff, and then OVR staff will enter their responses into the online system hosted by Mathematica to conduct random assignment. For some youth, OVR staff may read the questions to the youth in person or on the phone and enter their responses directly into the online system.

Follow-up survey: The follow-up survey will yield information about the outcomes of all 700 demonstration enrollees one year after they enroll in the demonstration. Mathematica will field the survey and expects to achieve an 80 percent response rate.

Qualitative data from the implementation and operations staff: Mathematica will visit sites in the service area during the first quarter of 2026 and 2027. During the visits, Mathematica will collect data from a review of program documentation and conduct semi-structured interviews with OVR staff, its community partners, and UMD staff. Mathematica will design the semi-structured interview guide to collect information about staff's experiences and the changes made to the intervention during implementation. The interview guide will include questions tailored to the experiences of individual staff and a set of core questions to enable the evaluation team to systematically collect information across all staff. Mathematica will determine the format of the site visits (in-person or telephone) based on discussions with the leads at UMD and OVR. Mathematica will decide based on the extent to which key informants are geographically dispersed, their availability to schedule an in-person site visit, and the degree to which meeting in person is needed to establish rapport. Mathematica expects to interview up to 18 people per site visit. Mathematica will select interview respondents based on their role and knowledge of YTED at each stage of implementation. Mathematica will interview all Vocational Rehabilitation Counselor-YTED (VRC-YTED), and other select staff at OVR who can provide the desired information, because OVR is expected to provide the information needed to assess program implementation and fidelity, Mathematica expects a 100 percent response rate for qualitative interviews among OVR staff playing key roles on YTED.

Qualitative data from YTED treatment group members: Mathematica will conduct in-depth interviews by telephone with a convenience sample of twelve total enrollees drawn from the universe of the YTED treatment group. Mathematica will conduct the

three rounds of four interviews in the second quarter of 2025, 2026, and 2027. The goals of the interviews with treatment group members are to (1) understand treatment group members' motivations to enroll in YTED, (2) learn about their perceptions and experiences with TE services, and (3) describe their goals and experiences with employment before and after YTED. Topics for the interviews are in Attachment C.

Mathematica will recruit interviewees who have high or low levels of service engagement to learn about a range of experiences with YTED. Mathematica also conduct outreach until we reach four completed in-depth interviews in each year. Mathematica will send an invitation to potential interviewees by mail. The invitation will describe the purpose of the interview and ask enrollees to call a toll-free number to schedule an appointment to complete an interview.

Interviewers will reach out to respondents by telephone to remind them of the interview before the scheduled appointment using the enrollees' preferred mode of contact. During the interview, the team member leading the interview will obtain the respondents' verbal consent to participate and will also request consent to digitally record the interview. The interviewer will explain the benefits and risks associated with participation confidentiality of the information shared during the interview, and the voluntary nature of participation. The interviewer will inform respondents that they may request that the interviewer suspend recording at any time and will assure them that the interviewer will not request personally identifying information during the interview. The interviewer will use an interview guide, based on the topic list in Attachment C, to conduct the interviews. Each interview will last up to 45 minutes. Mathematica will mail a thank you letter with a \$50 gift card to each respondent who completes the interview.

2. Procedures for Collecting the Information

Mathematica expects that the estimation methods differ more substantively by analysis type than by data collection effort. The following subsections describe the methods that Mathematica plans to use for each of the major analysis components of the YTED:

Process analysis. As discussed in Part A, Mathematica will use site visit and semi-structured interview data to provide a detailed description of the TE model, how it is implemented, the context in which it operates, and the program operations and their fidelity to design. The detailed description will assist in interpreting program impacts, identifying program features, and highlighting necessary conditions for effective program replication or improvement. Mathematica will gather information using a range of data sources to fully describe the programs and activities. Mathematica plans to use the Consolidated Framework for Implementation Research (Damschroder et al. 2009) to guide the collection, analysis, and interpretation of qualitative data. The Consolidated Framework for Implementation Research guides systematic assessment of the multilevel and diverse contexts in interventions implementation and describes the myriad factors that might influence intervention implementation and effectiveness. Using the framework will allow Mathematica to structure the analyses of YTED implementation to produce results based on objective, reliable qualitative data across the key domains related to the program environment, program operations, and strategies to support implementation. For each of these domains, Mathematica will develop measurable constructs that align with research questions for the YTED process analyses shown in

Part A and Attachment C. Based on this framework, Mathematica will be able to readily produce tables that summarize the major process findings in the study's reports. The Consolidated Framework for Implementation Research might also allow Mathematica to make systematic use of qualitative data as part of the impact analysis, enabling an examination of how impacts vary with certain implementation constructs.

Impact analysis. The objective of the impact analysis is to provide statistically valid and reliable estimates of the effects of TE on the employment and benefit-related outcomes of YTED enrollees. Mathematica will rely on randomized experimental design to estimate the causal impacts of TE services available to enrollees through the demonstration. Random assignment will enable Mathematica to estimate the net impact of those services by comparing average outcomes across the treatment and control groups. Mathematica's analysis will focus on intent-to-treat estimates that measure how the offer of TE services affected outcomes of enrollees after participating in the demonstration.

Primary outcomes. To avoid concerns about multiple comparisons and reduce the extent of false positives, Mathematica will prespecify a parsimonious set of primary outcomes. A preliminary list of those outcomes includes employment status, earnings above a substantive threshold, enrollment in school or job training programs, and applications for or receipt of SSI or SSDI benefits. The evaluation will also include results for secondary outcomes related to employment, earnings, OVR services, education, receipt of benefits, expectations, satisfaction, and health. These results, however, will be considered exploratory. This approach strikes a balance between addressing the multiple comparisons problem and maintaining the evaluation's ability to detect policy-relevant impacts. By limiting the number of primary outcomes tested in the impact analysis, this approach reduces the likelihood of false positives without undermining the evaluation's statistical power to detect true impacts on any single outcome.

Model. The main impact model will be a weighted linear regression model as follows:

$$(1) y_i = \alpha + \beta X_i' + \gamma T_i + \varepsilon_i,$$

where i is an index for demonstration enrollees; y is an outcome measure; X is a vector of baseline or pre-baseline enrollee characteristics such as demographic characteristics, disabling condition, socioeconomic characteristics, and employment; T is an indicator of treatment group membership (1 for treatment and 0 for control); α , β , and γ are parameters to be estimated; and ε is a random error term. The value of the coefficient γ will indicate the impact of the TE intervention on the outcome. The regression models will also incorporate elements, such as geographic indicators and nonresponse weights (for survey-based outcomes), that add precision to the impact estimates and minimize the chance of bias because of survey nonresponse.

Mathematica will use a model similar to equation (1) to estimate impacts for select subgroups defined by baseline characteristics and compare impacts across these subgroups.

Benefit-cost analysis. Mathematica will conduct the benefit-cost analysis for YTED using an approach that Mathematica successfully adopted in other evaluations, including

the Benefit Offset National Demonstration (BOND) (Bell et al. 2011), the Youth Transition Demonstration (Fraker et al. 2014), and the Promoting Opportunity Demonstration (POD) (Wittenburg et al. 2021). To do this, Mathematica will develop a comprehensive accounting framework that incorporates a range of perspectives to guide benefit-cost data collection, analysis, and reporting. These perspectives include those of the treatment group members, OVR, SSA, and other government entities.

The accounting analysis will focus on monetized impacts. It is not feasible to monetize the value of some benefits and costs, such as quality of life or social integration stemming from greater employment. Most inputs to the benefit-cost analyses will come directly from the impact estimates, such as the estimated effects on earnings and benefit amounts. Mathematica will supplement the impact estimates in the benefit-cost analysis with other sources of data, including program cost information collected during site visits to OVR.

Degree of accuracy needed for the purpose described in the justification: The expected sample size of eligible enrollees supports an analysis that can reliably distinguish the impact of TE on outcomes from other factors shaping the outcomes of enrollees. Calculating minimum detectable impacts (MDIs) is a standard way to characterize the expected precision of the evaluation's results given the sample sizes and research design. MDIs quantify the smallest true impact that is likely to be found to be significantly different from zero, based on one and two-sided statistical tests of differences.

MDIs for outcomes measured in administrative data for all subjects. Mathematica expects that impact estimates will have sufficient precision to reliably assess policy-relevant impacts of YTED on outcomes. Exhibit B.1 summarizes MDIs for the binary employment status outcome over a range of sample sizes.

With 700 enrollees, the evaluation will be well powered to detect employment rate increases of at least 9.2 percentage points. This is smaller than the analogous impacts observed in four of the six Promoting Readiness of Minors in Supplemental Security Income (PROMISE) programs, which showed employment rate impacts of at least 11 percentage points 18 months after random assignment (Mamun et al. 2019).

The impacts of YTED would need to be somewhat larger than in the main analysis to be reliably detected for subgroups, but Mathematica expects these subgroup analyses will still have enough precision to be informative. For example, with a sample half the size of the enrollment target (350 enrollees), the impacts would have to be at least 13 percentage points. Impacts would need to be correspondingly larger to be reliably detected for more focused subgroups containing smaller percentages of demonstration enrollees.

MDIs for outcomes measured using survey data for subsets of subjects. Mathematica expects to be able to detect modest-sized impacts overall and for all but the smallest subgroups. Because of survey nonresponse, Mathematica expects to have less precision for survey-based outcome measures than for outcomes measured in administrative data. Based on the expected response rate of 80 percent, Mathematica expects a respondent sample size of 560 for the follow-up survey. The MDI for employment with a sample of 560 is 10.3 percentage points.

Exhibit B.1. Minimum detectable impacts for the YTED evaluation

Sample size			Minimum detectable impact (percentage points)	
Treatment	Control	Total	Two-sided test	One-sided test
350	350	700	0.092	0.082
280	280	560	0.103	0.091
250	250	500	0.109	0.097
200	200	400	0.122	0.108
175	175	350	0.130	0.116

Note: Requires at least an 80 percent chance of correctly identifying true impacts as statistically significant using statistical tests with a 5 percent significance level.

Assumes an employment rate in the control group of 30 percent and an impact regression model that explains 10 percent of the variation in employment outcomes.

3. Methods to maximize response rates and data reliability

Designing an informative recruitment process. Informed consent from YTED enrollees will help the evaluation produce more reliable estimates of YTED’s impacts on a study population for whom the treatment is salient. Consequently, the recruitment plan centers around providing demonstration-eligible youth, and their parent or guardian if the youth is younger than age 18 with the information they need to make informed choices. Recruitment will primarily occur in person when OVR staff are helping youth apply for pre-ETS or VR services, visiting schools to deliver pre-ETS or attend IEP meetings, or attending events to promote YTED. OVR staff will begin the recruitment process by explaining YTED and answering any questions. If youth indicate they would like to enroll, OVR staff will confirm the youth’s eligibility and then help them complete the consent form, release form, and baseline survey. Finally, OVR staff will enter information from the consent form, release form, and baseline survey into an online system hosted by Mathematica to conduct random assignment. A secondary recruitment method is for OVR staff to conduct outreach to youth who are referred by a community partner, received pre-ETS in the past, or contact OVR in response to receiving a letter from the Social Security Administration or learning about YTED through the media or social media. OVR staff will contact these youth to confirm their eligibility and then ask them whether they would prefer to meet in person or have the consent form, release form, and baseline survey mailed or emailed to them. For youth who would prefer a meeting, OVR staff will meet with the youth and help them complete the consent form, release form, and baseline survey. For youth who would prefer mail or email, OVR staff will send the consent form, release form, and baseline survey to the youth for them to complete on their own and then mail back to OVR. As noted previously, OVR staff will conclude the recruitment process by entering information from the consent form, release form, and baseline survey into an online system hosted by Mathematica to conduct random assignment.

Data reliability. Mathematica developed the baseline survey instrument using materials developed and fielded on recent similar SSA demonstrations, including Way2Work Maryland (W2WMD), Next Generation of Enhanced Employment Strategies Project (NextGen), and PROMISE and SSA’s current ICAP project with the Kessler Foundation. Several experts on the YTED team including, Mathematica economists; disability policy researchers; survey researchers; information systems professionals and staff at SSA,

UMD; and OVR reviewed the draft instrument and helped refine them further. Mathematica also field tested the instrument with a small number of OVR affiliated youth, as described in Section B.4.1.

Addressing item nonresponse. SSA will not use unit-level nonresponses for the baseline survey by construction. Mathematica's experience suggests that item-level nonresponse will be low for the baseline survey, although some item nonresponse is inevitable. To address this when running subsequent analysis that draw on the baseline survey, Mathematica will consider using one of several standard imputation techniques, as described in Allison (2001), depending on the pattern item nonresponse observed in the final study sample.

Balance of study sample and integrity of random assignment. Mathematica will use an online system to create balanced study groups and minimize on-site contamination risks. Built-in eligibility and duplication checks will prevent staff from enrolling youth who have not completed a baseline survey and provided written consent or who might have previously enrolled. Mathematica will use block random assignment to improve the extent to which the distribution of baseline characteristics is similar in each experimental group. The procedure will create a series of blocks based on important baseline characteristics. Enrollees within each block will be equally (randomly) assigned into the treatment and control groups. Block random assignment will help ensure balance across experimental groups for all the baseline characteristics used to create the blocks. Mathematica will monitor the random assignment process, checking the balance of study groups using data on subjects recorded in the baseline survey. Such characteristics could include the same ones used to block random assignment (as discussed in Section B.2.1) and other measures determined to be of substantive importance during the design phase of the evaluation. Substantial or statistically significant differences (based on *t*-tests and chi-square tests, as appropriate) in characteristics across subgroups and assignment status could prompt Mathematica to adjust or correct the random assignment procedure.

Response rates. Mathematica's approach to the follow-up survey addresses several challenges that can depress response rates. Mathematica will offer the follow-up survey in two modes (web and telephone) to encourage response. Enrollees will receive an advance letter by mail with a link to the web survey and their unique survey login information so they can easily access the web survey. Mathematica designed the survey interview to be brief to encourage response and full completion. As discussed in greater detail in Part A, the evaluation team will offer \$55 in incentives to encourage participation. Mathematica will take several actions to help enrollees complete the follow-up survey. To promote response among Spanish speakers, Mathematica will develop a Spanish-language version of the instrument. Mathematica will train bilingual telephone interviewers to address enrollee questions clearly and effectively and to gain enrollees' cooperation and avoid refusals. For sample members who do not complete the survey by web or respond to telephone calls or who refuse to participate in the survey, Mathematica will mail reminder letters, reminder postcards, and refusal conversion letters, as appropriate (see Attachment B). Finally, Mathematica will train data collection staff on techniques for locating enrollees who are no longer at the address and telephone

number provided at enrollment. The impairments of some sample members will make responding to the survey problematic, especially by telephone. To facilitate responses to the CATI interview, Mathematica will offer the use of several assistive devices (for example, amplifiers and Telecommunications Relay Service) and will instruct interviewers to remain patient, repeat answers for clarification, and identify signs of respondent fatigue.

Data reliability. Mathematica developed the follow-up survey instrument and contact materials using materials developed and fielded on recent similar SSA demonstrations such as W2WMD, NextGen, PROMISE, and SSA’s current ICAP project with the Kessler Foundation. Several experts on the YTED team including Mathematica economists, disability policy researchers, survey researchers, and information systems professionals and staff at SSA, UMD, and OVR, also reviewed the draft follow-up instrument and contact materials and helped refine them further. The evaluation team field tested the instrument with a small number of OVR-affiliated youth, as described in Section B.4.2.

Item nonresponse. Although Mathematica’s past experience conducting surveys for similar evaluations suggests that rates of item nonresponse on the follow-up survey will be very low, some item nonresponse is inevitable. The follow-up survey primarily collects data on outcome measures to be used in the impact analysis. Depending on the pattern of missing data observed, Mathematica will consider alternative multivariate imputation techniques or omitting subjects with missing data on a given outcome when analyzing that outcome.

Individual-level nonresponse. As with almost any survey, some nonresponse in the follow-up survey is inevitable. Some sample members will not be located, and others will not be able or willing to respond to the survey. Mathematica expects to attain a response rate of at least 80 percent based on its experience with prior SSA demonstrations. In the event that response rates are lower, Mathematica will analyze nonresponse using various data items from administrative data records and baseline survey. The nonresponse bias analysis will consist of the following steps:

- *Compute response rates for key subgroups.* Mathematica will compute the response rate for the subgroups using the American Association for Public Opinion Research definition of the participation rate for a nonprobability sample: the number of respondents who provided a usable response divided by the total number of people from whom participation in the survey is requested (American Association for Public Opinion Research 2016). Mathematica will compare the response rate across key subgroups, including most notably the treatment group and the control group, as well as subgroups used for block random assignment. The goal of this analysis is to determine whether response rates in specific subgroups differ systematically from those of other subgroups or from the overall response rate. This could inform the development of nonresponse weights for use in the analysis.
- *Compare the distributions of respondents’ and nonrespondents’ characteristics.* Again, using data from administrative records and the baseline survey, Mathematica

will compare the characteristics of respondents and nonrespondents. They will assess the statistical significance of the difference between these groups using *t*-tests or chi-squared tests. This type of analysis can help identify patterns of differences in observable characteristics that might suggest nonresponse bias. This approach, however, has low power to detect substantive differences when sample sizes are small, and the large number of statistical tests conducted can also result in high rates of Type I error. Consequently, the results of this item-by-item analysis will be interpreted cautiously.

- *Identify the characteristics that best predict nonresponse and use this information to generate nonresponse weights.* This is a multivariate generalization of the subgroup analysis described previously. Mathematica will use logistic regression models to assess the partial associations between each characteristic and response status; propensity scores obtained from such models provide a concise way to summarize and correct for initial imbalances (Särndal et al. 1992). Because of the rich set of program and baseline survey data available for this analysis, Mathematica will use a mixture of substantive knowledge and automated machine learning methods to identify covariates to include in the final weights. Examples of automated procedures they could use to produce these weights efficiently include (1) using prespecified decision rules, such as those described by Imbens and Rubin (2015) and Biggs et al. (1991) to select covariates and interactions between them, and (2) identifying and addressing outliers by, for example, trimming weights in a way that minimizes the mean-square error of the estimates (Potter 1990).
- *Compare the nonresponse-weighted distribution of respondent characteristics with the distribution for the full random assignment sample.* In this last step, Mathematica will compare the weighted distribution of respondents' baseline characteristics with the unweighted distribution of the full set of study subjects that went through random assignment. Mathematica will make these comparisons for the whole sample and for subgroups, as described earlier in this subsection. This step will include validation of the nonresponse weights using outcomes measured in the program data for the full sample (but not used in the construction of the weights). This analysis can highlight the measures with the greatest potential for nonresponse bias, even after weighting, in which case greater caution should be exercised in interpreting the observed findings.

Qualitative data from the implementation and operations staff

Response rates. Mathematica expects that OVR will provide the information required to assess program implementation. OVR and Mathematica collaborated on the proposal plan for site visits, so Mathematica anticipates high levels of cooperation for the qualitative interviews. Mathematica will hold a phone call with the OVR principal investigator to describe the information it will gather from OVR staff and its service providers during site visits and interviews. To minimize burden on site staff and maximize staff availability, Mathematica will work with the OVR principal investigator to determine the most convenient times to convene the interviews. Mathematica will limit the interviews to about one hour so that the data collection imposes only a modest burden on respondents. To facilitate a smooth interview process and improve the

completeness of the data collected, Mathematica will share an information packet to the OVR principal investigator containing the final site visit and interview schedule. Mathematica will send out the packet about two weeks before the site visit. The packet will contain the lead site visitor's contact information so the respondents can reach the visitors in case the schedule changes or other issues arise before the interviews. Mathematica will also send email reminders to the OVR principal investigator several days before the site visit confirming the day and time of the interviews. Providing the local sites with adequate information ahead of time in a professional manner will help build rapport, facilitate a more fluid interview process, and establish that interviewees are available and responsive.

Data reliability. Mathematica interviewers will use an interview guide, based on the interview topic list provided in Attachment C, to conduct the semi-structured staff interviews. They will use separate guides for each potential respondent type (for example, OVR staff and its service providers) so they do not ask respondents about activities or issues that do not apply to them. The interviewers will review program documentation ahead of the interviews to minimize burden and supplement information provided by respondents to the interviews, rather than ask respondents to answer detailed questions about specific operations. They will take notes and obtain permission to record each interview. After the interviewers complete all interviews for a site visit, they will develop a summary of the information collected.

Qualitative data from treatment group members

Response rates. Because Mathematica will draw interviewees from a convenience sample of volunteers, target response rates to ensure a representative population are not at issue. To mitigate interview nonresponse, Mathematica will contact each individual who has an appointment on the day before the interview, using the person's preferred mode of communication. A team member will confirm the best telephone number to use to reach each enrollee so the appointment takes place when the enrollee is not distracted by other responsibilities. Because Mathematica will conduct the interviews by telephone, participants will not face barriers related to transportation to an interview location. Finally, the \$50 gift card will encourage interview participation.

Data reliability. Interviewers will use an interview guide, based on the interview topic list provided in Attachment C, to conduct the semi-structured interviews with treatment group members. Mathematica will train interviewers on best practices for interviewing youth with disabilities. The interviewers will take notes and obtain permission to record each interview.

4. Tests of procedures

Pretesting the baseline survey. Mathematica pretested the baseline questionnaire with six respondents. The instrument did not require revisions based on findings from the pretest. The pretest provided an accurate estimate of respondent burden (15 minutes) as required by the Office of Management and Budget (OMB), and Mathematica also assessed flow and respondent comprehension. Participants in the pretest of the baseline survey received an incentive for their participation.

Refining and testing random assignment procedure. As noted previously (Sections B.1.1 and B.3.1), random assignment will be at the individual level, and Mathematica will use block random assignment to improve the balance of the treatment and control groups. Before starting the recruitment phase, Mathematica will evaluate the usefulness of this randomization approach with fabricated data to (1) verify that it balances appropriately on the stratifying variables using the anticipated batch sizes and (2) assess the risk of imbalance on other variables. Based on this testing, Mathematica might adapt the randomization procedure. Further, after the pilot phase of recruitment is complete, Mathematica might refine the stratification factors and random assignment batch schedule (as discussed in Section B.3.1) based on the flow of study subjects and the prevalence of their characteristics observed to date. Mathematica would evaluate such refinements by again testing the procedure using fabricated data in the manner described previously.

Follow-up survey: Mathematica pretested the follow-up survey instrument with six respondents. The instrument did not require revisions based on findings from the pretest. The pretest provided an accurate estimate. In addition, Mathematica assessed flow and respondent comprehension by debriefing each respondent to determine whether any words or questions were difficult to understand and answer. Similar to actual study subjects, participants in the pretest of the follow-up survey received an incentive for their participation.

Qualitative data from the implementation and operations staff: Mathematica will base site visit protocols on those used for similar evaluations, and it will use the first interview with the OVR principal investigator as a pilot to test the interview protocol. During this pilot test, the interviewer will conduct a cognitive test of the semi-structured interview protocol to establish that the respondent interprets questions as intended and has the information necessary to answer the questions. The interviewer will ask the respondent to explain how they arrived at their answers and whether any items were difficult to answer. At the end of the pilot interview, Mathematica will modify or clarify the interview questions when necessary to improve the data collection tools and procedures. Senior research staff will also assess the site visit agenda, including the data collection activities they will conduct and how these activities are structured, to check that the interviewers can feasibly conduct all interviews as part of the site visits and yield the desired information.

Qualitative data from YTED treatment group members: Mathematica will base participant interview protocols on those used for similar evaluations. Mathematica will conduct no pre-tests of the interview protocols. Mathematica will make minor modifications to the data collection procedures and discussion guides, if necessary, based on the experiences of the early interviews.

5. Statistical Agency Contact for Statistical Information

For further information you can communicate with the following staff members:

David Mann, Ph.D., Senior Researcher
Telephone: 609-275-2365
Email: dmann@mathematica-mpr.com

Anna Hill, Ph.D., Senior Researcher
Telephone: 617-715-6957
Email: ahill@mathematica-mpr.com

Sarah Croke, Senior Researcher
Telephone: 734-205-3083
Email: scroake@mathematica-mpr.com

Stacie Feldman, Survey Researcher
Telephone: 609-936-3250
Email: sfeldman@mathematica-mpr.com

Geraldine Haile, Senior Survey Researcher
Telephone: 202-838-3563
Email: ghailel@mathematica-mpr.com

Karen Katz, Senior Managing Consultant
Telephone: 312-585-3352
Email: kkatz@mathematica-mpr.com

Jody Schimmel Hyde, Ph.D., Principal Researcher
Telephone: 202-554-7550
Email: jschimmel@mathematica-mpr.com

Jeffrey Seabury, Statewide Specialist
Telephone: 724-320-7542
Email: jseabury@pa.gov

Kelli Thuli Crane, Ph.D., Assistant Research Professor
Telephone: 240-418-2684
Email: kcrane@umd.edu

References

- American Association for Public Opinion Research, *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. Ninth edition. Oakbrook Terrace, IL: AAPOR, 2016.
- Allison, Paul D. "Multiple Imputation: Basics," *Sage University Papers Series on Quantitative Applications in the Social Sciences*. Thousand Oaks, CA: Sage Publications, 2001.
- Bell, Stephen, David C. Stapleton, Daniel Gubits, David Wittenburg, Michelle Derr, David Greenberg, Arkadipta Ghosh, and Sara Ansell. "BOND Implementation and Evaluation: Evaluation Analysis Plan." Cambridge, MA: Abt Associates, and Washington, DC: Mathematica Policy Research, 2011.

- Biggs, David, Barry de Ville, and Ed Suen. "A Method of Choosing Multiway Partitions for Classification and Decision Trees." *Journal of Applied Statistics*, vol. 18, no. 1, 1991, pp. 49–62.
- Damschroder, L.J., D.C. Aron, R.E. Keith, S.R. Kirsh, J.A. Alexander, and J.C. Lowery. "Fostering Implementation of Health Services Research Findings into Practice: A Consolidated Framework for Advancing Implementation Science." *Implementation Science*, vol. 4, no. 50, 2009.
- Eichel, L., and K. Martin. "Disability Rate in Philadelphia Is Highest of Largest U.S. Cities." Philadelphia, PA: The Pew Charitable Trust, 2018. Available at <https://www.pewtrusts.org/en/research-and-analysis/articles/2018/07/17/disability-rate-in-philadelphia-is-highest-of-largest-us-cities>.
- Fraker, Thomas, Arif Mamun, Todd Honeycutt, Allison Thompkins, and Erin Jacobs Valentine. "Final Report on the Youth Transition Demonstration." Washington, DC: Mathematica Policy Research, October 16, 2014.
- Imbens, Guido, and Donald Rubin, *Causal Inference in Statistics, Social, and Biomedical Sciences*. New York: Cambridge University Press, 2015.
- Mamun, A., A. Patnaik, M. Levere, G. Livermore, T. Honeycutt, J. Kauff, K. Katz, A. McCutcheon, J. Mastrianni, and B. Gionfriddo. "Promoting Readiness of Minors in SSI (PROMISE) Evaluation: Interim Services and Impact Report." Washington, DC: Mathematica, July 3, 2019.
- Potter, Francis J. "A Study of Procedures to Identify and Trim Extreme Sampling Weights." In *Proceedings of the American Statistical Association, Section on Survey Research Methods*. Alexandria, VA: American Statistical Association, 1990, pp. 225–230.
- Särndal, Carl-Erik, Bengt Swensson, and Jan Wretman. *Model-Assisted Survey Sampling*. New York: Springer-Verlag, 1992.
- School District of Philadelphia. "Special Education in the School District of Philadelphia: Recognizing the Landscape, 2019–2020." Philadelphia, PA: The School District of Philadelphia Office of Research and Evaluation, 2021. Available at <https://www.philasd.org/research/wp-content/uploads/sites/90/2021/10/Special-Ed-Landscape-2019-20-Research-Report-October-2021.pdf>.
- U.S. Census Bureau. "American Community Survey. 2020: Disability Characteristics, Philadelphia." U.S. Census Bureau, 2020. Available at <https://data.census.gov/cedsci/table?q=disability%20philadelphia%20city>.
- Wittenburg, David, Michael Levere, Sarah Croake, Stacy Dale, Noelle Denny-Brown, Denise Hoffman, Rosalind Keith, David R. Mann, Rebecca Coughlin, Monica Farid, Heather Gordon, Rachel Holzward, and Shauna Robinson. "Promoting Opportunity Demonstration: Final Evaluation Report." Report submitted to the Social Security Administration. Washington, DC: Mathematica, December 30, 2021.