The Long-Term Impacts of the Technical and Vocational Vouchers Program (TVVP) and Start-up Capital for Youth (SCY) Program in Kenya

AEA RCT Title: Start-up Capital for Youth

RCT ID: AEARCTR-0000632

Principal Investigators: Joan Hamory (University of Oklahoma), Michael Kremer (Harvard University), Isaac Mbiti (University of Virginia), Edward Miguel (University of California, Berkeley), and Michael Walker (University of California, Berkeley)
Additional Contributor: Michelle Layvant (Center for Effective Global Action)

Date: August 20, 2020

Summary: This document outlines the analysis used to estimate the impacts of the Technical and Vocational Vouchers Program (TVVP) and Start-up Capital for Youth (SCY) program in Kenya using data from the Kenya Life Panel Survey (KLPS). The TVVP randomly allocated vouchers to support individuals through a vocational training course of their choice, and SCY subsequently randomly allocated unconditional cash grants to support the start-up or expansion of small businesses, among the same sample of individuals who were part of the TVVP. The near term impacts of these programs have been analyzed using earlier data collection rounds.

This document describes the analysis we propose to undertake to study the longer-term impacts of these programs based on a new round of data collection (KLPS Round 4, which is ongoing) to estimate effects on the vocational training voucher and unconditional cash grant programs both alone and in combination with each other. Our estimating equations follow the Hicks et al. 2017 pre-analysis plan which outlines the analysis of the TVVP and SCY programs using data collected through 2016, and Fernald et al. 2019, which describes related intergenerational effects. We look at an analogous set of outcomes as to the long-run effects of deworming, pre-specified in Baird et al. 2017 and Baird et al. 2019.

1. Introduction

1.1. Summary

Youth unemployment is a major economic and social problem in Africa. Many countries have tried various programs to support youth transitions to the workforce, including supporting vocational training programs and providing cash grants; however, there is mixed evidence on the impact of such programs. In 2008, 2,163 individuals from the KLPS participated in a RCT that provided vocational training vouchers (the TVVP), and in mid-2013, these same individuals participated in a second RCT that awarded unconditional cash grants to support entrepreneurship (the SCY program). Medium-term results of these interventions show that labor market gains are small. After up to 7 years, the vocational training voucher treatment group did not experience any significant impacts on hours worked, total earnings, self-employed profits, or wage earnings with some evidence of decreasing agricultural earnings over time and somewhat higher unemployment. Receipt of the cash grant led to large short-run gains to hours worked and self-employment profits (for both genders), but these dissipated over time and are close to zero two years later; however, there is suggestive evidence that agricultural earnings are persistently higher for cash grant recipients (Hicks et al. 2018).

This document lays out the specific analyses we intend to conduct for analyzing the longer-term effects of these two RCTs, 2-5 years after our last survey round and roughly 5 and 10 years after the TVVP and SCY interventions. This plan captures our current thinking about analyses with this data but we anticipate carrying out some additional analyses beyond those included in this plan, so this document is not meant to be comprehensive nor to preclude additional analysis. Given the broad range of outcomes and their interest to different audiences, we expect the results to be published in multiple papers, though it is hard to know the breakdown in advance of results. In addition, we are interested in studying the intergenerational effects of these programs, which are described in further detail in Fernald et al. 2019.

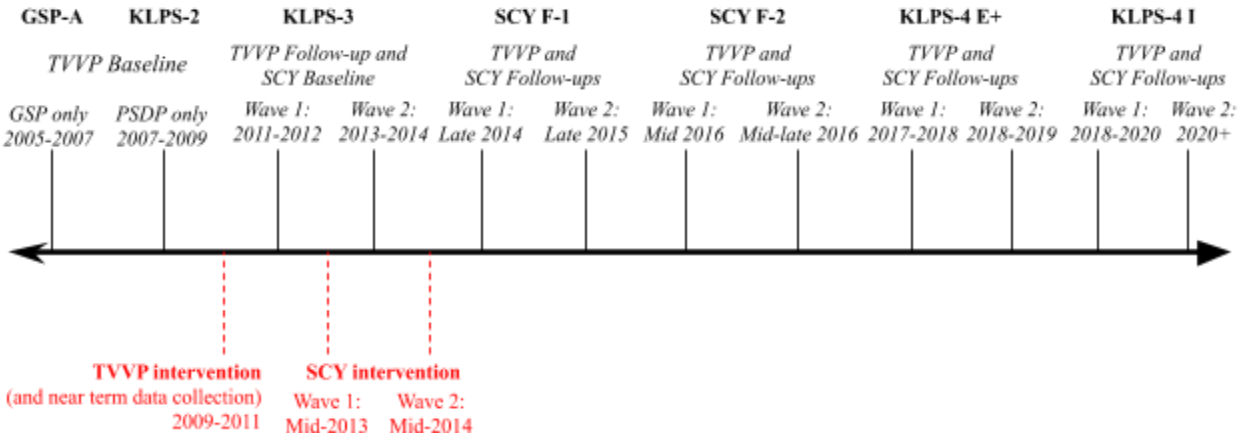
1.2. Experimental Design and Previous Surveys

The vocational training and cash grant program we study, which took place during 2009-2014, included 2,163 adolescents and young adults ranging from roughly 17 to 28 years of age who applied for vocational education tuition vouchers. Approximately 70% of these individuals were participants of the Primary School Deworming Program (PSDP, Miguel and Kremer 2004), and the others were participants in the Girls' Scholarship Program (GSP), a separate randomized education intervention that took place in a neighboring area (Kremer, Miguel and Thornton 2009). A randomly selected half of all training program applicants were awarded a vocational training voucher worth approximately 35,000 Kenyan shillings (about US $460), an amount sufficient to fully (or almost fully) cover the tuition costs for most public or private vocational

education programs in Kenya. Voucher winners attended courses during 2009-2011. In 2013 and 2014, a random half of voucher winners and voucher non-winners were given an unconditional cash grant worth Ksh 20,000 (about US $230 at the time). In the present analysis, we consider voucher winners as "treated" with respect to the vocational training program if they were randomly selected to receive a voucher, and the cash grant winners as "treated" with respect to the cash grant program if they were selected to receive a grant.

For further details on the experimental design and results from analyses of earlier data collection efforts, see Hicks et al. 2017, Hicks et al. 2015a, and Hicks et al. 2015b. The timeline below summarizes the relevant interventions and follow-up surveys.

Figure 1: Timeline of Surveys



Note: KLPS-3 is for both GSP and PSDP participants. SCY F1 and SCY F2 are for participants of the TVVP-SCY study only. KLPS-4 is for TVVP-SCY study participants and PSDP participants.

1.3. Analysis conducted to date

The main goal of this document is to commit to showing long-term outcomes pre-specified as part of other KLPS-4 pre-analysis plans (Baird et al. 2017, Baird et al. 2019) for the TVVP and SCY samples. As described in Section 2.2, the econometric approach closely follows analysis plans that have already been filed.

As shown in Figure 1, the KLPS-4 E+ Module (focusing on economic outcomes) was collected between 2017 and 2019. The first half (Wave 1) of the KLPS-4 I Module was collected from 2018-2020, the second wave is on hold due to COVID-19.

Primary analysis of KLPS-4 data to date has focused on the long-term effects of the PSDP program (Hamory et al. 2020). To date, in addition to tracking and summary statistics, the

authors have seen some results tied to the KLPS-4 I Module Wave 1. As the analyses included below are drawn from already-specified equations and outcomes, these results have not influenced the outcomes or specifications included in this plan.

2. Analysis

2.1. Sample for analysis

The TVVP and SCY sample includes 2,163 individuals -- 1,503 individuals participated in PSDP and 660 individuals participated in GSP. Approximately three quarters (1,541) of these individuals were randomly selected to receive either vocational training vouchers or cash grants (or both). Approximately one quarter (622) of these individuals, the double-control group, received neither the vocational training voucher nor the cash grants.

2.2. Econometric Approach

Approach 1 is a cross-sectional analysis of the KLPS-4 data, which estimates the long-term impacts of the TVVP and SCY interventions.[1] Approach 2 pools economic outcomes[2] from KLPS-3, SCY F1, SCY F2, and KLPS-4 E+ when available in order to estimate average effects across all rounds.[3] Approach 3 estimates effects for our primary outcomes in other domains[4] measured in KLPS-3, SCY F1, SCY F2, and KLPS-4, making use of the most recent data available.[5] Due to the nature of the tracking activity, there will be some respondents surveyed in one survey round that are not surveyed in another, or that did not provide usable data in a survey round. Because some outcome domains are expected to remain relatively stable across survey rounds, including all of these respondents in the analysis may increase statistical precision. Finally, Approach 4 is a cross-sectional analysis measuring the evolution of effects over time employing data from KLPS-3, SCY F1, SCY F2, and KLPS-4 when available.

See Hicks et al. 2017 for the list of treatment variables and list of covariates. As described in Baird et al. 2019, in cases where covariates are missing but outcomes are available, we will follow an approach based on Lin, Green and Coppock 2016: If a covariate is missing for no more than 10 percent of observations, then we will recode the covariate to the overall mean. If a covariate is missing for more than 10 percent of observations, then we will recode the covariate to the overall mean and add in an indicator equal to one for observations with the missing covariate.

---

[1] This approach is in line with Approach 1 in Baird et al. 2017, Baird et al. 2019, and Hicks et al. 2017.
[2] See Section 3 for a list of the economic outcomes. These are also prespecified in Baird et al. 2017.
[3] This approach is in line with Approach 2 in Baird et al. 2017.
[4] See Section 3 for a list of the outcomes in other domains. These are also prespecified in Baird et al. 2019.
[5] This approach is in line with Approach 2 in Baird et al. 2019.

In an extension of Approaches 1-4, which will be carried out separately for TVVP and SCY, we will also examine the complementary of being assigned to both the TVVP and SCY treatment groups. See Hicks et al. 2017 for the regression specification of Approach 1 with an interaction term between TVVP and SCY. We do not expect our study design to have sufficient statistical power to generate precise estimates for these interaction terms, and hence such analyses should be considered suggestive rather than definitive. The patterns that emerge will likely stimulate further exploratory analysis using the dataset.

As described in Hicks et al. 2017, we will make use of baseline controls (ANCOVA) when available to improve statistical precision.

2.3. Heterogeneous effects

We will run heterogeneity analyses on all dimensions prespecified in Hicks et al. 2017. We estimate effects separately by gender, and will statistically test for differences by gender by interacting treatment status with an indicator for gender. In addition to gender, we will estimate heterogeneous treatment effects in the following dimensions:

- Age (at baseline in 2008)
- School attainment (at baseline)
- Indicator for enrolled in vocational training (by 2008 at baseline)
- English test score (for the subset of individuals with this information at KLPS-2 or GSP-A)
- Ravens Matrices test score (for the subset of individuals with this information at KLPS-2 or GSP-A)
- Urban status (at baseline)
- Stated preference for a particular vocational training course/occupation (at baseline)
    - The primary occupation groups that we will focus on include: (i) construction and related trades, (ii) textiles and tailoring, (iii) mechanics and driving, (iv) beauty, (v) computers/secretarial/business, and (vi) other
- Indicators for participation and treatment status in the earlier PSDP and GSP programs
- Risk aversion at baseline (available for KLPS-2 and -3, not for GSP-A)
- Parental educational attainment

In addition, we may conduct exploratory analyses on a number of other dimensions of heterogeneity.

2.4. Multiple Testing Adjustment

As described in Baird et al. 2019, we plan to report the the standard “per comparison” p-values for TVVP and SCY treatment variables, as well as the False Discovery Rate (FDR) adjusted

q-values separately for each treatment over the primary outcomes within a domain, across all of the primary outcomes included within a single paper, across all components of an index, and on interaction terms for outcomes within each domain when testing for heterogeneous effects.

2.5. Tracking and Attrition

We will report effective tracking rates by TVVP and SCY treatment arm, both overall and by gender. Second, we will estimate average baseline differences in terms of baseline covariates using standard two-sample t-tests between those found and not found during the KLPS-4 E+ and KLPS-4 I-Module tracking activity. These covariates include:

- Indicators for treatment status in the earlier PSDP and GSP programs
- Gender
- Baseline preferred training course (see Hicks et al. 2017)
- Age (at baseline in 2008)
- School attainment (at baseline)
- Indicator for any prior enrollment in vocational education (at baseline)
- Stated preference for a private (versus public) training institution (at baseline)

If we observe differential rates of attrition across the treatment and control groups, we will investigate the robustness of our primary results by presenting i) estimates reweighted to account for attrition (IPW), ii) upper and lower bounds on impact estimates (Lee, 2009), as well as iii) adjustments using the techniques of Kling and Liebman 2004 as described in Baird et al. 2019.

3. Main Outcomes

The primary outcomes of interest include all economic outcomes pre-specified in Baird et al. 2017, as well as all additional domains pre-specified in Baird et al. 2019. The six economic outcome domains pre-specified in Baird et al. 2017 include: (1) household consumption, (2) household wealth, (3a) individual earnings, (3b) household earnings, (4) labor supply, and (5) occupational choice. The eleven outcome domains pre-specified in Baird et al. (2019) include: (1) education, training, and human capital, (2) health and well-being, (3) migration, (4) financial participation, (5) marriage and fertility, (6) gender equity attitudes and behaviors, (7) ethnic identity and religious identity, (8) political behavior, (9) political attitudes, (10) time use, and (11) sleep.

In addition to the outcomes pre-specified in Baird et al. 2017 and Baird et al. 2019, there are additional outcomes pre-specified in Hicks et al. 2017 that will be considered for exploratory analyses. Finally, we may explore occupational choice within self-employment (in addition to wage employment) as additional exploratory outcomes.

As mentioned above, given the broad range of outcomes and their interest to different audiences, we expect the results to be published in multiple papers, though it is hard to know the breakdown in advance of results.

References

Baird, S., J. Hamory Hicks, M. Kremer and E. Miguel (2017). Pre-Analysis Plan for "The 20-year Impacts of Child Deworming in Kenya." AEA RCT registry. RCT ID: AEARCTR-0001191.

Baird, S., J. Hamory Hicks, M. Kremer, E. Miguel, and M. Walker (2019). Pre-Analysis Plan for "The 20-year Impacts of Child Deworming in Kenya: Additional Domains." AEA RCT registry. RCT ID: AEARCTR-0001191.

Fernald, L., J. Hamory Hicks, P. Kariger, E. Miguel and M. Walker (2019). Pre-analysis plan for "Estimating Causal Intergenerational Impacts of Parent Human Capital Interventions in Kenya," AEA Trial Registry: AEARCTR-0003995.

Hamory, Joan, Edward Miguel, Michael Walker, Michael Kremer, and Sarah Baird. (2020). "Twenty Year Economic Impacts of Deworming," unpublished working paper.

Hicks, J. Hamory, M. Kremer, I. Mbiti, and E. Miguel. (2015a). Start-up Capital for Youth: Assessing the Potential of Small Business Grants and Vocational Training in Kenya. Presentation at the 6th IZA/Kauffman Foundation Workshop on Entrepreneurship Research, September 17, Washington, D.C.

Hicks, J. Hamory, M. Kremer, I. Mbiti and E. Miguel. (2015b). Vocational Education in Kenya: A Randomized Evaluation. Unpublished Manuscript.

Hicks, J. Hamory, M. Kremer, I. Mbiti and E. Miguel. (2017). Pre-analysis Plan for the Evaluation of the Technical and Vocational Vouchers Program (TVVP) and Start-up Capital for Youth (SCY) Program in Kenya.

Hicks, J. Hamory, M. Kremer, I. Mbiti, and E. Miguel. (2018). Impacts of Vocational Education and Cash on Kenyan Youth. Unpublished Presentation.

Kling, J. and J. Liebman. (2004). "Experimental Analysis of Neighborhood Effects on Youth," Working Paper Series 04-034, Harvard University, John F. Kennedy School of Government.

Kremer, M., E. Miguel and R. Thornton (2009). Incentives to Learn. *Review of Economics and Statistics*, 91(3): 437-456.

Lee, D.S. Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, 76.3: 1071-1102 (2009).

Lin, W., D.P. Green and A. Coppock (2016). "Standard operating procedures for Don Green's lab at Columbia." Version 1.05, June 7, 2016. Available at: https://github.com/acoppock/Green-Lab-SOP .

Miguel, E., and M. Kremer (2004). "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica* 72.1: 159-217.