

Pre-Analysis Plan for:

What are the benefits of mentorship for aspiring micro-entrepreneurs? An examination of mixed gender, same gender, and refugee-host and host-host mentoring pairs and the effects of mentoring and mentoring with perspective-sharing over simple cash transfers.

Georgetown University, Princeton University,
and the International Rescue Committee

March 22, 2023

Abstract

This project works with aspiring micro-entrepreneurs from vulnerable host and refugee groups in Nairobi, Kenya, and with more established Kenyan entrepreneurs recruited as mentors for these aspiring micro-entrepreneurs. This project's interventions offer capital support to the aspiring micro-entrepreneurs, and for a random subsample, a mentoring program. The mentoring program aims to improve economic outcomes, but also social outcomes—the mentee's confidence, social attitudes, and well-being. For the mentor-mentee combinations featuring mixed gender or refugee-host mentor-mentee pairings, we also test for social cohesion improvement—increased economic interactions across identity groups outside of the mentorship context, and improved attitudes toward refugees and women, and toward policies aimed at improving those groups' welfare.

Previous laboratory and field experiments suggest that perspective-sharing, i.e. hearing directly from the other about their perspective and experience, leads to improved attitudes toward and increased support for policies related to minoritized populations. The mentoring intervention uses this psychological insight and is randomly administered to half of the mentoring pairs. We examine any added effect of perspective-sharing informed mentorship in business mentorship pairs on economic, psychological, and social cohesion outcomes.

First, we construct cohorts based on geographical proximity. Then, by cohort, stratifying on gender and country of origin, we randomly assign aspiring micro-entrepreneurs to one of four research arms: (i) a control group, which will receive treatment after completion of the study; (ii) business grant only; (iii) business grant and mentorship; (iv) business grant and mentorship with perspective-sharing. Participants in the active treatment arms receive a business grant one month after the start of the program. Participants in the mentorship treatment arms engage in nine pairwise meetings between aspiring and more established micro-entrepreneurs over the course of two and a half months. These meetings are centered around either just economic/business content or economic/business content and perspective-sharing, i.e. building a shared understanding of one another's experiences and circumstances, between the mentor and the mentee. We examine whether mentorship acts as a complement to the business grant, thereby improving economic and psychological outcomes, and whether these effects are more pronounced for those in the perspective-sharing informed mentorship research arm.

Taking into consideration geographical proximity, gender, and industry of interest, aspiring micro-entrepreneurs in groups (iii) and (iv) are randomly matched with eligible mentors.

By cohort, stratifying on gender, recruited mentors are similarly randomly assigned to one of three research arms: (i) control; (ii) mentorship; and, (iii) mentorship with perspective-sharing. We then randomly specify aligned vs. misaligned nationality/residency status mentee matches for mentors assigned to groups (ii) and (iii). For mentors, we examine whether perspective-sharing fosters greater interpersonal cohesion and whether this translates into improved social and political attitudes towards refugees and women for misaligned matches. We further examine whether misaligned matches play a role in expanding and diversifying the mentors' business interactions and networks. Across both populations, we measure outcomes 1, 3, 6, and 9 months after the first meeting.

Contents

1	Overview	1
2	Experimental Design	3
2.1	Intervention	3
2.2	Timeline	5
2.3	Random Assignment	7
2.4	Re-Assignment and Mop-Ups	7
3	Final Study Sample	8
4	Hypotheses and Outcomes for Aspiring Microentrepreneurs/Mentees	9
4.1	Primary Outcome Domains	9
4.2	Secondary Outcome Domains	12
4.3	Outcomes and Hypotheses Summary Tables	15
5	Hypotheses and Outcomes for Mentors	18
5.1	Primary Outcome Domains	18
5.2	Secondary Outcome Domains	19
5.3	Outcomes and Hypotheses Summary Tables	22
6	Data collection and post-processing	25
6.1	Data collection	25
6.2	Data post-processing	25
7	Statistical analysis	26
7.1	Estimation	26
7.2	Testing	26
7.3	Multiple hypotheses	27
7.4	Compliance	27
7.5	Attrition	28

1 Overview

This study goes through 8 major steps:

1. **Registration of prospective mentees and mentors** : Our implementation partner, the International Rescue Committee (IRC), registered approximately 2000 aspiring microentrepreneurs from vulnerable host and refugee groups in Nairobi, Kenya, who expressed interest in receiving microenterprise support services from the IRC. The IRC aimed to register roughly equal numbers of refugees, hosts, men and women. The IRC then recruited approximately 1800 established Kenyan entrepreneurs to serve as mentors for these aspiring microentrepreneurs. Mentors were required to speak English and Swahili; have an existing business; have at least 3 years experience running a business; and, express a willingness to participate in and commit to the program requirements. The IRC aimed to recruit roughly equal numbers of men and women.
2. **Sampling and cohort assignment** : At the time of registration, all participants were told that the IRC was partnering with external researchers to conduct a research study and were asked for their consent to share their registration data with the researchers. The IRC shared the data of all those who consented to participate in the study. These participants constitute the full study sample and were divided into 6 cohorts based on geographical proximity. Cohort size varied depending on the number of participants registered in any given location or site. The final cohort, corresponding to the site of Rongai, had an excess registration of mentees compared to the study target sample size. Some participants were thus pre-emptively excluded from the study sample at that stage in order to respect budget constraints. The random sampling of mentees in that final cohort was stratified by region of origin, in order to maintain the original target split of Kenyans vs refugees.

From this point onward, the study was rolled out on a cohort by cohort basis to take into account the implementation capacity constraints of the IRC.

3. **Baseline survey** : All registered and eligible mentees and mentors were visited by Georgetown University's field team at their place of residence for a 90-minute baseline survey collecting demographics, some economic, social, and psychological information, as well as running a simplified version of the dictator game. About 93% of the registered mentees and 90% of the registered mentors were successfully reached at that stage. Individuals who were not successfully reached at that stage (attrition at baseline) were dropped from the study sample.
4. **Research arm assignment** : Cohort by cohort, potential mentors and mentees were assigned to one of the research arms. The assignment probability for potential mentees to each research arm is consistent across cohorts. However, assignment probability for mentors to the different research arms varies by cohort and is contingent on the number of mentees assigned to the mentorship treatment arms per cohort (See Section 2.3 for more details). On top of being assigned to one of the research arms, participants in the mentorship treatment arms are also assigned a specific pair (mentor/mentee respectively) for the study, and mentors are assigned to a specific class for their mentor training.
5. **Implementation of the program** : For each cohort, the research team would communicate the random assignments to the IRC, and the IRC rolled out the program for all research arms. Depending

on the treatment assignment, the program for any given individual consisted of one or more of these components : an informational phone call, a training (for mentors), an information session, an introductory meeting with their assigned pair, a business grant, and eight 1:1 meetings between mentor and mentee. Components of the program for each treatment arm are summarized in Figure 1.

- For the aspiring microentrepreneurs, the IRC informed the control group that they had not been selected to receive services at this time, but were still registered at the IRC and will receive services as soon as possible. All aspiring microenterpreneurs in the active treatment arms received a business grant one month after the start of the program. Participants in the business grant only arm were invited for an information session held by the IRC. During this information session, participants were provided additional details about the program and a timeline for when to expect their business grant. Participants in the mentorship arms were also invited to an information session in small pre-assigned groups where they are briefed about the program and what to expect in the coming months. After these information sessions, they were introduced to their mentors and had their first 1:1 meeting with their mentor on the same day. These participants were then expected to engage in eight additional weekly 1:1 meetings with their mentors.
 - For the mentors, the IRC informed the control group that they had not been selected to be a mentor with the IRC at this time but that their name will be placed on a wait-list in case there is a need for mentors in the future. Mentors assigned to the mentorship arms participated in a half day training in small pre-assigned groups. The training for the mentors in the perspective-sharing mentorship arm had a different focus from the regular training in that it emphasized active non-judgmental listening and the value of understanding the perspective and life circumstances of the other when offering them advice. This difference in focus is also reflected in the outline of their subsequent meetings with their mentees. On the day of the training and after the training was completed, mentors were matched with their mentees and held their first 1:1 meetings with the mentee on the same day. These mentors were then expected to engage in eight additional weekly 1:1 meetings with their mentees.
6. **Mop-ups** : All participants (mentors and mentees) were given two opportunities to participate in the program. If they missed their first training or information session, they could participate in a mop-up organized 1-2 weeks later for their cohort. This first set of mop-up meetings were managed by the IRC as part of implementation directly and kept existing mentor-mentee pairs. On top of that, if their assigned pair missed the class or information session and the mop-up opportunity or was no longer interested in the program, participants could be re-assigned to another person and begin their program during a second set of mop-up meetings scheduled for the last week of implementation for a given cohort. The re-assignment process is detailed in section 2.4 of the Experimental Design section of this document.
 7. **Follow up surveys** : We invited all aspiring microentrepreneurs and mentors who were randomized into the study sample to take follow-up surveys 1, 3, 6 and 9 months after the start of the intervention. We made up to three attempts to reach any given participant for any given survey round.
 8. **Data analysis** : The analysis of the baseline and follow up survey data was done on an ongoing basis. Prior to the publication of the pre-analysis plan, the only group for which any follow up data was analyzed is the control group. This allowed the research team to make better informed decisions on

how to construct the outcomes of interest, given the experimental nature of many measurements used, while not jeopardizing the need for blindness prior to defining the full analytical strategy. Data storing was compartmentalized to ensure no accidental analysis of individuals assigned to any treatment arm was conducted.

Overall, the study included two pilots and a randomized controlled trial (RCT). The pilots ran from February-June 2022 and were intended to refine the intervention and the implementation procedures and to pilot the measures we were expecting to field as part of the RCT. These pilots only included specific aspects of the full design that required pre-testing. The pilot sample and data are not considered a part of the study and have not been analyzed for any treatment/control differences.

Given the cohort by cohort roll out of the program, we outline the full timeline per site and cohort below. Each site corresponds to one cohort, except Pangani which was split into three cohorts given its size. We will pool the data from all cohorts for our analyses. Detailed cohort-by-cohort timeline is described in the Experimental Design section of this document in section 2.2.

Table 1: Data Collection and Program Roll-out Timeline Per Site and Cohort

Cohort	Baseline	Implementation	1-month follow-up	3-month midline	6-month midline	9-month midline
Kawangware	July 2022	Sept 2022	Oct 2022	Dec 2022-Jan 2023	Mar 2023	June 2023
Pangani 1	Aug-Sept 2022	Sept 2022	Oct-Nov 2022	Jan-Feb 2023	Apr 2023	June-July 2023
Pangani 2	Sept 2022	Oct 2022	Nov 2022	Feb-March 2023	Apr-May 2023	July 2023
Pangani 3	Sept-Oct 2022	Nov 2022	Nov-Dec 2022	March 2023	May 2023	July 2023
Kitengela	Oct 2022	Jan 2023	Feb 2023	Apr 2023	July 2023	Oct 2023
Rongai	Oct-Nov 2022	Jan-Feb 2023	Feb-Mar 2023	May 2023	July-Aug 2023	Oct 2023

2 Experimental Design

2.1 Intervention

The intervention components are summarized in the Figure 1 below.

Mentees in the program are randomly assigned to one of four research arms:

- Research Arm 0: Delayed treatment (Control). Participants assigned to this research arm are provided services after completion of the research, approximately 1 year after the start of the intervention for any given cohort.
- Research Arm 1: Business grant. Participants assigned to this research arm are invited for one information session and given a labeled business grant of KSH 56,000 to start a microenterprise. Participants are explicitly told that this grant is for their business, however, there is no subsequent monitoring of how they spend the funds.
- Research Arm 2: Business grant and mentorship focused on business content. Participants assigned to this research arm receive a labeled business grant of KSH 56,000 to start a microenterprise and eight weekly 1:1 mentorship sessions with an experienced entrepreneur, on top of their introductory meeting combined with the initial information session. This mentoring relationship focuses on a transfer of business/economic knowledge.

Timeline	Program Component	Mentees				Mentors		
		Control	Grant Only	BC Mentorship	PS Mentorship	Control	BC Mentorship	PS Mentorship
Week -1	Assignment Phone Call	Yes	Yes	Yes	Yes	Yes	Yes	
Day 1	Mentorship Training	-	-	-	-	No	Yes	Yes
Day 1	Stipend for Attending Mentorship Training	-	-	-	-	No	Ksh 3,600	Ksh 3,600
Day 1	Information Session	No	Yes	Yes	Yes	-	-	-
Day 1	Intro Meeting with Assigned Pair	No	No	Yes	Yes	No	Yes	Yes
Day 1	Transport Stipend for Info Session and Intro Meeting	No	Ksh 400	Ksh 400	Ksh 400	No	Ksh 400	Ksh 400
Weeks 2-9	Weekly Meetings with Assigned Pair	No	No	Yes (x8)	Yes (x8)	No	Yes (x8)	Yes (x8)
Week 2	Data for Downloading Mentorship Videos	No	No	Ksh 400	Ksh 400	No	Ksh 400	Ksh 400
Week 2 & 5	Stipend for Mentorship Time	-	-	-	-	No	Ksh 2,600 (x8)	Ksh 2,600 (x8)
Week 2 & 5	Transport Stipend for Attending Mentorship	-	-	Ksh 400 (x8)	Ksh 400 (x8)	-	-	-
Week 4	Business Grant	No	Ksh 56,000	Ksh 56,000	Ksh 56,000	-	-	-
Total # Meetings		0	1	9	9	0	9	9
Total Cash Transfers		0	Ksh 56,400	Ksh 60,000	Ksh 60,000	0	Ksh 25,200	Ksh 25,200

Figure 1: Components of the Intervention for each Research Arm

- Research Arm 3: Business grant and mentorship focused on business content and perspective-sharing. Similar to research arms 1 and 2, participants assigned to this research arm also receive a labeled business grant of KSH 56,000 to start a microenterprise. In addition, they receive eight weekly 1:1 mentorship sessions with an experienced entrepreneur, on top of their introductory meeting combined with the initial information session. This mentoring relationship focuses on a transfer of business/economic knowledge and includes a perspective-sharing component that seeks to build trust and a shared understanding of experiences and circumstances of both the mentor and the mentee.

Mentors in the program are randomly assigned to one of five research arms detailed below. Research arms 1 and 2 both consist of mentorship, however, research arm 1 focuses on a transfer of business knowledge as the primary purpose of the 1:1 mentoring sessions, whereas research arm 2 has a relationship building component through perspective-sharing. a denotes an aligned/shared nationality for the mentor-mentee pair and b denotes a misaligned nationality for the mentor-mentee pair.

- Research Arm M0: Control. Mentors assigned to this research arm are told that the IRC is not able to match them with mentees at this time and that they will be added to a list and considered for future opportunities with the IRC.
- Research Arm M1a: Mentorship focused on business content with an aligned nationality mentee. Mentors assigned to this research arm held 1:1 mentoring sessions focused on transferring business content and knowledge to the mentee. Mentors in this research arm were randomly assigned to a mentee of the same nationality as them.
- Research Arm M1b: Mentorship focused on business content with a misaligned nationality mentee. Mentors assigned to this research arm held 1:1 mentoring sessions focused on transferring business

content and knowledge to the mentee. Mentors in this research arm were randomly assigned to a refugee mentee of a different nationality as them.

- Research Arm M2a: Mentorship focused on transferring business knowledge and on perspective-sharing with an aligned nationality mentee. Mentors assigned to this research arm held 1:1 mentoring sessions focused on transferring business knowledge and on building trust and a shared understanding of experiences and circumstances between the mentor and the mentee through perspective-sharing. Mentors in this research arm were randomly assigned to a mentee of the same nationality as them.
- Research Arm M2b: Mentorship focused on transferring business knowledge and on perspective-sharing with a misaligned nationality mentee. Mentors assigned to this research arm held 1:1 mentoring sessions focused on transferring business knowledge and on building trust and a shared understanding of experiences and circumstances between the mentor and the mentee through perspective-sharing. Mentors in this research arm were randomly assigned to a refugee mentee of a different nationality as them.

2.2 Timeline

The detailed timeline of each step of the research is presented in Figure 2 below.

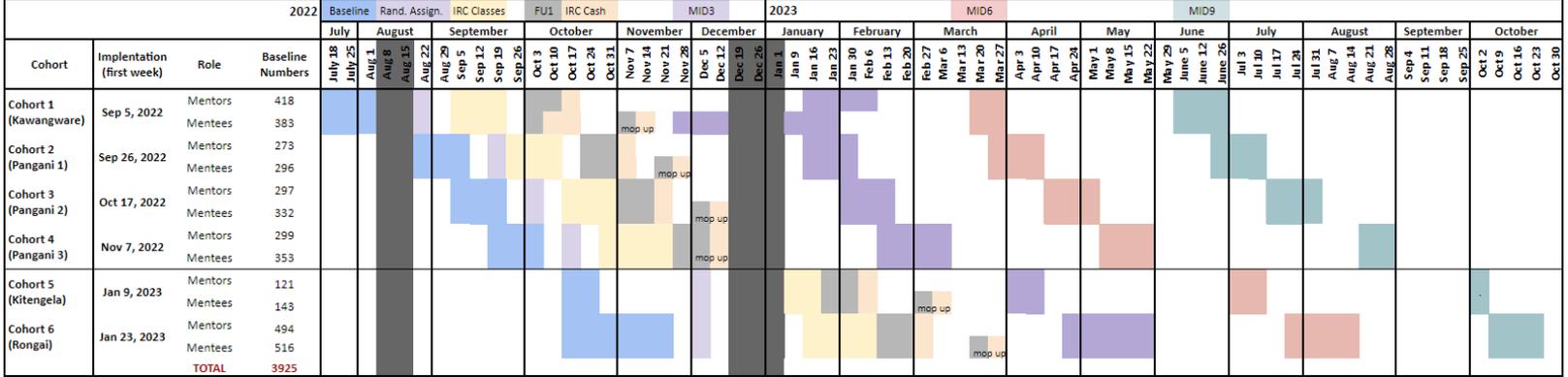


Figure 2: Detailed Timeline for Each Cohort

2.3 Random Assignment

For each cohort, the random assignment is done as follows :

- Among mentees/aspiring microentrepreneurs who completed the baseline survey: stratifying on gender and country of origin, random assignment to one of four research arms using the following assignment probability:

Control: 0.2

Business grant only: 0.2

Business grant and mentorship with basic business content: 0.3

Business grant and mentorship with basic business content and perspective sharing: 0.3

- Among mentors who completed the baseline survey: stratifying on gender, random assignment to one of three research arms:

Control; Mentorship with basic business content; Mentorship with basic business content and perspective sharing.

Assignment probability to the different research arms varied by cohort and was contingent of the number of mentees assigned to the mentorship arms per cohort and the total number of mentors recruited within that cohort.

For each cohort, a random order waitlist was generated in case of noncompliance. That waitlist was used during the second mop-up to assign a new mentor to mentees whose mentor had not complied with the study; any remaining mentor on the waitlist after that was assigned to the control group.

Mentors assigned to the mentorship arms were further randomly assigned to an aligned or a misaligned nationality mentee match.

- All random assignment took place by cohort and at the individual level, within blocks defined by residency status (where applicable) and gender. Specific mentor and mentee pairs were also randomly assigned in advance.

2.4 Re-Assignment and Mop-Ups

In the mentorship arms, after two opportunities to participate in the program, if one member of the pair is no longer interested in the program or cannot be reached, there is a re-assignment process carried out by the research team. First, the IRC team generates a list of all mentors and mentees who are interested in participating in the program but are unmatched because their matched partner dropped out of the program or cannot be reached. The research team first carries out reassignment within this list ensuring that each individual's original treatment assignment is maintained (i.e. within a given research arm and aligned versus misaligned pairing). If there are no eligible matches, and there is a waitlist of mentors, mentors are drawn in numerical order from the waitlist to complete the matching. All remaining waitlist mentors are rolled into the control group.

There is no waitlist of mentees for this project. Therefore, additional recruitment is undertaken by the IRC to complete the pairs of unmatched mentors who are interested in participating in the program but their

assigned mentees were no longer interested or could not be reached. Mentees recruited for this purpose are not a part of the research sample but serve the purpose of minimizing noncompliance among mentors who have been randomly assigned to the mentorship arms and remain interested in the program. There is no data collection for these mentees except the information captured by the IRC at the time of registration.

3 Final Study Sample

Figures 3 and 4 are consort diagrams detailing separately for both study populations (mentees and mentors) the sample sizes and exclusion cases at each step of the study from the registration, the baseline survey and the randomization, which combined lead to the final study sample.

The registration step was conducted by the IRC, which made sure registered participants followed the study’s eligibility criteria. All registered participants formed the sample for the baseline survey, except for 69 mentees in Rongai which were excluded from the study due to over-subscription compared to the target sample size. Most of the exclusions happen at baseline, where we drop people who a) couldn’t be found, b) declined to participate, c) were found ineligible after completion of the baseline, or d) were erroneously dropped during data processing. From that baseline sample is run the randomization into the relevant treatment arms.

For mentors, three mentors are excluded from the study post randomization, because they had mistakenly been included in the randomization script (two cases of no consent and one case of non eligibility). These were dropped in the programming after the randomization and replaced through the waitlist mechanism, and these observations are thus removed from the study sample. This post-randomization exclusion explains the discrepancy between the number of mentors and mentees in the mentorship arms in the study sample (at the stage of first randomization). Finally, during the mop-up process, the waitlist mentors are re-randomized into either the control group or one of the mentorship arms. This process finalizes the sample.

For mentees, the final sample is the same as the one generated by the initial randomization, as no mentees are excluded post-randomization and no mentees are re-assigned to a different treatment arm during the mop-up process.

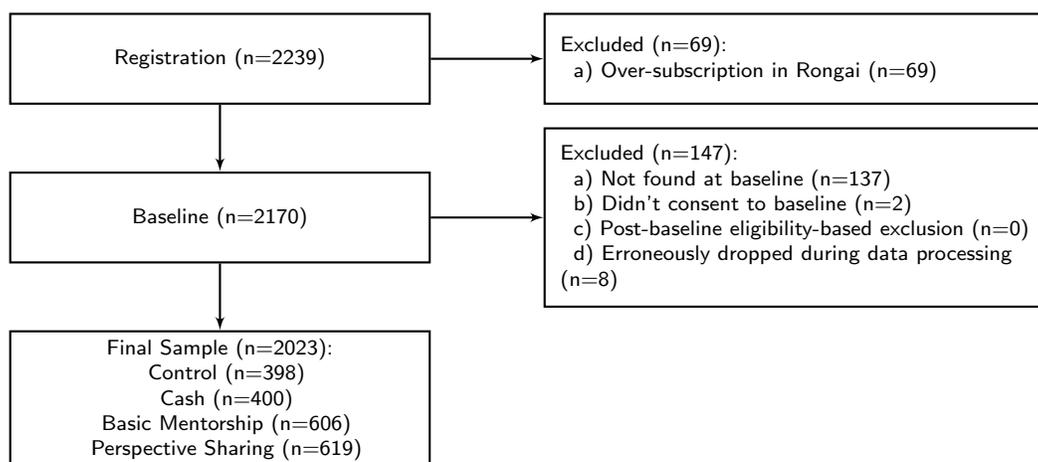


Figure 3: Consort Diagram for Mentees

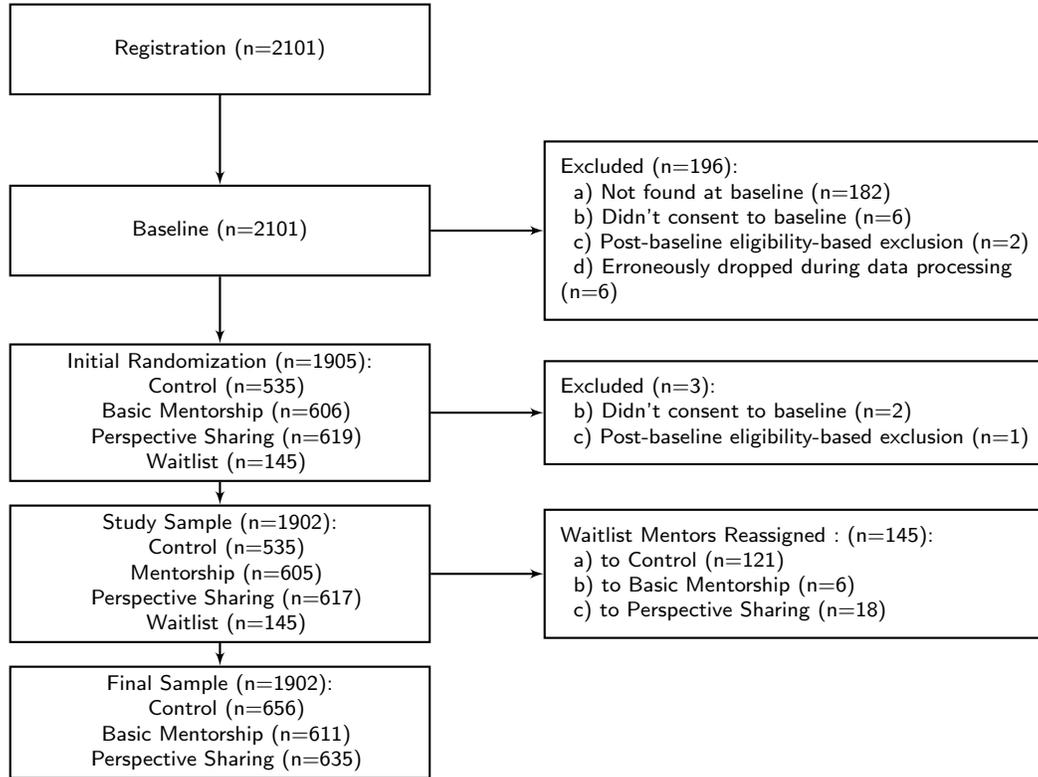


Figure 4: Consort Diagram for Mentors

4 Hypotheses and Outcomes for Aspiring Microentrepreneurs/Mentees

4.1 Primary Outcome Domains

Domain 1: Relationship with mentor

Measures: Twenty-three (23) questions that capture the mentee’s perception of their relationship with their mentor across 5 different sub-domains:

1. Competency (6 questions)
2. Feeling Heard (3 questions)
3. Trust (3 questions)
4. Identification (7 questions)
5. Satisfaction (4 questions)

All responses are measured on a 1-5 scale ranging from *Strongly Disagree* to *Strongly Agree* and are captured at the 1 and 3 month follow-up.

Items will be analyzed in the following way: First, we will construct an index for each sub-domain (competency, feeling heard, trust, identification, and satisfaction) by taking the average of all non-missing values for the items in that sub-domain per respondent. We will then construct a *relationship with mentor* index by taking the average of all 5 sub-domains. We will standardize the final index by the reference group mean (in this case, the cash+mentoring arm) and standard deviation.

This domain will be contrasted specifically between the treatments cash+mentoring vs. cash+perspective-getting and mentoring. This domain will only be tested at times 1 and 3 months, after which the mentoring relationship is not expected to continue.

Domain 2: Perceptions and beliefs about self as a micro-entrepreneur

Measures: Sixteen (16) questions that capture the mentee's perceptions and beliefs about themselves as micro-entrepreneurs across 3 different sub-domains:

1. Belonging in profession (4 questions)
2. Confidence and Resilience (4 questions)
3. Entrepreneurial Self-efficacy (4 questions)
4. Meta-perceptions about mentor's beliefs about the self (4 questions)

At the 1-month follow-up, all responses are measured on a 1-5 scale ranging from either *Strongly Disagree* to *Strongly Agree* or *Not at all true* to *Always true*. At the 3-month follow-up and in subsequent rounds of data collection, all responses are measured on a 1-10 scale from *Not at all* to *Absolutely/Very*.

Items will be analyzed in the following way: First, we will construct an index for each sub-domain (belonging in profession, confidence and resilience, entrepreneurial self-efficacy, and meta-perceptions about mentor's beliefs) by taking the average of all non-missing values for the items in that sub-domain per respondent. We will then construct a *perceptions and beliefs about self as a micro-entrepreneur* index by taking the average of the *belonging in profession*, *confidence and resilience*, and *entrepreneurial self-efficacy* sub-domains as a primary outcome. We will standardize the final index by the control group mean and standard deviation.

An item that theoretically belongs in this domain but was only measured once is *meta-perceptions about mentor's beliefs about the self*. We will separately analyze this outcome as a secondary outcome because it was only captured in the 1 month follow-up survey.

This domain will be tested as a test of joint equality of all treatment arms. We are further interested in constructing a test of the two mentoring arms vs. the cash-only arm.

Domain 3: Business and economic success

Measures: Multiple dimensions of business and economic success are captured during the follow-up surveys. We will use the continuous measure of business profits as a primary outcome for business success, and use other measures as secondary outcomes within this domain.

- Primary outcome
 1. **Business profits** (continuous) : self-reported business profits in the past 30 days (in KES) summed across all of the respondent’s businesses, converted to USD and winsorized at the 1st and 99th percentiles. We will check the sensitivity of the analysis by comparing it to a constructed measure of profits calculated from itemized revenues and costs in the past 30 days and summed across all of the respondent’s businesses, as well as measures focusing on the respondent’s main business only, all similarly converted and winsorized.
- Secondary outcomes
 1. **Business ownership** (binary) : a binary indicator taking the value 1 if the respondent spends any time doing any micro-enterprise work in the past seven days. We will check the sensitivity of the analysis by comparing it to other binary self-employment variables using a higher number of hours as a cutoff point for self-employment, as well as a self reported business ownership status in the last 30 days.
 2. **Hours spent working on business** (continuous) : number of hours spent doing any microenterprise or other self-employment activity during the seven days before today.
 3. **Value of productive assets** (continuous) : sum of asset values that the respondent or their business owns (tools, machinery, furniture, inventories, and other physical assets.) This outcome will be winsorized at the 1st and 99th percentiles. We will estimate this outcome via poisson (quasi maximum likelihood) regression.
 4. **Hours spent on any employment** (continuous) : number of hours spent doing any wage labor, salaried work, microenterprise or other self-employment activity during the seven days before today. This is constructed as a sum of the relevant time use questions.
 5. **Total individual productive income** (continuous) : self-reported income from any income-generating activity in the past 30 days (in KES), including wages and self-employment income, converted to USD and winsorized at the 1st and 99th percentiles. This is constructed as a sum of each of the relevant self-reported incomes. We will check the sensitivity of the analysis by comparing it to a self-reported measure of overall income in the past 30 days, similarly converted and winsorized.

We capture these measures for all rounds of data collection at the 1, 3, 6 and 9 month follow-up.

This domain will be tested as a test of joint equality of all treatment arms. We are further interested in constructing a test of the two mentoring arms vs. the cash-only arm.

Domain 4: Business networks and interactions

Measures: The business/economic networks module captures engagement with other businesses, suppliers, collaborators and other individuals who the respondent discusses economic/income related things with. We capture the size, strength and diversity of the business/economic networks for each respondent.

Size of network and the strength of the ties in the network:

1. Number of contacts listed in a business/economic networks module
2. Number of contacts listed as individuals from whom the respondent seeks advice in a business/economic networks module

Diversity of network:

1. Number of contacts of another gender or residency status (as different from their own) listed in a business/economic networks module. If a contact is different from the respondent on dimensions of gender and residency status, the contact is counted twice.

We capture the measures in the business/economic networks module at the 3, 6 and 9 month follow-up.

Items will be analyzed in the following way: We will analyze the size and diversity of contacts as a primary outcome. We will analyze the strength of the ties in the network as a secondary outcome.

This domain will be tested as a test of joint equality of all treatment arms. We are further interested in constructing a test between the two mentoring arms, and between the pooled mentoring arms vs. the cash-only arm.

Domain 5: Network game

Measures: The 3-month follow-up survey consists of a network game that requires the respondent (mentee) to find a Kenyan supplier in their broader geographical area through their network (i.e. people they directly know or are connected to by their personal/professional network). Our primary outcome of interest for this measure is whether the respondent identified the supplier within 30 minutes (i.e. successfully completed the game).

Items will be analyzed in the following way: For our primary outcome (completion), we will conduct a test of joint equality of all treatment arms. We are further interested in examining features of the search process that resulted in successful outcomes. In order to do so, we will descriptively look at the following measures:

1. number of initial contacts generated (“seeds”) in order to find the supplier
2. Length of chains (i.e. contacts/calls generated through the same “seed”)
3. Diversity of initial contacts on dimensions of nationality, gender, profession, etc

4.2 Secondary Outcome Domains

Domain 1: General Well-Being

Measures: Fourteen (14) questions that capture the mentee’s overall mental well-being across 3 different sub-domains:

1. Anxiety (7 questions) captured on a scale of 1-4.
2. Depression (5 questions) captured on a scale of 1-4.

3. Well-being and happiness (2 questions) - one item captured on a scale of 0-10 and the other item on a scale of 1-4.

We capture these measures for all rounds of data collection at the 1, 3, 6 and 9 month follow-up.

Items will be analyzed in the following way: We will construct indices for the *anxiety* and *depression* sub-domains by taking the average of all non-missing values for the items in each of those sub-domains per respondent. Given the different response scales for the well-being and happiness questions, we will construct an index for the *well-being and happiness* sub-domain by first calculating a z-score for each item by subtracting the control group mean and dividing it by the control group standard deviation. We will then sum the two z-scores for all non-missing components and divide by the number of non-missing components for each observation. Then, we will construct a *general well-being* index in the following way:

1. For each sub-domain (anxiety, depression, and, well-being and happiness), we will calculate a z-score by subtracting the control group mean and dividing it by the control group standard deviation
2. Sum the z-scores for all non-missing components and divide by the number of non-missing components for each observation
3. Standardize the final index by the control group mean and standard deviation

This domain will be tested as a test of joint equality of all treatment arms as a secondary outcome. We are further interested in constructing a test of the two mentoring arms vs. the cash-only arm.

Domain 2: Attitudes towards women

Measures: Thirteen (13) questions that capture attitudes towards women across 4 different sub-domains:

1. Attitudes towards women as mentors and women in the workplace (6 questions)
2. Benevolent Sexism (2 questions)
3. Hostile Sexism (2 questions)
4. Attitudes toward equal rights and opportunities for women in Kenya (Afrobarometer, 3 questions)

All responses are measured on a 1-5 scale ranging from *Strongly Disagree* to *Strongly Agree* and are captured for all rounds of data collection at the 1, 3, 6 and 9 month follow-up.

Items will be analyzed in the following way: First, we will construct an index for each sub-domain by taking the average of all non-missing values for the items in that sub-domain per respondent. We will then construct an *attitudes towards women* index by taking the average of all four sub-domains. We will standardize the final index by the control group mean and standard deviation.

For this domain, we are interested in two design based and one observational comparison:

1. First, this domain will be tested as a test of joint equality of all treatment arms as a secondary outcome.

2. Second, we are interested in constructing a difference-in-difference test between the cash+mentoring versus the cash+perspective-getting and mentoring arms in terms of the difference in attitudes between those who are assigned to female vs. male mentors. Within each treatment arm, we will take the difference in attitudes between those assigned to female versus male mentors and compare this difference between the two treatment arms.
3. Third, we are interested in comparing the attitudes of those who are assigned female mentors versus those who are assigned male mentors, pooling data from both the mentorship arms

Domain 3: Contact with Kenyans (refugees only)

Measures: Two (2) questions that capture frequency of contact with Kenyans at home or at social events outside the home:

1. Contact at home
2. Engagement in a social activity outside the home

Responses are captured on a 1-7 scale ranging from *Never* to *Several times a day* and are captured for all rounds of data collection at the 1, 3, 6 and 9 month follow-up.

Items will be analyzed in the following way: We will construct a *contact* index by taking the average of all non-missing values for the two items per respondent. We will standardize the index by the control group mean and standard deviation.

This domain will be tested as a test of joint equality of all treatment arms. We are further interested in constructing a test between the pooled mentoring arms vs. pooled cash and control arms.

Domain 4: Belonging in Kenya (refugees only)

Measures: Four (4) questions that capture refugees' sense of belonging in Kenya.

Items will be analyzed in the following way: We will construct a *belonging in Kenya* index by taking the average of all non-missing values for the four items. We will standardize the final index by the control group mean and standard deviation.

This domain will be tested as a test of joint equality of all treatment arms as a secondary outcome. We are further interested in constructing a test of the two mentoring arms vs. the cash-only arm.

Domain 5: Feelings towards Kenyans (refugees only)

Measure: A single item capturing feelings towards Kenyans using the feelings thermometer. Responses are measured on a 0-100 scale ranging from Cold to Warm.

This domain will be tested as a test of joint equality of all treatment arms as a secondary outcome. We are further interested in constructing a test of the two mentoring arms vs. the cash-only arm.

Domain 6: Social trust

Measures: The 3-month follow-up survey includes a simplified version of a dictator game in which the respondent is given 500 KES and asked to allocate the amount between themselves and two other participants from the same program. The two other participants they are asked to split the money 3-way with are always of the same gender as themselves. For refugees, they consist of one person of the same nationality as them, and one Kenyan. For Kenyans, they consist of one refugee (of unspecified nationality) and one Kenyan. After they split three-way the 500 KES between themselves and the two hypothetical partners, they are randomly assigned either the sender or receiver role. In practice, one third of the participants are assigned to a sender role, and two thirds to a receiver role.

Prior to making their allocation choices, participants receive the following information :

Our research team has an extra 500 KES available. You will decide what to do with this money. In particular, we ask you to allocate this money between yourself and two others who are participating in the reBuild program, like you are. One of those others is a [insert same gender] refugee [insert same country of origin if participant is a refugee]. The other is a [insert same gender] from Kenya. We want to emphasize that your choices will be completely anonymous and will not affect the provision of IRC services to you in the future. After you go through the activity, you will be paired with other participants and randomly assigned the role of either the 'Giver' or 'Receiver'. If you are assigned to the role of the 'Giver', whatever amount you decide to keep for yourself, you will receive by mobile money within 10 days. If you are assigned to the role of the 'Receiver', whatever amount someone else decided to give you, you will receive by mobile money within 10 days.

Items will be analyzed in the following way: We will use the amount allocated to the (outgroup) as a measure of trust between communities. The amount will be winsorized at the 1st and 99th level.

This domain will be tested as a test of joint equality of all treatment arms as a secondary outcome. We are further interested in constructing a test of the two mentoring arms vs. the cash-only arm.

Domain 7: Household finances

Measures: Stock of household debt, savings and other income at the time of survey.

1. Stock of debt : self-reported, converted to USD and winsorized at the 1st and 99th percentiles
2. Stock of savings : self-reported, converted to USD and winsorized at the 1st and 99th percentiles

These measures are collected at each round (1, 3, 6 and 9-month follow up) and will be tested as a test of joint equality of all treatment arms.

4.3 Outcomes and Hypotheses Summary Tables

Table 2: Primary Outcomes and Hypotheses : Mentees

Population of Interest	Domain	Primary or Secondary Domain	Outcomes	Primary or Secondary Hypothesis	Relevant research arms	Reference group for primary test	Primary test	Reference group for secondary test	Secondary test	Reference group for tertiary test	Tertiary test	Survey round(s)
Mentee	1. Relationship with mentor	Primary	Relationship with mentor	Primary	Cash+Mentorship; Cash+Mentorship+Perspective-sharing	Cash+Mentors hip	Test of the cash+mentorship arm vs. cash+mentorship+perspective-	N/A	N/A	N/A	N/A	1, 3
Mentee	2. Perceptions and beliefs about self as a micro-entrepreneur	Primary	Perceptions and beliefs about self as a micro-entrepreneur	Primary	Control; Cash; Cash+Mentorship; Cash+Mentorship+Perspective-sharing	Control	Test of joint equality of all treatment arms	Cash	Test of the two mentoring arms pooled vs. cash	N/A	N/A	1, 3, 6, 9
Mentee	3. Business and Economic Success	Primary	Business Profits Business Ownership Hours spent working on business Value of productive Hours spent on any employment Total individual productive income	Primary Secondary Secondary Secondary Secondary Secondary	Control; Cash; Cash+Mentorship; Cash+Mentorship+Perspective-sharing	Control	Test of joint equality of all treatment arms	Cash	Test of the two mentoring arms pooled vs. cash	N/A	N/A	1, 3, 6, 9
Mentee	4. Business networks and interactions	Primary	Size of network Diversity of network Strength of the ties in	Primary Primary Secondary	Control; Cash; Cash+Mentorship; Cash+Mentorship+Perspective-sharing	Control	Test of joint equality of all treatment arms	Cash	Test of the two mentoring arms pooled vs. cash	Cash+Mentorship	Test of the cash+mentorship arm vs.	3, 6, 9
Mentee	5. Network game	Primary	Completion of network game	Primary	Control; Cash; Cash+Mentorship; Cash+Mentorship+Perspective-sharing	Control	Test of joint equality of all treatment arms	N/A	N/A	N/A	N/A	3

Table 3: Secondary Outcomes and Hypotheses : Mentees

Population of Interest	Domain	Primary or Secondary Domain	Outcomes	Primary or Secondary Hypothesis	Relevant research arms	Reference group for primary test	Primary test	Reference group for secondary test	Secondary test	Reference group for tertiary test	Tertiary test	Survey round(s)
Mentee	6. General well-being	Secondary	General well-being	Secondary	Control; Cash; Cash+Mentorship; Cash+Mentorship+Perspective-sharing	Control	Test of joint equality of all treatment arms	Cash	Test of the two mentoring arms pooled vs. cash	N/A	N/A	1, 3, 6, 9
Mentee	7. Attitudes towards women	Secondary	Attitudes towards women	Secondary	Control; Cash; Cash+Mentorship; Cash+Mentorship+Perspective-sharing	Control	Test of joint equality of all treatment arms	Cash+Mentorship	Test of the cash+mentorship arm vs. cash+mentorship+perspective-sharing arm	N/A	N/A	1, 3, 6, 9
Mentee	8. Contact with Kenyans (refugees only)	Secondary	Contact with Kenyans	Secondary	Control; Cash; Cash+Mentorship; Cash+Mentorship+Perspective-sharing	Control	Test of joint equality of all treatment arms	Control	Test of the two mentoring arms pooled vs. the pooled cash and control arms	N/A	N/A	1, 3, 6, 9
Mentee	9. Belonging in Kenya (refugees only)	Secondary	Belonging in Kenya	Secondary	Control; Cash; Cash+Mentorship; Cash+Mentorship+Perspective-sharing	Control	Test of joint equality of all treatment arms	Cash	Test of the two mentoring arms pooled vs. cash	N/A	N/A	1, 3, 6, 9
Mentee	10. Feelings towards Kenyans (refugees only)	Secondary	Feelings towards Kenyans	Secondary	Control; Cash; Cash+Mentorship; Cash+Mentorship+Perspective-sharing	Control	Test of joint equality of all treatment arms	Cash	Test of the two mentoring arms pooled vs. cash	N/A	N/A	1, 3, 6, 9
Mentee	11. Social trust	Secondary	Social trust	Secondary	Control; Cash; Cash+Mentorship; Cash+Mentorship+Perspective-sharing	Control	Test of joint equality of all treatment arms	Cash	Test of the two mentoring arms pooled vs. cash	N/A	N/A	3
Mentee	12. Household finances	Secondary	Stock of debt Stock of savings	Secondary	Control; Cash; Cash+Mentorship;	Control	Test of joint equality of all	N/A	N/A	N/A	N/A	1, 3, 6, 9

5 Hypotheses and Outcomes for Mentors

5.1 Primary Outcome Domains

Domain 1: Relationship with Mentees

Measures: Twenty-six (25) questions that capture the mentor’s perception of their relationship with their mentee across 3 different sub-domains:

1. Identification (6 questions)
2. Satisfaction (4 questions)
3. Perceived strength of relationship (15 questions)

All responses are measured on a 1-5 likert scale ranging from *Strongly Disagree* to *Strongly Agree* and are captured at the 1 and 3 month follow-up.

Items will be analyzed in the following way: First, we will construct an index for each sub-domain (identification, satisfaction, and perceived strength of relationship) by taking the average of all non-missing values for the items in that sub-domain per respondent. We will then construct a *relationship with mentee* index by taking the average of all 3 sub-domains. We will standardize the final index by the reference group mean (in this case, the aligned nationality mentors in the mentorship only treatment arm) and standard deviation.

This domain will be contrasted specifically between the treatments mentoring vs. perspective-getting and mentoring. This domain will only be tested at times 1 and 3 months, after which the mentoring relationship is not expected to continue.

Domain 2: Feelings towards refugees

Measures: Ratings on a feelings thermometer for specific refugee groups by nationality and for the overall refugee category. All responses are measured on a 0-100 scale ranging from *Cold* to *Warm*.

1. Feelings thermometer rating of specific refugee groups by nationality.
2. Feelings thermometer rating of refugees as a category.

Items will be analyzed in the following way: We will construct an index that will contain all items that measure feelings toward specific refugee groups by nationality. We will construct this index by taking the average of all non-missing values for each refugee nationality per respondent. We will separately analyze the index (consisting of all refugee groups by nationality) as a primary outcome and the rating of feelings towards refugees as a category as a secondary outcome.

This domain will be tested as a test of joint equality of all treatment arms.

Domain 3: Attitudes towards refugees

Measures: Six (6) items measuring attitudes towards refugees with respect to their contribution to and desire to fit in to the host society (Kenya):

1. Attitudes towards refugees as a category. These items are measured in all rounds of data collection
2. Attitudes towards specific refugee groups defined as refugees from East African countries or from non East African countries. These items are measured at the 1 month follow-up
3. Attitudes towards specific refugee groups defined as refugees from the DRC or from Somalia. These items are measured at the 6 and 9 month follow-up

All responses are measured on a 1-5 scale ranging from *Strongly Disagree* to *Strongly Agree*.

Items will be analyzed in the following way: For this domain, we will construct 3 separate indices detailed below. We will construct each index by taking the average of all non-missing values for the items that correspond to that index per respondent. We will standardize the index by the control group mean and standard deviation. We will analyze the attitudes towards refugees as a category index as a primary outcome and the attitudes towards specific refugee groups indices as secondary outcomes.

1. Attitudes towards refugees as a category: an index consisting of 6 items asking about refugees as a category
2. Attitudes towards EAC and non-EAC refugees: an index consisting of 12 items asking about EAC and non-EAC refugees
3. Attitudes towards Congolese and Somali refugees: An index consisting of 12 items asking about Congolese and Somali refugees

This domain will be tested as a test of joint equality of all treatment arms.

5.2 Secondary Outcome Domains

Domain 1: Perceptions and beliefs about self as a micro-entrepreneur

Measures: Twelve (12) questions that capture the mentor's perceptions and beliefs about themselves as micro-entrepreneurs across 3 different sub-domains:

1. Belonging in profession (4 questions)
2. Confidence and Resilience (4 questions)
3. Entrepreneurial Self-efficacy (4 questions)

At the 1-month follow-up, all responses are measured on a 1-5 scale ranging from either *Strongly Disagree* to *Strongly Agree* or *Not at all true* to *Always true*. At the 3-month follow-up and in subsequent rounds of data collection all responses are measured on a 1-10 scale from *Not at all* to *Absolutely/Very*.

Items will be analyzed in the following way: First, we will construct an index for each sub-domain (belonging in profession, confidence and resilience, and entrepreneurial self-efficacy) by taking the average of all non-missing values for the items in that sub-domain per respondent. We will then construct a *perceptions and beliefs about self as a micro-entrepreneur* index by taking the average of all three sub-domains. We will standardize the final index by the control group mean and standard deviation.

This domain will be tested as a test of joint equality of all treatment arms.

Domain 2: Contact with refugees

Measures: Two (2) questions that capture frequency of contact with refugees at home or at social events outside the home:

1. Contact at home
2. Engagement in a social activity outside the home

Responses are captured on a 1-7 scale ranging from *Never* to *Several times a day* and are captured for all rounds of data collection at the 1, 3, 6 and 9 month follow-up.

Items will be analyzed in the following way: We will construct a *contact* index by taking the average of all non-missing values for the two items per respondent. We will standardize the index by the control group mean and standard deviation.

This domain will be tested as a test of joint equality of all treatment arms.

Domain 3: Support for policies promoting the well-being of refugees

Measures: Nine (9) items describing policies supporting refugees as a whole in terms of the actions that the Kenyan government should or should not take with respect to refugees. Seven (7) items describing policies supporting specific refugee groups (EAC refugees, non-EAC refugees, Congolese refugees, and Somali refugees) in terms of the actions that the Kenyan government should or should not take with respect to refugees.

1. Support for policies related to refugees as a category. These items are measured in all rounds of data collection
2. Support for policies related to specific refugee groups defined as refugees from East African countries or non East African countries. These items are measured at the 1 month follow-up
3. Support for policies related to specific refugee groups defined as refugees from the DRC or from Somalia. These items are measured at the 6 and 9 month follow-up

All responses are measured on a 1-5 scale ranging from *Strongly Disagree* to *Strongly Agree*

Items will be analyzed in the following way: For this domain, we will construct 3 separate indices detailed below. We will construct each index by taking the average of all non-missing values for the items that

correspond to that index per respondent. We will standardize the index by the control group mean and standard deviation. We will analyze each index separately.

1. Support for policies related to refugees as a category: an index consisting of 9 items asking about refugees as a category
2. Support for policies related to EAC and non-EAC refugees: an index consisting of 14 items asking about EAC and non-EAC refugees
3. Support for policies related to Congolese and Somali refugees: An index consisting of 14 items asking about EAC and non-EAC refugees

This domain will be tested as a test of joint equality of all treatment arms.

Domain 4: Business networks and interactions

Measures: The business/economic networks module captures engagement with other businesses, suppliers, collaborators and other individuals who the respondent discusses economic/income related things with. We capture the size, strength and diversity of the business/economic networks for each respondent.

Size of network and the strength of the ties in the network:

1. Number of contacts listed in a business/economic networks module
2. Number of contacts listed as individuals from whom the respondent seeks advice in a business/economic networks module

Diversity of network:

1. Number of contacts of another gender or residency status (as different from their own) listed in a business/economic networks module. If a contact is different from the respondent on dimensions of gender and residency status, the contact is counted twice.

We capture the measures in the business/economic networks module at the 3, 6 and 9 month follow-up.

Items will be analyzed in the following way: We will analyze each item separately as secondary outcomes.

This domain will be tested as a test of joint equality of all treatment arms.

Domain 5: Attitudes towards women

Measures: Thirteen (13) questions that capture attitudes towards women across 4 different sub-domains:

1. Attitudes towards women as mentors and women in the workplace (6 questions)
2. Benevolent Sexism (2 questions)
3. Hostile Sexism (2 questions)

4. Attitudes toward equal rights and opportunities for women in Kenya (Afrobarometer, 3 questions)

All responses are measured on a 1-5 scale ranging from *Strongly Disagree* to *Strongly Agree* and are captured for all rounds of data collection at the 1, 3, 6 and 9 month follow-up.

Items will be analyzed in the following way: First, we will construct an index for each sub-domain by taking the average of all non-missing values for the items in that sub-domain per respondent. We will then construct an *attitudes towards women* index by taking the average of all four sub-domains. We will standardize the final index by the control group mean and standard deviation.

For this domain, we are interested in two design based and one observational comparison:

1. First, this domain will be tested as a test of joint equality of all treatment arms as a secondary outcome.
2. Second, we are interested in constructing a difference-in-difference test between the cash+mentoring versus the cash+perspective-getting and mentoring arms in terms of the difference in attitudes between those who are assigned to female vs. male mentees. Within each treatment arm, we will take the difference in attitudes between those assigned to female versus male mentees and compare this difference between the two treatment arms.
3. Third, we are interested in comparing the attitudes of those who are assigned female mentees versus those who are assigned male mentees, pooling data from both the mentorship arms

Domain 6: Social trust

Measures: We will be looking at two behavioral measures of social trust between groups : a simplified dictator game at the 3-month mark, and a mystery shopper exercise conducted around the 6-month mark. The simplified dictator game is the same as the one explained in Domain 6 of section 4.2 for Mentees on page 12. The mystery shopper exercise will measure directly whether mentors (who are all Kenyans) are willing to extend buyer's credit for a low-cost item to Kenyan and refugee mystery shoppers respectively.

1. Inter-group peer trust : the winsorized dollar amount allocated to the outgroup participant in the simplified dictator game
2. Inter-group client trust : the likelihood to extend buyer credit to an outgroup mystery shopper

5.3 Outcomes and Hypotheses Summary Tables

Table 4: Primary Outcomes and Hypotheses : Mentors

Population of Interest	Domain	Primary or Secondary Domain	Outcomes	Primary or Secondary Hypothesis	Relevant research arms	Reference group for primary test	Primary test	Reference group for secondary test	Secondary test	Survey round(s)
Mentor	1. Relationship with mentee	Primary	Relationship with mentee	Primary	Mentorship; Mentorship+Perspective-sharing	Mentorship	Test of the mentorship arm vs. mentorship+perspective-sharing arm	N/A	N/A	1, 3
Mentor	2. Feelings towards refugees	Primary	Feelings towards specific refugee groups by nationality Feelings towards refugees as a	Primary Secondary	Control; Mentorship; Mentorship+Perspective-sharing	Control	Test of joint equality of all treatment arms	N/A	N/A	1, 3, 6, 9
Mentor	3. Attitudes towards refugees	Primary	Attitudes towards refugees as a Attitudes towards specific refugee groups (defined as refugees from EAC and non-EAC) Attitudes towards specific refugee groups (defined as refugees from Somalia and DRC)	Primary Secondary Secondary	Control; Mentorship; Mentorship+Perspective-sharing	Control	Test of joint equality of all treatment arms	N/A	N/A	1, 3, 6, 9 1 6, 9

Table 5: Secondary Outcomes and Hypotheses : Mentors

Population of Interest	Domain	Primary or Secondary Domain	Outcomes	Primary or Secondary Hypothesis	Relevant research arms	Reference group for primary test	Primary test	Reference group for secondary test	Secondary test	Survey round(s)
Mentor	4. Perceptions and beliefs about self as a micro-entrepreneur	Secondary	Perceptions and beliefs about self as a micro-entrepreneur	Secondary	Control; Mentorship; Mentorship+Perspective-sharing	Control	Test of joint equality of all treatment arms	N/A	N/A	1, 3, 6, 9
Mentor	5. Contact with refugees	Secondary	Contact with refugees	Secondary	Control; Mentorship; Mentorship+Perspective-sharing	Control	Test of joint equality of all treatment arms	N/A	N/A	1, 3, 6, 9
Mentor	6. Support for policies promoting the well-being of refugees	Secondary	Support for policies related to refugees as a category Support for policies related to specific refugee groups (defined as refugees from EAC and non-refugees from Somalia and	Secondary Secondary	Control; Mentorship; Mentorship+Perspective-sharing	Control	Test of joint equality of all treatment arms	N/A	N/A	1, 3, 6, 9 1 6, 9
Mentor	7. Business networks and interactions	Secondary	Size of network Diversity of network Strength of the ties in the network	Secondary Secondary Secondary	Control; Mentorship; Mentorship+Perspective-sharing	Control	Test of joint equality of all treatment arms	N/A	N/A	3, 6, 9
Mentor	8. Attitudes towards women	Secondary	Attitudes towards women	Secondary	Control; Mentorship; Mentorship+Perspective-sharing	Control	Test of joint equality of all treatment arms	Mentorship	Test of the mentorship arm vs. mentorship+perspective-sharing arm	1, 3, 6, 9
Mentor	9. Social trust	Secondary	Social trust	Secondary	Control; Mentorship; Mentorship+Perspective-sharing	Control	Test of joint equality of all treatment arms	N/A	N/A	3

6 Data collection and post-processing

6.1 Data collection

There are five rounds of data collection for this study: baseline; 1-month follow-up; 3-month follow-up; 6-month follow-up and 9-month follow-up. The timing for the follow-up rounds is determined relative to the information session that the IRC will hold with all participants during which they are told about the services they receive from the IRC, which is referred to as Day 1 in Figure 1. We anticipate collecting data from 2000 program participants and 1800 potential mentors at baseline. The follow up surveys will be conducted on the full study sample as defined at the randomization stage, regardless of participants compliance with program components.

6.2 Data post-processing

In general, prior to estimation and testing, we will transform variables in the following ways:

- Constructed indices and their components :
 - Where applicable, we will reorient variables that are components of indices such that a higher score corresponds to a more favorable outcome, i.e. higher equals more positive.
 - When the scales of indices and their components differ, they will be standardized based on the distribution among the comparison group relevant for the hypothesis they are testing. The comparison group and the standardization process is specified in each outcome construction.
- Monetary measures including profit, revenue, assets, debt, & savings:
 - Values will be winsorized at the 1st and 99th percentiles within each survey round and major treatment arm (control, cash, basic mentorship, perspective-sharing). Winsorization will be done within treatment arm, using the first level of randomization (i.e. ignoring the sub-randomization of aligned vs mis-aligned matches).
 - Nominal values will be converted to real values using the CPI from the Kenya National Bureau of Statistics.
 - If the respondent declines to provide an exact amount but provides a range (pre-specified on the survey), we will use the midpoint of the range. For firms in the top range, the median of firms in the top range with reported point estimates will be used. For firms in the bottom range of measures that can take on negative values (such as profit), we will use the median of firms in the bottom range with reported point estimates.
 - For firms that are not operating, values of profits, revenues, assets, etc., will be recorded as 0.

7 Statistical analysis

7.1 Estimation

The starting point for all analyses of intent-to-treat effects will be one of two empirical models of outcomes. Our choice of model depends on the nature of the outcome under study.

For outcomes that may contain negative values (including profits, and all normalized indices) or for outcomes that are bounded from both above and below (such as binary variables), our estimate of intent-to-treat effects will be derived from the following ANCOVA specification:

$$Y_{ict} = \sum_{t \in \mathcal{T}} \sum_{w \in \mathcal{W}} \tau_t^w D_i^w + \rho^w D_i^w Y_{ic,0} + \eta^w D_i^w M_{i0} + \delta^w D_i^w X_{ic,0} + \gamma_t + \mu_c + e_{it} \quad (1)$$

for individuals i in cohort c , as observed at time period $t \in \mathcal{T}$, and D_i^w an indicator for assignment to treatment $w \in \mathcal{W}$. The set of time periods considered, \mathcal{T} , and feasible treatments, \mathcal{W} , will differ by outcome (as discussed above) and by estimating samples (aspiring microentrepreneurs/mentees vs mentors), as discussed below. For outcomes measured at baseline, $Y_{ic,0}$ reflects this baseline value (or analogously construct based on subcomponents collected at baseline), de-means and imputed at its mean in cases of missingness; $M_{ic,0}$ is an indicator for such missing/imputed cases.

Baseline covariates are selected by post-double lasso, from a candidate set of covariates including:

- Baseline values of all primary and secondary outcomes, where available;
- Demographics: age, gender, country of origin.

In all cases, we will impute missing values of baseline covariates at their mean, and we will include an indicator for missingness/imputation among the set of candidate controls. Selected covariates will be centered at zero in the control group and interacted with treatments following Lin (2013).

Where outcomes are by construction weakly greater than zero and unbounded from above¹ we will replace the model of equation (1) with an analogous Poisson Quasi-Maximum Likelihood Estimate (QMLE) specification:

$$E[y_{ict}] = \exp \left\{ \sum_{t \in \mathcal{T}} \sum_{w \in \mathcal{W}} \tau_t^w D_i^w + \rho^w D_i^w Y_{ic,0} + \eta^w D_i^w M_{i0} + \delta^w D_i^w X_{ic,0} + \gamma_t + \mu_c \right\}. \quad (2)$$

7.2 Testing

Our primary approach to testing hypotheses of equality between pairs of treatment arms will be to use randomization inference, permuting only those treatments for which equality is implied by the specific hypothesis in question. We will use 2,000 such feasible randomizations, each allowable under the assignment mechanism.

Testing will reflect the fact that equations (1) and (2) allow treatment effects to differ across time periods, and, in general, we wish to test equality between (two or more) treatment arms within a given period, without implying constant effects across periods.

For any given hypothesis, within each time period, t in the set of post-intervention time periods $t \in \mathcal{T}$, we will construct a test statistic as follows:

- For pairwise comparisons between two arms, hypothesis testing will be based on a t statistic that studentizes the estimated difference between these arms in a specific time period, τ_t^w , in the relevant

¹Examples include (but are not limited to) business revenues, productive asset values, savings, and debt.

outcomes regression. We will re-define reference category in the treatment vector W to allow this to be read off directly from regression results.

- For tests of equality between multiple arms, hypothesis testing will be based on an F statistic that encodes the hypothesis of equality.

Our general test is against the sharp null of no difference between arms in *any* of the considered time periods (with relevant time periods, \mathcal{T} , different by outcome domain and mentor/mentee sample). For a given test statistic ϕ_t^r (either a t or F statistic, as described above) of the difference between treatment arms (w, w') in time period t for feasible randomization r , we take as our omnibus test of the hypothesis of no difference the *mean* of the values of this test statistic across all time periods considered, $t \in \mathcal{T}$: $\phi^r = \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \phi_t^r$. We compute two-sided p -values against the distribution of this test statistic when it is derived from a t statistic, and one-sided p -values against distribution of this test statistic when it is derived from an F statistic.

Use of studentized regression coefficients has the advantage that it affords the asymptotic interpretation of rejections of the “sharp null” as implying non-zero regression coefficients (Chung and Romano, 2013, DiCiccio and Romano, 2017). However, when outcomes are far from normally distributed, alternative test statistics may be better powered against the null, though these sacrifice the asymptotic interpretation of t statistics.

As a secondary specification, therefore, we use a residualized KS statistic as the basis for randomization inference. We compute this in the following steps:

1. To take advantage of the ability of regression controls to increase power, we first residualize outcome y_{ict} based on the ANCOVA outcomes regression in Equation (1) (or modify the below analogously for Equation (2) as appropriate). Using estimated coefficients from this regression we define the residualized outcome

$$\ddot{y} = y_{ict} - (\rho Y_{ic,0} + \eta M_{i0} + \delta X_{ic,0} + \gamma_{ct}) \quad (3)$$

2. For a hypothesized difference in outcomes between treatments $w, w' \in \mathcal{W}$, our test statistic for the difference between these arms in period t is given by

$$KS_t = \sup_{\ddot{y}} \left| \hat{F}_w(\ddot{y}) - \hat{F}_{w'}(\ddot{y}) \right| \quad (4)$$

3. We take the mean of such test statistics across time periods, t , as the relevant test statistic for a given potential assignment.

When the null hypothesis in question is a test of equality between *all* treatment arms, then instead of a single KS statistic for period t , we consider the maximum of the KS statistics for all pairwise comparisons between arms in that time period.

7.3 Multiple hypotheses

For our primary hypothesis tests, naive p -values reported have correct coverage for those with a priori interest in a particular hypothesis. Beyond this, we control the false discovery rate in two ways (Haushofer et al., 2017, Egger et al., 2022). First, we report Anderson (2008) sharpened q values that represent the minimum q value at which each primary hypothesis is rejected, pooling all mentee-related and all mentor-related hypotheses in turn. Second, within each outcome domain, we report q values that control the FDR for all reported p -values within that domain. In each case, we report both standard p -values and the accompanying minimum q -values.

7.4 Compliance

Compliance with the assigned treatment encompasses both reception of cash and attendance of meetings with the mentor, depending on treatment arm. We will measure and document how many of the mentees

received their cash and how many meetings (if assigned to the mentorship arm) they attended with their mentor. We will test for heterogeneity in treatment adherence taking into account identity alignment between the mentor and mentee on dimensions of residency status and gender.

7.5 Attrition

We address attrition from the estimating sample at 1-month, 3-month, 6-month, and 9-month follow-up surveys as follows:

1. For all primary and secondary outcome estimating samples, we will test for differential attrition by estimating the probability of attrition as a function of treatments and controls, in a linear probability model that replaces the outcome of equation (1) with a binary measure for individual i 's inclusion in the estimating sample in post-intervention time period t .
2. If this test for differential attrition is statistically significant at the 5 percent level, then we modify the estimating equation in (1) or (2) to use inverse probability weights. To generate IPW weights, the probability of retention in each round will be modelled (via lasso) as a function of treatments, candidate baseline covariates, and their interactions.
3. In the case of statistically significant differential attrition, we will also estimate Lee (Lee, 2009) bounds as a robustness check.

References

- Anderson, Michael L.**, “Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects,” *Journal of the American Statistical Association*, 2008, *103* (484), 1481–1495.
- Chung, EunYi and Joseph P Romano**, “Exact and asymptotically robust permutation tests,” *The Annals of Statistics*, 2013, *41* (2), 488–507.
- DiCiccio, Cyrus J and Joseph P Romano**, “Robust permutation tests for correlation and regression coefficients,” *Journal of the American Statistical Association*, 2017, *112* (519), 1211–1220.
- Egger, Dennis, Johannes Haushofer, Edward Miguel, Paul Niehaus, and Michael Walker**, “General equilibrium effects of cash transfers: experimental evidence from Kenya,” *Econometrica*, November 2022, *90* (6), 2603–2643.
- Haushofer, Johannes, Edward Miguel, Paul Niehaus, and Michael Walker**, “GE effects of cash transfers: Pre-analysis plan for household welfare analysis,” American Economic Association RCT Registry, Trial No. AEARCTR-0000505 September 2017.
- Lee, David S**, “Training, wages, and sample selection: Estimating sharp bounds on treatment effects,” *The Review of Economic Studies*, 2009, *76* (3), 1071–1102.
- Lin, Winston**, “Agnostic notes on regression adjustments to experimental data: Reexamining Freedman’s critique,” *Annals of Applied Statistics*, March 2013, *7* (1), 295–318.