Worker Assignment and National Unity: Are All Stable Matches Socially Stable?*

Willa Friedman[†] Guthrie Gray-Lobe[‡] Michael Kremer[§]

PLEASE DO NOT CIRCULATE
September 13, 2024

We study the effects of alternative worker assignment policies in the context of a Kenyan government program that deployed recent college graduates to run literacy camps in schools with the explicit purpose of promoting interaction across ethnic groups. On average, workers favor locations with ethnic compositions similar to their home regions, which conflicts with the goal of promoting inter-group social interactions. Using random variation in assignment we show that placing workers in less preferred regions fosters trust in local communities, reduces home-biased preferences among workers as evidenced by lab-in-the-field tasks, and results in higher levels of worker effort as evidenced by the number of children reached by the literacy camps and literacy skills. However, ignoring worker preferences leads to substantial welfare losses and can lead to underserved communities when participation is voluntary. Giving workers choices within constrained sets such that they cannot be assigned to counties near their home county flattens the trade-off between social interactions and worker welfare.

*The researchers thank Evidence Action, especially Sasha Gallant and Samantha Bastian, for sharing anonymized data for this study with the researchers. The data for this study were collected by Evidence Action and REMIT with permission from the National Commission for Science, Technology, and Innovation (NACOSTI). Additional approvals from relevant County commissioners/County Education officers were obtained. IRB approval was received from Maseno University Ethics Review Committee (MUERC) (reference 00643/18). The evaluation received support from the World Bank Strategic Impact Evaluation Fund (SIEF), JPAL Governance Initiative, and the International Growth Centre (IGC). We are grateful to Nidhila Adusumalli, Shreya Bhattacharya, Ronak Jain, and Rashmi Menon for invaluable research assistance. We also thank the Busara Center for Behavioral Economics, especially Mark Millrine, for supporting applicant matching and data analysis.

[†]University of Houston, email: willa.friedman@gmail.com
[‡]University of Chicago, email: graylobe@uchicago.edu
[§]University of Chicago, email: kremer.m@gmail.com

1 Introduction

People often prefer to reside and work in communities with which they identify. When there are external benefits of levaing ones "homeland, even well-functioning labor markets that generate efficient contracts between workers and employers may be globally inefficient. This paper considers two concerns. First, in environments where groups are segregated, contact across ethnic lines may be underprovided relative to the social optimum when residential choices are based on individual utility maximization. A large literature demonstrates that contact across ethnic and class lines can reduce animosity (e.g. Okunogbe, 2018; Rao, 2019; Lowe, 2018; Mousa, 2020). Second, it can undermine efforts to reduce inequality between groups. Historically marginalized communities will tend to have low skill endowments, so the quality of local services will tend to be low. Building up the skill endowment through in-migration will tend to be very costly.

Governments can and sometimes do internalize the social consequences of homophily in worker assignment. First, governments sometimes take steps to encourage interactions between diverse groups. Youth service programs encourage young, often urban and highly-educated youth to spend time in more disadvantaged communtiies. Teacher assignment policies sometimes encourage teachers to work outside of their home communities to foster national unity. In many cases these programs severely constrain the choice that people have over where they can work, potentially having large negatie impacts on worker welfare. However, despite these potentially large social costs, little is understood about the benefits of these types of efforts on local communities. Second, governments often use hazard pay to compensate those who work in those regions that are less desirable to highly-educated civil servants. However, direct financial incentives to work in a region only influence the decisions of the marginal worker, with large rents going to those inframarginal workers who, for idiosyncratic reasons, may be more willing to work in a particular area.

We examine these questions in the context of a Kenyan government program designed to promote national unity and reduce inequality in education. Greatness-United (G-United) was a high-profile program of the Government of Kenya that recruited and deployed recent graduates of Kenyan universities to low-performing public schools to administer an early-grade literacy program with an explicit goal of promoting national unity through inter-regional contact. As part of the program, workers were provided temporary residence with a homestay family during their six-month deployment period in the community.

We leverage rich data and experimental variation in assignments of workers to job positions to evaluate the consequences of different assignment policies. Applicants were invited to submit their preferences over locations of job positions and were assigned using an incentive-compatible mechanism. The mechanism first randomly partitioned the market of applicants and job positions into a market assigned at random one that respected worker location preferences. Applicants and job positions in the random assignment market were assigned at random to counties, and then randomly assigned to schools within counties. Applicants and job positions in the worker location preference market were assigned algorithmically using a a version of the Random Serial Dictatorship of Abdulkadiroğlu and Sönmez (1998)) to counties according to their stated preferences. After assignment to counties, applicants were randomly assigned to schools within counties.[1]

Application data reveal heterogeneity in preferences across counties among the self-selected sample of university graduates who apply to the program. Applicants who come predominantly from wealthier parts of the country – prefer being assigned to locations that are geographically closer and have similar ethnic compositions to their home counties.[2] To be indifferent between marginalized and non-marginalized coun-

[1] Assignment algorithms are described in more detail in Section 2.2.2.

[2] An error in the preference survey instructions may have generated some errors in measurement of preferences. We investigate evidence of this reporting error, discuss the ways it could bias our results, and use preference data from a later cohort in which

ties, applicants need 62 percent higher salaries when posted in marginalized areas. Applicants avoid those assignments that would result in more novel social interactions. We define "ethnic similarity as the probability that two randomly drawn residents from two counties are from the same ethnic group. To compensate applicants for an assignment that with 10 points lower ethnic similarity, they must be paid 28 percent higher salaries. While these results do not necessarily mean that applicants are averse to interacting with other groups *per se*, they show that applicant preferences are, on average, at odds with the objectives of addressing geographic inequality in skill endowments and generating inter-group contact.

We exploit experimental variation in the assignment of workers to counties to examine the impact of assigning workers to more preferred locations on a variety of outcomes, including applicants themselves and members of the community where they are stationed. We find that when workers are assigned to less preferred, more distant, and more ethnically dissimilar locations it promotes attitudes associated with national unity among workers and the communities they serve. Workers reveal less home bias in an incentivized task, and local community members (head teachers in the school where they work and their home-stay families) express greater generalized trust, and children are less likely to reveal fear towards the workers' group in a face-anger attribution task. Workers assigned against their preferences also produce larger learning gains among children, and children are more likely to report having taken part in the program, suggesting greater effort.

Assignment using worker location preferences has a large impact on worker welfare and a moderate effect on retention.[3] The anticipated welfare gain – based on an incentivized auction mechanism – from assigning workers to their location preferences compared to random assignment is equivalent to increasing workers' salary by about two-thirds. Eighty-six percent of worker-location-preference-assigned applicants received their first choice, compared to 6 percent of randomly assigned applicants. Assignment using worker location proferences increased the share of applicants who attended a mandatory training (a measure of initial retention) by 5 percentage points from a base of 53 percent. The boost in participation from assigning using worker location preferences is largest in poorer and margainlized areas.Applicants assigned using worker location preferences to counties with over 50 percent poverty are 10 percentage points more likely to be retained initially (from a base of 50 percent), and those assigned to marginalized counties using worker location preference sare 17 percentage points more likely to be initially retained (with a base of 42 percent).

This study contributes to an empirical literature on the effects of exposure to other groups on attitudes and behaviors (Allport, 1954; Key, 1949; Scacco and Warren, 2018; Corno, Ferrara and Burns, 2019; Rao, 2019; Lowe, 2018; Chen et al., 2020; Boisjoly et al., 2006; Mo and Conn, 2018; Billings, Chyn and Haggag, 2020; Kaplan, Spenkuch and Tuttle, 2019). A subset of the research on the contact hypothesis has examined the effects of youth service programs on volunteers (Ramiah and Hewstone, 2012; Dobbie and Fryer Jr, 2015; Okunogbe, 2018; Mo and Conn, 2018; Chen et al., 2020). Beyond contributing another test of the contact hypothesis among pupils, volunteers, and adults in the community, we contribute in two important ways. First, the literature on youth service and volunteer programs has mostly focused on impacts on the social attitudes and behaviors of deployed youth (Okunogbe, 2018). However, in many cases, these policies are motivated by a concern about the beneficiary communities. We extend this analysis to examine impacts on adults who are likely to interact with the volunteer and on the direct program beneficiaries (students who received the program). Second, by centering the study around a centralized worker assignment system, we

the survey was simplified as well as other sources of random variation to remove the influence of the measurement error on our estimates.

[3]We measure welfare based on applicants' reports of the amount they would need to prefer one option to another, with an expectation that these would be used. We explain this in more detail below.

are able to examine how different assignment allocations would affect program outcomes.

The results of this study potentially have applications to a range of civil servant assignment policies. The delivery of many government services involves the deployment of workers – such as judges, public defenders, health care workers, teachers, and tax collectors – with similar skills, throughout a large geographic area, posing a worker assignment problem. Allowing workers to choose where they are deployed can improve worker's expected welfare and participation. However, governments often want civil servants to serve outside their home regions to support nation-building and avoid favoritism. Botswana has deliberately used teacher assignment to promote national unity since independence (Berbner, 2020). In Kenya, there has been a push for a similar policy, known as 'delocalisation,' which transfers headmasters outside their home areas. This policy has provoked resistance from civil service teachers who have strong preferences over where they are assigned. A high level of teacher turnover - one potential consequence of workers' dissatisfaction with their posts - is a recognized barrier in providing quality instruction in developing countries (Zeitlin, 2021).

Our study contributes to an emerging literature on the importance of match-specific effects of workers in delivering public services, and how worker preferences may be at odds with full exploitation. Bates et al. (2022) shows that teachers are averse to serving disadvantaged populations where they often have larger impacts, and examine the impacts of reallocating teachers. Ba et al. (2021) examine how bidding procedures used in assigning police affect neighborhood crime, showing that these procedures tend to assign less experienced workers to police neighborhoods with more crime. Our study was conducted in parallel and adds to this literature by examining the impact of worker preferences on segregation forces and their social impacts.

Our study also contributes to the literature on the potential efficiency gains from market design solutions to labor market matches (e.g. Roth, 1984; Davis, 2019). Our study is the first, to our knowledge, to experimentally evaluate the impacts of alternative assignment systems. We show that such improvements may be especially important to ensure workers are deployed to less desirable locations. Using novel data from a mechanism that elicited values of compensating variation, we are able to assign the welfare gains from alternative assignment systems.

The remainder of this paper is organized as follows: Section 2 describes the setting, including relevant background on ethnic groups and segregation in Kenya, the G-United program, and the application and assignment process; Section 3 describes the data that will be used in the analysis; Section 4 describes the empirical framework used to analyze the data; Section 5 presents the results; Section 6 presents the results from simulations of alternative assignment mechanisms; and Section 7 discusses the interpretation of the results and limitations of the study.

2 Background and program overview

2.1 Kenyan ethno-political context

Kenya is an ethnically diverse country with over 40 distinct groups, which are geographically segregated. The average Kenyan lives in a county in which their ethnic group comprises 63 percent of the population while at the national level, this figure is only 12 percent.[4] A single group represents the majority of residents in 38 Kenyan counties have a majority ethnic group while the largest single group accounts for less than 20 percent of the population,[5] an artefact of colonial era policies that deliberately sought to maintain ethnically

[4] Authors' calculation using Demographic and Health Surveys data from 2014.
[5] Authors calculations based on Kenya National Bureau of Statistics and ICF International (2014).

homogeneous geographic districts (Burgess et al., 2015; Nderitũ, 2017). The concept of a "home county" represents the region of an individual's ancestry and is closely linked to an ethnic identity.[6]

Several studies have proposed that isolation of groups from one another promotes mistrust between groups. Kasara (2017) shows that inter-ethnic tolerance is greater in locations with less segregation, and Berge et al. (2019) find that more recent migrants to Kibera, an ethnically diverse slum in Nairobi, exhibit more ethnic bias in behavioral games.

Kenya has experienced episodes of ethnic violence around national elections.[7] In 2007, after Mwai Kibaki, a Kikuyu, was announced the winner of a presidential election, based on contested vote counts, violent conflict erupted. Approximately 1,500 people were killed. The Government of Kenya has undertaken major initiatives to support national integration and counter regional and ethnic favoritism, including the establishment of the National Cohesion and Integration Commission (NCIC). Additionally, the Ministry of Education has initiated efforts to transfer headmasters outside their home area to support national unity, in a process called "delocalisation." This delocalisation is explictly designed to encourage "cultural exchange," and "promote national cohesion."[8]

2.2 Greatness-United

Greatness United (G-United) was created in 2014 as a youth service program, with an aim of fostering national unity through interactions across groups, among other goals.[9] The program was an initiative of the government of Kenya and had a line item in the national budget, high-level oversight from the Ministry of Education (MOE) and the Presidency through the Executive Office of the Deputy President (ODP), and a committed national steering committee.

The program recruited and trained recent university graduates to assist in low-performing public primary schools throughout the country. While in their host communities, workers – referred to as volunteer graduate assistants in the program – were also encouraged to initiate programs for community development. G-United was intended to increase literacy for primary school pupils, build national unity through social interactions between volunteers and host communities, and provide volunteers with useful skills that would help them in the labor market.

Volunteer wages were paid by the government of Kenya. While G-United volunteers were employed by the Ministry of Education, they were not hired as civil service teachers and did not replace civil service teachers in schools.[10]

Evidence Action, an NGO, assisted in the planning and implementation of the program starting in the second year of the program. The organization provided technical assistance to the program, managing the recruitment, training, and placement of G-United volunteers. Evidence Action raised funds to support technical assistance, program improvements, and activities related to monitoring and evaluation.

Before 2018, G-United assigned volunteers to counties at random. In the spirit of promoting national unity through social interactions of diverse groups, applicants could be assigned to any county, excluding their

[6] In many cases, a person might never have lived in their home county, and may identify with the county solely through their ethnic identity.

[7] Ethnicity is highly salient in political discourse in Kenya, with ethnic identity playing a large role in voting behavior (Bratton and Kimenyi, 2008) and political patronage (Burgess et al., 2015; Kramon and Posner, 2016; Muhula, 2009). Political alliances between ethnic groups have changed over time.

[8] Githinji, Irene, "Report wants teachers delocalised," *People Daily*. October 22, 2020.

[9] The program was originally known as the Kenya Future Leaders Programme (KFLP).

[10] The program is deliberately distinct from the assignment of regular teachers to schools, done through the Teachers Service Commission (TSC). Volunteers in the G-United program only teach outside of normal school hours.

home county (defined as their county of ancestry). In 2017, the third year of the program, only 48 percent of the 1,008 applicants who were invited to training arrived at training. Motivated by low volunteer retention rates, Evidence Action began eliciting applicants' preferences over counties and assigning a random subset of volunteers using a mechanism that took into account worker location preferences for the 2018 cohort.

In 2018, the year of the study, the G-United program selected 658 schools in 20 counties to receive volunteers. Each county held 30 and 34 program schools, with two positions per school.[11] Figure 2 shows the counties that received workers. A complete timeline of activities, each described in more detail below, can be found in Figure 1.

2.2.1 Recruitment and Screening

The program posted advertisements on Facebook and in online and print media outlets. The program also received publicity through news coverage and opinion pieces written in the popular press. Evidence Action screened initial applicants, limiting to those within the age and education restrictions. In March 2018, the program invited 4,457 eligible applicants to submit their preferences through an online survey. Applicant needed to rank all 20 counties participating in G-United, except their home county. A total of 2,332 applicants completed the preference elicitation survey.[12]

2.2.2 Assignment

Assignment took place in March 2018. In the first round, 1,285 volunteers were assigned directly to one of the 658 program schools.[13]

After submitting preferences, applicants were partitioned randomly into two treatment arms: Constrained-worker-optimal stable match (CWOSM) assignment and at-random assignment (ARA).The CWOSM arm assigned applicants to counties using their stated location preferences using the worker-optimal the random serial dictatorship algorithm. It is constrained insofar as workers were could not be assigned to their home county.

Individual slots in schools (two per school) were also divided into CWOSM and ARA slots. Schools therefore were randomly selected to receive either zero (25 percent chance), one (50 percent chance), or two (25 percent chance) CWOSM-assigned applicants.

Applicants received a text message informing them of their county assignment and inviting them to one of two trainings. Training was mandatory for deployment. As the program received information on declined offers, additional rounds of assignment were conducted to fill empty positions. Appendix Table A2 describes the number of assignments/invites and retention in each round.[14] Ultimately, 1,228 applicants attended the mandatory training. Out of these, 1,125 had completed the preference survey. The analysis in this study includes only those who did complete the preference survey.

[11] A group of stakeholders based in each county, including representatives of various government agencies and teachers unions, selected the set of schools eligible to receive volunteers. The County Directors of Education and the County Coordinators chaired the teams. The program instructed the teams to select high-needs schools, but left interpretation to the county teams.

[12] Completion of the preference survey gave applicants priority in assignment, but some who did not complete it were still assigned in the final round of assignment. These applicants are not included in this analysis.

[13] In the first round, 30 applicants were deterministically assigned to satisfy requirements of the Ministry of Education. One of these requirements was that an applicant from every county should be assigned. Applicants from counties with very few applicants were therefore assigned outside the process used for other applicants. These applicants are not included in the analysis of effects of the program.

[14] It is important to note that the types of counties that remain unassigned in each subsequent round are potentially endogenous to the assignment mechanism used in earlier rounds. The positions available at the end of a round of assignment may not be representative of the positions available in the first round. This will be considered in the analysis below.

2.2.3 Deployment

Volunteers were deployed to schools in April 2018 and served for a term of six months. Volunteers stayed with host families during their service. In schools, volunteers administered literacy tests and organized up to five two-week-long literacy "camps" for students. These camps were modeled off of previous Teaching at the Right Level programs, involving individual and small-group instruction targeted at each student's needs. The camps took place before and after normal school hours, on regular school days. Volunteers were expected to maintain regular contact with G-United administrative staff regarding the organization of the camps and the number of pupils that progressed according to the results of the literacy tests. In addition to the literacy camps, volunteers were encouraged to work with their host communities and initiate "community projects." For examples, volunteers organized tree plantings, taught about new agricultural techniques, and organized life-skills workshops with older students. The details of the projects were left open to the volunteers.

3 Data

This study links data on applicants and communities served by the G-United program. We compile application data from potential volunteers from the 2018 G-United cohort, including 1) background demographic characteristics and interest in the program, 2) preferences over geographic placements while volunteering, and 3) the magnitude of these preferences relative to potential changes in salaries. Data on applications and preferences over locations are also available for the eventually-cancelled 2019 cohort applicants. Additionally, follow-up data were collected from three categories of people. For all applicants, two rounds of follow-up surveys collected information about where they were living, their employment, and their beliefs on a range of questions dealing with national unity. No follow-up was conducted with the 2019 cohort of applicants. Among pupils at schools where volunteers did or would have been assigned, two rounds of literacy tests were administered, in both English and Kiswahili, and one round of surveys to measure students' feelings toward others. Finally, two surveys with adult community members (head teachers and homestay hosts) where volunteers served were conducted, regarding their experience with the program and national-unity-related beliefs. We describe each of these in more detail below.

3.1 Application data

Applicants submitted a survey online to apply for the program. The initial application collected basic background data, including the applicant's national ID, home county, current residence, education level and diploma status, field of study, and employment status. Applicants also answered questions regarding aptitude, motivations to participate in the program, and willingness to serve in the conditions available to volunteers.[15]

Table 2 provides descriptive statistics for the sample of applicants who completed the preference survey and are therefore included in the analysis to follow. We present mean characteristics in the ARA group (Column 1) and differences for those assigned via CWOSM (Column 2), above and below the lottery cut-off (Column 3), and by differences in Assignment characteristics (Columns 4-8). Applicants qwew on average 24.6 years old, and 40 percent are women. Only 4 percent of applicants reported that they had a job at baseline. Applicants come from a broad range of academic backgrounds. The most common field of

[15]Specifically, applicants were asked if they would be "comfortable in a rural setup."

study is business/commerce/economics (33 percent), followed by humanities/social sciences (23 percent) and education (13 percent).

Applicants disproportionately represent the largest ethnic groups. Among applicants 23 percent are Kikuyu, 18 percent are Kalenjin, 16 percent are Luhya, and 16 percent are Luo, and 8 percent are Kamba. Of the five largest ethnic groups in Kenya, only Kambas are under-represented.[16]

Most applicants reported pro-social motivations for applying. Seventy-three percent report that their primary motivation is "Giving back to the community." Thirteen percent reported that they are primarily motivated by a desire to encourage national cohesion. Only 4 percent reported personal development, and less than 0.1 percent reported that they are primarily motivated by the opportunity to receive a stipend.

At baseline, the majority of applicants (54 percent) were already residing outside of their home county. On average, applicants lived in counties where about 47 percent of the population shares their ethnic identity, and 48 percent are in the majority group in their county of residence.[17]

3.1.1 County Preference Survey (CPS)

One month after the *initial* application, eligible applicants were invited to submit their geographic preferences through another web-based survey via SMS. Each eligible applicant was asked to complete the county preference survey (CPS) followed by one of two surveys to elicit monetary-equivalent magnitudes of their preferences.

In the county preference survey (CPS), preference data were elicited after informing applicants that truthful reporting of preferences was an optimal strategy.[18] Applicants first ranked all available options. After that, they continued to a screen where they were asked in which counties they would be unwilling to serve, given the monthly stipend of 6,000 Ksh.[19] Only 25.7 percent of applicants ranked any counties, including their least preferred, as "unacceptable" at a salary of 6,000 KSh.

[16]The population shares of these groups in the general population of Kenya are 16 percent Kikuyu, 14 percent Kalenjin, 12 percent Luhya, 10 percent Luo, and 9 percent Kamba.

[17]According to DHS data from 2014, the average Kenyan adult between the ages of 20 and 30 lived in a county where their ethnic group was 59 percent of the population, and 62 percent were the majority group in their residence county. Adults between the ages of 20 and 30 who have more than a secondary school degree live in counties where their group accounts for 50 percent of the population and 51 percent were in the majority in their county of residence.

[18]The form specifically said

> We are now going to ask you to rank the counties where you are eligible to volunteer in the order of your most preferred county of deployment [rank 1] to your least preferred county of deployment [rank 19].
>
> To make the matching process work best for you, please rank the counties in order of **your true preferences**.
>
> Whilst we cannot guarantee that everyone will be allocated to one of their most preferred counties, we will try our best within the constraints that we face.

[19]This information was collected using "tick boxes" allowing applicants to check each county where they were unwilling to work. We assume that someone is unwilling to work at all counties ranked below the highest ranked county they listed as "unacceptable," whether they checked the boxes on all subsequent counties or not. A subset of applicants listed that they would not be willing to work at their $(n-1)$th or $(n+1)$th choices, but would be willing to work at their nth choice. A likely explanation is that some applicants did not fully understand the given instructions, and they did not feel the need to check all boxes. We, therefore, infer that someone is unwilling to work at all counties ranked below the highest ranked county they listed as "unacceptable" at the stipend amount of 6,000 Ksh. There are other options for handling the data on "unacceptable" locations. For example, we could use the reported data directly and not impute at all. This leads to a smaller gap in retention between those assigned to acceptable vs. unacceptable counties, suggesting that this choice increases the measurement error contained in the characteristic "unacceptable."

4 Empirical Framework

We are interested in understanding the impact on workers and members of the community of assignments that are less preferred by workers.

Empirically, we want to know to estimate specifications of the the following form:

$$Y = \alpha + \beta M + X'\Gamma + \varepsilon \tag{1}$$

where Y may represent an attribute of a volunteer or a member of the community, M represents some characteristic of the match between the volunteer and the community, and X is some set of controls on which potential outcomes of Y are conditionally independent. In our study, identification depends on whether we are estimating impacts on volunteers, for whom assignment to a particular M varies endogenously with stated preferences and the applicants' home county, and the communities, for whom M depends on the pool of workers assigned to the community's county. We describe the details below.

We also examine the impacts of the program itself on applicants and communities. Our interest in this arises from two sources. First, for applicants, participation is sensitive to M. Participation itself is a potential mechanism through which M could affect Y. Second, the impact of the program provides useful context for the interpretation of the results. In many cases, the interpretation of the impact of match characteristics M on an outcome may depend on the average impact of the program. If the program reduces literacy skills among children, a positive estimate of the impact of a characteristic M may indicate that matches with more M are less harmful. If the program raises literacy, greater M means they are more impactful.

IS1: For applicants, we evaluate the impact of the program using the lottery tiebreaker. In addition to determining whether an applicant received an offer, the lottery tiebreaker also determined the applicants' priority in receiving any offer. Specifically, we use as the instrument an indicator for whether they were assigned a lottery number lower than the cut-off for assignment in the first three rounds. The cut-off varies across applicants depending on their home county.

IS2: For members of the community, the impact is estimated using controlled linear regression using a strategy similar to Kline and Moretti (2014). We compare outcomes in communities that received the program in 2018 to those that were scheduled to receive the program in 2019 (cohort 5) but did not receive it because the 2019 program was unexpectedly canceled. The types of communities targeted in each cohort may have changed over time. However, schools appear to be similar in terms of observable characteristics. If anything, cohort 5 schools are positively selected, having slightly higher average KCPE scores and more night lights, so comparisons of cohort 5 to 4 schools could *understate* any positive impact of the program. In specifications using this identification strategy, we also include controls for the school's average KCPE score.

IS3: To evaluate the impact of assigned match characteristics M on outcomes at the community level we estimate specifications where M is a characteristic that varies at the level of the county (e.g., distance of the applicant's home county to the community's county). Because each county was assigned a different pool of applicants, X must contain a county fixed effect. Identification comes from the randomized assignment of applicants to communities within counties.

IS4: The impact of match characteristics M on applicants is identified using an approach based on Angrist et al. (2022). Match characteristics of assignments are endogenous to individual applicants' stated preferences and home county. If an applicant's rank order list r, where K is the number of counties to choose from, and their home county is indicated by the categorical variable h, then applicants can be distinguished

by types $\theta = (r, h)$. For applicants with the same θ the expected match characteristic, $E[M\theta]$ before running the assignment algorithm, is equal. Variation in $M - E[M|\theta]$ among applicants with the same θ is induced by the random variation in the lottery tiebreaker. The residual variation in M controlling for the expectation of $E[M|\theta]$ is therefore independent of potential outcomes (Robinson, 2017; Borusyak and Hull, 2020; Angrist et al., 2022). In our estimates of the impact of match characteristics on applicants, we, therefore, control linearly for the expected match characteristic over 10,000 simulations of the assignment algorithm described above (including the initial step of randomly partitioning the applicant sample into a CWOSM and ARA group).

5 Results

We report impacts first on workers, adult community members, and then students. We examine the impact of the program and assignment of workers and communities to different matches.

5.1 Impacts on workers

5.1.1 Worker preferences

This section shows that worker preferences are at odds with the objectives of fostering social interactions between diverse groups and serving poorer regions. We also find that there is considerable heterogeneity in preferences.

We examine preferences using three estimation strategies 1) multinomial demand estimation (McFadden, 1973), 2) rank-order-logit estimation (Beggs, Cardell and Hausman, 1981), and 3) a novel hedonic approach using the elicited compensating variation amounts. The latter approach uses OLS regression of elicited compensating variation on characteristics of the county. For each estimation strategy, we report univariate specifications with a single regressor and multivariate specifications. The univariate specifications may be more informative about how giving way to worker welfare might affect assignments in aggregate, while multivariate specifications are more useful for understanding, for example, whether a preference for a particular characteristic reflects a correlated, excluded characteristic.

Hedonic specifications focus on data obtained using the "top value auction" described above, as we feel there is compelling evidence that these data are higher quality.

Data. In estimating preferences, we examine five characteristics of county assignments: poverty rates, an indicator for whether the county is a marginalized county as classified by the Commission on Revenue Allocation (2013)., the distance between the applicant's home or residence county and the county of assignment, and the ethnic similarity between the applicant's home county and the county of assignment. Physical distance between two counties is defined as the linear distance between county centroids. We measure ethnic similarity between counties c and d as the probability that two residents, drawn at random, are of the same ethnic group.[20] Poverty rates are those published in the Basic Report on Well-Being in Kenya (page 46)

[20] That is, $\sum_g s_{cg} s_{dg}$, where g indexes ethnic groups.

(Kenya National Bureau of Statistics, 2016).[21][22]

The data used for estimation come from preference surveys.

Preference estimates Broadly, the results can be characterized – with a few exceptions – as showing an aversion to working in counties that are marginalized counties, have higher poverty, and are at a greater distance from an applicants' home (whether their home county or residence county), and a preference for counties that are more ethnically similar. For Cohort 4 (Panel A Table 3)), coefficient estimates for poverty rate in univariate specifications are all negative and highly statistically significant. In Cohort 5 (Panel B), the univariate conditional logit estimate is positive and highly statistically significant, although rank-ordered logit results are more consistent with the Cohort 4 results.[23]

Across all specifications and for all cohorts, applicants reveal strong preferences for counties that have similar ethnic group compositions to their home counties: the univariate hedonic estimate indicates an applicant needs a 1,727 shilling raise (almost 1/3 of their wage) to be indifferent between one county and another that is 10 percentage points more ethnically similar.

Hedonic estimates also indicate that applicants, on average, would need a 3.7 percent increase in their wage to compensate for assignment to a county with a 10 percentage point higher poverty rate, a 61 percent wage increase to be assigned to an ASAL county, and an 30 percent increase to be assigned a county 100 kilometers further from their home county.

5.1.2 The effect of CWOSM on retention and assignment characteristics

This section examines how assignment by CWOSM (Constrained Worker-Optimal Stable Matching), compared with at-random assignment (ARA), influences the characteristics of the positions that applicants receive and retention. We examine the effects using all applicants who were assigned in the first round. We also report results specifically for high-poverty counties (those with above median food poverty according to KNBS – 36 percent) and marginalized counties. We also report results for a sample of applicants whose reported preferences we trust more, those who – based on estimates using the procedure of Revelt and Train (2000) – are determined to be averse to assignments that are far from their residence county. This is consistent with the observation that all applicants in a more recent round of preference data collection were averse to distance. Finally, we report results separately for female and male applicants.

CWOSM assignment increased participation in the program. The overall rate of retention at the training stage in the first round of assignment was 53 percent. The CWOSM sample had higher retention by 5 percentage points (Panel A, Table 4). Only 41 percent of applicants completed the full six-month service

[21] Alternative measures could be constructed using information on the applicants' ethnic group. In practice, we find that the two measures are highly collinear. We opt for using the applicants' home county both because this may more fully capture novel social interactions and because applicants' ethnic groups were not observed until post-assignment follow-up. Prior work suggests that observed individual ethnic identification could be endogenous to the exposure to other groups (Okunogbe, 2018; Bazzi, Rothenberg and Wong, 2019).

[22] The DHS data do not contain some minority languages, and in those cases, respondents are identified as "other ethnic group." We treat respondents identified as "other" as belonging to distinct, locally concentrated, and county-specific ethnic groups. In practice, this choice has little effect on the results, as only Marsabit and Tana River have more than 6 percent identified as "other."

[23] The positive coefficient appears to be driven by a strong tendency among Cohort 5 applicants to rank Busia county first, which has one of the highest poverty rates of the 22 counties (0.595). Busia was not included among program counties in Cohort 4. Busia City is a commercially important border crossing with potentially a good labor market for university graduates. Another explanation is that Busia would have appeared early in the list of options. We do find a small preference for alphabetical order of county names. This could have arisen from some applicants ranking the options in the order of their appearance. Excluding Busia from choice sets in estimating the conditional logit specification yields a coefficient of -1.49 for county poverty.

period. CWOSM-assigned applicants are 3 percentage points more likely to complete service, although this difference is not statistically significant. Retention in high-poverty counties is 50 percent in the ARA group, 10 percentage points higher under CWOSM. The retention effect is especially stark in marginalized counties where only 42 percent in the ARA showed up for training compared to 59 percent in the CWOSM sample, more than eliminating the 9 point gap in retention between marginalized and non-marginalized counties in the retention rate. CWOSM also has a 14 percentage point effect on completion of service rates, suggesting that giving workers the ability to select their postings may be especially important for ensuring service delivery in the most marginalized areas.

Estimates of the effect of CWOSM on retention may understate the potential benefit of allowing for worker choice if there is error in the preference data. We also report results using only the subset of distance-averse applicants who appear to have expressed preferences that are more similar to later cohorts for which the preference elicitation process was clearer and streamlined. Retention among this group in the ARA condition is 52 percent and CWOSM assignment increases retention by 8 percentage points.[24]

Assignment by CWOSM is associated with an increase in worker welfare equivalent to approximately two-thirds of their salary. Relative to the ARA group, workers are assigned to postings associated with 4,230 KSh lower compensating variation as indicated by their stated "auction values. Most CWOSM applicants received their first choice (86 percent compared with 6 percent in the ARA group) (Column 2, Panel B, Table 4). Assignment by CWOSM increases the distance between an applicant's home and assignment counties and has only a small positive effect on the ethnic similarity of the assignment.

The effects on welfare and first choices are similar when restricting the sample to high-poverty or marginalized counties (Columns 3-6).

For applicants who are identified as distance averse, the CWOSM assignment increases an applicant's expected welfare by 4,680 KES and increases the probability that an applicant is assigned to their top choice by 80 percentage points. As would be expected, given that they have expressed a relatively strong preference to be close to their residence county, these applicants are assigned closer to their residence county under CWOSM.

5.1.3 Impact of the program on applicants

We estimate the effect of participating in the G-United program using random variation in the timing of offers induced by a random lottery number. As part of the assignment process, all applicants were given a random lottery number r_i, which determined the round in which an applicant received an offer, and subsequently influenced the likelihood of attending training.[25] Only 32 percent of applicants who were given offers in the fourth round of assignment attended the training, compared over 50 percent in the first three rounds (see Appendix Table A2). Figure 5 illustrates the relationship between the random lottery number, the round of assignment, and the likelihood of attending the mandatory training. Those with higher lottery numbers are much less likely to participate in the program than those with lower numbers, despite having the same expected unobservable background characteristics.

The relationship between the lottery number and the probability of serving forms the basis of a natural experiment that can be used to estimate the causal effect of serving in the program. We form a binary

[24]Part of this effect could be because these applicants are systematically assigned to more desirable counties under CWOSM because other applicants have reported their preferences with errors. However, this result is robust to including controls for the specific county to which the applicant was assigned (not reported).

[25]Later offers may not be representative of earlier offers in that they are "picked over." This "picked over" effect could represent a violation of the exclusion restriction. For this reason, most of our analysis relies only on those assigned in the first round.

instrument $Z_i = \mathbb{1}\{r_i \leq 0.8\}$ which is then used to predict service in the program.[26] Applicants with a lottery number below 0.8 are 42 percentage points more likely to attend the mandatory training and the F-statistic for the first stage is 105.

We use two measures of program participation: attendance at the mandatory training (short-run) and completion of service (long-term). We examine effects on social attitudes and county of placement using OLS and 2SLS using the randomly assigned lottery number that increased the likelihood of being assigned early as an instrument.

Applicants who attend training score 0.06 SDs higher on an index of national unity attitudes according to OLS estimates, although this difference is not statistically significant (Table 5).[27] Those who complete service score 0.16 SDs higher on the national unity index, and this difference is highly statistically significant. These results may reflect selection into participation if more socially motivated applicants are also more likely to participate. However, 2SLS estimates have similar point estimates, but in all cases, they are statistically insignificant due to larger standard errors.

Overall, 30 percent of applicants voted for their home county in the voting game. Applicants who attend training are 14 percentage points less likely to vote for their home county, and applicants who complete their service are 11 percentage points less likely to vote for their home county. The 2SLS estimates show that applicants induced to attend training by variation in the timing of the offer are 19 percentage points less likely to vote for their home county, and applicants who are induced to complete their service by variation in the timing of the offer are 24 percentage points less likely to vote for their home county. We also find more votes for more distant and dissimilar counties, although the results are statistically insignificant.

In contrast to voting patterns, applicants who attended training are more likely to be living in their home county. 2SLS estimates are similar, but not statistically significant. These results are in contrast to those in a related study by Okunogbe (2018). One potential hypothesis is a "lock-in effect similar to that observed in many active labor market programs (Card, Kluve and Weber, 2017) whereby applicants who otherwise would have found more permanent arrangements outside of their home county are prevented from searching for work outside their home.[28]

5.1.4 Impact of match characteristics on applicants

This section leverages random variation in assignment to estimate the effect of being assigned to locations that are associated with higher levels of applicant representative utility on applicant social attitudes and geographic preferences.Preferences estimates above indicated that applicants prefer locations that are more ethnically similar, closer to their home. They also prefer non-marginalized and lower-poverty counties. We examine the impacts of these characteristics of assignments directly. We also estimate the effect of assignments to those locations with greater representative utility that incorporates all of the characteristics that appear to influence applicant preferences. To do so, we estimate a discrete choice conditional logit model including these parameters (as well as indicators for whether the county is classified as "arid or semi-arid and

[26] It is also possible to construct multiple instruments from the running lottery number r_i. We do not find that these have a significant effect on the precision of the resulting estimates. We therefore adopt the parsimonious option of identifying a single cut-off.

[27] This index is constructed as the first principal component of responses to a feeling thermometer (a set of indicators from a ten-point scale), reporting of ethnic compared with national identity (a set of indicators from a five-point scale), agreement with the statement, "It is valuable to interact with people from other ethnic groups," and an indicator for whether the respondents' first two reported friends share their mother tongue.

[28] Consistent with this we find some evidence that participation in the program is associated with slightly lower rates of employment, although the estimates are not statistically significant.

whether it is physically adjacent to their home county) and use the resulting estimates to predict applicants representative utility from assignment to each county. .

We focus on estimates of the intent-to-treat effect of assignment to a county with given characteristics, noting that these may combine separate impacts on program participation and the effects of working in a location with those characteristics. In each case, we examine whether assignment to less preferred locations is associated with lower retention. We discuss interpretation of the effects in light of the participation effects below. As noted in section 5.1.3, volunteering is associated with changes in geographic preferences measured by the voting task (Table 5)..

Assignment to marginalized or high-poverty locations seems to have a large negative effect on retention (Columns 1 and 2, Table 6). When an applicant is assigned to a marginalized county, they are 8.7 percentage points less likely to attend training. The effect on completion is smaller (4 percentage points) and is not statistically significantly different from zero. This suggests that applicants on the training attendance margin may be likely to drop out later, even when assigned to more favorable locations.

Assignment to more ethnically similar counties increases retention. Assigning an applicant to a county that is 10 percentage points more ethnically similar increases the retention rate by 2.3 percentage points. The corresponding effect on completion is similar at around 2 percentage points. Although in this setting, assignment to an applicant's home county was impossible, this result suggests that if applicants were able to serve in their home counties (where ethnically similarity could be 5-6 times higher than in their choice set in this program), retention rates could be substantially higher. The effect of assigning applicants to counties that are more preferred to the representative applicant are associated with similar retention effects to ethnic similarity, which is expected given the strong influence of ethnic similarity in the estimates of representative utility.

Overall, variation in assignment characteristics has little or no effect on the general social attitudes of applicants (Columns 3 and 4 of Table 6). None of the assignment characteristics considered has any notable or statistically significant effect. We also do not find any statistically significant effect of assignment characteristics on home bias in the voting game (Column 5 of Table 6).

Applicants preferences for the geographic distribution of resources appears to be highly sensitive to the types of locations to which they are assigned. Despite being more likely to volunteer and complete their service, applicants assigned to more preferred locations are slgihtly more likely to vote for their home county in the voting task, although the estimate is not statistically distinguishable from zero. Assignment to more preferred locations is associated with a preference to distribute resources to more ethnically similar locations, and a lower preference for distributing resources to marginalized counties. We find comparable effects using measures of the ethnic similarity between the applicants home county and the county of deployment, as well as home county distance. Applicants assigned to marginalized and higher-poverty counties are also more likely to vote to distribute resources to those counties.

Assignment characteristics may also have some influence on where volunteers choose to reside after the program. Volunteers assigned marginalized and higher-poverty counties are more likely to be living in marginalized counties after the program. This may be especially surprising given the negative effect of assignment to a marginalized county on program participation. We do not find any evidence that other characteristics matter for residential location.

5.2 Impacts on communities

5.2.1 Impacts on pupils

We first examine whether the program, compared to its absence, benefits pupils as measured by learning gains. To provide empirical evidence on the impact of the program, we compare schools that received the program in Cohort 4 to those that were scheduled but did not receive the program in Cohort 5. More details of this data collection efforts with pupils are included in Appendix B.

Evidence Action collected follow-up data twice in both 2018 program schools and another set of schools scheduled to participate as 2019 program schools. Schools were selected for the 2019 cohort using the same deliberative process described for 2018. Both rounds of data collection included pupil test scores, and the second survey also included extensive surveys of pupils and community members regarding national unity outcomes.

The first round, in January 2019, included 1,626 schools, of which 274 had previously received a volunteer.[29] In each school, 8 pupils in Standards 2 and 3, the target grades for G-United, were selected to be tested using the Uwezo assessment tool, first in Kiswahili and then in English.

In the second round, in February and March 2020, Evidence Action collected school and community follow-up data in all 658 schools that participated in the 2018 cohort of G-United, along with a sub-sample of 150 schools, randomly selected from the planned 2019 cohort. This data collection included another round of pupil assessments in English, Kiswahili, and civics, along with pupil and community member surveys about inter-ethnic trust. Again, 8 students were selected, this time sampling two each from the rosters for Standards 1 to 4 in 2018.

To estimate the effects of the G-United program on pupil learning outcomes, we estimate a linear regression controlling for the school's KCPE score and enrollment levels, and the student's gender and age.[30] Identification of the effect of the program relies on the assumption that, after controlling for observable characteristics, within this set of treated and almost-treated schools, assignment to G-United is independent of potential outcomes.[31]

Pupils in schools that received G-United in 2018 (Cohort 4) outperformed the comparison schools in Kiswahili in both the short and the medium term (Figure 3). These effects are larger for targeted standards, although the difference is only significant for the short term. In civics, which was measured only in the medium term, we see small and positive point estimates, but they are not statistically significantly different from zero. Estimates indicate gains in English as well, although these are not statistically distinguishable from zero. Appendix Table A1 reports impacts on additional outcomes, showing positive but statistically insignificant effects on civics assessment and marginally significant effects on reported university aspirations.

For all learning outcomes, the effect of the program appears to be larger for pupils in targeted standards. Consistent with the intended design in which Standards 2 and 3 were the focus of volunteers' instruction and literacy camps, Figure 4 show that pupils in targeted standards received the program in greater numbers.[32] However, consistent with survey evidence, these results point to non-targeted pupils receiving the intervention in high numbers.

[29] This round of data collection was initially intended as a baseline for an evaluation of the program in 2019.

[30] We do not find evidence that grade progression and subsequently age in grade is influenced by assignment to G-United. The results are robust to including age as a control.

[31] This is consistent with what we see comparing school characteristics in Cohort 4 (Column 1) and Cohort 5 (Column 6) in Table 1. If anything, we find that older students who recently took the KCPE exam - who were unlikely to have been influenced by the program - are higher achieving in comparison schools.

[32] Appendix Table A10 provides the same results in table form.

5.2.2 Effects of match characteristics on pupil literacy outcomes

In this section, we estimate the effect of assignment to workers who are more ethnically similar to the county. The identification strategy leverages the randomized assignment of volunteers within counties. We estimate Equation **??** where the outcome variable is a pupil's literacy outcome. We report results separately for the full sample, targeted standards, and untargeted standards.

Assignment of more ethnically similar volunteers is associated with substantial negative effects on pupil outcomes (Table 8). We also find that these effects are most pronounced for Kiswahili. We do not see a difference in learning outcomes as a function of the distance between the volunteer's home county and assignment county. These results may partly be explained by lower effort of workers when serving in locations that they find to be more desirable. We find some evidence that more ethnically similar volunteers generate less recalled exposure to the program, although this effect is only marginally statistically significant. This effect is larger in magnitude for pupils in untargeted populations.

5.2.3 Effects of match characteristics on pupil interactions and trust

In this section, we examine whether pupils report greater exposure to other ethnic groups. One concern with interpreting these results is that children know the ethnicity of the volunteers, and they also may not have developed associations with different ethnic groups. Ethnic distinctions are learned over time. If ethnic identities become more salient with early life socialization, children may not attend to specific information about ethnic identity.

Pupils report low levels of exposure to other ethnic groups. Pupils report having met a person from a randomly selected ethnic group (excluding responses where the ethnic group in question is the pupil's own) only 15.9 percent of the time. Pupils are about 7.7 percentage points more likely to report having interacted with a person from the ethnic group of their round 1-assigned applicant (Table 7). This effect is concentrated among those pupils who were in standards that were targeted for the intervention: pupils in targeted standards were 8.5 percentage points more likely to have ever met a person from their round-1-assigned volunteer's ethnic group.

On the one hand, this is likely an underestimate of the exposure of a pupil to a volunteer through this program as not all volunteers assigned in round 1 ended up volunteering. Compliance with round 1 assignments was about 50 percent. Still, this may seem like a small effect as 50 percent of pupils in targeted standards in schools that did receive a G-United volunteer report having taken part in the program. Young pupils may simply not know or notice the ethnicity of the volunteer.

Pupils also report low levels of baseline exposure to people from other counties. Pupils report having met a person from a randomly selected county (excluding responses where the home county is the county of the pupil's school) only 9.5 percent of the time (Table 7). Pupils are about 1.3 percentage points more likely to report having met someone from a county when it is the home county of one of the assigned applicants.

Pupils report slightly higher distrust toward the assigned applicants' ethnic group. Pupils are 1.8 percentage points more likely to identify the applicants' ethnic group as the least trusted group, although this effect is not statistically significant. Among targeted standards, pupils identify the assigned applicant's county as the least trusted 3.6 percentage points more often. This may suggest that a small subset of pupils were induced to trust less the group represented by their volunteer. However, an alternative explanation - based on the way the question was asked - may be more likely given that this result is not corroborated when asking about applicants' counties and contradicts the increase in trust among adult respondents. Given this

inconsistency with other results, it is possible that the way the least trusted ethnic group was elicited induced this response. Pupils first identified the most trusted and then were asked to identify the least trusted. This design may have induced some pupils to name the group with which they were most familiar out of the remaining groups, and they may have not recognized the other groups listed.

5.2.4 Effects on adult community members

In this section we first examine whether the program had any impact on adults in the community, then examine how match characteristics affect community members' social attitudes.

Impact of the program on adult community members To evaluate whether the program affected adult community members, we first examine whether the program produced novel interactions and connections between adults in the community. We then turn to comparisons of the Cohort 4 schools to those in Cohort 5.

Each adult was asked a series of questions about the ethnic group of one – randomly selected in the case of communities with multiple applicants – applicant who was assigned to the community as well as two additional ethnic groups – randomly selected from the set of 20 ethnic groups represented by the applicants. For each of the three ethnic groups, respondents were asked if they had any friends from the ethnic group, to rate their level of trust toward the group on a five-point scale, and to predict the share of G-United applicants from that ethnic group who reported identifying more with their ethnic group than with a national Kenyan identity.

We find that the program increases the development of friendships between community members and volunteers. Among all adult respondents, the probability that someone is friends with a randomly selected non-coethnic person is 52.8 percent (Table 9). Respondents were 4.6 percentage points more likely to report being friends with a non-coethnic if one of the two volunteers in the community was from that ethnic group (Table 9). This effect is largest for colleagues (headteachers), for whom the effect is 8.8 percentage points.

In parallel with increasing friendships, distrust of the assigned applicant's ethnic group falls. Overall, 12.2 percent of adult respondents report distrusting a randomly selected non-coethnic person. Assigning an applicant to a community the probability that an adult reports distrusting a non-coethnic group by 3.2 percentage points, although this effect is not statistically significant. This effect is especially pronounced (and marginally significant) among homestay household members who also report slightly more distrust for other groups (12.9 percent compared to 11.2 percent for head teachers).

Impact of the program on adult social attitudes Homestay members could only be identified and interviewed in Cohort 4 schools, so the study can only evaluate the impact of the program on professional colleagues (head teachers). One concern with using headteachers is that a large share reported being newly placed in their schools, likely a reflection of Kenya's policy of 'delocalization'. To give the estimates a causal interpretation, one must assume that headteacher transfers were independent of the G-United program.

Relative to those cohort 5 headteachers (who never received a volunteer), those in cohort 4 are 10 percentage points more likely to report trusting people in their community and 8 percentage points more likely to report trusting people in Kenya in general, although only the latter is statistically significant, and that only marginally so (Panel A of Table 10).

Impact of match characteristics on adults We examine the effect of each assignment characteristic on levels of generalized trust and identification with a national Kenyan identity (as opposed to ethnic group) of assignments to volunteers based on match characteristics. All specifications include school county dummy controls, so that identification is from random assignment of applicants to schools within counties.

Overall, we find that assignment of more ethnically similar applicants is associated with lower levels of trust, especially for people in Kenya broadly (Panel B of Table 10). We also find that when applicants are assigned from more distant locations, this is associated with higher trust, although the effect is not statistically significant. Assigning applicants to their first choice and to counties that are predicted to be associated with higher levels of applicant utility also appears to reduce the impact of the program on trust outcomes. The effect of assigning workers to locations at which they have higher predicted utility has a large negative and highly statistically significant effect on the generalized trust of homestay members.

These effects are largely driven by the homestay families who may have lower levels of baseline social interaction. In all cases, the effect of assigning applicants to more similar, less distant, or more preferred locations seems to have a larger impact on trust toward people in Kenya generally.

6 Policy simulations

In this section, we examine the effect of different choices of assignment mechanisms. The results above have shown that there may be important trade-offs in terms of retention and applicant welfare on the one hand and generating meaningful inter-regional contact and program effectiveness on the other. We, therefore, consider whether there exist improvements in the assignment mechanism that can improve the assignment mechanism performance in terms of generating contact and applicant welfare. We consider using at-random assignment and CWOSM combined with either restricting the applicant's choice set to remove options that would generate less inter-regional contact and assigning preferences for counties over applicants so that applicants receive higher priority at more distant locations. In each case, we simulate the assignment mechanism 1,000 times to estimate the overall expected performance of the mechanism.

We start by simulating the assignment using ARA and CWOSM separately but for the full sample of applicants. This simulation reveals similar results to those reported in Table 4. Assignment by CWOSM increases applicant welfare by about 5,500 KES (Table (12)), more than 5/6 the applicant wage. We see that 89 percent of applicants receive their first choice when assigned by CWOSM. This compares to only 5 percent when assigned by ARA. The simulations also show that CWOSM assignment would, in expectation, reduce the average distance between applicant home counties and their assigned counties, but only by approximately 1 km. CWOSM assignment also generates more ethnically similar assignments (0.11 compared to 0.09 with ARA).

Graphically, we see very similar distributions of distance from residence county, distance from home county, and ethnic similarity between the ARA and CWOSM groups, yet very different distributions of the applicants' rank of the assigned county (Figure 6).

One concern with using this sample is that the preferences may contain substantial error. Pure noise in applicant preferences may give the impression that there is more heterogeneity in preferences, and therefore more scope to improve welfare through allocative efficiency than is truly the case. With this concern in mind, we conduct policy simulations with two different samples that remove this concern. First, we use the sample of applicants who, according to the procedure from Revelt and Train (2000), reported distance aversion in their preferences. Second, we use the Cohort 5 data, for which we have all of the useful information except

auction values.

Focusing on distance-averse applicants, we find that the expected welfare gain from using CWOSM is about 4,500 KES compared to ARA, smaller than when using the full sample. Only 68 percent of CWOSM-assigned applicants receive their first choice when restricting to distance-averse applicants. For this sample, CWOSM assignment reduces the average distance between applicants' home counties and their deployment counties and also increases the ethnic similarity of assignments.

Restricting choice sets to exclude options that are less than 100 km from the applicants' home county reduces the welfare of CWOSM-assigned applicants by less than 1,000 KES. Importantly it also increases the expected distance between applicant home county and assignment county by 274 km on average. This amount exceeds the average distance of the random algorithm which only yields an average home county distance of 262. CWOSM assignment also results in lower levels of ethnic similarity in matches.

Giving counties preferences over applicants so that applicants from more distant home counties are prioritized in assignment produces much more distant assignments with much lower applicant welfare. However, it is remarkable that CWOSM assignment with county distance preference for more distant applicants still yields higher welfare than ARA assignment using the status quo mechanism.

Replicating the simulation using Cohort 5 data yields similar comparisons. We again see a large improvement in applicants' ranking of their assigned spot (60% compared with 6% received their first choice). We again see only small differences in distance to assignment and ethnic similarity. Graphically, we see a similar pattern (Figure 7) in which assignment via CWOSM dramatically alters the rank of assignment but generates similar distributions of distances to home and residence counties as does at-random assignment.

7 Conclusion

Policymakers often have significant influence over where workers are posted. A substantial literature documenting the benefits of inter-ethnic contact is consistent with policymakers' attempts to use this authority to influence social interactions between diverse communities. In many cases, such as a recent program of "de-localization" of teachers in Kenya and longstanding policies of teacher assignment in Botswana, these policies may override the preferences of workers themselves. Consistent with this, we do find that providing volunteers with more preferred options may have important costs in terms of the program's effectiveness at increasing national unity and promoting literacy. Yet, this study has provided empirical evidence that policies that ignore workers' preferences can reduce the size of the labor pool and limit the strength of connections between workers and host communities, even across groups. As such, policies that assign workers without regard for their preferences over locations could have both private and social costs that should be carefully considered. Importantly, we see that a policy that incorporates workers' preferences within a limited choice set that aligns with government aims can simultaneously serve the governments' aims can increase worker welfare (and thereby effort and continued participation) while simultaneously contributing to the policymakers' goals. Future work can continue to illuminate these trade-offs to understand additional strategies that may be able to yield the greatest benefits with the fewest costs.

Our results add to a growing literature demonstrating that intergroup contact can influence trust, social identity, and national cohesion. Previous work has shown that exogenously generated interaction across groups yields more positive attitudes towards others (Key, 1949; Scacco and Warren, 2018; Corno, Ferrara and Burns, 2019; Rao, 2019; Lowe, 2018; Chen et al., 2020; Boisjoly et al., 2006; Mo and Conn, 2018), supporting the contact hypothesis (Allport, 1954). Recent research further suggests that increases in interactions across

groups early in life can even lead to very long-term changes in attitudes (Billings, Chyn and Haggag, 2020; Kaplan, Spenkuch and Tuttle, 2019).

Our study does not claim that promoting intergroup contact on its own can eliminate distrust between ethnic groups. Although there is substantial evidence that inter-ethnic contact can promote trust, it is also associated in some cases with greater conflict in the case of existing background conflict (Key, 1949). This risk may be especially salient in Kenya, where colonial-era land policies may have increased contact in the Rift Valley, the site of considerable inter-ethnic tension today.

A limitation of this study is that applicants are never assigned to their home counties, or their home communities. The relationship between worker effort could be non-linear in the degree of social connection to a place.

While we have shown a lot of heterogeneity of preferences, some counties are still broadly less popular assignments. We see that counties that have high poverty rates and counties that are remote are lower-ranked than wealthier and more centrally located ones. As we see in Table 5, assignment to a remote county reduces the likelihood of attending training by 14.5 percentage points. A high poverty rate similarly reduces participation. This analysis does not tell us that applicants dislike poverty or remoteness per se. Applicants may be responding to a distaste for characteristics that happen to be correlated with poverty. However, regardless of why applicants express these preferences, these locations can be harder to serve. This is similar to a standard difficulty faced by governments in serving less popular locations.

Our results suggest that incorporating worker preferences could in some cases be a very cost-effective way to improve social programs. We conservatively estimate the costs of implementing the CWOSM treatment arm to be 7,526 USD for the full cohort.[33] This includes the costs of administering the preference survey, SMS and email reminders to applicants to complete the survey, and personnel costs associated with administering the preference survey and running the assignment mechanism. We do not include costs of implementing G-United that are required regardless of the assignment mechanism or costs of the evaluation. This works out to just over 6 USD per applicant assigned via CWOSM. Of this, only 144 USD (for SMS messages) reflects a per-applicant cost. Therefore, the total cost would be similar with a larger sample of applicants assigned via CWOSM, and the per-applicant cost would be even smaller. Especially given the small cost of implementation, CWOSM assignment is likely to be on-net cost-saving. By increasing participation, even modestly, this assignment mechanism reduces the costs associated with assignment of workers in later rounds and tracking and replacing those who do not show up. Furthermore, the social value of eliciting preferences may be large due to increases in worker welfare. As noted above, these gains may be as much as two-thirds of the wage bill or over 50 percent of the program's total cost.

[33]The details of this calculation can be found in the Appendix.

References

Abdulkadiroğlu, Atila, and Tayfun Sönmez. 1998. "Random Serial Dictatorship and the Core From Random Endowments in House Allocation Problems." *Econometrica*, 66(3): 689–701.

Allport, Gordon. 1954. *The Nature of Prejudice.* Reading, MA: Addison-Wesley.

Angrist, Joshua, Guthrie Gray-Lobe, Clemence M Idoux, and Parag A Pathak. 2022. "Still Worth the Trip? School Busing Effects in Boston and New York." National Bureau of Economic Research Working Paper 30308, July.

Ba, Bocar, Patrick Bayer, Nayoung Rim, Roman Rivera, and Modibo Sidib. 2021. "Police Officer Assignment and Neighborhood Crime." National Bureau of Economic Research Working Paper 29243, September.

Bates, Michael D, Michael Dinerstein, Andrew C Johnston, and Isaac Sorkin. 2022. "Teacher Labor Market Equilibrium and Student Achievement." National Bureau of Economic Research.

Bazzi, Samuel, Alexander D. Rothenberg, and Maisy Wong. 2019. "Unity in Diversity? How Intergroup Contact Can Foster Nation Building." *American Economic Review*, 109(11).

Beggs, Steven, Scott Cardell, and Jerry Hausman. 1981. "Assessing the potential demand for electric cars." *Journal of Econometrics*, 17(1).

Berbner, Bastian. 2020. "Botswana's Radical Experiment in National Unity." *We Are Not Divided*. Link.

Berge, Lars, Ivar Oppedal, Kjetil Bjorvatn, Simon Galle, Edward Miguel, Daniel Posner, Bertil Tungodden, and Kelly Zhang. 2019. "Ethnically Biased? Experimental Evidence from Kenya." *Journal of the European Economic Association.*

Billings, Stephen, Eric Chyn, and Kareem Haggag. 2020. "The long-run effects of school racial diversity on political identity." National Bureau of Economic Research.

Boisjoly, Johanne, Greg J. Duncan, Michael Kremer, Dan M. Levy, and Jacque Eccles. 2006. "Empathy or Antipathy? The Impact of Diversity." *American Economic Review*, 96(5): 1890–1905. Link.

Borusyak, Kirill, and Peter Hull. 2020. "Non-Random Exposure to Exogenous Shocks: Theory and Applications." National Bureau of Economic Research Working Paper 27845, September.

Bratton, Michael, and Mwangi S. Kimenyi. 2008. "Voting in Kenya: Putting Ethnicity in Perspective." *Afrobarometer Working Papers*. Link.

Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria, and Gerard Padro i Miquel. 2015. "The value of democracy: Evidence from road building in Kenya." *American Economic Review*, 105(6): 1817–1851.

Card, David, Jochen Kluve, and Andrea Weber. 2017. "What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations." *Journal of the European Economic Association*, 16(3): 894–931.

Chen, Yi, Ziying Fan, Xiaomin Gu, and Li-An Zhou. 2020. "Arrival of Young Talent: The Send-down Movement and Rural Education in China." *American Economic Review*, 110(11): 3393–3430.

Commission on Revenue Allocation. 2013. "Policy on the Criteria for Identifying Marginalised Areas and Sharing of the Equalisation Fund, Financial Years 2011 to 2014." *Policy Document*, Reference Number: CRA/RF/Vol. I/44.

Corno, Lucia, Eliana La Ferrara, and Justine Burns. 2019. "Interaction, stereotypes and performance. Evidence from South Africa." Institute for Fiscal Studies. Link.

Davis, Jonathan. 2019. "The Short and Long Run Impacts of the Deferred Acceptance Algorithm: Evidence from Matching Teach For America Teachers to Schools." *Working Paper*. Link.

Dobbie, Will, and Roland G Fryer Jr. 2015. "The impact of voluntary youth service on future outcomes: Evidence from Teach For America." *The BE Journal of Economic Analysis & Policy*, 15(3): 1031–1065.

Kaplan, Ethan, Jörg L Spenkuch, and Cody Tuttle. 2019. "School Desegregation and Political Preferences: Long-Run Evidence from Kentucky."

Kasara, Kimuli. 2017. "Does Local Ethnic Segregation Lead to Violence?: Evidence from Kenya." *Quarterly Journal of Political Science*. Link.

Kenya National Bureau of Statistics. 2016. Link.

Kenya National Bureau of Statistics and ICF International. 2014. "Kenya Demographic and Health Survey 2014." *https://microdata.worldbank.org/index.php/catalog/2544/get-microdata*, Accessed: December, 15 2019.

Key, V.O. 1949. *Southern Politics in State and Nation.* Knoxville, TN: University of Tennessee Press.

Kline, Patrick M., and Enrico Moretti. 2014. "Local economic development, agglomeration economies, and the Big Push: 100 Years of evidence from the Tennessee Valley Authority." *The Quarterly Journal of Economics*.

Kramon, Eric, and Daniel N. Posner. 2016. "Ethnic Favoritism in Education in Kenya." *Quarterly Journal of Political Science*.

Lowe, Matt. 2018. "Types of contact: a field experiment on collaborative and adversarial caste integration." *Dissertation*.

McFadden, Daniel. 1973. "Conditional Logit Analysis of Quantitative Choice Behavior." In *Frontiers in Econometrics*. eds. P. Zarambka. New York: Academic Press.

Mo, Cecilia Hyunjung, and Katharine M Conn. 2018. "When do the advantaged see the disadvantages of others? A quasi-experimental study of national service." *American Political Science Review*, 112(4): 1016–1035.

Mousa, Salma. 2020. "Building social cohesion between Christians and Muslims through soccer in post-ISIS Iraq." *Science*, 369.

Muhula, Raymond. 2009. "Horizontal Inequalities and Ethno-regional politics in Kenya." *Kenya Studies Review*.

Nderitũ, Alice Wairimũ. 2017. *Kenya: Bridging Ethnic Divides.* Nairobi, Kenya: Mdahalo Bridging Divides Limited.

Okunogbe, Oyebola Motunrayo. 2018. "Does Exposure to Other Ethnic Regions Promote National Integration: Evidence from Nigeria." World Bank Group. Link.

Ramiah, Ananthi Al, and Miles Hewstone. 2012. "'Rallying around the flag': Can an intergroup contact intervention promote national unity?" *British Journal of Social Psychology*, 51: 239–256.

Rao, Gautam. 2019. "Familiarity does not breed contempt: Generosity, discrimination, and diversity in Delhi." *American Economic Review*, 109(3): 774–809.

Revelt, David, and Kenneth Train. 2000. "Customer-specific taste parameters and mixed logit." *Working Paper*.

Robinson, Amanda Lea. 2017. "Ethnic Diversity, Segregation and Ethnocentric Trust in Africa." *British Journal of Political Science*.

Roth, Alvin E. 1984. "The evolution of the labor market for medical interns and residents: A case study in game theory." *Journal of Political Economy*, 92: 991–1016.

Scacco, Alexandra, and Shana S. Warren. 2018. "Can Social Contact Reduce Prejudice and Discrimination? Evidence from a Field Experiment in Nigeria." *American Political Science Review*, 112(3): 654–677. Link.

Zeitlin, Andrew. 2021. "Teacher turnover in Rwanda." *Journal of African Economies*, 30(1): 81–102.

Tables

Table 1: Covariate balance of identification strategies for effects on communities

	Cohort 4 Mean (1)	Ethnic similarity (2)	Home distance (3)	First choice (4)	CWOSM assigned (5)	Comparison to Cohort 5 schools (6)
School characteristics						
Average KCPE score	219.12	0.52	-0.00	-4.22	-1.27	4.79
		(11.08)	(0.02)	(4.17)	(4.16)	(3.33)
		621	621	621	621	755
Night lights	-0.04	-0.22	0.00	0.14	0.08	0.21**
		(0.27)	(0.00)	(0.09)	(0.09)	(0.10)
		526	526	526	526	640
Number teachers	11.53	-2.32***	0.00	1.03**	1.08**	0.14
		(0.91)	(0.00)	(0.49)	(0.49)	(0.43)
		621	621	621	621	756
Pupil-teacher ratio	35.23	-2.31	-0.01	-1.56	-2.09	-0.62
		(3.50)	(0.01)	(1.60)	(1.53)	(1.20)
		621	621	621	621	756
Head teacher characteristics						
Delocalized head teacher	0.31	-0.00	-0.00	0.04	0.01	0.04
		(0.16)	(0.00)	(0.06)	(0.07)	(0.05)
		445	445	445	445	542
Living in home county	0.57	0.04	-0.00	-0.03	-0.05	-0.05
		(0.92)	(0.00)	(0.24)	(0.24)	(0.10)
		39	39	39	39	112
Male	0.77	-0.10	-0.00	0.02	0.03	-0.05
		(0.12)	(0.00)	(0.05)	(0.05)	(0.04)
		621	621	621	621	756
Homestay family characteristics						
Household head	0.78	-0.25*	0.00*	-0.01	0.02	
		(0.15)	(0.00)	(0.05)	(0.05)	
		626	626	626	626	
Living in home county	0.73	0.03	0.00	-0.17	-0.31*	
		(0.20)	(0.00)	(0.18)	(0.17)	
		39	39	39	39	
Male	0.45	0.15	0.00	0.01	-0.02	
		(0.13)	(0.00)	(0.06)	(0.06)	
		629	629	629	629	

Notes: Table reports balance on covariates of independent variables used for identification of effects of volunteer characteristics and program participation on communities. All specifications include county fixed effects. Standard errors (reported in parentheses) are clustered at the school level where there are multiple observations per school. Home stay families were not interviewed in Cohort 5 communities. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 2: Balance on covariates for applicants

				Assignment characteristics				
	ARA mean (1)	CWOSM assignment (2)	Lottery cutoff (3)	Marginalized county (4)	Poverty rate (5)	Distance from home county (6)	Distance from residence county (7)	Ethnic similarity (8)
General characteristics								
Age	24.60	0.071	0.041	0.169	0.014***	0.054	0.052	0.036
Female	0.40	-0.003	-0.016	-0.018	-0.001	-0.005	-0.001	0.155*
Employed at baseline	0.04	-0.008	0.008	0.017	-0.000	0.006	0.002	-0.007
Marginalized home county	0.02	0.008	-0.002	0.007	0.000	-0.003	-0.004	0.010
Home county poverty rate	31.26	1.589***	-0.748	-0.506	0.022	-0.589***	-0.527***	0.120
Marginalized residence county	0.02	0.004	-0.002	0.005	-0.000	-0.002	-0.008**	0.016
Residence county poverty rate	26.21	1.120*	-0.025	-0.876	-0.027	0.013	-0.457**	-0.135
Field of study								
Agriculture/Veterinary Sciences	0.06	0.022	0.006	0.020	0.000	0.003	0.001	-0.003
Biological/Physical Sciences	0.10	0.000	0.003	0.009	-0.001	0.017***	0.015*	-0.122***
Business/Commerce/Economics	0.33	0.011	-0.022	0.047	0.002	-0.007	0.000	0.011
Education	0.13	-0.022	-0.001	-0.006	-0.000	-0.002	-0.003	-0.056
Engineering/Architecture	0.05	-0.007	0.015	0.012	0.000	-0.002	-0.002	0.002
Environmental Studies	0.07	-0.014	0.012	-0.023	0.000	-0.008	-0.009	-0.004
Health Sciences/Medicine	0.05	-0.010	0.014	0.002	0.001	-0.005	0.001	0.033
Humanities/Social Sciences	0.23	0.019	-0.027	-0.060**	-0.002*	0.003	-0.003	0.139*
Stated primary motivation								
Giving back to the community	0.73	-0.025	-0.044*	-0.043	-0.001	-0.006	-0.015	0.002
Helping encourage national cohesion	0.13	-0.006	0.002	0.024	0.000	0.006	0.010	-0.037
Personal development	0.04	0.007	0.005	-0.004	0.000	-0.003	0.002	0.009
Receiving a stipend	0.00	-0.000	-0.001	-0.005	-0.000	-0.001	-0.001	0.010
Supporting primary school children	0.08	0.025	0.035**	0.023	0.001	0.002	0.001	0.018
Traveling and discovering more of Kenya	0.01	-0.000	0.003	0.005	0.000	0.002	0.003	-0.002
Ethnicity								
Kikuyu	0.23	-0.025	0.018	-0.005	-0.001	0.005	0.001	-0.064
Kalenjin	0.18	0.051**	0.002	-0.036	-0.002**	0.006	0.003	-0.111**
Kamba	0.08	-0.018	-0.010	0.026	0.000	0.002	0.002	0.047
Kisii	0.09	0.021	-0.004	-0.012	0.001	-0.001	-0.006	0.049
Luhya	0.16	0.010	-0.013	0.012	0.002**	-0.020**	-0.008	0.013
Luo	0.16	-0.018	-0.008	-0.009	-0.001	0.008	0.008	-0.024
Baseline contact								
Resides in home county	0.46	0.015	0.036	-0.002	0.001	0.009	0.001	0.072
Observations	640	1,285	2,251	1,285	1,285	1,285	1,285	1,285
Round 1 only	X	X		X	X	X	X	X
All rounds			X					
Controls for expectated value				X	X	X	X	X

Notes: Table tests for covariate balance on independent variables used to identify the effects of RSD assignment, program participation, and match characteristics. The lottery cut-off identification strategy is used to identify program participation. Except for the lottery cut-off identification strategy, all samples restrict to applicants assigned in the first round. All tests for the effect of assignment characteristics include linear controls for the expected value, given uncertainty in the random lottery, of the assignment characteristic conditional on the applicants' preference list. Robust standard errors are reported in parentheses. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 3: Preference estimates for different demand specifications

	Conditional logit		Rank-ordered Logit		Hedonic	
	Univariate (1)	Multivariate (2)	Univariate (3)	Multivariate (4)	Univariate (5)	Multivariate (6)
Panel A: Cohort 4 applicants						
County characteristics						
Poverty rate	-0.611***	-0.507***	-0.143***	-0.042	-20.235***	-2.358
	(0.152)	(0.209)	(0.037)	(0.051)	(4.687)	(2.721)
Marginalized county	-0.115***	-0.006	-0.052***	-0.034*	-3.741***	0.886
	(0.046)	(0.076)	(0.011)	(0.018)	(1.417)	(0.946)
Relationship to home county						
Distance (00s km)	-0.243*	0.416*	-0.120***	0.031	-17.148***	-0.092
	(0.138)	(0.237)	(0.034)	(0.058)	(4.528)	(2.600)
Ethnic similarity	0.523***	0.556***	0.088***	0.067**	17.272***	5.995***
	(0.101)	(0.116)	(0.028)	(0.032)	(4.007)	(1.967)
Relationship to residence county						
Distance (00s km)	-0.043***	-0.031	-0.015***	-0.007	-2.092***	-0.690***
	(0.015)	(0.022)	(0.004)	(0.005)	(0.489)	(0.286)
N	45,576		45,576		22,597	
N groups	2,336		2,336		1,160	
Panel B: Cohort 5 applicants						
County characteristics						
Poverty rate	0.409***	2.783***	-1.456***	0.676***		
	(0.107)	(0.134)	(0.030)	(0.042)		
Marginalized county	-0.704***	-0.573***	-0.864***	-0.656***		
	(0.036)	(0.055)	(0.009)	(0.016)		
Relationship to home county						
Distance (00s km)	-0.376***	-0.053*	-0.341***	-0.077***		
	(0.012)	(0.028)	(0.003)	(0.010)		
Linguistic/Ethnic similarity	3.724***	2.967***	2.250***	0.981***		
	(0.174)	(0.181)	(0.056)	(0.060)		
Relationship to residence county						
Distance (00s km)	-0.413***	-0.265***	-0.355***	-0.174***		
	(0.013)	(0.028)	(0.003)	(0.010)		
N	65,126		73,615			
N applicants	5,058		5,058			

Notes: Table shows estimates of applicant preferences over program counties (20 in Panel A, 22 in Panel B). For univariate specifications, each row represents a separate specification. For the multivariate specificatons, a column represents a single specification. Results from conditional multinomial logit specifications (McFadden, 1973) are reported in the first two columns. Results from rank order multinomial logit (Beggs, Cardell and Hausman, 1981) specifications are reported in columns 3 and 4. Hedonic estimates restrict to the sample of applicants who participated in the *top choice auction*. *Marginalized counties* are those identified as historically marginalized by the Commission on Revenue Allocation (Commission on Revenue Allocation, 2013). Results from hedonic estimation using elicited compensating variation are in units of thousands of Kenyan Shillings (approximately 10 US dollars).

Table 4: Effect of CWOSM assignment on match characteristics

	Full sample		High poverty		Marginalized county		Distance averse		Female		Male	
	RA mean (1)	Coef. (2)	RA mean (3)	Coef. (4)	RA mean (5)	Coef. (6)	RA mean (7)	Coef. (8)	RA mean (9)	Coef. (10)	RA mean (11)	Coef. (12)
Panel A: The effect of CWOSM assignment on retention												
Attended training	0.53	0.05*	0.50	0.10***	0.42	0.17***	0.52	0.08**	0.49	-0.01	0.56	0.09***
		(0.03)		(0.04)		(0.05)		(0.04)		(0.04)		(0.03)
Completed service	0.45	0.03	0.44	0.07*	0.38	0.14***	0.45	0.05	0.39	-0.00	0.49	0.06
		(0.03)		(0.04)		(0.05)		(0.04)		(0.04)		(0.04)
Panel B: The effect of CWOSM assignment on assignment characteristics												
Auction value (000s KES)	-4.90	4.22***	-4.97	4.24***	-5.50	4.38***	-5.21	4.68***	-4.33	3.45***	-5.32	4.72***
		(0.36)		(0.49)		(0.68)		(0.57)		(0.43)		(0.53)
First choice	0.06	0.80***	0.06	0.81***	0.06	0.77***	0.05	0.80***	0.07	0.80***	0.06	0.80***
		(0.02)		(0.02)		(0.03)		(0.02)		(0.03)		(0.02)
Home county distance (00s km)	2.57	0.10	3.17	0.18	3.91	0.21	2.57	-0.08	2.50	0.12	2.61	0.08
		(0.09)		(0.15)		(0.16)		(0.13)		(0.14)		(0.12)
Ethnic similarity	0.10	0.00	0.06	0.01	0.04	-0.01	0.11	0.02	0.12	-0.01	0.09	0.01
		(0.01)		(0.01)		(0.01)		(0.02)		(0.02)		(0.01)
Residence county distance (00s km)	2.58	-0.08	3.24	-0.09	3.81	-0.15	2.58	-0.39***	2.48	0.01	2.65	-0.14
		(0.09)		(0.13)		(0.15)		(0.12)		(0.13)		(0.12)
Predicted utility	0.13	0.02	0.02	0.03	-0.05	-0.01	0.17	0.06	0.16	-0.00	0.11	0.03
		(0.02)		(0.03)		(0.02)		(0.04)		(0.04)		(0.03)
N	1,261		573		381		628		508		753	

Notes: Table shows the estimated effect of assignment by CWOSM on match-specific characteristics between the applicant and the assignment county. Odd columns show the mean of the control (RA) sample for the dependent variable. Even columns show the results of a regression of the match characteristic on an indicator for whether the applicant was assigned by CWOSM. Robust standard errors are reported in parentheses. Columns (1) and (2) present results for the full sample of assigned applicants. Columns (3) and (4) report results restricting to counties with above median poverty rates. Columns (5) and (6) report results for historically marginalized counties as determined by the Commission on Revenue Allocation (2013). *Auction value* reports the results only for the sub-sample of applicants selected to receive the *top auction value* form of the elicitation survey. Columns (7) and (8) restrict to the sample of applicants who, based on their reported preferences, were estimated to have individual preference parameters that indicate an aversion for distance ($\lambda_i < 0$). ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 5: Effect of participation on attitudes

	General social attitudes (Indices)		Geographic preferences Characteristics of county voted for			Residence county			
	National unity (1)	Poverty empathy (2)	Home (3)	Ethnic similarity to home (4)	Marginalized county (5)	Home county (6)	Ethnic similarity to home (7)	Marginalized county (8)	Same as baseline county (9)
Panel A: OLS estimated effect of program participation									
Attended training	0.06	-0.04	-0.14***	-0.09***	0.13***	0.09***	0.02	0.01	-0.01
	(0.06)	(0.07)	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)	(0.01)	(0.03)
Completed service	0.16***	-0.08	-0.11***	-0.07***	0.09***	0.10***	0.03	0.01	0.01
	(0.06)	(0.06)	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)	(0.01)	(0.03)
Observations	1,111	1,111	1,111	1,097	1,111	1,004	991	1,004	1,004
Panel A: 2SLS estimated effect of program participation									
Attended training	0.11	-0.13	-0.19*	-0.11	0.03	0.14	0.01	0.02	0.14
	(0.25)	(0.24)	(0.11)	(0.08)	(0.10)	(0.14)	(0.09)	(0.06)	(0.14)
Completed service	0.13	-0.16	-0.24*	-0.14	0.04	0.17	0.02	0.02	0.17
	(0.30)	(0.30)	(0.14)	(0.10)	(0.12)	(0.16)	(0.11)	(0.07)	(0.17)
Observations	1,931	1,931	1,931	1,902	1,931	1,745	1,717	1,745	1,745
Mean of dependent variable	-0.00	0.00	0.30	0.30	0.24	0.35	0.37	0.04	0.45

Notes: Table shows the effect of program participation on social attitudes. Attending training is a measure of short-term participation. Completing service is a measure of long-term participation. Panel A restricts to applicants assigned in round 1. Panel B reports results that include applicants assigned in later rounds. The 2SLS instrument is receiving a random lottery tiebreaker less than 0.8. Applicants with higher lottery numbers were less likely to participate in the program due to the delay in invitation. The first stage effect of the random tiebreaker cutoff on attending training and completing service varies between 0.19 and 0.23 due to variation in sample due to outcome availability. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 6: Effect of assignment on applicant outcomes

	Participation		General social attitudes		Geographic preferences (Voting)			Residence at endline		
	Trained (1)	Completed (2)	National unity (3)	Poverty empathy (4)	Home county (5)	Ethnic similarity to home (6)	Marginalized county (7)	Home county (8)	Ethnic similarity to home (9)	Marginalized county (10)
Marginalized county	-0.087***	-0.040	-0.040	0.008	-0.006	-0.028	0.287***	0.028	0.010	0.028*
	(0.032)	(0.033)	(0.072)	(0.073)	(0.033)	(0.025)	(0.029)	(0.036)	(0.024)	(0.015)
Poverty rate	-0.003***	-0.002	0.001	-0.002	-0.002	-0.001*	0.006***	0.002*	0.001	0.001*
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Ethnic similarity	0.228***	0.200***	0.036	-0.114	0.059	0.184***	-0.178**	0.063	-0.002	-0.018
	(0.080)	(0.081)	(0.182)	(0.186)	(0.084)	(0.062)	(0.079)	(0.091)	(0.061)	(0.037)
Predicted utility	0.338***	0.289***	-0.129	-0.080	0.049	0.195***	-0.352***	0.102	0.006	-0.050
	(0.092)	(0.094)	(0.208)	(0.212)	(0.096)	(0.071)	(0.090)	(0.103)	(0.069)	(0.042)
First choice	0.045*	0.024	-0.039	0.101*	-0.029	0.003	0.044*	0.030	-0.006	-0.001
	(0.027)	(0.027)	(0.060)	(0.061)	(0.028)	(0.021)	(0.026)	(0.030)	(0.020)	(0.012)
Distance from home county (00s km)	-0.024***	-0.019**	-0.023	-0.004	-0.010	-0.019***	0.055***	0.006	-0.001	-0.001
	(0.010)	(0.010)	(0.022)	(0.022)	(0.010)	(0.007)	(0.009)	(0.011)	(0.007)	(0.004)
Observations	1,261	1,261	1,106	1,106	1,106	1,092	1,106	1,000	987	1,000
Mean of dependent variable	0.50	0.41	-0.00	0.00	0.31	0.30	0.24	0.35	0.36	0.04

Notes: Table shows the effect on applicant retention and social attitudes of characteristics of their assignment. Sample restricts to applicants assigned in the first round. All specifications include linear controls for the expectation of the characteristic across alternative random assignments derived from their preference lists. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 7: The effect of receiving volunteer on pupil's reported exposure to volunteer's ethnic group or home county

	Ethnic group			County		
	All Cohorts (1)	Targeted cohorts (2)	Untargeted cohorts (3)	All Cohorts (4)	Targeted cohorts (5)	Untargeted cohorts (6)
Panel A: Effect on having met someone from assigned applicant's [ethnic group/county]						
Assigned applicant from [ethnic group/county]	0.024**	0.032**	0.018	0.005	0.005	0.002
	(0.011)	(0.014)	(0.013)	(0.006)	(0.008)	(0.008)
Observations	13,756	6,879	6,869	14,900	7,440	7,428
Mean for non-assigned [ethnic group/county]	0.135	0.129	0.142	0.095	0.097	0.093
Panel B: Effect on reporting assigned applicant's [ethnic group/county] is least trusted						
Assigned applicant from [ethnic group/county]	-0.005	0.014	-0.025	-0.011	-0.000	-0.019
	(0.014)	(0.019)	(0.019)	(0.010)	(0.015)	(0.015)
Observations	13,756	6,879	6,869	14,900	7,440	7,428
Mean for non-assigned [ethnic group/county]	0.344	0.338	0.350	0.340	0.337	0.344
Panel C: Effect on face anger attribution task when primed with applicants' [ethnic group/county]						
Assigned applicant from [ethnic group/county]	-0.018	-0.038**	0.006			
	(0.012)	(0.017)	(0.016)			
Observations	13,756	6,879	6,869			
Mean for non-assigned [ethnic group/county]	0.500	0.504	0.496			

Notes: [group] indicates either an ethnic group or a county. Pupils were asked to report their exposure and impressions of three randomly selected ethnic groups and counties. The set of ethnic groups included at least one of the ethnic groups of the volunteers assigned to the school in the first round. We exclude observations about the pupil's own ethnic group and county from all tests. All specifications include fixed effects for the interaction between the respondent's county of residence and the ethnic group or county in question. Effects have a reduced form interpretation due to non-compliance of volunteers with assignment (retention). Standard errors are clustered at the school level. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 8: Effects of volunteer similarity of pupil outcomes

	All cohorts (Standards 1-4)			Targeted cohorts (Standards 2 & 3)			Un-targeted cohorts (Stamdards 1 & 2)		
Volunteer characteristic	Kiswahili (1)	English (2)	Took part in G-United (3)	Kiswahili (4)	English (5)	Took part in G-United (6)	Kiswahili (7)	English (8)	Took part in G-United (9)
Ethnic similarity	-0.211***	-0.137**	-0.116**	-0.245***	-0.152**	-0.108*	-0.167*	-0.114	-0.122**
	(0.085)	(0.062)	(0.053)	(0.097)	(0.071)	(0.064)	(0.092)	(0.080)	(0.054)
Predicted utiltiy	-0.402***	-0.305***	-0.189**	-0.405***	-0.311***	-0.203**	-0.394***	-0.296**	-0.176**
	(0.140)	(0.107)	(0.083)	(0.160)	(0.122)	(0.100)	(0.149)	(0.132)	(0.085)
Volunteer's first choice	-0.047	-0.024	-0.009	-0.004	0.005	-0.006	-0.096***	-0.058*	-0.016
	(0.034)	(0.027)	(0.020)	(0.037)	(0.031)	(0.023)	(0.037)	(0.033)	(0.021)
Home county distance (00s km)	0.016	0.003	-0.002	0.021	0.010	-0.008	0.010	-0.003	0.004
	(0.014)	(0.011)	(0.008)	(0.015)	(0.013)	(0.009)	(0.015)	(0.013)	(0.008)
Observations	8,877	8,819	9,402	4,504	4,452	4,709	4,373	4,367	4,693
Pupils	4,679	4,651	4,968	2,376	2,349	2,486	2,303	2,302	2,482

Notes: Specifications stacks observations for each school-applicant combination. Targeted cohorts were in Standards 2 and 3 in 2018. Untargeted cohorts were in standards 1 and 4 in 2018. Across schools, one standard deviation of the ethnic similarity measure is 0.194, and one standard deviation of the distance between the volunteers' home counties and the school's county is 1.660. All specifications include controls for the school's average KCPE score, pupil gender, and pupil age. Standard errors are clustered at the school level. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 9: The effect of assignment on local adults' reported interactions with ethnic group of assigned worker

	All adult respondents (1)	Homestay HH members (2)	Head teachers (3)
Panel A: Effect on reported friendship with someone in ethnic group			
Assigned applicant from ethnic group	0.046*	0.027	0.088**
	(0.026)	(0.035)	(0.042)
Observations	2,642	1,687	856
Mean for non-assigned ethnic groups	0.528	0.464	0.647
Pane B: Effect on distrusting ethnic group			
Assigned applicant from ethnic group	-0.032	-0.047*	0.036
	(0.020)	(0.025)	(0.043)
Observations	2,642	1,687	856
Mean for non-assigned ethnic groups	0.122	0.129	0.109
Panel C: Effect on predicted share of ethnic group who identify as group over na			
Assigned applicant from ethnic group	0.014	0.012	0.028
	(0.014)	(0.018)	(0.028)
Observations	2,642	1,687	856
Mean for non-assigned ethnic groups	0.549	0.544	0.559

Notes: Adults (head teachers and members of homestay families) were asked to report their friendships and impressions of two randomly selected ethnic groups, as well as the ethnic group of at least one of the volunteers assigned to the school in the first round. Respondents rated the degree to which they agreed with the statement that people from the ethnic group could be trusted on a five point scale. We code applicants as distrusting an ethnic group if they disagree with the statement. Panel C reports the effect of assigning an applicant from the ethnic group on the respondents' expectations of the level of ethnic group identification of people in the applicants' ethnic group over a national Kenyan identity. All specifications include fixed effects for the county of the respondents' residence and the ethnic group under question. Results exclude respondents who were not head teachers at the school in 2018. Effects have a reduced form interpretation due to non-compliance of volunteers with assignment (retention). Standard errors, clustered at the school level, are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 10: Effects of volunteer similarity of adult community members national unity attitudes

	All respondents			Head teachers			Homestay HH members		
	People can be trusted.. ..in community	..in Kenya	Identifies more as Kenyan	People can be trusted.. ..in community	..in Kenya	Identifies more as Kenyan	People can be trusted.. ..in community	..in Kenya	Identifies more as Kenyan
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Effect of program									
Cohort 4 school	0.10*	0.08	-0.05	0.10*	0.08	-0.05			
	(0.05)	(0.07)	(0.07)	(0.05)	(0.07)	(0.07)			
Respondents	414	414	413	414	414	413			
Panel B: Effect of volunteer characteristics (Volunteer characteristics)									
Ethnic similarity	-0.12*	-0.13**	-0.03	-0.00	-0.01	0.07	-0.17**	-0.20***	-0.07
	(0.07)	(0.06)	(0.07)	(0.08)	(0.11)	(0.11)	(0.08)	(0.07)	(0.09)
Observations	1,938	1,938	1,938	686	686	686	1,252	1,252	1,252
Respondents	1,054	1,054	996	362	362	355	692	692	641
Predicted utility	-0.13	-0.24***	-0.03	-0.10	-0.04	0.22	-0.14	-0.33***	-0.12
	(0.09)	(0.10)	(0.11)	(0.12)	(0.16)	(0.16)	(0.11)	(0.12)	(0.13)
Observations	1,938	1,938	1,938	686	686	686	1,252	1,252	1,252
Respondents	1,054	1,054	996	362	362	355	692	692	641
Volunteer's first choice	-0.03*	-0.04*	-0.02	-0.04*	-0.03	-0.01	-0.03	-0.04*	-0.02
	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)	(0.04)	(0.02)	(0.03)	(0.03)
Observations	1,938	1,938	1,938	686	686	686	1,252	1,252	1,252
Respondents	1,054	1,054	996	362	362	355	692	692	641
Home county distance (00s km)	0.01	0.01	-0.01	-0.01	-0.00	-0.02	0.02*	0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Observations	1,938	1,938	1,938	686	686	686	1,252	1,252	1,252
Respondents	1,054	1,054	996	362	362	355	692	692	641
Mean of dep var	0.79	0.61	0.47	0.83	0.59	0.49	0.77	0.62	0.45

Notes: Panel A estimates the effect of the G-United program using OLS by comparing the level of the outcome in schools that received the G-United program in 2018 and those that were scheduled to received the program in 2019 (which was cancelled). Homestay households were only interviewed in 2018 G-United schools, so these are excluded from this test. Panel B stacks observations for each school-applicant combination. Across schools, one standard deviation of the ethnic similarity measure is 0.19, and one standard deviation of the logarithm of the distance from the volunteers' home counties and the school's county is . All specifications include controls for the county of the school, the ethnic group and gender of the respondent. Standard errors are clustered at the school level. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 11: The effects of mechanisms on assignment characteristics

	Cohort 4								Cohort 5	
	CWOSM (1)	RA (2)	CWOSM (3)	RA (4)	CWOSM (5)	RA (6)	CWOSM (7)	RA (8)	CWOSM (9)	RA (10))
Auction value (000s KES)	-0.51	-6.03	-4.19	-6.45	-5.17	-6.65	-1.96	-6.41		
First choice	0.89	0.05	0.29	0.06	0.16	0.05	0.60	0.04	0.60	0.06
Assignment location in list	1.31	10.24	4.39	10.35	6.11	10.29	2.10	10.52	2.84	8.78
Distance from home county (00s km)	2.61	2.64	4.11	4.11	3.66	3.66	2.86	2.95	2.66	2.90
Distance from residence county (00s km)	2.53	2.58	4.30	4.30	3.25	3.25	2.57	2.77	2.63	2.90
Ethnic similarity	0.10	0.09	0.01	0.01	0.01	0.01	0.05	0.04	0.05	0.04
Counties rank by disutility			X	X	X	X	X	X		
Counties rank by ethnic dissimilarity							X	X		
Home distance $>$ 100 km					X	X	X	X		

Notes: Table shows the average match characteristic over 100 simulations of the assignment. The mechanism used for assignment is indicated in the column header. Distance averse applicants are those that are identified as having above median averse to distance from their residence county using the procedure of Revelt and Train (2000). Columns 5 and 6 restrict applicant choice sets so that they can only rank counties that are more than 100 km from their home county. Columns 7 and 8 assign preferences to counties so that they prioritize applicants who are from more distant counties first. Columns 9 and 10 simulate assignment using the Cohort 5 data.

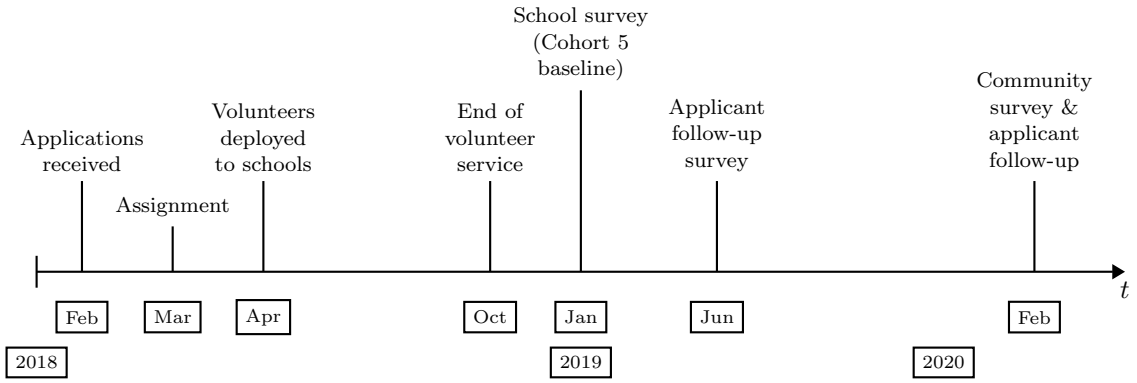
Table 12: The effects of mechanisms on assignment characteristics

	Cohort 4								Cohort 5	
	CWOSM (1)	RA (2)	CWOSM (3)	RA (4)	CWOSM (5)	RA (6)	CWOSM (7)	RA (8)	CWOSM (9)	RA (10))
Auction value (000s KES)	-0.53	-6.02	-0.83	-6.33	-2.02	-6.45	-3.80	-6.72		
First choice	0.89	0.05	0.84	0.05	0.60	0.04	0.33	0.04	0.60	0.06
Assignment location in list	1.32	10.24	1.49	10.23	2.10	10.54	3.47	10.73	2.85	8.78
Distance from home county (00s km)	2.62	2.62	2.45	2.64	2.86	2.95	3.95	3.97	2.65	2.90
Distance from residence county (00s km)	2.54	2.57	2.30	2.60	2.57	2.77	3.28	3.36	2.63	2.90
Ethnic similarity	0.11	0.09	0.12	0.09	0.05	0.05	0.02	0.02	0.05	0.04
Distance averse only			X	X	X	X	X	X		
Home distance > 100 km					X	X				
Counties rank by home distance							X	X		

Notes: Table shows the average match characteristic over 100 simulations of the assignment. The mechanism used for assignment is indicated in the column header. Distance averse applicants are those that are identified as having above median averse to distance from their residence county using the procedure of Revelt and Train (2000). Columns 5 and 6 restrict applicant choice sets so that they can only rank counties that are more than 100 km from their home county. Columns 7 an 8 assign preferences to counties so that they prioritize applicants who are from more distant counties first. Columns 9 and 10 simulate assignment using the Cohort 5 data.

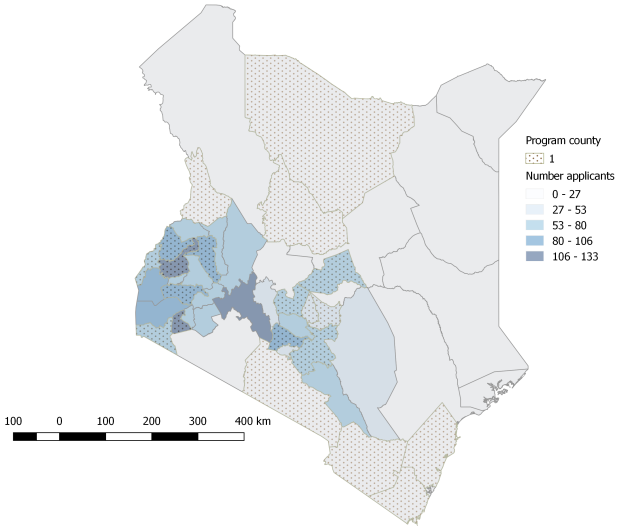
Figures

Figure 1



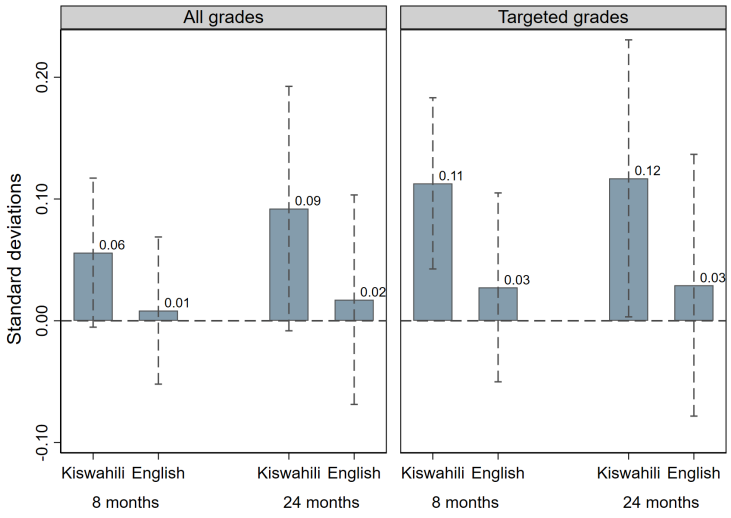
Note: Timeline of events related to the study. Applications were received in February 2018. There were four rounds of assignment in March 2018 and two rounds of training. In April 2018, volunteers were deployed to their assigned schools. Volunteers served for 6 months in their host communities. Starting in June 2019, applicants who completed the preference survey were contacted as part of a phone call follow-up survey.

Figure 2



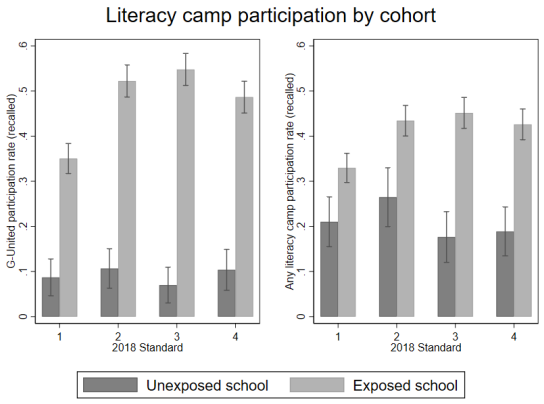
Note: This map shows counties participating in the G-United program and applicant home counties.

Figure 3



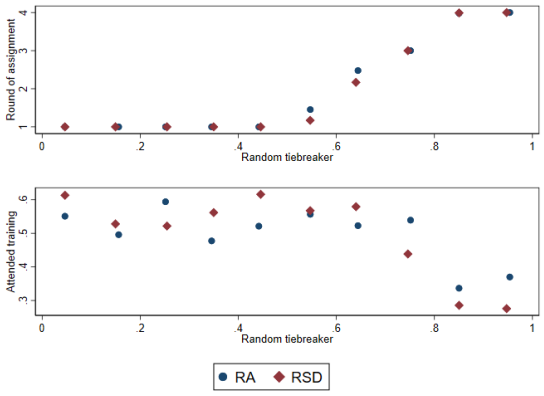
Note: This figure shows measured literacy levels in treated and comparison schools.

Figure 4



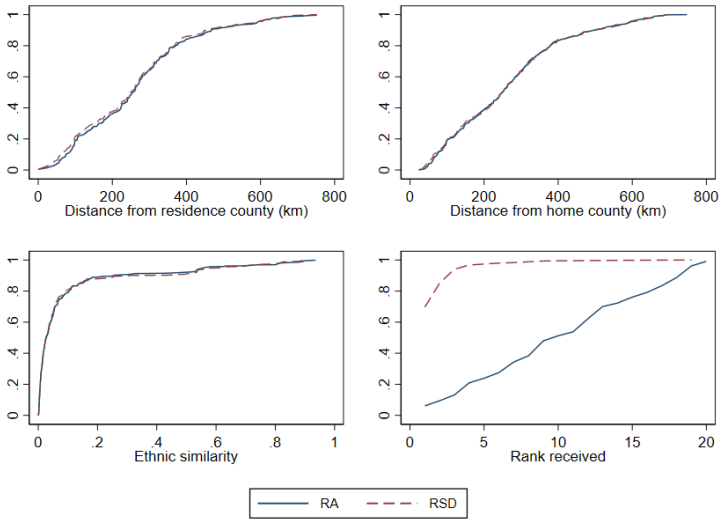
Notes: Figure shows pupils' recalled participation rates in the G-United program and in literacy camps generally. The figure shows results separately for schools that were unexposed (scheduled to received the G-United program in 2019), and schools that were exposed as part of the fourth cohort of G-United.

Figure 5



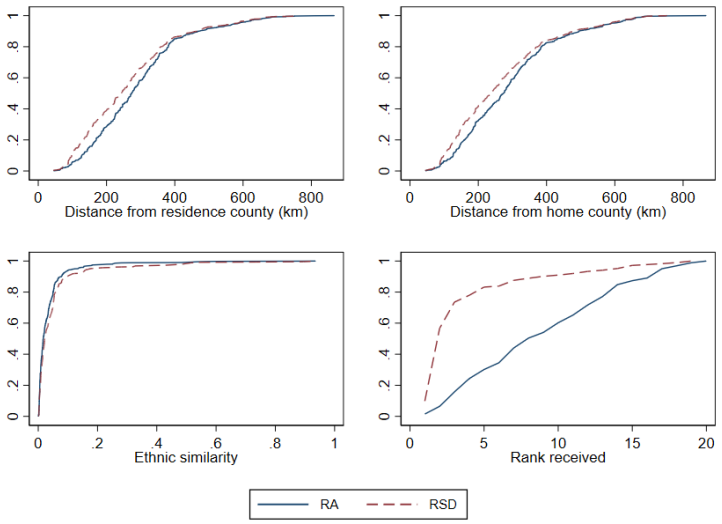
Notes: Binned scatterplot show the relationship between random tiebreaker, round of assignment, and attendance at the mandatory training - a measure of program participation.

Figure 6



Note: Cohort 4. Cumulative density function plots of match characteristics in ARA and CWOSM samples.

Figure 7



Note: Cohort 5. Cumulative density function plots of match characteristics in ARA and CWOSM samples.

A Costs

A.1 Instrument development and deployment.

Applicants geographic preferences were elicited through the Country Preference Survey (CPS), a tool developed in-house by Evidence Action, with support from the research team. The CPS was coded and run through Open Data Kit (ODK), a free, open source platform. To easily deploy this survey and ensure it could be conducted without a connection, on any device, Evidence Action bought a short-term subscription to Enketos platform services. The survey data was hosted on Evidence Actions existing servers.

Costs for development and hosting of the CPS include:

1. Enketo platform subscription: 29 USD/month; 3-month subscription. **Subtotal: 87 USD**

A.2 Instrument deployment and data receipt.

As the CPS was administered as part of the standard G-United recruitment process, applicants were not financially incentivized to take the survey. However, Evidence Action notified applicants via email of their eligibility to take the survey, and reminded applicants to take the survey via SMS. Evidence Action was able to leverage its large-scale subscription for mass SMS services, bringing SMS costs down to 3 KES (approximately .03 USD) per SMS. Evidence Action was also able to leverage its existing subscription to an email service to send applicants notification and an invitation to take the survey.

Costs for deployment of the CPS, including receiving and hosting volunteer data include:

1. SMS reminders: approximately 2,400 applicants received an average of two SMS notifications / reminders to take the CPS; 2,400 volunteers * 2 SMS * .03 USD. **Subtotal: 144 USD**

2. Email notifications: Evidence Actions subscription costs for Mailchimps services came to 7,500 KES (75 USD) per month; Evidence Action was not charged per email. Conservatively, we assume that deploying the CPS accounted for one third of one months emails. **Subtotal: 25 USD**

A.3 Administrative costs.

No additional staff is required to deploy the CPS as compared to other assignment instruments. However, in order to implement either assignment strategy, we should assume some portion of administrative costs are required to oversee and run the design, implementation, data collection and assignment of volunteers. These administrative costs would be split between two roles: (1) a Research Manager and (2) a Research Associate. This work was implemented over the course of three months.

Administrative costs for design, implementation, data collection, volunteer assignment and related oversight include:

1. Research Manager. The monthly all-in cost for the time and expenses of a Research Manager at Evidence Action is approximately 650,000 KES or 6,500 USD. In steady state, we would expect a Research Manager to dedicate one month at 50% LOE to applicant assignment; 6,500 USD * 1 month * .5 LOE. **Subtotal: 3,250 USD**

2. Research Associate. The monthly all-in cost for the time and expenses of a Research Associate at Evidence Action is approximately 375,000 KES or 3,750 USD. In steady state, we would expect a Research Associate to dedicate three months at 33% LOE to applicant assignment; 3,750 USD * 3 months * .33 LOE. **Subtotal: 3,750 USD**

In total, the annual cost of implementing the CPS for volunteer assignment (either alongside or in place of alternative assignment strategies) is approximately **7,256 USD**.

All costs detailed above were collected on a monthly basis by Evidence Action.

B Pupil and School Community Surveys

Evidence Action collected follow-up data in communities served by the G-United program twice, once at the beginning of 2019, and again at the beginning of 2020. Both rounds included data collection from communities that never received the G-united program, but were selected for 2019 cohort of G-United which was never deployed. These schools serve as a comparison group.[34] Both of these included pupil test scores, and the second survey also included extensive surveys of pupils and community members regarding national unity outcomes.

First, in January 2019, Evidence Action administered literacy tests to pupils in program schools and a set of pseudo-control schools. This data collection was designed as a baseline for a planned 2019 cohort, and it included 1,626 schools. Of these, 274 had received the G-United program in the previous year and the rest had not.[35] All schools were set to be included in the 2019 cohort, which was eventually canceled. In each school, 8 pupils in Standards 2 and 3, the target grades for G-United, were selected to be tested using the Uwezo assessment tool. The Uwezo tool is closely related to the ASER test that is widely used in India. ASER is used as part of the TaRL intervention developed by Pratham and is used as a diagnostic tool to guide the instructional level. A similar tool is used as part of the G-United TaRL intervention. The Kiswahili portion of the assessment was administered first, followed by the English portion. Students were

[34] Schools were selected for the 2019 cohort using the same deliberative process described for 2018.

[35] Some of the schools that did not receive a volunteer in 2018 had received volunteers in previous rounds of the program.

selected randomly from rosters, stratified by grade and gender, so that 2 students of each gender and grade were selected.

In February and March 2020, Evidence Action collected a second-round of school and community follow-up data. All schools that participated in the 2018 cohort of G-United were selected for follow-up, along with a sub-sample of 150 schools, randomly selected from the planned 2019 cohort of schools within the 20 counties that participated in 2018. This data collection included another round of pupil assessments, along with pupil and community member surveys.

In each school, 8 students were selected for both the assessments and the survey. Enumerators sampled from the 2017 rosters, identifying pupils who were likely to be in Standards 1 to 4 in 2018, the year of the intervention.[36] [37]

The pupil assessments again included English and Kiswahili tests. The pupil survey focused on inter-ethnic trust. Pupils were asked about their general level of trust in other ethnic groups. They were also asked specific questions about a set of three ethnic groups. The ethnic groups were selected at random, but always included the ethnic group of at least one of the applicants assigned to the school in the first round of assignment. Pupils were asked whether they had ever met someone from each of the three ethnic groups and to identify the county where most people from each ethnic group live on a map. Pupils were asked a similar series of questions about three different counties.

The 2020 school and community data collection also included interviews with the head teacher of the school and at least one member of a household of the volunteer homestay family. In comparison communities, these homestay families were prospective homestay families.

These surveys collected data on the respondent's level of local and generalized trust (how much can people be trusted in your community and in Kenya), degree of identification with their ethnic group, demographic characteristics, and interactions with the volunteer. The survey also included several tasks to measure beliefs about others, using incentivized prediction games. Respondents were asked to make predictions about how people with different ancestral counties would behave in a dictator game setting if playing with someone from their home county. Respondents were informed that the field officer team, including the field officer conducting the interview, had played the dictator game with one another as part of training, and that the purpose of the activity was to accurately guess how people from each county had played against someone from the respondents' home county. Respondents were asked about four counties, including the county of the school and the home county of at least one of the applicants assigned to the school in the first round.

[36] In some cases, the 2017 roster was unavailable. In these cases, rosters from 2018 were used instead.

[37] Volunteers were not instructed to provide services for pupils in Standards 1 and 4, so these pupils were initially intended to provide a comparison group. We later found evidence that the volunteers included pupils from other standards in the learning camps. We will discuss this in more detail below.

C Tables

Table A1: Effect of program on student academic achievement

	Kiswahili		English		National civics		University aspiration	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Short-term effects								
Program school	0.06*	0.00	0.01	-0.00			0.03*	0.02
	(0.03)	(0.04)	(0.03)	(0.04)			(0.02)	(0.02)
Program school × Targeted standard		0.10***		0.02				0.01
		(0.04)		(0.04)				(0.02)
Total effect on targeted standards		0.11***		0.02				0.03*
Observations	11,775	11,775	11,775	11,775			11,775	11,775
R^2	0.16	0.16	0.19	0.19			0.04	0.04
Panel B: Medium-term effects								
Program school	0.09*	0.06	0.02	0.00	0.07	0.06	0.02	0.02
	(0.05)	(0.06)	(0.04)	(0.05)	(0.05)	(0.05)	(0.02)	(0.02)
Program school × Targeted standard		0.06		0.03		0.02		-0.00
		(0.05)		(0.05)		(0.06)		(0.03)
Total effect on targeted standards		0.12**		0.03		0.08		0.02
Observations	5,876	5,876	5,876	5,876	5,876	5,876	5,876	5,876
R^2	0.25	0.25	0.28	0.28	0.24	0.24	0.04	0.04

Notes: Panel A reports on outcomes observed approximately eight months after the end of the intervention. Panel B reports on outcomes observed two years after the end of the intervention. Kiswahili and English scores in the short term are from an five point Uwezo literacy assessment. Tests for medium term outcomes are designed to reflect the national curriculum. Note that Panel A includes only students who were in Standards 1 and 2 at the time of the assessment. Students in Standard 2 at that time are *Targeted standards*. Panel B includes students who were in Standards 1, 2, 3 and 4 at the time of the program. Students in Standards 2 and 3 at that time are *Targeted standards*. All specifications include controls for the school's average KCPE score, pupil age, gender, and county fixed effects. Standard errors are clustered at the school level. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table A2: Training invitations and retention by assignment round

	Round 1	Round 2	Round 3	Round 4
ARA & CWOSM samples				
Invited	1,285	217	302	447
Attended	709	117	151	144
Retention rate	0.552	0.539	0.500	0.322
Assigned randomly (ARA sample)				
Invited	640	108	164	210
Attended	341	54	89	76
Retention rate	0.533	0.500	0.543	0.362
Assigned by (CWOSM sample)				
Invited	645	109	138	237
Attended	368	63	62	68
Retention rate	0.571	0.578	0.449	0.287
Date	March 25	April 3	April 9	April 18

Notes: Table shows the numbers of applicants assigned in each round of assignment. Applicants assigned in rounds 2 and later were assigned to vacant seats after assigned applicants in earlier rounds dropped out.

Table A3: RSD assignment and recalled assignment to top choice county

	All rounds (1)	Assigned round 1 (2)
$\mathbb{1}\{\lambda_i \leq 0\}$	-0.030	-0.070*
	(0.033)	(0.041)
RSD assigned	-0.029	-0.042
	(0.037)	(0.046)
RSD assigned$\times\mathbb{1}\{\lambda_i \leq 0\}$	0.116***	0.168***
	(0.048)	(0.058)
RSD + RSD $\times\mathbb{1}\{\lambda_i \leq 0\}$ (*p-value*)	0.004	0.001
Observatons	1,694	1,121

Notes: Table shows the effect of RSD assignment on the applicant's end-line report of having received a top choice location. This table estimates heterogeneous effects of assignment by the RSD mechanism across applicants with λ_i (the applicants' preference for distance from residence county (log)) with different signs. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table A4: Mixed logit results

	Cohort 4 (1)	Cohort 5 (2)
Distance from residence county (Log)		
Mean	-0.039***	-0.801***
	(0.011)	(0.022)
SD	0.282***	1.289***
	(0.016)	(0.021)
N	350,640	619,705
N groups	2,336	5,058

Notes: Table shows the results of mixed logit estimates using ordinal choice data from Cohort 4 and Cohort 5. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table A5: Descriptive statistics for outcomes collected in the 2019 phone call survey

	All applicants	Residing in home county	Residing outside home county
	(1)	(2)	(3)
Migration results			
Living in home county	0.337	0.542	0.151
Dist moved from baseline (km)	86.1	75.5	96.0
Dist b/w res and home cnty (km)	127.7	75.5	176.2
National unity and pro-social attitudes			
Feelings thermometer (0 warm - 10 cold)	8.01	7.93	8.09
Self-identification (1 Kenyan - 5 ethnic grp)	2.42	2.39	2.45
Poor ppl don't work enough (1 agree -5 disagree)	3.92	3.87	3.98
Violence can be justified (1 agree - 5 disagree)	4.04	4.01	4.07
Geographic preferences for edu resources			
First vote for home county	0.303	0.337	0.272
Both votes for home county	0.038	0.034	0.043
Distance b/w first vote and home cnty (km)	199.1	183.5	213.5
First vote lang/ethnic similarity	0.311	0.325	0.298
First vote lang grp similarity	0.600	0.619	0.582
First vote political similarty	0.585	0.586	0.584
First vote poverty rate	0.441	0.440	0.442
Voted for deployment cnty	0.254	0.267	0.242
Employment outcomes			
Employed	0.389	0.376	0.400
Monthly emp income (000s KES)	27.0	24.3	29.2
Employed full time	0.158	0.145	0.169
Employed in govt job	0.072	0.088	0.058
Observations	2,242	1,067	1,175

Notes: Table shows mean responses for outcomes collected in the 2019 phone call survey with applicants. *Dist moved from baseline* indicates the distance between the applicant's residence county at baseline and residence county at endline. *Feelings thermometer* asked respondents to rate their feelings toward other ethnic groups on a scale of 0 (warm) to 10 (cold). Cold may have been interpreted as more favorable feelings. *Self-identification* asked applicants whether they identify as more Kenyan or more as their ethnic group. *Violence can be justified* asked respondents whether they agree or disagree with the statement "Sometimes violence is necessary to support a just cause". *Poor ppl don't work hard enough* asked respondents if they agreed or disagreed with the statement "Poor people don't work hard enough to escape poverty".

Table A6: Effect of applicant assignment characteristics on mediators

Assignment characteristic	Assignment characteristic mean (1)	Total workers in school (2)	Number of camps completed (3)	Total workers below median ethnic similarity (4)	Total workers above median home distance (5)	Total workers working in top choice county (6)	Total workers above median predicted utility (7)
Assigned by RSD	0.50	0.07**	0.48***	0.02	0.07*	0.57***	0.04
	[0.50]	(0.03)	(0.17)	(0.04)	(0.04)	(0.04)	(0.03)
Observations	1,285	1,285	1,285	1,285	1,285	1,285	1,285
Respondents	658	658	658	658	658	658	658
Below median ethnic similarity (0.03)	0.52	0.01	0.11	0.19***	0.45***	0.05	0.02
	[0.50]	(0.04)	(0.19)	(0.04)	(0.05)	(0.04)	(0.04)
Observations	1,292	1,292	1,292	1,292	1,292	1,292	1,292
Respondents	658	658	658	658	658	658	658
Above median predicted utility	0.50	0.04	-0.02	0.00	0.06	-0.04	0.04
	[0.50]	(0.05)	(0.21)	(0.05)	(0.05)	(0.05)	(0.05)
Observations	1,292	1,292	1,292	1,292	1,292	1,292	1,292
Respondents	658	658	658	658	658	658	658
Volunteer's first choice	0.47	0.08**	0.47***	0.05	0.06	0.60***	0.05
	[0.50]	(0.03)	(0.16)	(0.04)	(0.04)	(0.04)	(0.03)
Observations	1,292	1,292	1,292	1,292	1,292	1,292	1,292
Respondents	658	658	658	658	658	658	658
Above median home county distance (247 km)	0.50	0.03	-0.01	0.39***	0.23***	0.06	0.00
	[0.50]	(0.04)	(0.19)	(0.05)	(0.04)	(0.04)	(0.04)
Observations	1,292	1,292	1,292	1,292	1,292	1,292	1,292
Respondents	658	658	658	658	658	658	658
Dependent variable mean		1.71	4.81	0.86	0.90	0.65	1.57

Notes: Standard deviations are reported in brackets. Standard errors are reported in parentheses. Standard errors are clustered at the school level. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table A7: Effect of applicant assignment characteristics on mediators
Marginalized counties only

Assignment characteristic	Assignment characteristic mean (1)	Total workers in school (2)	Number of camps completed (3)	Total workers below median ethnic similarity (4)	Total workers above median home distance (5)	Total workers working in top choice county (6)	Total workers above median predicted utility (7)
Assigned by RSD	0.51	0.05	0.32	0.01	0.16**	0.54***	0.03
	[0.50]	(0.06)	(0.30)	(0.08)	(0.08)	(0.07)	(0.07)
Observations	387	387	387	387	387	387	387
Respondents	199	199	199	199	199	199	199
Below median ethnic similarity (0.03)	0.76	-0.08	-0.39	0.08	0.45***	0.10	-0.07
	[0.43]	(0.09)	(0.41)	(0.10)	(0.10)	(0.09)	(0.09)
Observations	389	389	389	389	389	389	389
Respondents	199	199	199	199	199	199	199
Above median predicted utility	0.19	0.09	0.79**	-0.12	-0.34***	-0.11	0.06
	[0.39]	(0.09)	(0.40)	(0.11)	(0.11)	(0.09)	(0.09)
Observations	389	389	389	389	389	389	389
Respondents	199	199	199	199	199	199	199
Volunteer's first choice	0.46	0.09	0.31	0.08	0.22***	0.57***	0.08
	[0.50]	(0.07)	(0.30)	(0.08)	(0.08)	(0.07)	(0.07)
Observations	389	389	389	389	389	389	389
Respondents	199	199	199	199	199	199	199
Above median home county distance (247 km)	0.80	-0.07	-0.89**	0.35***	0.16	0.07	-0.10
	[0.40]	(0.10)	(0.42)	(0.12)	(0.11)	(0.10)	(0.10)
Observations	389	389	389	389	389	389	389
Respondents	199	199	199	199	199	199	199
Dependent variable mean		1.68	4.71	1.36	1.32	0.65	1.54

Notes: Standard deviations are reported in brackets. Standard errors are reported in parentheses. Standard errors are clustered at the school level. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table A8: OLS estimates of the effect of G-United on reading levels in 2019

	Student can at least					
	Recognize letters		Recognize words		Read a paragraph or story	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: English			
Cohort 4 school	0.010	0.014	0.050***	0.036*	0.008	0.010
	(0.007)	(0.011)	(0.013)	(0.020)	(0.012)	(0.012)
Cohort 4 × STD3		-0.007		0.029		-0.004
		(0.010)		(0.020)		(0.020)
Standard 3	0.058***	0.059***	0.186***	0.181***	0.266***	0.266***
	(0.005)	(0.005)	(0.009)	(0.009)	(0.008)	(0.009)
Control mean	0.929		0.753		0.212	
Observations	11,775	11,775	11,775	11,775	11,775	11,775
R^2	0.025	0.025	0.079	0.079	0.115	0.115
			Panel B: Kiswahili			
Cohort 4 school	0.034***	0.012	0.048***	0.016	0.024	0.008
	(0.011)	(0.018)	(0.014)	(0.020)	(0.015)	(0.019)
Cohort 4 × STD3		0.045***		0.063***		0.032
		(0.019)		(0.021)		(0.023)
Standard 3	0.144***	0.136***	0.228***	0.217***	0.276***	0.271***
	(0.008)	(0.008)	(0.009)	(0.010)	(0.009)	(0.010)
Control mean	0.822		0.666		0.441	
Observations	11,775	11,775	11,775	11,775	11,775	11,775
R^2	0.058	0.058	0.086	0.086	0.097	0.097

Notes: All specifications include controls for the school's average KCPE score, enrollment in standards 2 and 3, pupil gender, and pupil age. Odd columns show the OLS relationship between 2018 assignment to the G-United program and 2019 reading levels. Even columns show estimates from a cohort difference-in-difference estimate, where the first difference comes from a comparison 2018 G-United schools and new schools assigned to receive G-United in 2019 and the second difference comes from a comparison of Standard 3 and Standard 2 students, where Standard 3 students were more likely to be exposed to the program. All reading levels are assessed using the Uwezo tool. Standard errors are clustered at the school level. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table A9: OLS estimates of the effect of G-United on reading levels in 2020

	Total score (std)		Student can at least: Recognize words		Read a sentence		Read a passage	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Panel A: English					
Cohort 4 school	0.028	0.007	-0.006	-0.016	0.017	-0.004	0.029	0.031
	(0.046)	(0.050)	(0.016)	(0.018)	(0.019)	(0.021)	(0.019)	(0.019)
Cohort 4 × Targeted		0.040		0.019		0.041*		-0.004
		(0.054)		(0.022)		(0.024)		(0.020)
Targeted cohorts	0.064***	0.031	0.045***	0.030	0.047***	0.013	-0.006	-0.002
	(0.024)	(0.049)	(0.009)	(0.020)	(0.010)	(0.022)	(0.009)	(0.018)
Control mean	0.093		0.752		0.700		0.416	
Observations	5,421	5,421	5,421	5,421	5,421	5,421	5,421	5,421
R^2	0.086	0.086	0.064	0.064	0.061	0.062	0.056	0.056
			Panel B: Kiswahili					
Cohort 4 school	0.091*	0.070	0.001	-0.018	-0.004	-0.015	0.038	0.047*
	(0.052)	(0.055)	(0.015)	(0.018)	(0.017)	(0.020)	(0.026)	(0.028)
Cohort 4 × Targeted		0.041		0.036		0.021		-0.017
		(0.051)		(0.022)		(0.023)		(0.025)
Targeted cohorts	0.073***	0.039	0.038***	0.008	0.062***	0.044**	0.023**	0.037
	(0.022)	(0.047)	(0.009)	(0.021)	(0.010)	(0.021)	(0.010)	(0.023)
Control mean	0.035		0.805		0.779		0.544	
Observations	5,452	5,452	5,452	5,452	5,452	5,452	5,452	5,452
R^2	0.079	0.079	0.044	0.045	0.055	0.055	0.047	0.047

Notes: Targeted cohorts were in Standards 2 and 3 in 2018. Untargeted cohorts were in standards 1 and 4 in 2018. All specifications include controls for the school's average KCPE score, pupil gender, and pupil age. Odd columns show the OLS relationship between 2018 assignment to the G-United program and test score outcomes. Even columns show estimates from a cohort diff-in-diff estimate, where the first difference comes from a comparison 2018 G-United schools and new schools assigned to receive G-United in 2019 and the second difference comes from a comparison of Standard 3 and Standard 2 students, where Standard 3 students were more likely to be exposed to the program. The total standardized test score is reported in the first two columns. The ability to recognize words, read sentences, and read a passage are assessed using an outcome that represents the percent correct of tasks related to each level of proficiency. Standard errors are clustered at the school level. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table A10: The effect of the program on exposure recall

	Took part in G-United		Has taken part in a literacy camp	
	(1)	(2)	(3)	(4)
Cohort 4 school	0.388***	0.324***	0.206***	0.186***
	(0.022)	(0.024)	(0.027)	(0.028)
Cohort 4 × Targeted		0.125***		0.038
		(0.020)		(0.026)
Targeted cohorts	0.084***	-0.020	0.046***	0.014
	(0.012)	(0.016)	(0.012)	(0.023)
Constant	0.121	0.173	0.222	0.238
	(0.153)	(0.153)	(0.171)	(0.171)
Observations	5,876	5,876	5,876	5,876
R^2	0.112	0.114	0.033	0.033

Notes: This table examines the first stage effect of assignment to the G-United program for pupils in cohorts and schools that were targeted for the G-United program. Targeted cohorts were in Standards 2 and 3 in 2018. Untargeted cohorts were in standards 1 and 4 in 2018. All specifications include controls for the school's average KCPE score, pupil gender, and pupil age. All specifications include controls for the school's average KCPE score, enrollment in standards 2 and 3, pupil gender, and pupil age. Standard errors are clustered at the school level. ***, **, and * indicate significance at 1%, 5%, and 10%.

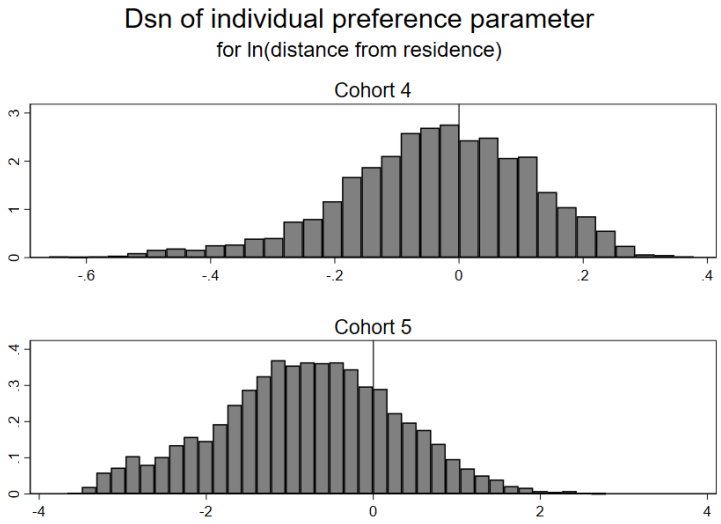
Table A11: Effect receiving volunteer on adult's impressions of people from volunteer's home county

		All adult respondents (1)	Homestay HH members (2)	Head teachers (3)
Dictator game play	Coef	-0.46	0.26	-9.23**
	SE	(2.00)	(2.61)	(4.44)
	N	3,791	2,305	1,093
	Mean of dep var	43.67	45.09	41.46
Honesty game play	Coef	4.60***	2.10	10.23***
	SE	(1.96)	(2.53)	(4.27)
	N	3,791	2,305	1,093
	Mean of dep var	56.40	56.46	56.31
Home cty bias in voting	Coef	2.78	3.51	0.89
	SE	(2.10)	(2.78)	(4.01)
	N	3,791	2,305	1,093
	Mean of dep var	66.57	65.19	68.73

Notes: Adults (head teachers and members of homestay families) were asked to report their impressions with different ancestral/home counties, including the home county of at least one of the volunteers assigned to the school in the first round. *Dictator game play* represents the respondent's prediction from an incentivized guess the average amount provided to someone from the respondent's home county. *Honesty game play* represents the respondent's prediction of the share of respondents from the county who report heads in an honesty game. *Home cty bias in voting* represents the respondents guess of the share of G-United applicants who voted for their home county. Results excludes respondents who were not head teachers at the school in 2018. All specifications include fixed effects for the school's county interacted fixed effects for the ethnic group about which the adult was asked. Effects have a reduced form interpretation due to non-compliance of volunteers with assignment (retention). Standard errors are clustered at the school level. ***, **, and * indicate significance at 1%, 5%, and 10%.

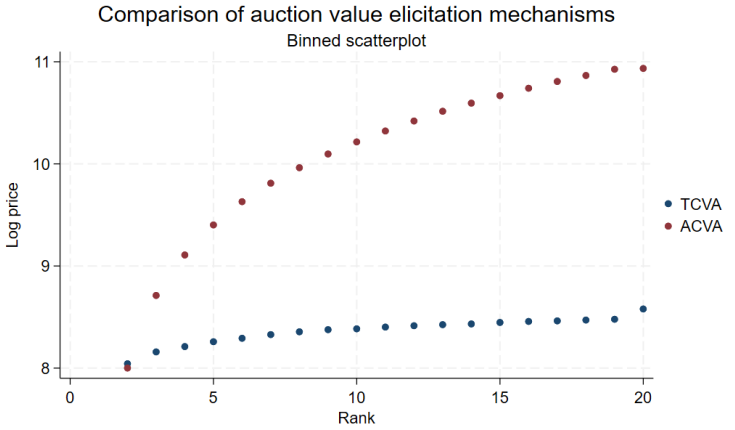
D Figures

Figure A1



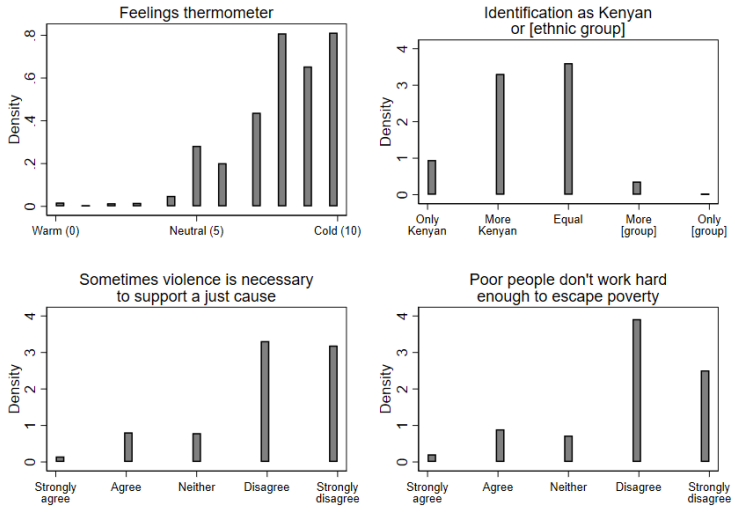
Note: Figure shows the distribution of individual preference parameters for (log) distance from residence county. Individual parameters are identified using a technique from Revelt and Train (2000) after estimated a mixed logit model.

Figure A2: Comparison of auction value elicitation mechanisms.



Notes: Binned scatterplot of the mean "price" required to make an applicant indifferent between lower ranked options and their top-ranked options. The ACVA mechanism imputes this value using "adjacent auctions".

Figure A3



Notes: Histograms of responses to questions about national unity and pro-sociality from first phone call interview.