

# How Refugee Resentment Shapes National Identity and Citizen Participation in Africa.\*

Yang-Yang Zhou<sup>†</sup>

November 29, 2019

## Supplementary Information

---

\*All replication material, including R code and data, will be made available via Harvard University's Dataverse.

<sup>†</sup>Assistant Professor, Department of Political Science, University of British Columbia. Email: [yangyang.zhou@ubc.ca](mailto:yangyang.zhou@ubc.ca),  
URL: [yangyangzhou.com](http://yangyangzhou.com)

# Supplementary Information

## Contents

<b>I</b>	<b>Supplementary Information</b>	<b>2</b>
<b>S1</b>	<b>Additional Context</b>	<b>3</b>
S1.1	Displacement in Africa . . . . .	3
S1.2	National Identification and Citizen Participation Trends from the Afrobarometer . . . . .	5
<b>S2</b>	<b>Focus Group Evidence</b>	<b>6</b>
S2.1	Sampling Strategy . . . . .	6
S2.2	Additional Focus Group Results by Proximity to the Refugee Camp . . . . .	8
S2.3	Land Disputes between Host Communities and Refugee Camps . . . . .	10
S2.4	Positive/Mixed Comments about Refugees . . . . .	10
S2.5	Comments about Public Goods Grievances . . . . .	12
<b>S3</b>	<b>Survey Evidence</b>	<b>13</b>
S3.1	Sampling Strategy . . . . .	13
S3.2	Survey Experimental Treatment priming Refugees . . . . .	15
S3.3	Measuring Identity Attachments . . . . .	17
S3.4	Survey Respondent Covariates . . . . .	18
S3.5	Attitudes about Refugees by Distance to the Refugee Camp . . . . .	20
S3.6	Survey Estimation Strategy for Analysis on National Identification . . . . .	21
S3.7	Results for other Identity Bins . . . . .	22
S3.8	Robustness Check using Distance by Road to the Refugee Camp . . . . .	23
S3.9	Placebo Results using Distances to Future Refugee Camps, Nduta and Mtendeli . . . . .	24
S3.10	Paired Conjoint Experiment . . . . .	25
<b>S4</b>	<b>Post-Influx Public Goods Observational Evidence</b>	<b>27</b>
S4.1	Descriptive Statistics . . . . .	27
S4.2	Public Goods Outcomes Estimation Strategy . . . . .	30
S4.3	Parallel Trends Assumption . . . . .	31
S4.4	Additional Primary School Outcomes Analysis . . . . .	32

# S1 Additional Context

This section offers additional descriptive and contextual information about refugees, national identification, and citizen participation across Africa.

## S1.1 Displacement in Africa

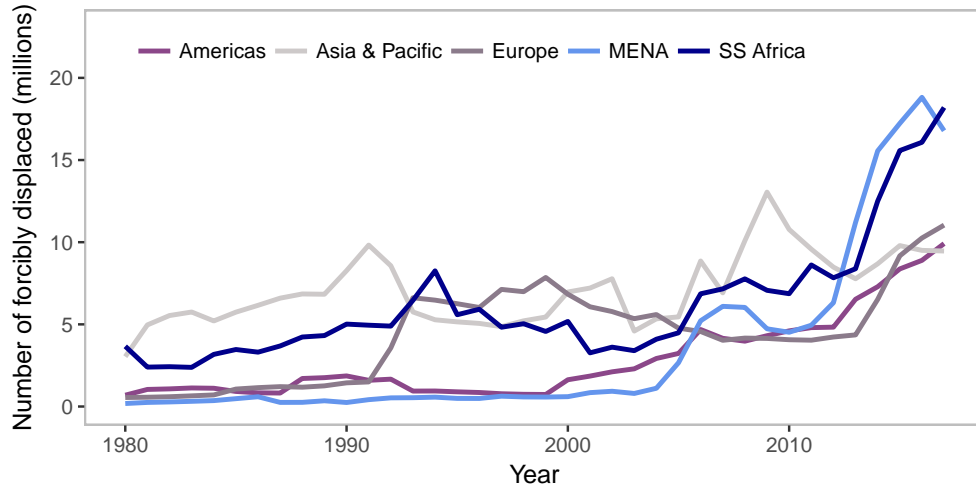


Figure S1: This figure compares the number of forcibly displaced people (including refugees, internally displaced, asylum-seekers, and stateless persons) across regions over time. In 2017, the region of sub-Saharan Africa hosts the most forcibly displaced, at historically unprecedented levels. Data source: UNHCR population statistics database.

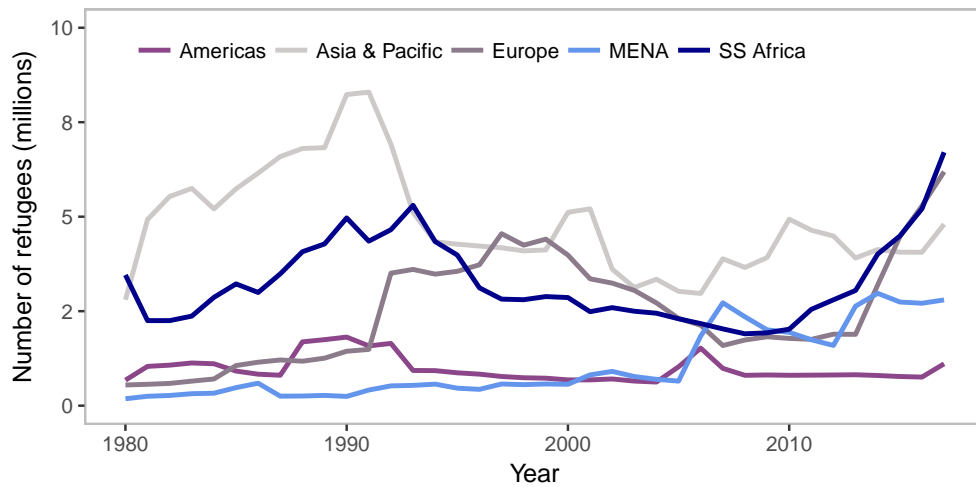


Figure S2: This figure compares the number of refugees across regions over time. In 2017, the region of sub-Saharan Africa hosts the most refugees, about a third of the overall refugee population. Data source: UNHCR population statistics database.

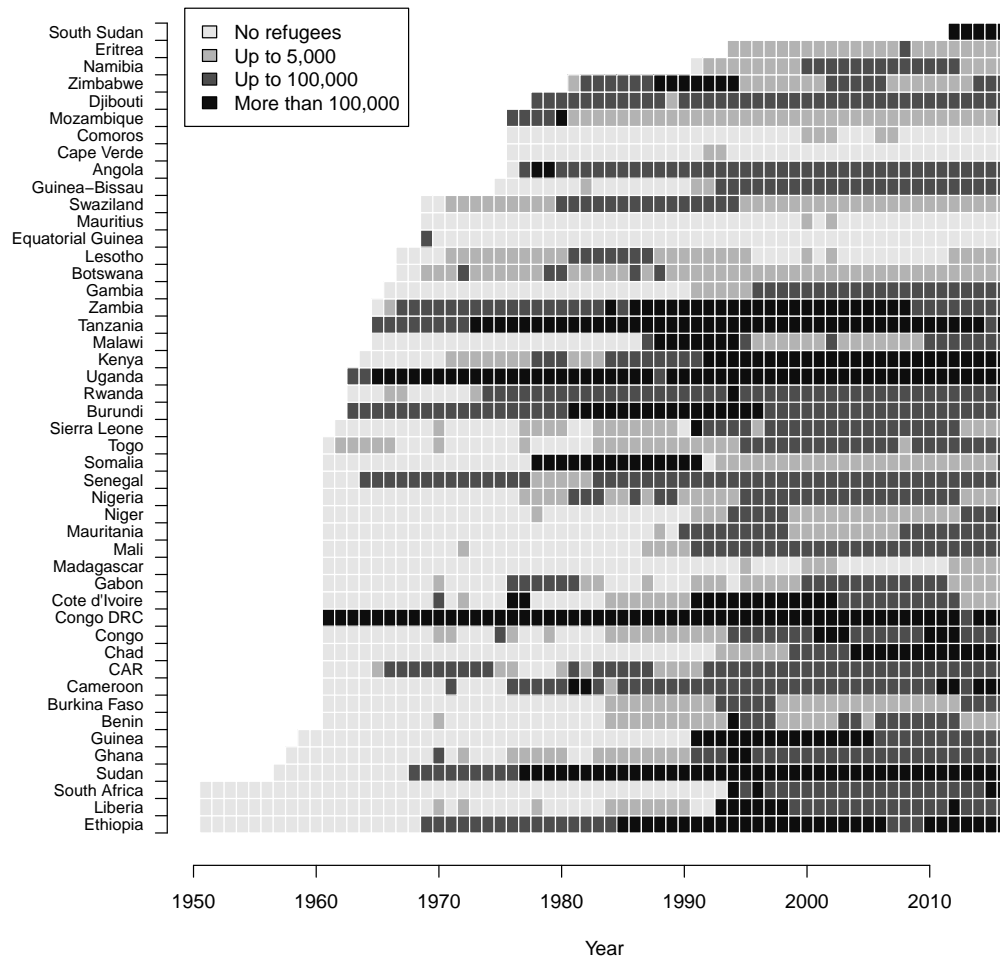


Figure S3: This figure shows whether a country has hosted refugees and the magnitude in any given year, beginning at the country’s independence (although cropped for South Africa (1931 from Britain), Liberia (1847 from the American Colonization Society), and Ethiopia (never colonized)). Every country has hosted refugees, and aside from islands, many have been doing so for the majority of their existence as nation-states. Data source: UNHCR population statistics database.

## S1.2 National Identification and Citizen Participation Trends from the Afrobarometer

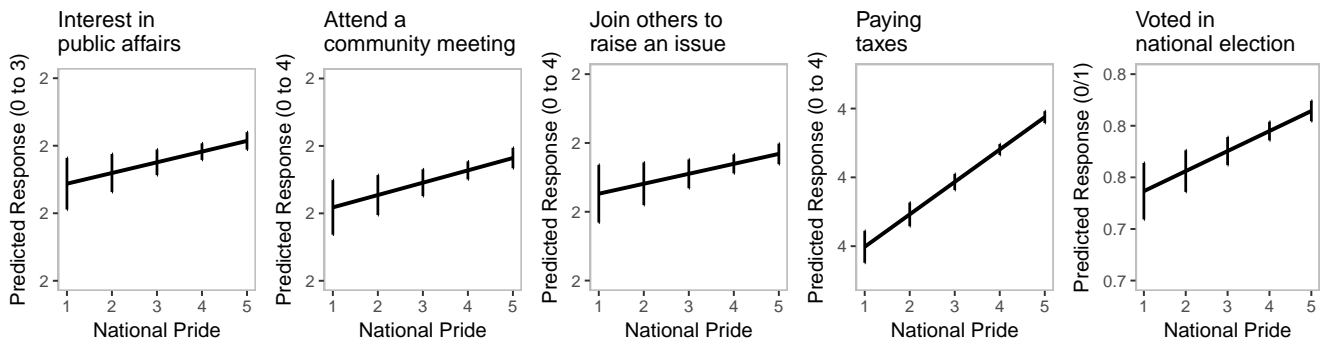


Figure S4: This figure shows that on the Afrobarometer round 5 ( $n = 34,772$  across 34 countries), measures of citizen participation are positively correlated with national pride. Each panel plots the predicted responses with 95% CIs from OLS regressing citizen participatory actions (scale from 0 Never to 3 Often) on national pride (scale of 1 to 5), controlling for age, education, rural, female, belongs to majority ethnic group, poverty, distances to the border, nearest city, and nearest road, with country fixed effects.

## S2 Focus Group Evidence

### S2.1 Sampling Strategy

This section describes the sampling strategy for the 10 experimental community focus groups conducted in July and August 2015 with 102 citizen participants in Kigoma, Tanzania. To explore how geography might contextualize the discussions, the selected communities needed to be varied in terms of their proximity to the Nyarugusu refugee camp – the only camp established at the time of the focus group research, and their proximity to the international border with Burundi.

While Kigoma region is administratively divided into 6 districts, due to travel and budget constraints, my research assistants and I selected 4 districts: Buhigwe (at the border, close to the camp), Kasulu rural (at the border, encompasses the camp), Kigoma rural (at the border, far from the camp), and Uvinza (farthest from the border, farthest from the camp). Within these districts we randomly selected 2/3 rural communities for a total of 10 communities. Figure S5 shows the locations of these communities as green squares, and figure S6 shows the distribution of their distances to Nyarugusu refugee camp.

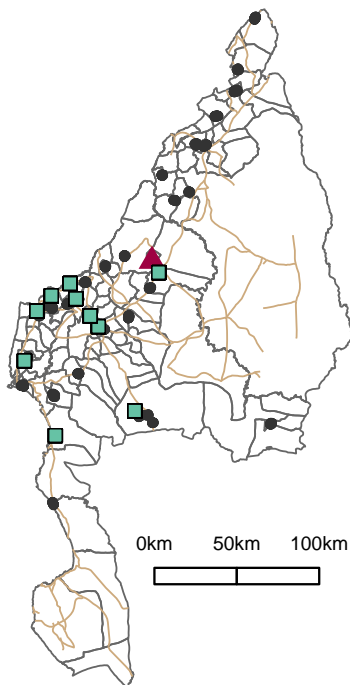


Figure S5: This map shows the 10 experimental community focus groups conducted in July-August 2015 (green squares) and the villages of the 2025 respondents surveyed in June-July 2016 (black points) in relation to the only established refugee camp in those periods, Nyarugusu (red triangle) and roads (brown lines) in the study region Kigoma. The empty wards towards the east and south are the Moyowosi Game Reserve and Mahale National Park, respectively.

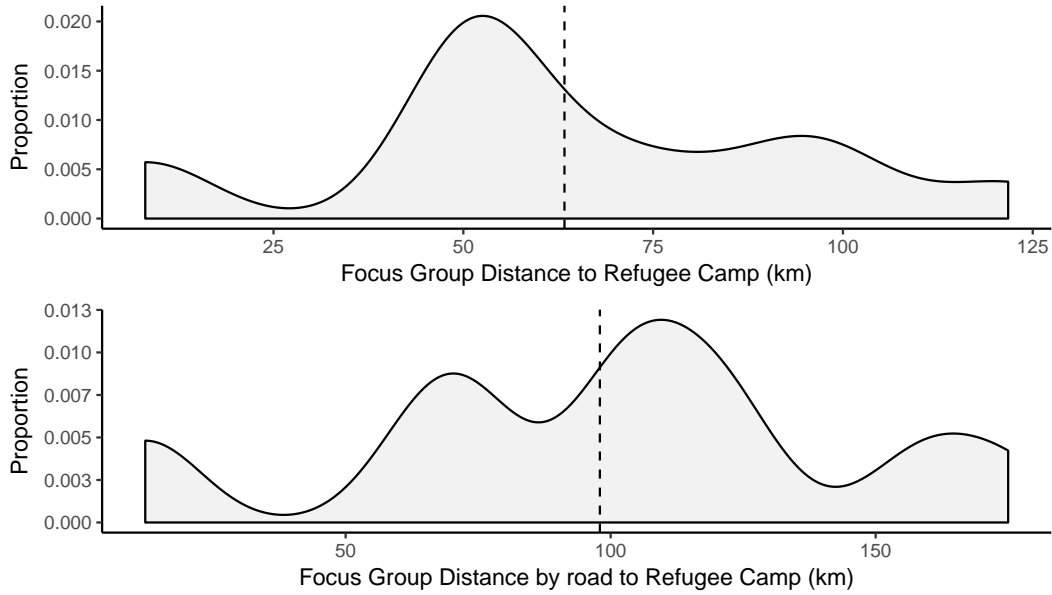


Figure S6: These figures show the distribution of community focus group distances to Nyarugusu refugee camp by direct path (top) and by road (bottom). The dashed lines show the mean distances.

For each community, we contacted the local officials, ward executive officers (WEO) and village leaders, requesting 15 adult citizens of voting age. We estimated 15 participants as a manageable number for a discussion group, so that anyone who wished to speak had an opportunity to do so. Within each community, my research assistants Bihaga Edward and Adrophinus Leopold asked village leaders to help find participants with a mix of ages, genders, and occupations. They also walked through town to ask people if they would like to participate. We balanced each group by gender with a mix in age. Most participants were farmers, the main occupation in the area.

I was not a part of the selection process because I am a foreigner, but I did participate in the discussions. We tried to minimize any social desirability impact by emphasizing that I was simply a student researcher from Princeton University, not working for an NGO or the government. I had piloted the focus group and survey with 10 participants from a poorer neighborhood in Kigoma town. I asked my research assistants to ask where the participants thought I was from and who had sent me. They all thought I was a Chinese or American student, not from the government, the UNHCR, or another development organization. I believe my background as a person of East Asian descent in an area with rarely any East Asians, at least at the time of the study, helped to dispel beliefs that I was an aid worker.

The focus group discussions lasted about 2 hours. This period did not factor in the time to obtain consent for each participant, making introductions, explaining the study, and debriefing afterwards. We provided refreshments and a small token of 1000 shillings (.50USD) to all participants. The discussions were conducted in Swahili and sometimes, Kiha the local dialect. My research assistants served as interpreters.

## S2.2 Additional Focus Group Results by Proximity to the Refugee Camp

This section shows supplementary results from the community focus groups on nationalism, citizenship, and public goods ranked by their proximity to Nyarugusu refugee camp.

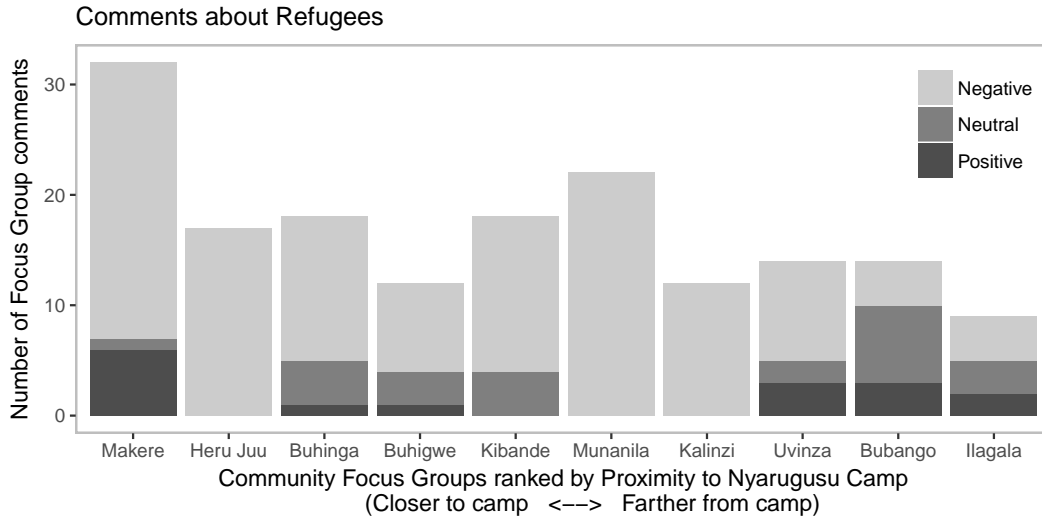


Figure S7: This figure shows the number of comments about refugees and whether those comments were positive, neutral, or negative for each community focus group, ordered by proximity to Nyarugusu refugee camp. Makere, the community closest to the camp (less than 10km away) made the most comments about refugees. While the majority of comments are negative, the communities closest and farthest away (approximately 120km away) also bring up positive effects of refugees.

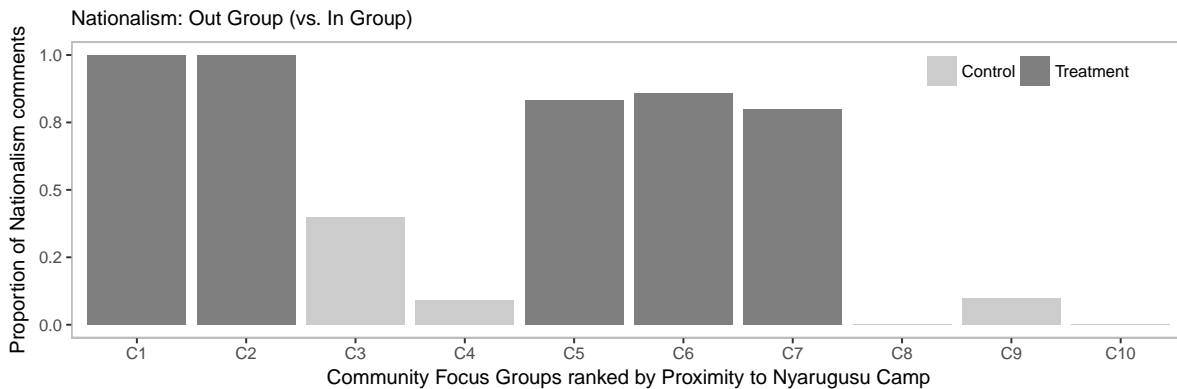


Figure S8: This figure shows for each focus group, ordered by proximity to Nyarugusu refugee camp, the proportion of comments about nationalism that make references to an out group, as opposed to in group values or pride. The treated focus groups (dark gray) make more references to out groups when speaking about nationalism, particularly the ones closest to the camp.



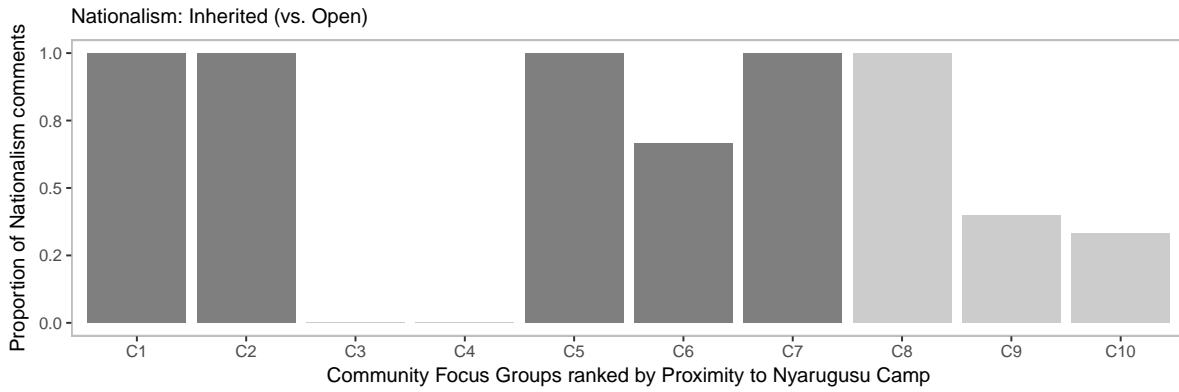


Figure S9: This figure shows for each focus group, ordered by proximity to Nyarugusu refugee camp, the proportion of comments about nationalism that describe it as innate or inherited, as opposed to being open to others. The treated focus groups (dark gray) make more references to an innate or inherited nationalism.

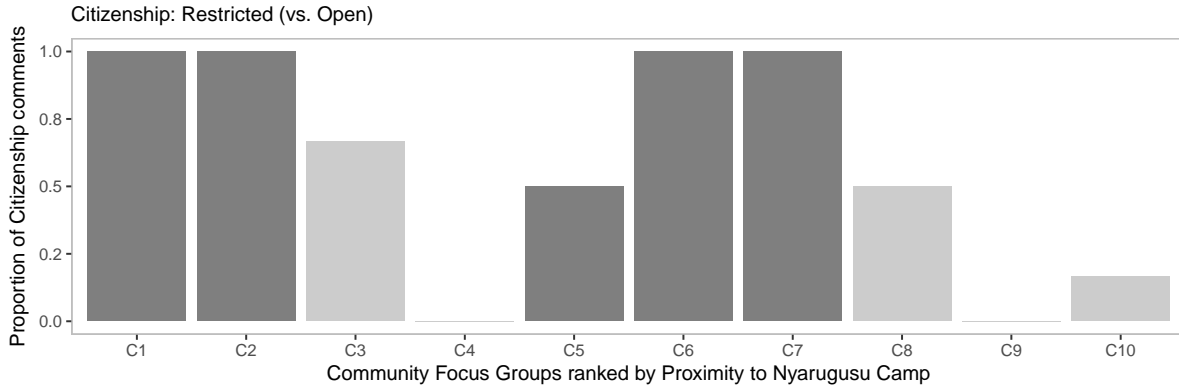


Figure S10: This figure shows for each focus group, ordered by proximity to Nyarugusu refugee camp, the proportion of comments about citizenship that say access should be restricted, as opposed to open to others. The treated focus groups (dark gray) are generally more likely to say that citizenship access should be restricted.

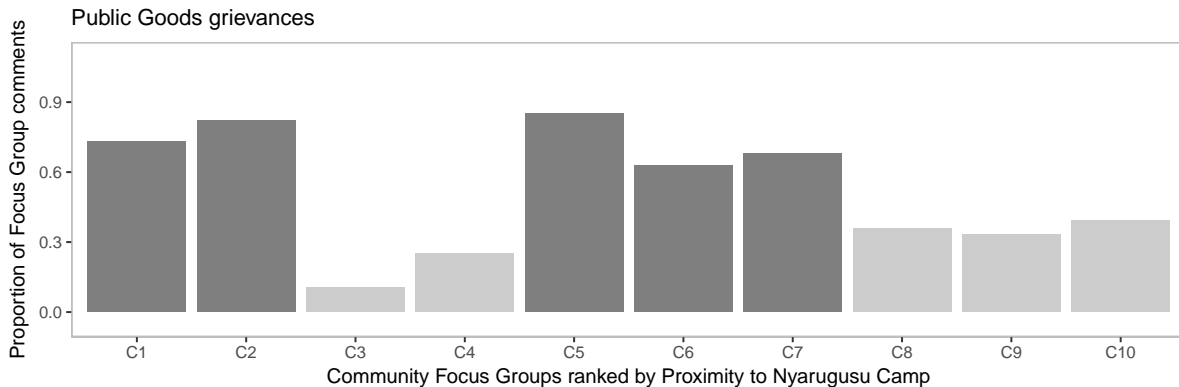


Figure S11: This figure shows for each focus group, ordered by proximity to Nyarugusu refugee camp, the proportion of comments about public goods grievances. The treated focus groups (dark gray) express more public goods grievances.

## S2.3 Land Disputes between Host Communities and Refugee Camps

Although according to government officials and humanitarian agencies working in these camps, many of the concerns expressed by local host communities about the refugees are unfounded, there are real land disputes near the camps. As the 2015 refugee influx expanded the existing refugee camp, Nyarugusu, to capacity and led to the creation of two additional camps, Mtendeli and Nduta, the land for these camps was originally farmland used by local citizens. Local citizens explained that they received no warning or compensation when the government appropriated the land for the camp. For example, the swath of land that became an extension of Nyarugusu camp shown in Figure 5 previously served as farms for the citizens of Makere. A participant from the Makere focus group recounted:

“In April one morning when we came to our cassava farms, we found people cutting down our crops and putting up the tents. They did not even tell us what was happening. They did not even let us harvest our crops first... Our village chairman went to the camp to ask about our farms and our cassava crops, why had they cut down our crops, but he was told that the land belongs to the UN and not to us. We have been using this land for many years, now it belongs to the UN? ...[When he went] to ask, just to speak with them on our behalf, they put him in the jail. We could not find him for days.”<sup>1</sup>

With the creation of Nduta and Mtendeli camps, the surrounding communities also reported similar land appropriations of their farmlands. Although there were initial plans by government authorities and the UNHCR to provide financial compensation,<sup>2</sup> most affected citizens received nothing (da Costa, 2018).

## S2.4 Positive/Mixed Comments about Refugees

With respect to positive comments, only Makere’s community members reported concrete positive, if mixed, economic externalities. First, the rapidly expanding refugee camps could increase opportunities for local employment, yet participants from Makere reported that “Some of the young men in the area are employed as security guards. But only some, most of the jobs in the camps do not go to locals.”<sup>3</sup> da Costa (2018) corroborated that the central government largely sends Tanzanian staff from the state center rather than employ locally.

Second, due to an influx of staff, there are more consumers and higher demand in the local markets: “It is better for the businessmen and the sellers in the market because they can do more business and they can sell for higher prices.”<sup>4</sup> Nevertheless, “increasing the prices of the goods in the markets here is good for the sellers, but it is very bad for the rest of us who need to buy from the market.”<sup>5</sup> The

---

<sup>1</sup>Female participant, Makere focus group, August 4, 2015.

<sup>2</sup>Interview with Director General, Refugee Services Dept., Ministry of Home Affairs, Dar es Salaam, July 1, 2016.

<sup>3</sup>Male participant, Makere focus group, August 4, 2015.

<sup>4</sup>Female participant, Makere focus group, August 4, 2015.

<sup>5</sup>Male participant, Makere focus group, August 4, 2015.

positive comments from other communities were directed towards ‘the 1972 refugees,’ another refugee group who had fled Burundi in 1972, settled in the neighboring region of Tabora, and recently naturalized as Tanzanian citizens.

## S2.5 Comments about Public Goods Grievances

This section shows the main topics discussed across the community focus groups with respect their public goods grievances.

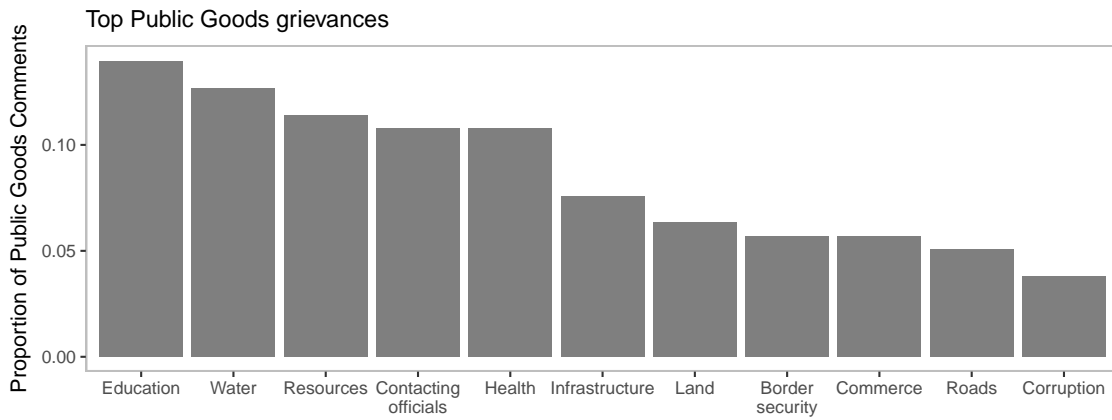


Figure S12: This figure shows the top public goods grievances brought up by the community focus group participants. 45.6% of these comments about public goods also mention refugees.

## S3 Survey Evidence

### S3.1 Sampling Strategy

This section describes the sampling strategy for the original geo-referenced survey conducted in June and July 2016 with 2025 citizen respondents in Kigoma, Tanzania.

First, using the 2012 Tanzanian Census, which includes population and geographic shapefile data at the smallest government administrative level, the ward, I sampled 25 of the 111 wards across the 6 districts in Kigoma region based on the distance (km) of their centroids' from the refugee camp and their population size.<sup>6</sup> These 25 sampled wards include 4 wards each in Kigoma, Uvinza, Kibondo, Buhigwe, and Kakonko, and 5 wards in Kasulu (which is both more densely populated and includes the refugee camp). Their representativeness in terms of distance to the refugee camp and population density are confirmed in figure S13.

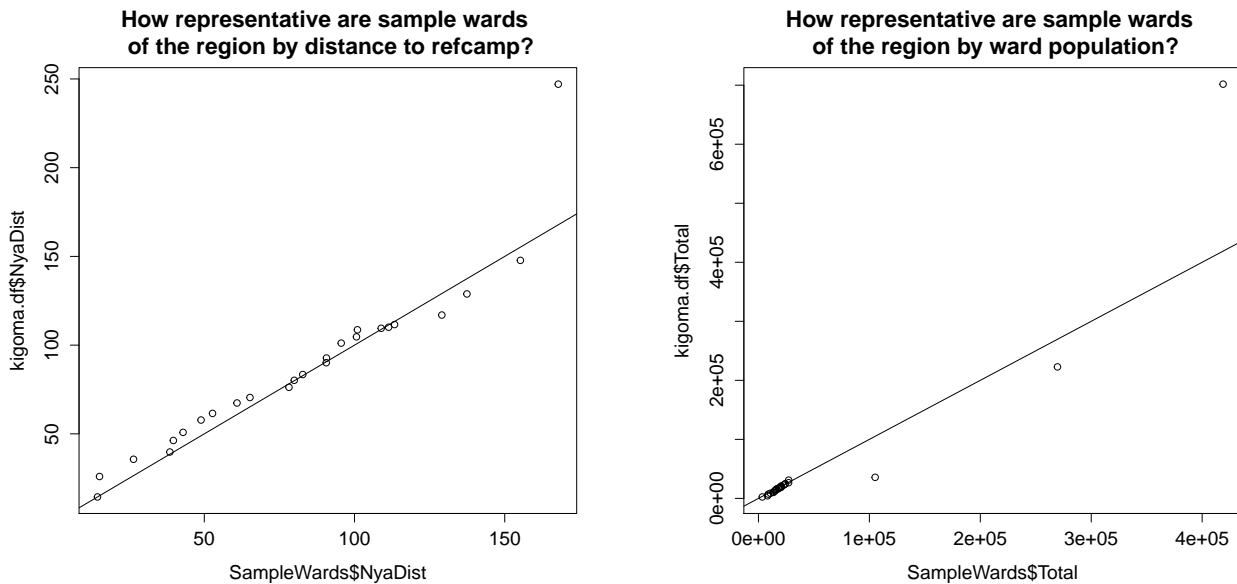


Figure S13: These QQplots show that the 25 sampled wards are representative of the range of distances of the 111 wards in Kigoma from Nyarugusu refugee camp calculated from the ward centroids (left) and are representative in terms of population density of the 111 wards in Kigoma.

Second, within each of the wards, I randomly selected two villages for a total of 50 villages. Once within a village, enumerators used a random-walk technique starting at village landmarks (e.g. church, primary school, pharmacy), taking different directions and choosing every other house on the left side of the road. Finally, within selected households, the person who answered the door was asked to list all eligible members – citizens age 18 or older alternating by gender – and the respondent was chosen randomly from that list.

---

<sup>6</sup>Specifically I partitioned the number of wards within each of these districts into those farther and closer (centroid distance is greater and less than the mean distance) to the camp. Then, within those two groups, I partitioned them again to greater and lower population density compared to the median ward within those groups.

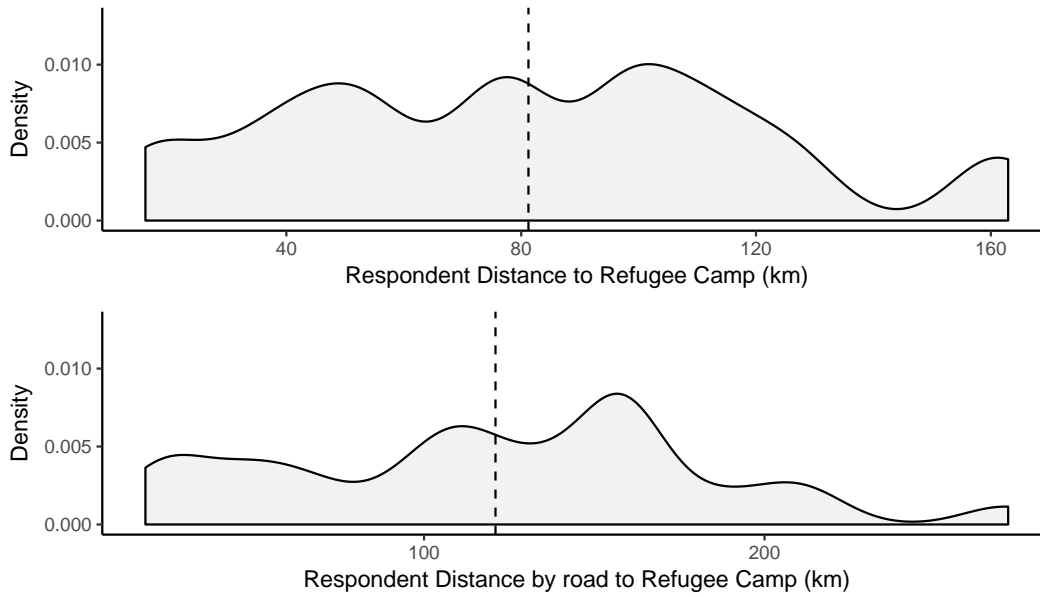


Figure S14: These figures show the distribution of survey respondent distances to Nyarugusu refugee camp by direct path (top) and by road (bottom). The dashed lines show the mean distances.

The survey team consisted of 2 supervisors and 10 enumerators. The enumerators were 5 women and 5 men; 4 were Muslim, 6 Christian. All team members were from Kigoma town and could converse in the local dialect, Kiha in addition to Swahili. Covering 2 villages within the same ward per day, each enumerator conducted a minimum of 8 surveys in Swahili on GPS and mobile data-enabled tablets for a total of 2025 survey respondents, shown as black points in figure S5, and figure S14 shows the distribution of their distances to Nyarugusu refugee camp.

### S3.2 Survey Experimental Treatment priming Refugees

Figure S15 shows the survey flow. The complete survey (English version) is included in section ??.

Respondents were randomly assigned to the refugee prime treatment or control with probability 1/6, 1/6, 1/6, 1/2 respectively. The three informational treatments are called *Refugee-BaseInfo*, *Refugee-AidInfo*, and *Refugee-ConflictInfo*.

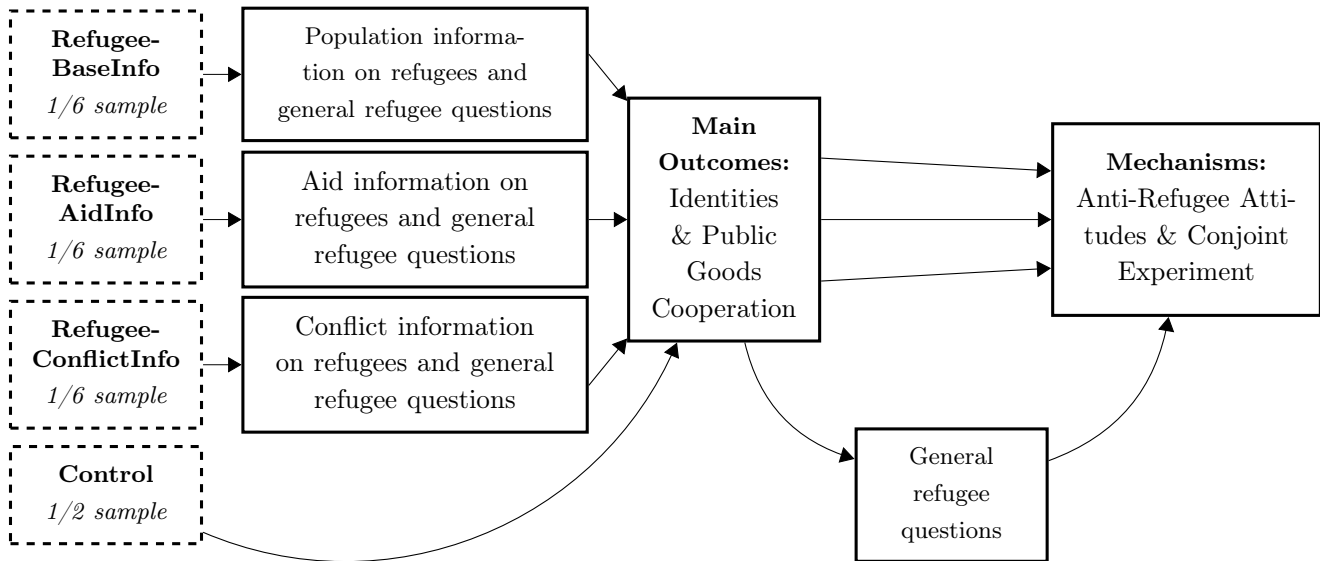


Figure S15: This figure shows the survey experiment flow beginning with the four experimental groups (dashed lines) moving through the informational treatments, main outcomes questions, and mechanisms questions. For the complete survey, please see the Appendix.

For the *Refugee-BaseInfo* group, respondents are first asked a few general questions assessing their prior knowledge and how salient refugees are to their lives:

Do you know which countries the refugees here in Tanzania come from?  
 Do you ever talk about refugees with your friends and family.  
 If so, how often in the past year?  
 Are there any refugees living in your village?  
 Do you personally know any refugees?  
 Do you know about the Nyarugusu refugee camp?

These respondents are then shown a map of where Nyarugusu camp is located, a few photos of the camp, and given factual population information about the refugees living there:

As you may know, these refugees have fled from Burundi and the DR Congo. Currently, there are about 270,000 refugees living in Tanzania. This means that there is 1 refugee here in Tanzania for every 180 Tanzanians. Most of the refugees in Tanzania are living here in Kigoma region at the Nyarugusu refugee camp, which is located in Kasulu district (show map). Nyarugusu has been open for more than 20 years. In 2014, there were 64,000 refugees in this camp. Today, that number has more than doubled; there are more than 160,000 refugees living in Nyarugusu.

This means that there is 1 refugee at Nyarugusu for every 13 Tanzanians from Kigoma. Nyarugusu is now one of the largest refugee camps in the world (show images).

For the Refugee-AidInfo group, respondents are asked the same general refugee questions and given the basic population information. Additionally, they are informed of the aid that the refugee camp receives to help disentangle the mechanism that aid inequality may drive anti-refugee sentiment. The aid information reads:

Many organizations provide assistance to the refugees at Nyarugusu camp. For example, they provide weekly rations of flour for porridge and ugali, schools, a health clinic and pharmacy, latrines and clean water. The Tanzanian government provides the land for this camp and natural resources like timber for construction and firewood.

For the Refugee-ConflictInfo group, again respondents are asked the same general refugee questions and given the basic population information. Additionally, they are informed of the conflict in Burundi which displaced the recent influx of refugees to test the mechanism that conflict may drive anti-refugee sentiment. The conflict information reads:

Most of these refugees have been arriving from Burundi starting in April of last year due to political protests, a failed military coup, violent police and military response, and fears of ethnic conflict. These events were set off when their president announced that he would run for a third term, which he later won.

Finally, for the Control Group, respondents receive no information and directly answer the main outcomes of interest. After these main outcomes are measured, they also answer the general refugee questions, but are not given any additional population, aid or conflict information.



### S3.3 Measuring Identity Attachments

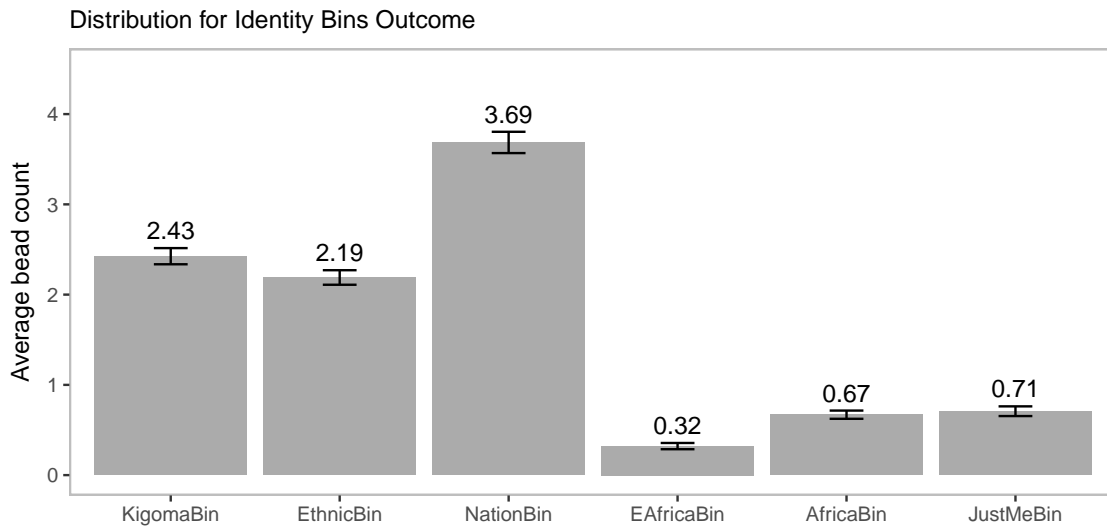


Figure S16: This figure shows the distribution of responses to the binned identity question over all respondents. Citizens feel greater attachment to National, Regional, and Ethnic Group.

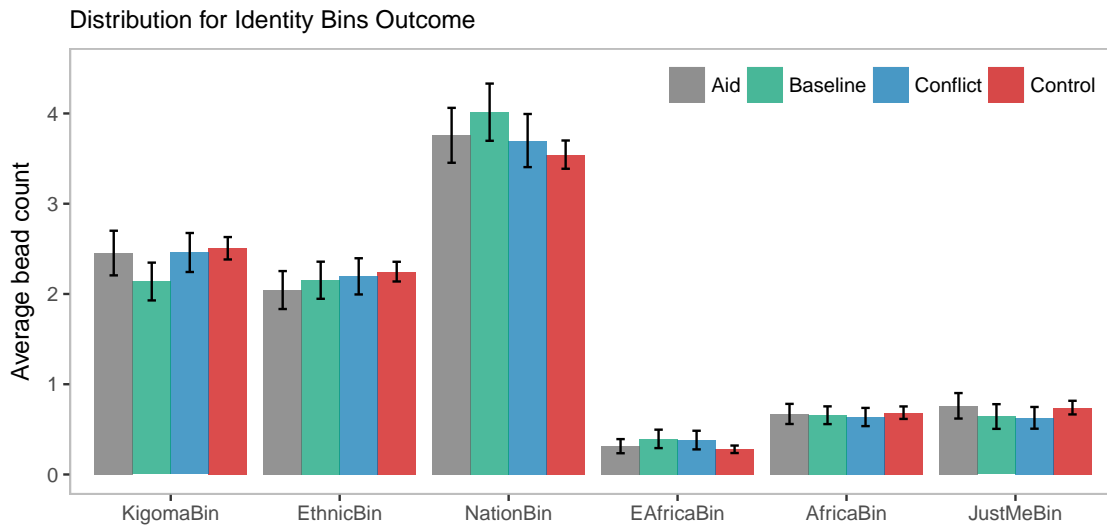


Figure S17: This figure shows the distribution of responses to the binned identity question by survey treatment group. Citizens feel greater attachment to National, Regional, and Ethnic Group.

### S3.4 Survey Respondent Covariates

Certain individual respondent characteristics such as formal education, poverty, urban or rural, ethnicity, political partisanship, media access, literacy, employment, personal migration and remittances, religion, gender, and age are reliable predictors of attitudes towards migrants and levels of nationalism (Espenshade and Hempstead, 1996; Citrin et al., 1997; Hood and Morris, 1998; Burns and Gimpel, 2000; Chandler and Tsai, 2001; Robinson, 2014). These demographic questions were included in the survey in two blocks, with more sensitive questions such as political partisanship asked at the end of the survey.

Table S1: Summary statistics of survey respondents.

Statistic	N	Mean	St. Dev.	Min	Median	Max
Rural	2,025	0.90	0.30	0	1	1
Female	2,025	0.50	0.50	0	0	1
Age	2,025	38.00	16.00	18	35	94
Household Size	2,025	6.00	3.00	1	6	28
Household Head	2,025	0.40	0.50	0	0	1
Education	2,022	3.00	1.00	1	3	7
Ethnic Majority	2,025	0.70	0.50	0	1	1
Employed	2,021	0.90	0.20	0	1	1
Economic Index	2,025	3.00	2.00	0	3	10
Media Access	2,025	0.50	0.50	0	1	1
Landowner	2,025	0.80	0.30	0	1	1
Literate	2,025	0.70	0.50	0	1	1
Christian	2,004	0.80	0.40	0	1	1
Political	1,985	0.80	0.40	0	1	1
CCM Supporter	1,550	0.80	0.40	0	1	1
Kigoma lifetime resident	2,020	0.80	0.40	0	1	1
Parents are Tanzanian	2,023	1.00	0.08	0	1	1
Receives remittances from abroad	2,023	0.01	0.08	0	0	1
Distance to Nyarugusu camp, direct (km)	2,025	81.00	40.00	16.00	78.00	163.00
Distance to Nyarugusu camp, by road (km)	2,025	121.00	62.00	18.00	122.00	272.00
Distance to border (km)	2,025	4.00	6.00	0.10	2.00	29.00
Distance to road	2,025	28.00	32.00	0.90	16.00	145.00

To briefly describe my survey respondents, practically all were born in Tanzania to at least one Tanzanian parent, and about 81.8% have lived in Kigoma region their whole life. 87.9% live in rural areas, practically no one receives foreign remittances. 78.3% support a political party, and 64.5% are CCM supporters. About 61.7% completed primary school, 67.1% are literate, and 53.8% have access to media (newspapers, radio, books, TV).

Since this survey was balanced on gender, 50% are female. 40.1% are heads of the household with an average of 6.2 people per household. The average age is 38.2. 71.5% are of the Ha ethnic group. 77.4% are farmers, and 80.6% are Christian.

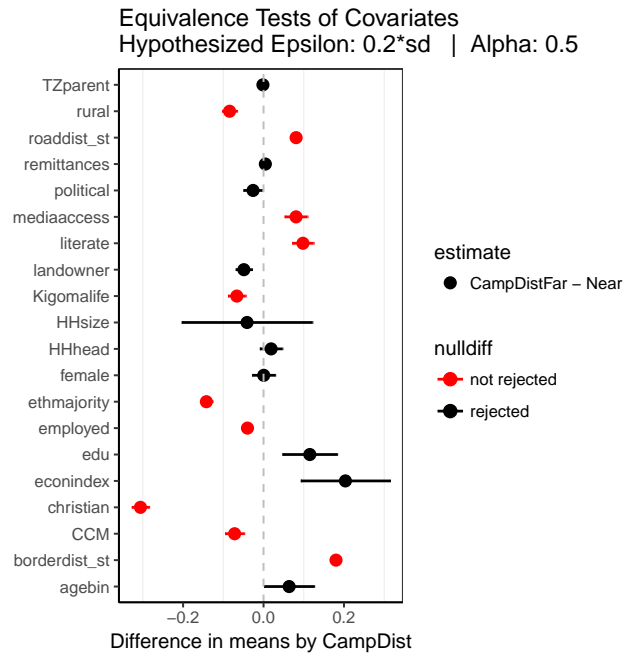
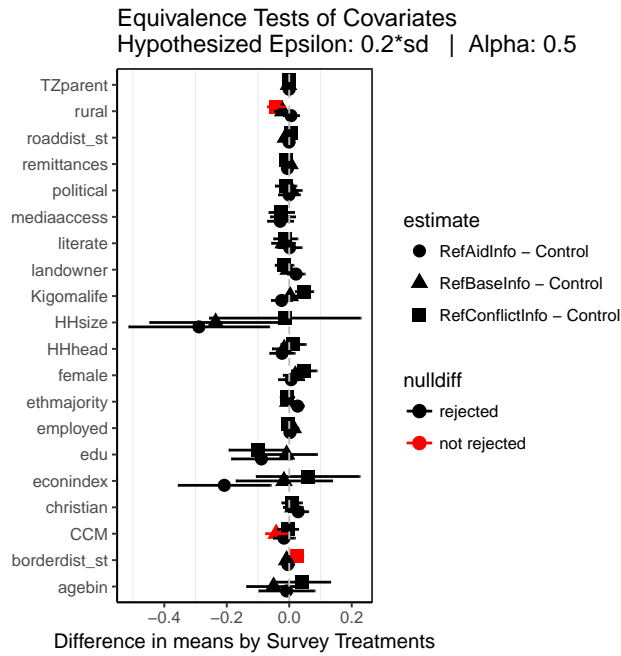


Figure S18: This figure shows the equivalence tests for in-survey treatments and distance to camp across demographic covariates. The covariates for which the null hypothesis of statistical difference was not rejected are controlled for in the regression analysis.

Using equivalence tests in lieu of less conservative balance tests (Hartman and Hidalgo, 2011), figure S18 shows that the null hypothesis of statistical difference was not rejected for the following covariates when comparing respondents who live close versus to far from the refugee camp: rural/urban, access to media, literacy, ethnicity, employment, Christian, political party, whether the respondent has lived her/his whole life in Kigoma. Thus, I control for these covariates.<sup>7</sup>

<sup>7</sup>Hartman and Hidalgo (2011) show that balance tests incorrectly specify the hypothesis test and can fail to reject a true difference between two samples due to low power. Equivalence tests correctly specify the hypothesis test so that difference is the null and equivalence is the alternative.

### S3.5 Attitudes about Refugees by Distance to the Refugee Camp

This section shows the trends of respondents' attitudes towards refugees by their distance to the Nyarugusu camp. Responses were on a scale from 1 to 5: 1 for "Absolutely disagree" and 5 for "Absolutely agree."

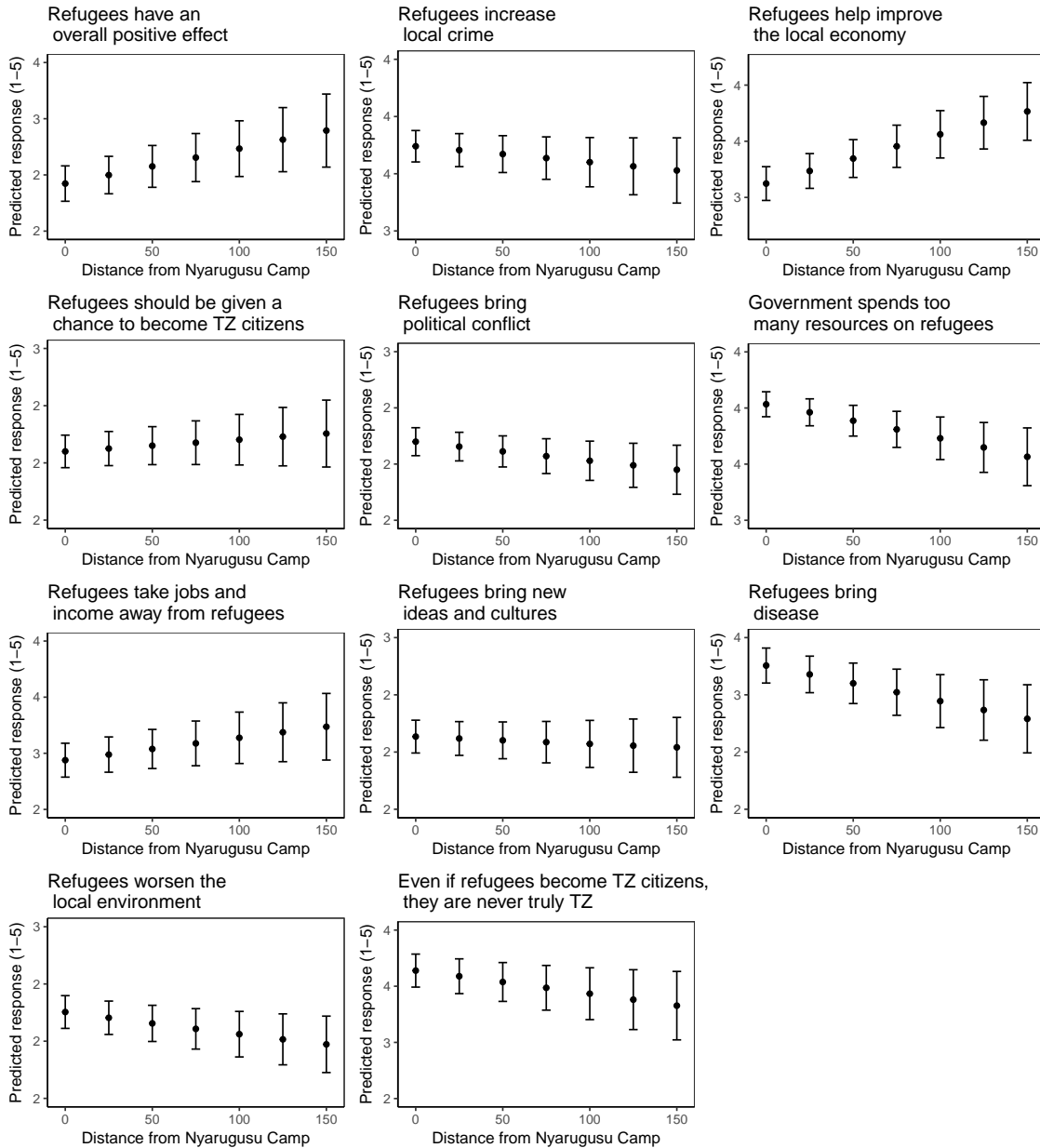


Figure S19: These figures show how respondent distances to Nyarugusu camp are related to the following attitudes about refugees.

### S3.6 Survey Estimation Strategy for Analysis on National Identification

For the main survey analysis on national identification, I pooled across the three refugee informational treatments – `Refugee-BaseInfo`, `Refugee-AidInfo`, and `Refugee-ConflictInfo` for efficiency, and use the following semiparametric generalized additive model using Wood (2011) to allow for more flexible, non-linear relationships between distance to the camp and the outcomes of interest:

$$Y_i = \beta_0 + \beta_1 \text{RefugeeInfo}_i + f(\text{NyaCampDistance}_i) + \text{RefugeeInfo}_i * g(\text{NyaCampDistance}_i) + \beta_4 Z_i + \epsilon_i$$

where  $Y_i$  is the survey response of respondent  $i$ ;  $\text{RefugeeInfo}_i$  is the binary indicator for receiving any of the three in-survey treatments giving refugee information;  $\text{NyaCampDistance}_i$  is the distance in kilometers of the respondent from Nyarugusu refugee camp;  $Z_i$  is the vector of demographic covariates,  $f(\cdot)$  and  $g(\cdot)$  are smooth functions of distance, and  $\epsilon_i$  is the respondent error term. The dimension of the splines are determined by cross-validation.

To calculate the predicted outcome of bead share by distance to the camp, I set the survey treatment condition to 0.

For the marginal effect of the survey treatment moderated by distance to the camp, I use the `inter.kernel()` function from the R package `Interflex` by Hainmueller, Mummolo and Xu (2018), with the bandwidths determined by cross-validation.

### S3.7 Results for other Identity Bins

This section shows the analysis on the effects of distance to Nyarugusu camp for the other five identity bins: Kigoma (subnational regional), Ethnic Group, East African, African, Just Me (no opinion).

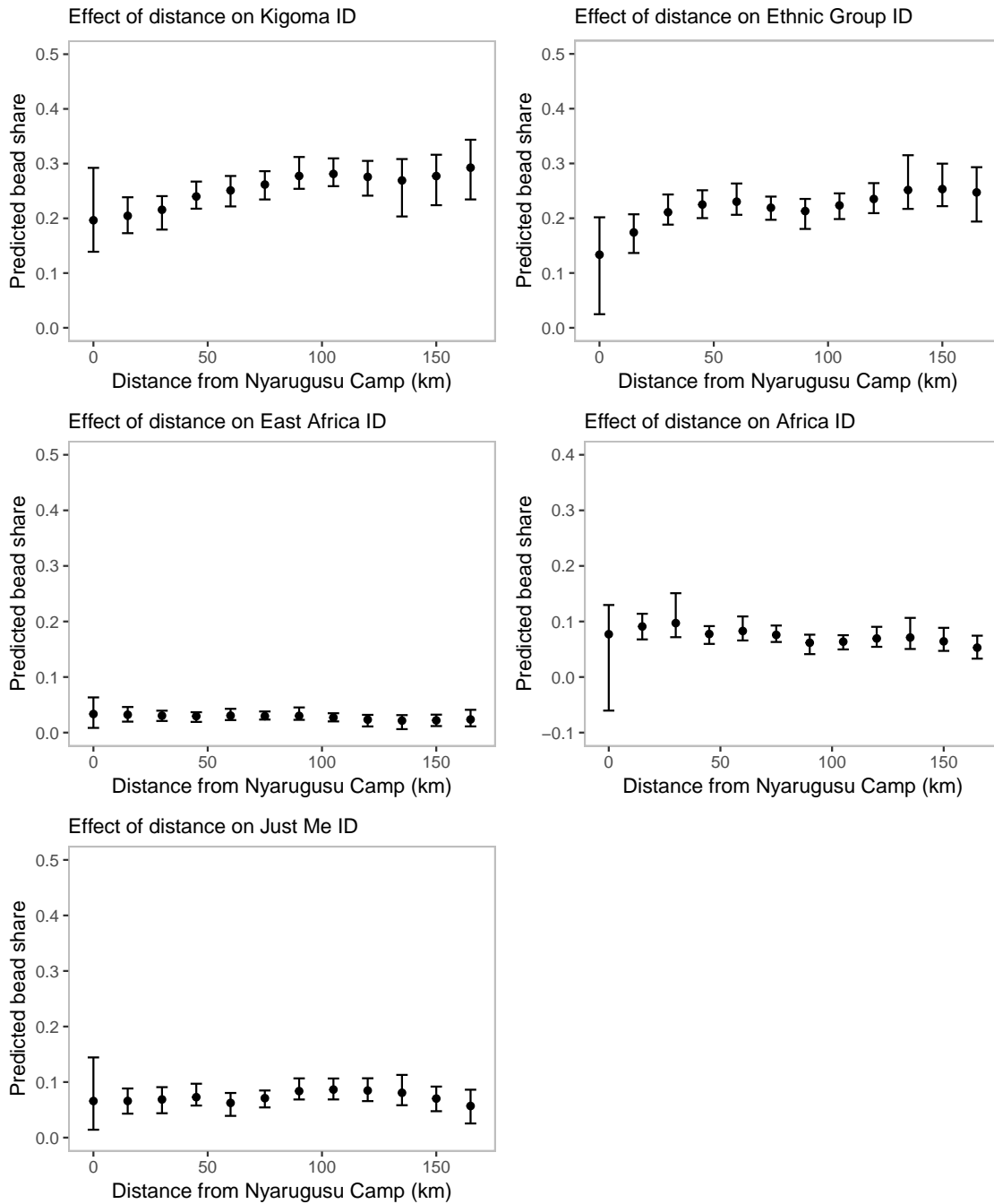


Figure S20: These plots show the predicted share of bead allocations (out of 10 total) for the other identities included in the survey measure by distance to Nyarugusu refugee camp (direct path) for the control condition, with 95% block bootstrapped CIs by village. Citizens closer to the camps feel slightly less regional and ethnic.

### S3.8 Robustness Check using Distance by Road to the Refugee Camp

This section shows the main survey results using an alternative specification to the direct distance to Nyarugusu refugee camp: distance to the camp by road. The results are robust.

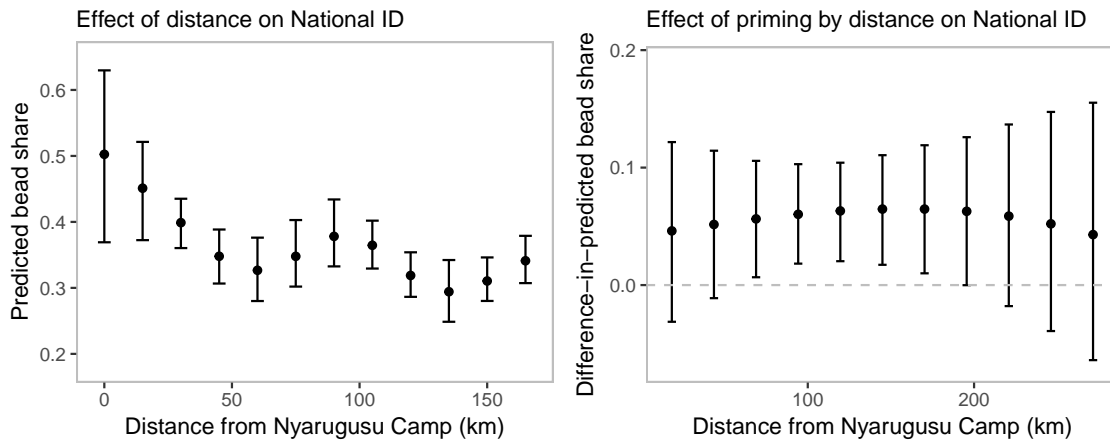


Figure S21: The plot on the left shows the predicted share of bead allocations (out of 10 total) for national identity by distance to Nyarugusu refugee camp (by road) for the control condition, with 95% block bootstrapped CIs by village. The plot on the right shows the treatment effect of the survey prime moderated by camp distance (by road), with 95% block bootstrapped CIs by village. Receiving the prime increases the bead allocation for national identity by about .7 of a bead. Results are fairly robust, except the survey treatment loses statistical significant closer to the camp when moderated by road distance to the camp.

### S3.9 Placebo Results using Distances to Future Refugee Camps, Nduta and Mtendeli

An obstacle inherent to research on migrant-host relations is the selection bias of where migrants are located vis-à-vis host citizens. To directly address concerns of selection bias, the following placebo tests show that there is no relationship between national identification and distance to the *future* refugee camps, Nduta and Mtendeli, which were established months after the survey. This suggests that there are no unobserved confounders biasing the results.

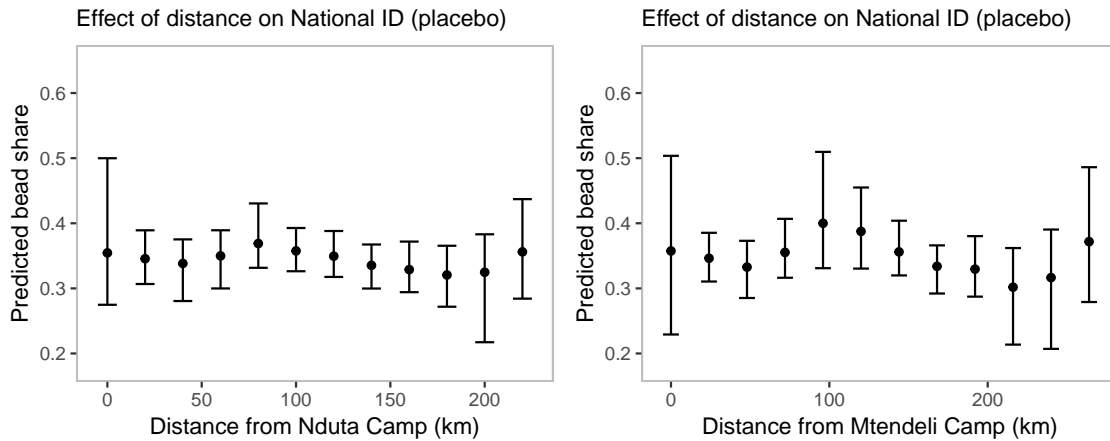


Figure S22: These plots show results for the placebo tests: the predicted share of bead allocations (out of 10 total) for national identity by distance to Nduta (left) and Mtendeli (right) refugee camps (direct path), which were created in late 2016 after the survey, for the control condition, with 95% block bootstrapped CIs by village. There is no effect of distance to these future camps and national identification.



### S3.10 Paired Conjoint Experiment

The following paired conjoint experiment reveals which hypothetical refugee attributes are more (less) preferred to native citizens. Conjoint survey experiments present respondents with profiles, for example, of immigrants (Hainmueller, Hopkins and Yamamoto, 2014), that are randomly assigned attributes and asks respondents to evaluate and choose between them. The random assignment of profile characteristics allows researchers to identify the causal effect of attributes on respondents' preferences. I implement a paired conjoint asking respondents to select between two refugee profiles.















Mkimbizi A	Mkimbizi B
	
	
	
	
	
	
	

Figure S23: This is an example of a hypothetical paired conjoint profiles: Refugee A receives aid from UNHCR, is from the DR Congo, speaks Tanzanian Swahili, is male, Christian, elderly, and a teacher. Refugee B does not receive aid from UNHCR, is from Burundi, speaks Tanzanian Swahili, is male, Muslim, able-bodied, and a farmer.

Due to low literacy in Tanzania, I use tools developed by Meyer and Rosenzweig (2016) to produce images shown in the survey tablet with symbols for attributes. See S23 for an example of a hypothetical paired conjoint profiles.

The randomized attributes for this conjoint experiment are: **Occupation** (farmer, fisherman, market seller, teacher, political leader, military); **Vulnerability** (able-bodied, elderly, injured, ill-health); **Country of Origin** (DR Congo, Burundi); **Gender** (female, male); **Religion** (Christian, Muslim), **Language** (English, French, not Tanzanian Swahili, Tanzanian Swahili, Kiha), and **Receiving aid** (Yes, No). Each respondent will repeat this exercise two times.

Please look at these two profiles side-by-side. They are comparing two potential refugees:

refugee A and refugee B, who both want to come live in Tanzania.

[Enumerator, looking at the profiles, go through each characteristic of refugee A first, then each characteristic of refugee B with the respondent, making sure they understand.]

If you had to choose between them, which would you prefer come to Tanzania? (repeat x2)

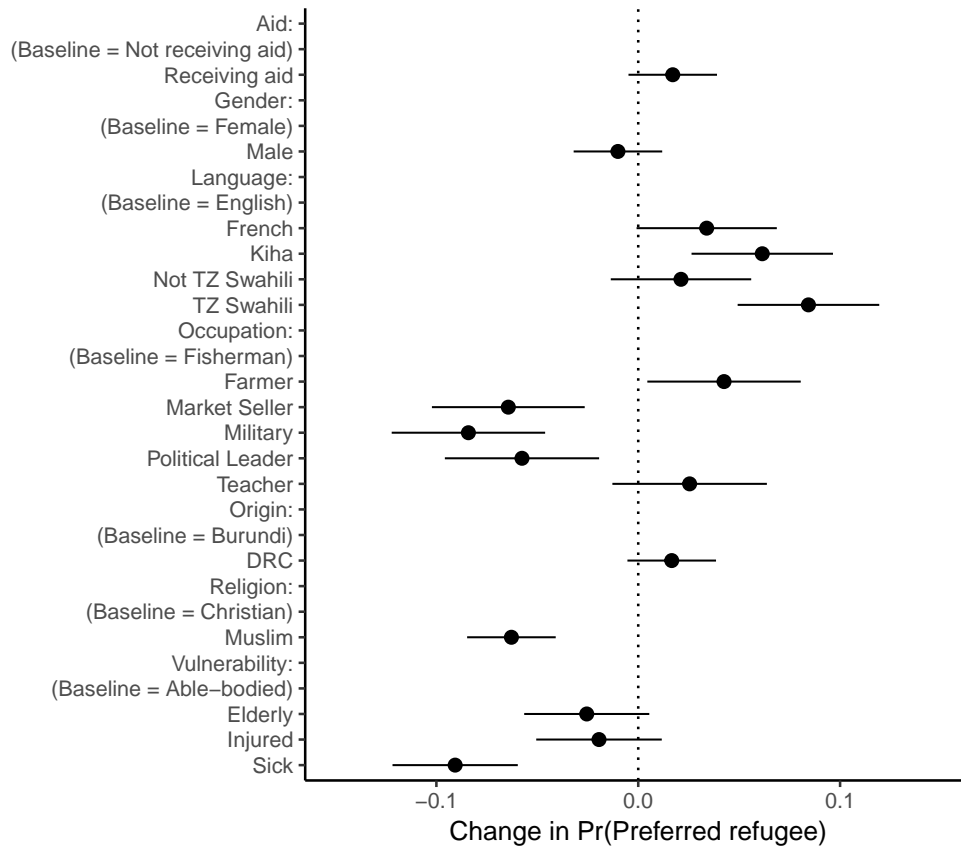


Figure S24: Paired conjoint experiment estimates with 95% CIs show which hypothetical refugee traits increase or decrease the probability of being selected.

Figure S24 shows that the most preferable characteristic is whether the refugee can speak Tanzanian Swahili, the national language which is distinct for example from the Swahili spoken in neighboring East African countries. This finding of the importance of national language as the key to integration speaks to the heightened national identity of citizens when refugees become contextually salient to their lives. Other preferred characteristics are being able to speak the local dialect Kiha, and being a farmer. Since the vast majority of citizen respondents are farmers, this suggests that citizens do not feel economically competitive with refugees in terms of occupation. Unfavorable attributes include being a market seller, having a military or political leadership background, illness, and being Muslim. Since the majority of respondents are Christian, this finding suggests shared religion is also important to citizens.

## S4 Post-Influx Public Goods Observational Evidence

This section shows the descriptive statistics and full regression results for the difference-in-differences analysis of the observational public goods data in Kigoma, Tanzania pre- and post-refugee influx.

### S4.1 Descriptive Statistics

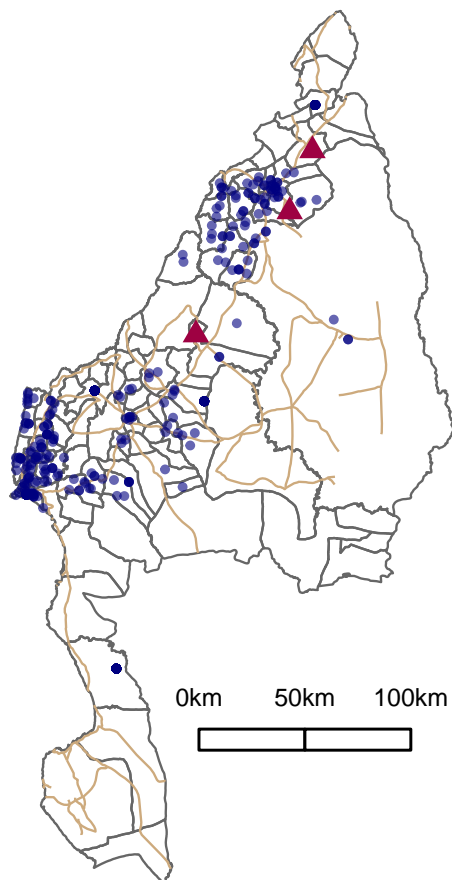


Figure S25: This map shows the 558 primary schools (blue points) in 2017 in relation to the refugee camps (red triangles) and roads (brown lines) in the study region Kigoma.

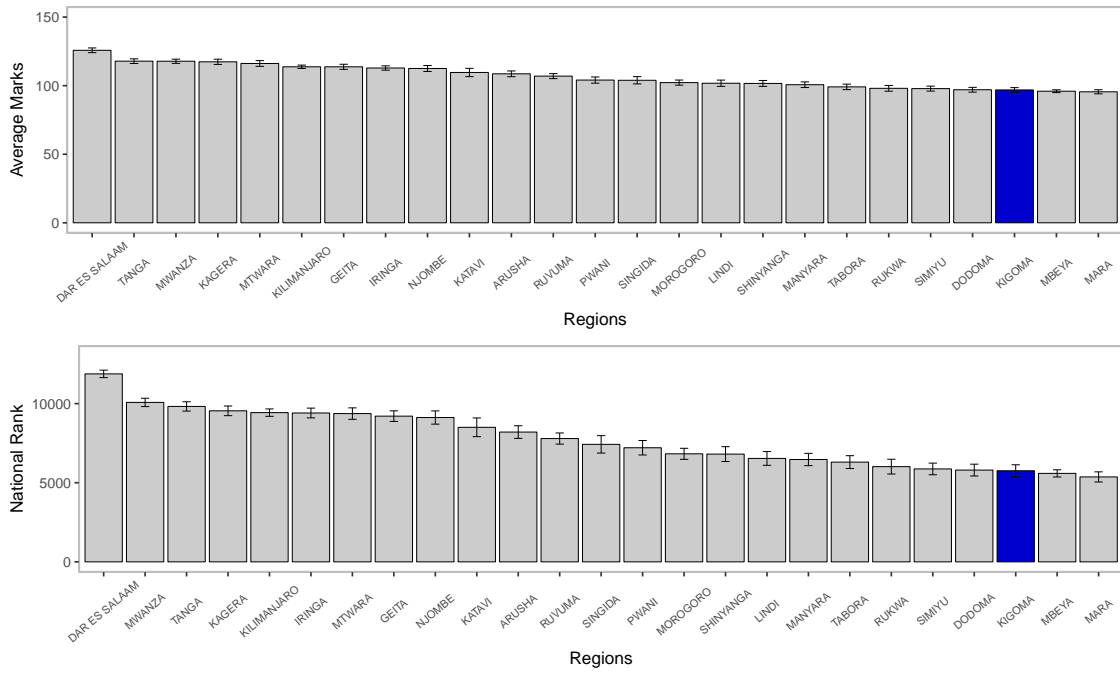


Figure S26: These figures show the region-level averages for the government primary school outcomes – average marks and national rank in 2014 (pre-influx). Compared to other regions in Tanzania, government primary schools in Kigoma generally performed poorly.

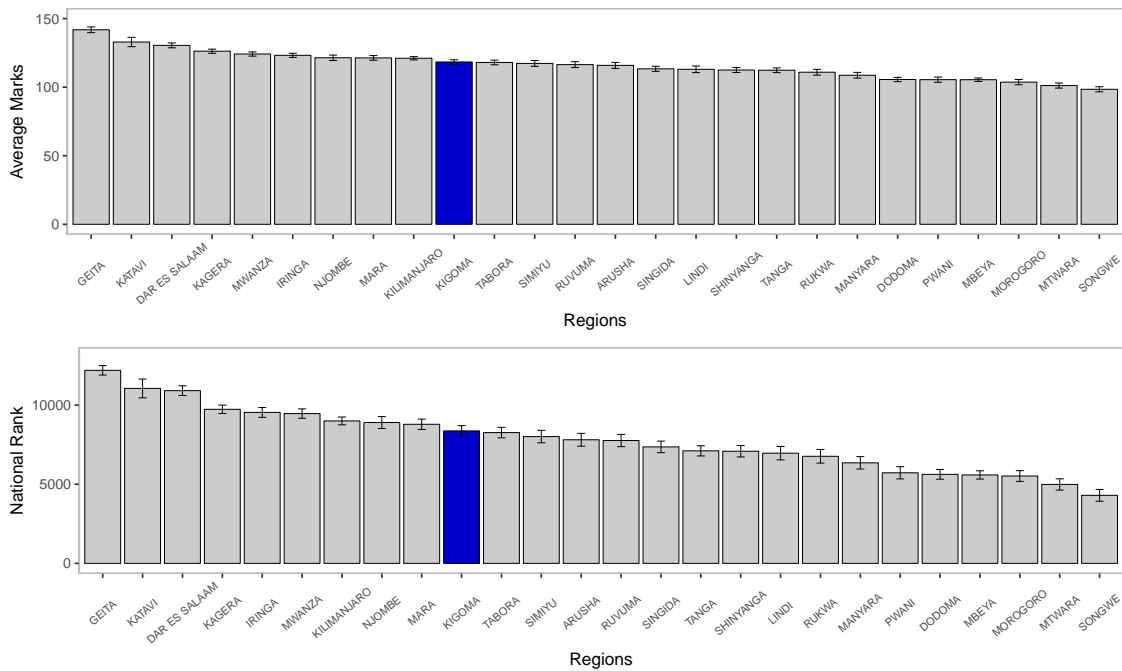


Figure S27: These figures show the region-level averages for the government primary school outcomes – average marks and national rank in 2016 (post-influx). Compared to other regions in Tanzania, government primary schools in Kigoma perform much better than in 2014.

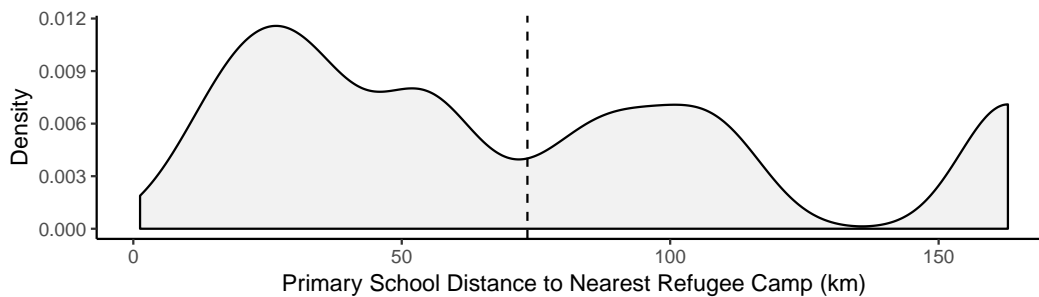


Figure S28: These figures show the distribution of primary school distances to the nearest refugee camp by direct path. The dashed lines show the mean distances.

Table S2: Summary statistics of primary school data.

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Average Scores (/250)	2,503	102.00	22.00	54.00	86.00	99.00	114.00	224.00
National Rank	2,503	6,598.00	3,943.00	42.00	3,212.00	6,194.00	9,600.00	16,122.00
Nearest Camp Distance (km)	3,061	73.00	49.00	1.00	32.00	55.00	107.00	163.00
Nyarugusu Camp Distance (km)	3,061	88.00	43.00	1.00	55.00	82.00	111.00	163.00
Nduta Camp Distance (km)	3,061	129.00	68.00	7.00	77.00	126.00	179.00	232.00
Mtendeli Camp Distance (km)	3,061	153.00	77.00	14.00	108.00	154.00	207.00	263.00
Rural	3,061	0.80	0.40	0	1	1	1	1
Ward Population (2012)	3,061	23,298.00	14,268.00	2,634	16,516	19,486	22,458	67,704
Poverty Rate (2012)	3,061	0.90	0.06	0.70	0.90	0.90	0.90	0.90
CCM Voteshare (2010)	2,977	0.50	0.10	0.30	0.40	0.50	0.60	0.80
Voter Turnout (2010)	2,977	3,595.00	1,434.00	251.00	2,680.00	3,404.00	4,034.00	9,835.00
Distance to road (km)	3,061	5.00	5.00	0.02	1.00	4.00	8.00	24.00
Distance to border (km)	3,061	30.00	17.00	2.00	12.00	33.00	49.00	86.00
Distance to regional capital (km)	3,061	88.00	65.00	0.90	36.00	94.00	111.00	229.00

## S4.2 Public Goods Outcomes Estimation Strategy

For the observational public goods analysis, I use the standard difference-in-differences OLS model controlling for baseline (pre-influx) time-invariant covariates:

$$Y_{it} = \beta_0 + \beta_1 \text{CampDistance}_i + \beta_2 \text{PostInflux}_{it} + \beta_3 \text{CampDistance}_i * \text{PostInflux}_{it} + \beta_4 Z_i + \epsilon_{it}$$

where  $Y_{it}$  is the public good outcome of school  $i$  measured at year  $t$ ;  $\text{CampDistance}_i$  is the distance in km to the nearest refugee camp;  $\text{PostInflux}_{it}$  is the binary indicator for pre- or post-influx;  $Z_i$  is the vector of time-invariant covariates, and  $\epsilon_{it}$  is the error term.

Additionally, I use the following semiparametric generalized additive model using Wood (2011) to allow for more flexible, non-linear relationships between distance to the camp and the outcomes of interest:

$$Y_i = \beta_0 + \beta_1 \text{PostInflux}_{it} + f(\text{CampDistance}_i) + \text{PostInflux}_{it} * g(\text{CampDistance}_i) + \beta_4 Z_i + \epsilon_i$$

where  $Y_i$  is the public good outcome of school  $i$  measured at year  $t$ ;  $\text{CampDistance}_i$  is the distance in km to the nearest refugee camp;  $\text{PostInflux}_{it}$  is the binary indicator for pre- or post-influx;  $Z_i$  is the vector of time-invariant covariates,  $f(\cdot)$  and  $g(\cdot)$  are smooth functions of distance, and  $\epsilon_i$  is the error term.

For the marginal effect of post influx moderated by distance to the camp, I use the `inter.kernel()` function from the R package `Interflex` by Hainmueller, Mummolo and Xu (2018), with the bandwidths determined by cross-validation.

### S4.3 Parallel Trends Assumption

This section shows the pre-influx trends of government primary school outcomes by year to visually inspect the parallel trends assumption.

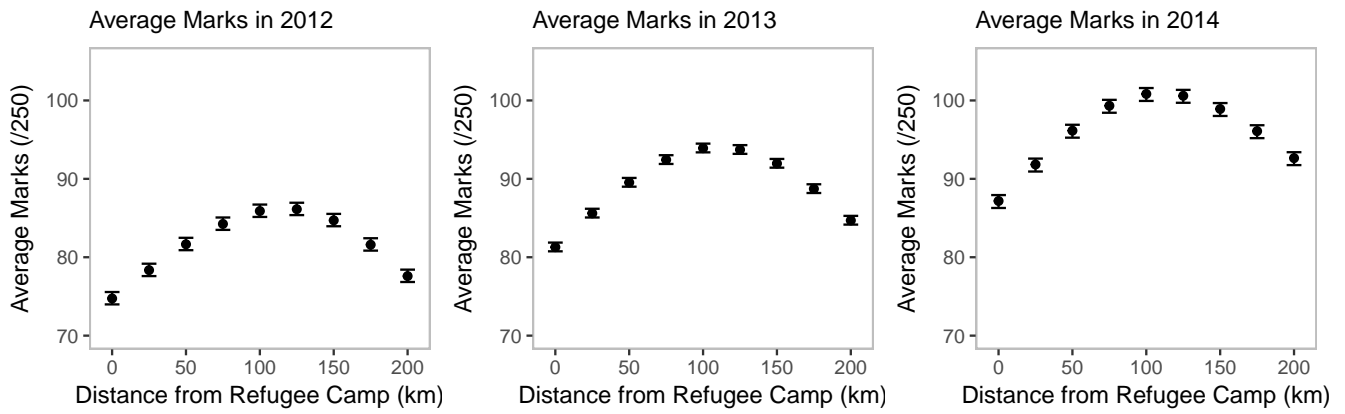


Figure S29: These figures show pre-refugee influx trends for government primary school outcomes across distances to the nearest refugee camp. Pre-influx, schools between 100 - 150km from a camp performed the best.

## S4.4 Additional Primary School Outcomes Analysis

This section shows the full regression tables with OLS models for the primary school outcomes, including the coefficients for the control variables.

Table S3: Comparing Government (columns 1-2) and Non-Government (columns 3-4) Primary School Outcomes on their Distance to the Nearest Refugee Camp pre- and post-influx using Difference-in-Differences OLS models, controlling for rural/urban, ward population in 2012, poverty rate (1 square km) in 2012, CCM voteshare and voter turnout in 2010, and distances to the nearest road, international border, and regional capital. Below each model are the means and standard deviations of the outcomes pre-influx.

	<i>Dependent variable:</i>			
	Average Marks Government Primary Schools	National Rank	Average Marks Non-Government Primary Schools	National Rank
Nearest Camp Distance (km):Post-Influx	-.10*** (.02)	-13.00*** (4.00)	.10 (.20)	22.00 (33.00)
Post-Influx	29.00*** (2.00)	3,435.00*** (347.00)	42.00** (20.00)	2,946.00 (3,518.00)
Nearest Camp Distance (km)	.07*** (.02)	-5.00 (4.00)	-214.00 (148.00)	-19,728.00 (26,614.00)
Rural	.40 (2.00)	-62.00 (307.00)	62,014.00 (43,273.00)	5,700,106.00 (7,760,513.00)
Ward Population (2012)	.0001** (0.0000)	.02* (.01)	.60 (.40)	58.00 (78.00)
Poverty Rate (2012)	-39.00*** (12.00)	-12,271.00*** (2,242.00)	-391,847.00 (273,095.00)	-36,056,424.00 (48,976,801.00)
CCM Voteshare (2010)	5.00 (5.00)	134.00 (884.00)	31,219.00 (21,639.00)	2,878,593.00 (3,880,698.00)
Voter Turnout (2010)	.001* (.0003)	.10* (.06)	10.00 (7.00)	946.00 (1,283.00)
Distance to road (km)	-.30** (.20)	-36.00 (32.00)	3,321.00 (2,306.00)	306,117.00 (413,634.00)
Distance to border (km)	-.04 (.03)	2.00 (7.00)	-394.00 (275.00)	-36,235.00 (49,338.00)
Distance to regional capital (km)	.0001 (.01)	-7.00*** (2.00)	16.00 (11.00)	1,391.00 (2,032.00)
Intercept	123.00*** (9.00)	17,427.00*** (1,814.00)	246,934.00 (172,038.00)	22,725,876.00 (30,853,222.00)
Pre-Influx Mean	102.02	6708.23	98.29	5821.99
Pre-Influx SD	21.42	3904.53	24.19	4130.98
Observations	2,122	2,122	311	311
Adjusted R <sup>2</sup>	.20	.10	.30	.30
Residual Std. Error	19.00	3,629.00	20.00	3,561.00

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01



## References

- Burns, Peter and James G Gimpel. 2000. "Economic insecurity, prejudicial stereotypes, and public opinion on immigration policy." *Political Science Quarterly* 115(2):201–225.
- Chandler, Charles R and Yung-mei Tsai. 2001. "Social factors influencing immigration attitudes: an analysis of data from the General Social Survey." *The Social Science Journal* 38(2):177–188.
- Citrin, Jack, Donald P Green, Christopher Muste and Cara Wong. 1997. "Public opinion toward immigration reform: The role of economic motivations." *The Journal of Politics* 59(03):858–881.
- da Costa, Diana Felix. 2018. You May Think He Is Not A Human Being: Refugee and Host Community Relations in and around Nduta and Mtendeli Refugee Camps, Western Tanzania. Technical report Danish Refugee Council.
- Espenshade, Thomas J and Katherine Hempstead. 1996. "Contemporary American attitudes toward US immigration." *International Migration Review* pp. 535–570.
- Hainmueller, Jens, Daniel J Hopkins and Teppei Yamamoto. 2014. "Causal inference in conjoint analysis: understanding multidimensional choices via stated preference experiments." *Political Analysis* 22(1):1–30.
- Hainmueller, Jens, Jonathan Mummolo and Yiqing Xu. 2018. "How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice."
- Hartman, Erin and F Daniel Hidalgo. 2011. "Whats the alternative?: An equivalence approach to balance and placebo tests." *Manuscript, Departments of Political Science, MIT and UC Berkeley* .
- Hood, MV and Irwin L Morris. 1998. "Give us your tired, your poor,... but make sure they have a green card: The effects of documented and undocumented migrant context on Anglo opinion toward immigration." *Political Behavior* 20(1):1–15.
- Meyer, Alexander and Leah R Rosenzweig. 2016. "Conjoint Analysis Tools for Developing Country Contexts."
- Robinson, Amanda L. 2014. "National versus Ethnic Identification in Africa: Modernization, Colonial Legacy, and the Origins of Territorial Nationalism." *World Politics* 66(04):709–746.
- Wood, S. N. 2011. "Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models." *Journal of the Royal Statistical Society (B)* 73(1):3–36.