Online Supporting Information for

Reducing Prejudice Towards Refugees in Uganda: Evidence that Social Networks Influence Attitude Change

Jennifer M. Larson* Janet I. Lewis†

February 14, 2024

Contents

*Associate Professor of Political Science, Vanderbilt University, jennifer.larson@vanderbilt.edu.

†Associate Professor of Political Science, George Washington University, janetilewis@gwu.edu.

Appendices

A Intervention Details

Following Broockman and Kalla (2016), our intervention proceeded in several steps and allowed for some flexibility to allow for a natural exchange rather than a fully scripted one. We trained the survey team to carry out the intervention with great care, to ensure that they thoroughly understood each component and had ample opportunity to practice the intervention. All research team members who conducted interviews were from West Nile (but not the villages we studied) and spoke the local language (Lugbara).

Step 1. Create Non-Judgmental Context

Intervention begins: Research team members engage in a series of strategies to elicit participants' opinions in a non-judgmental manner. Research team members should ask respondents about their opinion on the just-asked baseline questions about their own attitudes and ask them to explain their position in a non-judgmental manner, not indicating they were pleased or displeased with any particular answer, but rather to appear genuinely interested in hearing the subject think about the question. This is intended to encourage reflection and to build rapport.

For example, the research team member could say, "I see you said that you somewhat agreed that refugees would be a burden on the resources of your community. That's interesting. I'm curious to hear more. Could you please tell me more about why you feel that way?" After hearing the response, the research team member should not affirm that the respondent's opinion is "correct" but should convey understanding and empathy, for example: "I see. That sounds like a very difficult situation when (repeat back some of what you heard about what the respondent experienced or heard)."

Step 2. Exchange Narratives

The enumerator should reference back to an earlier baseline question: "Have you ever met a refugee living in Uganda?" If the respondent knows someone who is a refugee, the enumerator would have the respondent talk about how they know this person, their refugee story (e.g. how and why they came to Uganda), and how it must feel to be a refugee (e.g. "wow, do you think that was difficult for this person? What challenges did they face?"). Whether or not the respondent knows a refugee, the enumerator would always share a story about a refugee they know (we always used the name Gift).

For example, after hearing about the respondents' experiences with a refugee, the enumerator would say: "Oh, that is interesting. I know a refugee named Gift who lives in [name of nearest refugee settlement.] She left South Sudan a few years who when her village was burned by rebels. They cannot find her husband, and her mother was raped and killed. She and her 3 young kids walked the whole way to Uganda, hiding from rebels in the bush along the way. They are so grateful to be here in Uganda. Gift works as a translator now in the refugee camp where she lives, to earn enough food to for her family. Even still, she is usually only able to give her children one meal per day and life is very difficult.

The enumerator would end this section by asking the respondent if there is anything about the story that they can relate to, encouraging perspective taking. For example, “Have you ever experienced anything like that, such as not being able to feed your family as much as you would like? Or insecurity in your community?”

Step 3. Exchange narratives about a personal experience with compassion.

Enumerators ask respondents to share a time when someone showed them compassion. If necessary, enumerators should tell their own stories of being shown compassion in order to make respondents feel comfortable sharing a story of their own. Enumerator’s goal is for this non-judgmental exchange of narratives to end with individuals self-generating and explicitly stating aloud implications of the narratives that ran contrary to their previously stated exclusionary attitudes.

An example of a story an enumerator could share about experiencing compassion is: “Once, in school, I fell far behind because my mom was sick and I had to stay home and care for her. My teacher came to my home to tutor me every evening for 3 weeks to help me catch up. I never forgot what my teacher did for me. I appreciated this so much. Has such a thing ever happened to you – a time when someone showed you compassion and helped you?” Then: “Do you think this kind of compassion should apply to refugees? How should we do that?”

Step 4. Address Concerns.

At this point, the enumerator would return to any concerns about refugees that the respondent may have mentioned earlier. The enumerator would talk through these concerns and, where applicable, provide talking points to refute them. Enumerators will be trained not to address concerns until this point in the conversation so that respondents would not feel threatened by this section. Only after rapport had been established, stories shared, and the value of compassion activated would enumerators address concerns.

For example, if the respondent surfaced a concern about refugees using their land for gardening or firewood, the enumerator could say: “I hear you that sometimes it feels like they are using up our land. But did you know that during the 1980s, after Idi Amin was pushed out and many Ugandan people fled West Nile to South Sudan, many South Sudanese shared land with Ugandans? Also, I know that most of the South Sudanese refugees are very respectful and try not to overuse our land. They hoping to return back to their land as soon as it is safe.” See also the information sheet provided to the research team for additional factual information about refugees in Uganda.

If respondents surface a personal negative experience with refugees, it’s important to acknowledge that experience, and to share sympathy about the difficult experience. This can be followed by statements about most refugees not behaving in that negative way.

Step 5. Make the Case.

The enumerator should then reiterate for the enumerator why they hoped the respondent would become more supportive of refugees.

For example, the enumerator could say, "I wanted to exchange stories about refugees because I have this knowledge about them – what they have been through and how they are trying to do good things for their families and how they respect Uganda – and I felt that if I shared it with you, it may help you come to understand them and support them better."

B Theory of Attitude Change

B.1 A Model of Attitude Formation with Social Processing

Suppose a person updates her attitude in response to an experience by averaging the result of the experience with her prior, baseline attitude, weighted by how sure she is that her original attitude was correct. Then we could say that a treatment that aims to change i's attitude to be $y^\star$ results, in the short-term, in an attitude of:

$$y_{i,st} = (1 - s_i)y^\star + s_i y_{i,bl}$$

where s_j is how sure i was in her baseline attitude $y_{i,bl}$ and $0 \leq s_i \leq 1$. When i is quite sure of her original attitude, her short-term attitude will remain close to her baseline attitude. As long as $s_i \neq 1$, the larger the difference between the target attitude and i's original one, $y^\star$ and $y_{i,bl}$, the larger the change in i's attitude will be (that is, the absolute difference between y_{bl} and y_{st} will be larger).[1]

After the treatment, suppose i has the chance to return to her normal life before her long-term attitude is formed. We can account for two additional, possible sources of pressure on her attitude there: the normal stream of information coming from her environment in day to day life, and reactions to her (potentially new) attitude by peers in her social network. Let w^e and w^{nw} be the weights that i places on these environmental and network components of pressures on her attitudes, respectively, such that $0 \leq w_i^{nw} \leq 1$, $0 \leq w_i^e \leq 1$, and $w_i^{nw} + w_i^e \leq 1$. Now we can represent her long-term attitude as comprised of:

$$y_{i,lt} = (1 - w_i^{env} - w_i^{nw})\text{NEW ATTITUDE} + (w_i^e)\text{ENVIRONMENT} + (w_i^n)\text{NETWORK}$$

To fill in some pieces, we already have notation to represent i's new attitude in response to the treatment: $y_{i,st}$. To account for the effect of a return to i's environment, suppose whatever information it provides, it continues to reinforced her initial attitude, $y_{i,bl}$. To capture the reaction of i's network neighbors to i's new attitude, let $y_{i,nw}$ stand for some aggregation of the reactions i receives through the village network. Below we consider specific functional forms that this aggregation could take; for now, these pieces are enough to represent i's long-term attitude as:

[1] This setup is a simplification of standard Bayesian learning models, and captures the same intuition about the role of new information as a function of the mean and precision of the prior.

$$y_{i,lt} = (1 - w_i^e - w_i^{nw})y_{i,st} + w_i^e y_{i,bl} + w_i^{nw} y_{i,nw}$$

In this setup, i's long-term attitude is a weighted average of i's new attitude formed in response to the treatment in the short-term, her baseline attitude that her environment continues to reinforce, and the attitude that the reactions of people in i's network suggests is best.

This simple framework leads to a number of insights. First, consistent with standard models of incorporating new information, some people may react more or less strongly to the treatment promoting $y^\star$ in the short-term depending on how different it was from their prior, baseline attitude and how sure they were about that attitude. People for whom $y^\star$ is quite different than their baseline attitude (large $|y^\star - y_{i,bl}|$) and who are quite willing to take in new information (small s_i) will react most strongly to the treatment in the short-term.

Second, when $w_i^{nw} = 0$, i does not take into account the reactions of people in her network. We call such a situation one that exhibits no social processing. If social processing were not present, long-term responses would be constrained to feature either durable attitude change or attenuation. To see this, suppose $w_i^{nw} = 0$. We would have

$$y_{i,lt} = (1 - w_i^e)y_{i,st} + (w_i^e)y_{i,bl}$$

where

$$y_{i,st} = (1 - s_i)y^\star + s_i y_{i,bl}.$$

Now we can more clearly see the options for i's long-term attitude change without social processing. First, suppose she receives no further doses of information from her environment after receiving x (or equivalently, values these at zero), $w_i^e = 0$. Then wherever her attitude landed in the short-term, it will remain in the long-term ($y_{i,lt} = y_{i,st}$). In such a case, we could say the effect of $y^\star$ is **durable**. If instead $w^e \neq 0$ and her environment provided more of the same signals that formed i's baseline attitude in the first place, then her long-term attitude will be pulled closer to her baseline. In other words, $|y_{i,lt} - y_{i,bl}| < |y_{i,st} - y_{i,bl}|$. Long-term change would be smaller than short-term change, classic treatment response **attenuation**. For someone who reacted to the treatment ($y_{i,st} \neq y_{i,bl}$), these are the only two options: effect durability or attenuation.

Claim 1 *If no social processing were present ($w_i^{nw} = 0$), long-term attitude change would feature only effect durability ($y_{i,lt} = y_{i,st}$) or effect attenuation ($|y_{i,lt} - y_{i,bl}| < |y_{i,st} - y_{i,bl}|$).*

B.1.1 Representing Social Processing

The attitude that the network recommends to i, $y_{i,nw}$, is the result of some process of reactions spreading through the network and being aggregated by i. There are lots of ways this could happen. One simple way is that everyone to whom i is directly linked in the network (i's network neighbors) reacts to i's new attitude, expresses that reaction as a function of their baseline attitudes, and i averages these.[2] For instance, we could have

$$y_{i,nw} = \frac{1}{\#N_i} \sum_{j \in N_i} r(y_{j,bl} | y_{i,st}) \tag{1}$$

N_i is the set of i's neighbors in the network and $r(\cdot)$ outputs network neighbor j's reaction to i's short-term attitude as a function of j's own baseline beliefs. A straightforward representation of i's reaction is $r(y_{j,bl}|y_{i,st}) = y_{j,bl}$. Then we would have:

$$y_{i,lt} = (1 - w_i^e - w_i^{nw}) y_{i,st} + w_i^e y_{i,bl} + w_i^{nw} \left(\frac{1}{\#N_i} \sum_{j \in N_i} y_{j,bl} \right)$$

so that i's long-term attitude is a function of her new attitude in the short-term, her baseline attitude that the environment reinforces, and the average baseline attitudes of her network neighbors. It should be clear that even in this simple representation of social processing taking place in the network– where i's neighbors let i know their reaction to her new attitude by conveying their own attitudes– it is no longer the case that effect durability and attenuation are the only options. Depending on the views of her network neighbors, her long-term attitude could be pulled even farther from her baseline ($|y_{i,lt} - y_{i,bl}| > |y_{i,st} - y_{i,bl}|$), **effect acceleration**. Her long-term attitude could even flip directions, moving the long-term change in the opposite direction from her short-term change ($y_{i,lt} - y_{i,bl} > 0$ while $y_{i,st} - y_{i,bl} < 0$, or vice versa), an example of **effect flipping**.

Claim 2 *If social processing were present ($w_i^{nw} > 0$), long-term attitude change could, in addition to featuring effect durability and effect attenuation, feature effect acceleration ($|y_{i,lt} - y_{i,bl}| > |y_{i,st} - y_{i,bl}|$) or effect flipping ($y_{i,lt} - y_{i,bl} > 0$ while $y_{i,st} - y_{i,bl} < 0$, or vice versa).*

The setup above also makes clear that social processing could pull an individual's long-term attitude towards the attitudes of her network neighbors.

Claim 3 *Social processing can result in an individual's long-term attitude moving closer to the average of her network neighborhood's baseline attitudes. That is,* $\left| y_{i,lt} - \left(\frac{1}{\#N_i} \sum_{j \in N_i} y_{j,bl} \right) \right| < \left| y_{i,st} - \left(\frac{1}{\#N_i} \sum_{j \in N_i} y_{j,bl} \right) \right|$.

The above representation of social processing assumes that all neighbors of i give i their reactions. We could say it assumes that i's new attitude activates all of i's neighbors. In a

[2] We could imagine much more complicated forms of neighbors' reactions. For instance, they could react as a function of i's extremity of change, or the difference between their and i's positions at the end of the short-term, etc.

setting where multiple people are treated with $y^\star$, or where i's treatment and new attitude is especially noteworthy, it is possible that more than just i's neighbors are activated. Of course, it is also possible that not every one of i's neighbors are activated. We can stipulate a more general setup that captures the situation of some people within i's network becoming activated to offer their reaction. In this setup, the activated may or may not be directly linked to i. We make the standard network assumption that those farther away from i can still offer reactions that reach i, but with some decayed potency based on how far through the network the reaction would need to travel to reach i.

For this setup, call A the set of individuals in i's community who are activated to offer their reaction. The reaction of an activated $j \in A$ travels through the network and reaches i if there is a finite path from j to i. Call the length of the shortest path between j and i $\ell(j,i)$. If j is i's neighbor, they are connected in a path of length 1, and so $\ell(j,i) = 1$. If j is a neighbor of a neighbor of a neighbor of i's, $\ell(j,i) = 3$. Suppose i averages these reactions, weighting the reactions of those closer in the network more highly than those farther away. Again for simplicity, suppose that their reaction can be captured by their baseline view. We can represent this version of network reaction as

$$y_{i,nw} = \frac{\sum_{j \in A} y_{j,bl} \delta^{\ell(j,i)-1}}{\sum_{j \in A} \delta^{\ell(j,i)-1}}$$

where $0 < \delta \leq 1$ is the decay parameter; for larger δ, reactions of activated people farther away in the network count more heavily. We can see that this is a generalization of the first setup which averages network neighbors. If we assume that all network neighbors are activated, $A = N_i$, and these are the only people in the network activated to give i a reaction, then we would have

$$y_{i,nw} = \frac{\sum_{j \in N_i} y_{j,bl} \delta^{1-1}}{\sum_{j \in N_i} \delta^{1-1}} = \frac{\sum_{j \in N_i} y_{j,bl}}{\# N_i},$$

the same formulation as above. Using this more general statement of aggregated network reactions, we can see that a person's long-term attitude would move towards that of an activated person's in the network, but the movement would be greater if the activated person were closer.

Claim 4 *Social processing can result in an activated person j reacting and influencing i's long-term attitude; that influence will be greater the closer that person is to i in the network (the smaller is $\ell(j,i)$).*

B.1.2 Understanding the control

We can also use the above representation of short-term and long-term attitude formation to understand the reactions of individuals who did not receive the treatment. Consider a person k in the control condition. At the time of treatment, person k was not given $y^\star$. Instead, we can think of person k receiving his normal signal received in his day to day life, which we have been representing as $y_{k,bl}$. Person k's short-term attitude can then be represented as

$$y_{k,st} = (1 - s_k)y_{k,bl} + s_k y_{k,bl} = y_{k,bl},$$

which is to say k does not change attitudes in the short-term. In the long-term, however, whether k's attitude moves is a function of whether network processing is at play, and whether any reactions kicked off in the network reach k. Of course if there is no social processing, then in the long-term we would have

$$y_{k,lt} = (1 - w_k^e)y_{k,bl} + w_k^e y_{k,bl} = y_{k,bl},$$

and again see no change in attitude. Person k in the control would retain his baseline view into the long term.[3].

If social processing is present, though, the long-term attitudes of the control could change if views of the activated reach them. Using the more general representation from above, we would have

$$y_{k,lt} = (1 - w_k^e - w_k^{nw})y_{k,bl} + w_k^e y_{k,bl} + w_k^{nw} \frac{\sum_{j \in A} y_{j,bl} \delta^{\ell(j,k)-1}}{\sum_{j \in A} \delta^{\ell(j,k)-1}}$$

where again δ is the decay as a function of the distance between person k and the activated person reacting, $\ell(j,k)$. This representation makes clear that in the long-term, as long as some people in the network were activated to react, that there is a finite path between at least one person who reacted and k, and k places non-zero weight on the attitude recommended by his network, then k's attitude can change in the long-term even when k was in the control.

Claim 5 *If no social processing were present ($w_k^{nw} = 0$), then the long-term attitudes of the control would not change ($y_{k,lt} = y_{k,bl}$).*

It is also clear that social processing can have the same types of impacts on the attitudes of the control as described above for the treated:

Claim 6 *Social processing can result in long-term attitude change in the control of the same forms as the treated: the long-term attitude of the control can to move closer to the attitudes of his network neighbors and to the attitudes of the activated elsewhere in the network with decaying influence proportional to distance.*

[3]This assumes that the environment continues to be the same as whatever it was when it supported k's baseline views. If the environment changes, then it could output something other than $y_{k,bl}$ to k and change k's attitude (and the same is true for individuals receiving treatment as well).

B.2 Assessing Information Updating Assumptions

	DV: Short-term Change in Pro-refugee Score		
	(All T)	(T < 30)	(T < 28)
Baseline Atts	−0.488***	−0.522***	−0.516***
	(0.037)	(0.044)	(0.051)
Constant	13.511***	14.169***	14.049***
	(0.862)	(0.984)	(1.105)
Observations	289	264	240
R^2	0.374	0.350	0.301
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 1: How the treated responded to the treatment as a function of their baseline views. Biased reactions such as motivated reasoning would appear as a positive relationship between baseline views and the treatment. Instead, we see a negative relationship, and this does not seem to be driven by the scale capped at 30: the second model excludes the treated who start at the cap, the third excludes those who start close to it with scores greater than 27. People who had higher baselines, and so were more aligned with the message of the treatment, did not respond more strongly to the treatment, consistent with Bayesian updating.

C Baseline Attitudes and Individual Response to Treatment

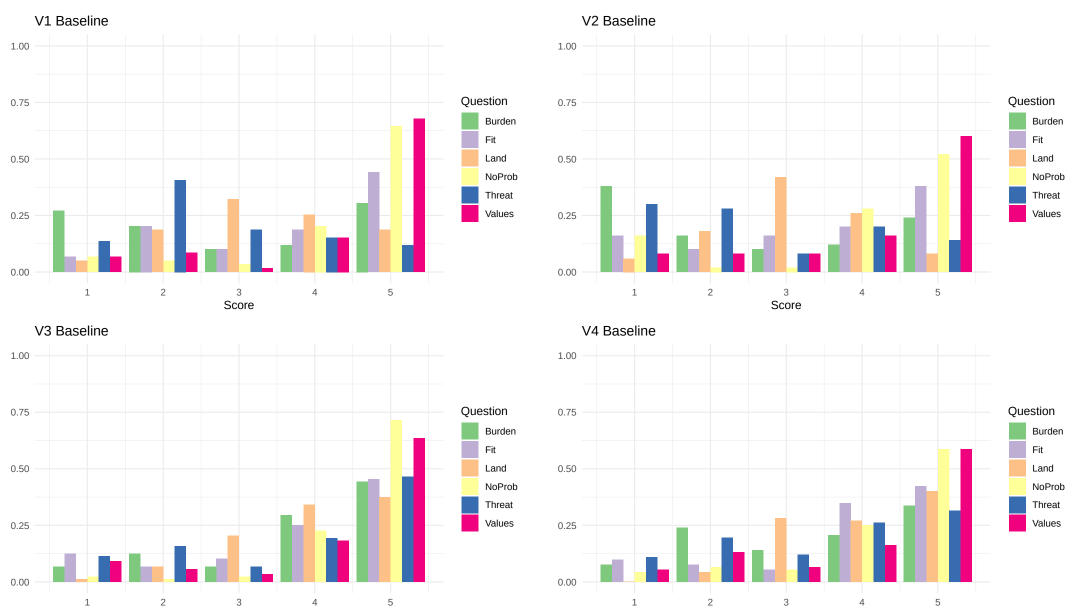

Figure 1: Baseline attitudes for the treated in each of the four villages.

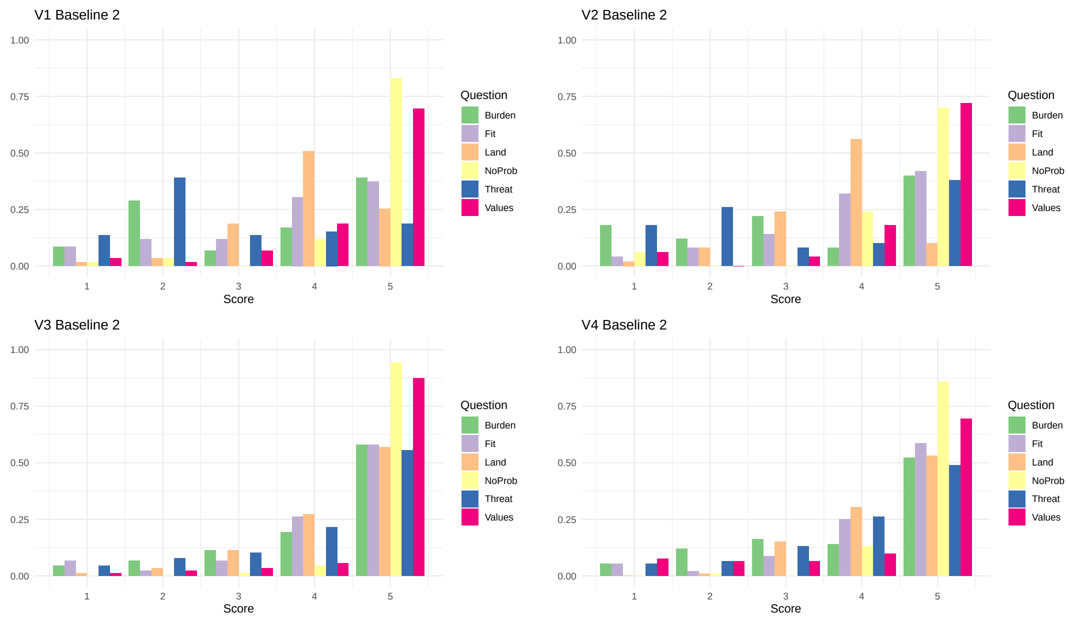

Figure 2: Baseline 2 (post-treatment) attitudes for the treated in each of the four villages.

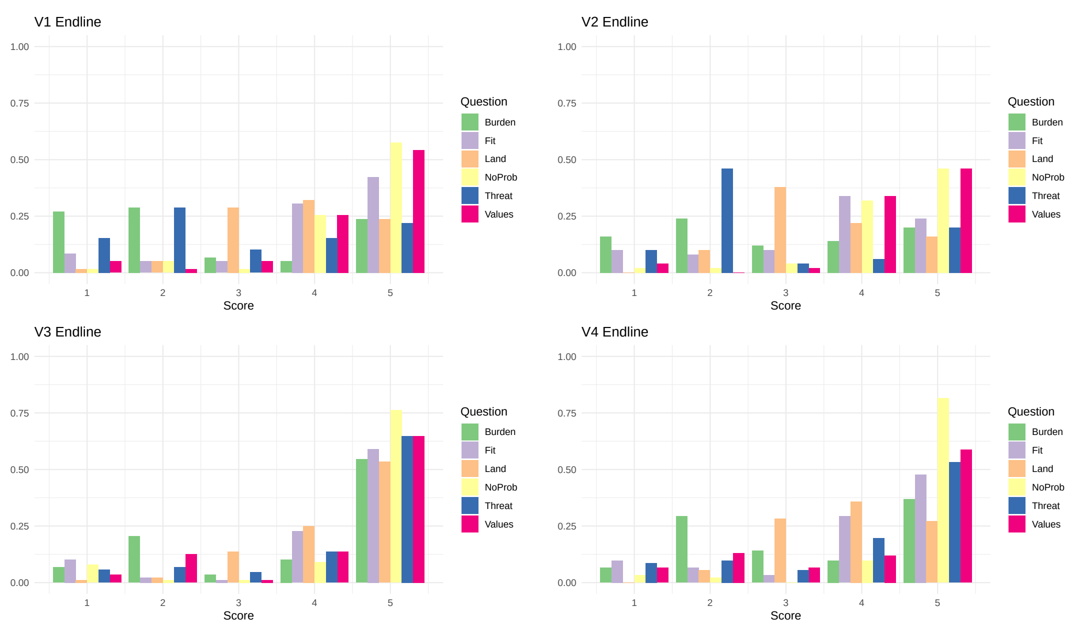

Figure 3: Endline attitudes for the treated in each of the four villages. Notice when comparing across baseline, baseline 2 and endline that neither attitude warming nor attenuation is driven by any one constituent question.

	V1 T	V1 C	V2 T	V2 C	V3 T	V3 C	V4 T	V4 C	All T	All C
Pro-ref bl	21.4	20.2	20.0	19.5	24.3	21.6	23.3	23.6	22.6	21.3
Pro-ref bl2	23.3		23.3		26.7		25.8		25.1	
Pro-ref el	22.2	22.4	21.6	20.5	25.6	24.6	24.1	24.4	23.8	23.1
Short-tm ch.	1.9		3.3		2.3		2.5		2.5	
Long-tm ch.	0.9	2.2	1.6	1.4	1.2	3.0	0.8	1.0	1.1	1.9
Prop s.t.=0	0.24		0.16		0.22		0.18		0.20	
Prop s.t.>0	0.59		0.70		0.65		0.70		0.66	
Prop s.t.<0	0.17		0.14		0.14		0.12		0.14	
Prop l.t.=0	0.09	0.10	0.09	0.00	0.19	0.09	0.13	0.02	0.14	0.06
Prop l.t.>0	0.50	0.68	0.53	0.62	0.46	0.62	0.52	0.66	0.50	0.65
Prop l.t.<0	0.41	0.23	0.37	0.38	0.35	0.29	0.35	0.32	0.36	0.30
n	59	68	50	48	88	58	92	58	289	232

Table 2: Response to treatment by village and treatment condition. Showing the average score in the baseline (bl), post-treatment baseline (bl2), and endline (el); the average difference between bl2 and bl for the treated (short-term change), the average difference between e and bl for everyone (long-term change); and the proportion of each treatment condition that experienced no, positive, or negative change in the short- and long-term.

C.1 Average Treatment Effect

	Control	Treated	p-val	Control	Treated Day 1	p-val
V1	20.22	21.36	0.159	20.22	19.87	0.781
V2	19.48	20.04	0.529	19.48	19.33	0.894
V3	21.55	24.32	0.003	21.55	23.29	0.201
V4	23.64	23.29	0.616	23.64	22.80	0.571
Pooled	21.25	22.65	0.001	21.25	21.15	0.874

Table 3: We see imbalance in the start-of-baseline views of the treated and control. Figure 7 below suggests the reason why: spillovers may have started shortly after some received treatment in the village, which could have raised subsequent baseline scores. We anticipated spillovers from treatment, and so our design randomized households into treatment and control but then completed all control baseline surveys first within each village. Consequently, we have better balance on baseline scores between control and treated if we compare all control to people treated early in the treatment wave, for instance on the first day of the treatment wave (Treated Day 1).

	V1	V2	V3	V4	Pooled
Naive ATE	3.05	3.82	5.11	2.14	3.85
p-value	0.000	0.000	0.000	0.001	0.000
Corrected ATE	1.98	2.92	5.16	2.56	2.91
p-value	0.114	0.005	0.000	0.013	0.000
n (control)	68	48	58	58	232
n (treated)	59	50	88	92	289
n (treated on day 1)	15	15	14	10	54

Table 4: ATE calculated in two ways. The naive ATE compares treatment and control at the end of baseline. This compares the treated's second baseline attitudes measure to the control's only baseline measure. We call this measure naive because it ignores the fact that spillovers from treatment could happen more quickly than we could survey the rest of the treated, during the few days that treatment is rolled out. Since all control were surveyed first within a village, treatment spillovers would effect the measurement of the treated's baseline, and that would be more true the farther into the treatment wave a treated was surveyed. The corrected ATE uses only respondents who were surveyed up to and including the first day in which treatment was administered. This includes all control respondents and the treated whose survey occurred on the first day of the treatment wave. This set of treated should be least susceptible to spillovers raising their baseline scores. Of course, SUTVA is only guaranteed for the within-treated change in response to treatment during the baseline, so we report the more conservative ATT instead of this corrected ATE in the main manuscript.

D Attrition

	V1 Attrit	V1 In	V2 Attrit	V2 In	V3 Attrit	V3 In	V4 Attrit	V4 In
n	11	116	13	85	4	142	5	145
Age	32	35	37	38	28	39	35	40
Protestant	0.27	0.47	0.15	0.12	0.00	0.00	0.00	0.00
Catholic	0.36	0.38	0.85	0.84	0.00	0.00	1.00	0.92
Muslim	0.27	0.13	0.00	0.04	1.00	0.99	0.00	0.06
Farmer	0.73	0.19	0.54	0.48	0.50	0.84	0.60	0.77
Trader	0.00	0.24	0.08	0.27	0.25	0.04	0.00	0.14
No Educ	0.00	0.03	0.15	0.06	0.25	0.33	0.20	0.12
Primary Educ	0.64	0.26	0.46	0.61	0.50	0.54	0.60	0.68
Secondary Educ	0.09	0.28	0.15	0.18	0.25	0.11	0.20	0.12
College Educ	0.27	0.42	0.23	0.15	0.00	0.02	0.00	0.08
Lived > 5yrs	0.73	0.63	0.54	0.76	1.00	0.80	0.80	0.83
Baseline 1 Proref	20.8	20.7	21.0	19.6	22.2	23.2	25.2	23.4
Baseline 2 Proref	24.2	23.2	23.6	23.3	27.0	26.6	23.0	25.9
Treated	0.45	0.47	0.54	0.51	1.00	0.59	0.60	0.61

Table 5: Comparing attritted from those who stayed in. It appears that those who left did so as if at random, at least with respect to demographics, treatment status, baseline attitudes, and response to treatment.

	Attritted	Stayed	p-vals
n	33	488	
Treated Neighbs	0.8	0.9	0.14
# Neighbs	6.3	7.8	0.13
Neighbs Bl Atts	21.2	22.0	0.11
Dist to Warmest	2.0	1.8	0.16
Dist to Coldest	2.3	2.1	0.21
Dist to Persuaded	2.2	2.2	0.80
Dist to Backlashed	2.1	2.0	0.88

Table 6: No significant differences between respondents who attritted and those who stayed.

E Social Processing

E.1 Network Differences

	V1	V2	V3	V4
AbsDifAvg1	5.48	4.58	5.92	4.75
AbsDifAvg_e	5.34	4.54	5.51	4.66

Table 7: Placebo test for network differences. Here, the true endline score is replaced with a simulated level shift equal to the mean village change. This placebo uses real baseline data. Endline proref scores are replaced with simulated ones that naively guess everyone has a level shift equal to their village's mean change (1.56, 1.52, 1.93, .88 respectively). Those who would be above the index ceiling of 30 have their score replaced with 30. Hitting the cap appears to be a small part of the increase in network similarity in V1 - V3. It could be most of the increase in V4, where lots of people were closer to the cap to start.

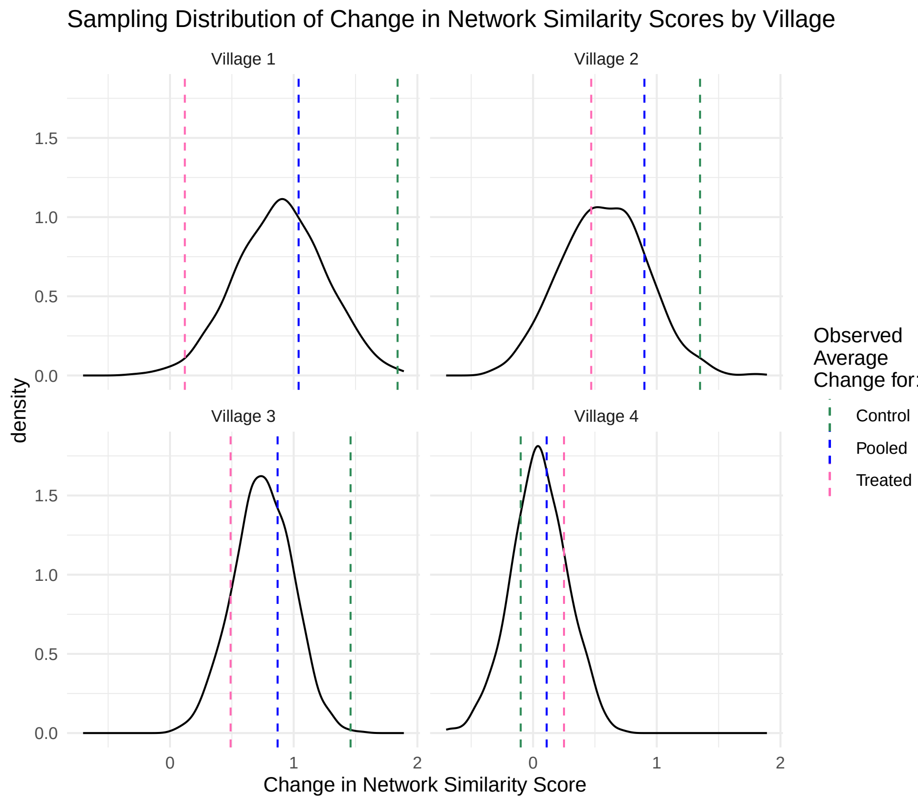

Figure 4: Sampling distribution of the change in network similarity score when the pairs of (baseline, endline) scores are randomly shuffled in our observed networks for villages 1 through 4. Specifically, these sampling distributions are constructed in the following way for each village: take the network and the treatment assignment as given. Randomly assign the pairs of observed baseline, endline scores in the network to new, randomly chosen nodes. That is, no baseline, endline pairs are broken, but the node to which they are assigned is selected at random. For each new attitude pair assignment, calculate the network similarity in the baseline, the network similarity in the endline, and the difference between the two. Repeat 1000 times. In villages 1-3, the control become much more similar to their network neighbors than would be expected by chance. Implied p-values: .018, .068, and .031, respectively. The same does not hold in village 4. (Of course this could be an artifact of some baselines being measured over a month later there, as does indeed appear to be the case in the next figure.)

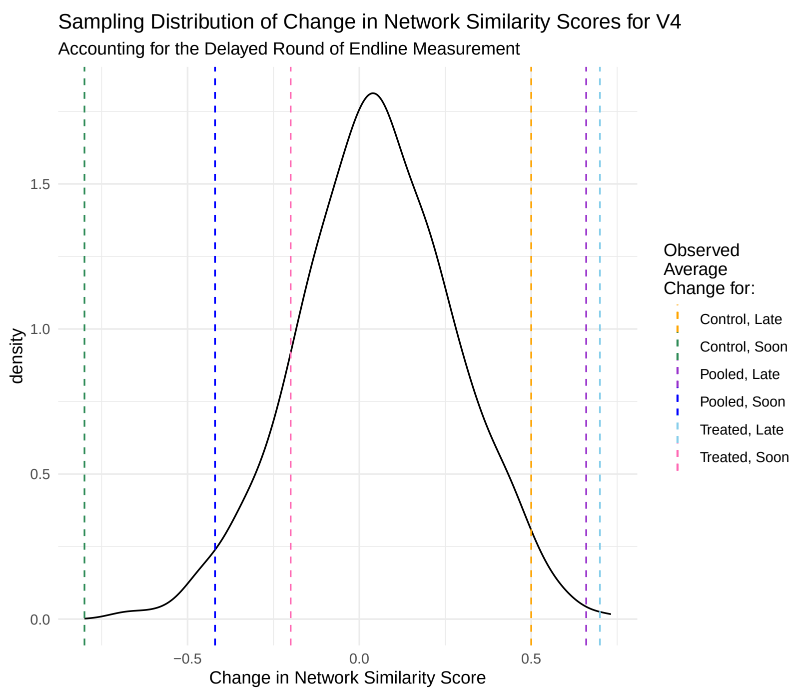

Figure 5: Indeed, in village 4, there is a large difference in the network similarity gains of those who were measured on schedule and those measured much later. Given more time, respondents, especially the control ones, adopted attitudes much more similar to their network neighbors.

E.2 Network Distance Measures

Our network distance measures are calculated as the length of the shortest path in the village network between each household and a reference set of households. We use four reference sets: the households of the respondents with the warmest baseline attitudes, the coldest baseline attitudes, the largest positive change in response to treatment, and the largest backlash to treatment measured as having the coldest second baseline score after treatment among those who reacted negatively to treatment.

Each village has its own reference sets. For each village and each reference set, we selected the cutoff value of the relevant variable such that at least three villagers were included. Then we accepted into the set all villagers with a value that extreme, which often resulted in more than three in the top set due to ties. For instance, in village 4, the third warmest baseline score was the max of 30. Collecting everyone who had a baseline of 30 into the reference set led to 9 members of village 4 serving as the reference set. In village 2, the third warmest score was 27. Collecting everyone with a score at least as large as 27 led to just the 3 included in the reference set.

Once the reference sets were established in each village, the minimum distance between each household and any member of this set is stored as the value of the distance variable. The maximum (finite) distance any household is to their nearest member of the reference set is 5. Villagers who are in their own component (only relevant for villages 1 and 2) are dropped because their paths to any member of the reference set are infinitely long. The values of the resulting distance variables become:

	0	1	2	3	4	5
Dist to Warmest	33	159	210	80	16	2
Dist to Coldest	14	82	258	123	21	0
Dist to Persuaded	16	87	219	152	26	1
Dist to Backlash	15	99	260	104	18	4

Table 8: Distribution of network distance variable values. Those with 0 distance are the reference set.

The observations with distance 0 are the members of the reference set. For the analyses above, we confirm that the results are not sensitive to leaving in the reference set. When we drop them from the full analyses, the results hold. The results also hold when we add an indicator variable for membership in the reference set, which allows us to distinguish the role of the distance 0-s (being a person with the warmest score, say) and the distance to them.

E.3 Robustness Checks for Regression Analysis

	Vlg 1	Vlg 2	Vlg 3	Vlg 4	All
Treated	0.46	0.51	0.60	0.61	0.55
# Treated Neighbs	2.18	3.11	6.45	4.06	4.09
Treated Neighbs	0.74	0.84	0.99	0.97	0.89
Baseline Atts	20.75	19.77	23.22	23.43	22.03
Neighbs Bl Atts	21.25	19.84	22.43	23.50	21.99
Dist to Warmest	2.04	2.21	1.26	1.85	1.79
Dist to Coldest	2.05	2.04	1.78	2.55	2.12
Dist to Persuaded	2.72	2.11	1.91	2.08	2.18
Dist to Backlashed	2.55	1.65	1.75	2.22	2.05
Warmest	3	3	18	9	33
Coldest	5	3	3	3	14
Most Persuaded	3	3	4	6	16
Most Backlash	3	4	4	4	15
Baseline hhs	127	98	146	150	521
Endline hhs	116	85	142	145	488

Table 9: Summaries of main variables in regression analyses. Mean values of: treatment status; a count of the number of treated neighbors; an indicator for having at least one treated neighbor; baseline pro-refugee score; mean neighborhood pro-refugee scores; network distance to one of the warmest households at baseline, one of the coldest households at baseline, one of the treated households most persuaded by treatment, andone of the treated households that responded most negatively to treatment. Also displays count of the reference households to which distance variables are calculated in the village, as well as the number of households in baseline and endline.

	DV: Endline Pro-Refugee Score				
	(1)	(2)	(3)	(4)	(5)
Treatment	4.403	4.163	3.926	4.533	4.593
	(3.258)	(3.274)	(3.255)	(3.266)	(3.213)
# Treated Neighbs	0.205	0.240	0.210	0.257	0.201
	(0.188)	(0.188)	(0.187)	(0.188)	(0.185)
# Neighbs	−0.133	−0.138	−0.144	−0.109	−0.135
	(0.114)	(0.115)	(0.114)	(0.114)	(0.113)
Baseline Atts	0.347***	0.376***	0.379***	0.364***	0.341***
	(0.039)	(0.039)	(0.038)	(0.038)	(0.040)
Neighbs Bl Atts	0.256**	0.314***	0.306***	0.298***	0.232**
	(0.109)	(0.109)	(0.108)	(0.108)	(0.109)
Dist to Warmest	−0.578**				−0.846***
	(0.236)				(0.260)
Dist to Coldest		−0.300			0.197
		(0.259)			(0.283)
Dist to Persuaded			−0.634***		−0.751***
			(0.235)		(0.243)
Dist to Backlashed				0.486**	0.788***
				(0.236)	(0.243)
Trt * # Treated Neighbs	−0.062	−0.063	−0.067	−0.108	−0.103
	(0.259)	(0.260)	(0.258)	(0.260)	(0.255)
Trt * # Neighbs	0.066	0.075	0.078	0.102	0.082
	(0.154)	(0.155)	(0.153)	(0.154)	(0.152)
Trt * Neighbs Bl Atts	−0.212	−0.199	−0.196	−0.213	−0.221
	(0.147)	(0.147)	(0.146)	(0.147)	(0.144)
Constant	11.508***	9.024***	10.149***	7.657***	12.325***
	(2.682)	(2.465)	(2.493)	(2.514)	(2.713)
Observations	470	470	470	470	470
R^2	0.216	0.208	0.218	0.213	0.245
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 10: Replicating the analyses in the paper using a count of the number of treated neighbors instead of an indicator for the presence of at least one treated neighbor. The social processing results remain the same.

	DV: Endline Pro-Refugee Score
Treatment	6.207* (3.207)
Treated Neighbs	−0.508 (1.067)
# Neighbs	−0.029 (0.044)
Baseline Atts	0.364*** (0.047)
Neighb Bl Atts	0.263** (0.113)
Warmest	−0.263 (0.964)
Coldest	−0.464 (1.381)
Most Persuaded	2.357* (1.302)
Most Backlash	−3.823*** (1.201)
Dist to Warmest	−0.811*** (0.307)
Dist to Coldest	0.118 (0.322)
Dist to Persuaded	−0.498* (0.281)
Dist to Backlashed	0.320 (0.278)
Trt * Treated Neighbors	−1.567 (1.516)
Trt * # Neighbs	0.039 (0.058)
Trt * Neighbs Bl Atts	−0.229 (0.143)
Constant	12.163*** (2.789)
Observations	470
R^2	0.270
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 11: Results continue to hold when controlling for membership in the reference categories from which the distance measures are constructed. By adding an indicator variable for being the reference households, we can better distinguish whether it matters to be close to the most negative household or to *be* the most negative household. Variables `Warmest` and `Coldest` are indicators for the respondents who have the warmest and coldest baseline scores (and to whom the distances in `Dist to Warmest` and `Dist to Coldest` are calculated). `Most Persuaded` and `Most Backlash` are indicators for respondents who responded most warmly and most coldly to the treatment at the end of the baseline. Distances to individuals in the reference categories is not driven by the individuals who themselves are the referents. Moreover, it suggests that for those who react most strongly to the treatment in the short-term, the response is durable. Those who reacted most positively to the treatment in the short-term have higher endline scores and those who reacted most negatively to the treatment have lower scores. Placebo test in paper replicates this analysis.

	DV: Endline Pro-Refugee Score
Treatment	5.580* (3.205)
Neighb Treat	−0.027 (1.064)
# Neighbs	−0.046 (0.043)
Baseline Atts	0.326*** (0.041)
Neighbs Bl Atts	0.233** (0.111)
Dist to Warmest	−0.765*** (0.260)
Dist to Coldest	0.336 (0.283)
Dist to Persuaded	−0.892*** (0.247)
Dist to Backlashed	0.837*** (0.248)
Age	0.002 (0.016)
Muslim	0.523 (1.654)
Catholic	−0.980 (1.629)
Protestant	−1.509 (1.717)
Some Primary	−0.946 (0.580)
Some Secondary	−1.406* (0.741)
Some College	−0.131 (0.846)
Farmer	−0.550 (0.451)
Male	−0.192 (0.424)
Been a Refugee	0.146 (0.501)
Lived > 5	−0.369 (0.505)
Trt * Neighb Treat	−1.984 (1.518)
Trt * # Neighbs	0.044 (0.059)
Trt * Neighbs Bl Atts	−0.188 (0.144)
Constant	14.392*** (3.275)
Observations	470
R^2	0.290
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 12: Main analysis with demographic controls added. The network results are unchanged. Omitted category for religion is "Other." Omitted category for education is "None."

	DV: Endline pro-refugee score				
	(1)	(2)	(3)	(4)	(5)
Treatment		0.112		−0.052	−0.062
		(0.380)		(0.384)	(0.383)
Treated Neighbs		−0.546		−0.854	−1.075
		(0.753)		(0.763)	(0.772)
# Neigbs		0.037		−0.006	−0.008
		(0.030)		(0.035)	(0.035)
Baseline Atts	0.371***	0.370***	0.337***	0.339***	0.343***
	(0.038)	(0.038)	(0.040)	(0.040)	(0.040)
Neighbs Bl Atts	0.214***	0.218***			0.141*
	(0.071)	(0.073)			(0.079)
Dist to Warmest			−0.990***	−1.052***	−0.893***
			(0.232)	(0.244)	(0.259)
Dist to Coldest			0.358	0.357	0.197
			(0.258)	(0.265)	(0.279)
Dist to Persuaded			−0.771***	−0.828***	−0.816***
			(0.230)	(0.245)	(0.244)
Dist to Backlashed			0.829***	0.803***	0.748***
			(0.234)	(0.241)	(0.242)
Constant	10.638***	10.713***	17.093***	18.203***	15.375***
	(1.680)	(1.709)	(1.054)	(1.519)	(2.194)
Observations	474	474	470	470	470
R^2	0.198	0.201	0.234	0.237	0.242
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 13: The variables used to assess social processing– the neighborhood attitude measure Neighbs Bl Atts and the four Dist measures– are robust to being included with the other social processing variables, and their relationship to endline scores remains unchanged regardless of whether information about treatment is included.

F Assessing the Possibility of Exogenous Changes to Attitudes During Our Study

Our design relies on an assumption that individuals in the treatment and control group's attitudes towards refugees warmed because of our intervention, and not because of other unobserved factors. One technique to probe the relevance of such factors could be to also survey "pure control" individuals living outside of our study villages. However, those living in villages proximate to our study villages could plausibly receive word of our treatment due to frequent (usually weekly) gatherings of residents from numerous clusters of villages on local market days. Moving beyond the catchment area of a market raises the risk of sampling from a context that could be influenced by a distinct set of external (unrelated to the study) events that could shift average, village-level attitudes.

Despite extensive focus on refugee issues in Uganda over several years, we know of no reason to expect a general time trend of attitudes warming towards refugees. Beyond that possibility, there could be two types of sources of attitudes shift in our villages other than our intervention: (1) region-wide shifts, which could be driven by major policy changes (on the part of the Ugandan government or major service providers such as UNHCR) or conspicuous, newsworthy acts such as a refugee publicly assisting (or harming) a host community member, or (2) smaller, village-specific (or cluster of local village-specific) shifts that created a secular warming of attitudes.

Our evidence against the occurrence of shifts of the first type comes from basic news media. Using the Factiva database, we conducted a search on *The Daily Monitor*, which is the more independent of the two major, widely-read newspapers in Uganda. A search on "refugee*" starting in the month prior to study until the end of our study (January 1 to August 31, 2021) yields 130 articles. None report an instance of major refugee-host cooperative acts nor any refugee-host violence in Uganda.[4] Nor do any articles report events in West Nile or policy shifts there during this period that could plausibly shift Ugandan attitudes towards refugees. No two (or more) articles focus on the same event or issue – except for several articles on the arrival of Afghan refugees in Uganda. Based on this reporting and our other local sources of information (described below), these Afghan refugees did not come to West Nile or become a salient issue there. In sum, these news articles indicate that there was no major focal point issue related to South Sudanese refugees in West Nile during our study.

Our evidence against trends of the second type comes from the fact that our research team in Uganda was in consistent communication with local sources throughout the study and conducted extensive qualitative follow up in our study villages, as described in the main paper, end of Section 6. Specifically, our enumeration team leader engaged with village leaders before our baseline and endline surveys were administered in their villages, and these

[4]There were five instances of Ugandan citizens killed in South Sudan in "ambushes" during this period: two instances in March, two instances in April, and one in August. Between four and nine Ugandans were killed in each instance; the home district or region of the victims is not specified. We know of no reason that these events would have influenced Ugandans' attitudes towards South Sudanese refugees, despite having probed about the relevance of such events with local interlocutors during the study. See Franklin Draku "Ugandan Nuns Killed in South Sudan Ambush" *The Daily Monitor.* August 17, 2021.

conversations specifically probed about, but did not unearth, events related to refugees that could plausibly influence study results. The study team also regularly consulted with NGO leaders and other local residents during the study with the same result. Our qualitative follow up inquired about villagers' reflections on the study results, and did not turn up evidence of external events driving the warming of attitudes we see in the data.

Our data also allow us to perform a limited investigation to corroborate the qualitative evidence. The study launched on different dates in each of the four villages, and in two waves, with villages 1 and 2 studied in March and then villages 3 and 4 studied in May and early June.

Figure 6 displays the baseline views of survey respondents organized by the day they were surveyed. In each pair of villages, we can see baseline views over 16 and 11 days of elapsed time. We can use these plots to see if there is evidence for a dramatic warming of attitudes in the area over time by looking at a linear smoother overlaid. Neither displays a significant time trend.

Figure 7 shows the same information, this time with a more flexible, Loess smoother overlaid. We can use these figures to look for evidence of an abrupt change in attitudes that could be a response to some external shock common to these villages. In each pair of villages there are two minor upticks in the trend of baseline views. We overlay dashed vertical lines which indicate the date on which the baseline surveys started to include treatment.[5] We see that in all four villages, the uptick in baseline views corresponds to the date that individuals started receiving the treatment in the village. To the extent that a shock boosted attitudes in this timeframe, it appears the shock was this study's treatment within the villages. This is further evidence that spillovers were in fact present in our study.

[5]Because we expected spillovers, we surveyed respondents randomized into the control condition before surveying the treated.

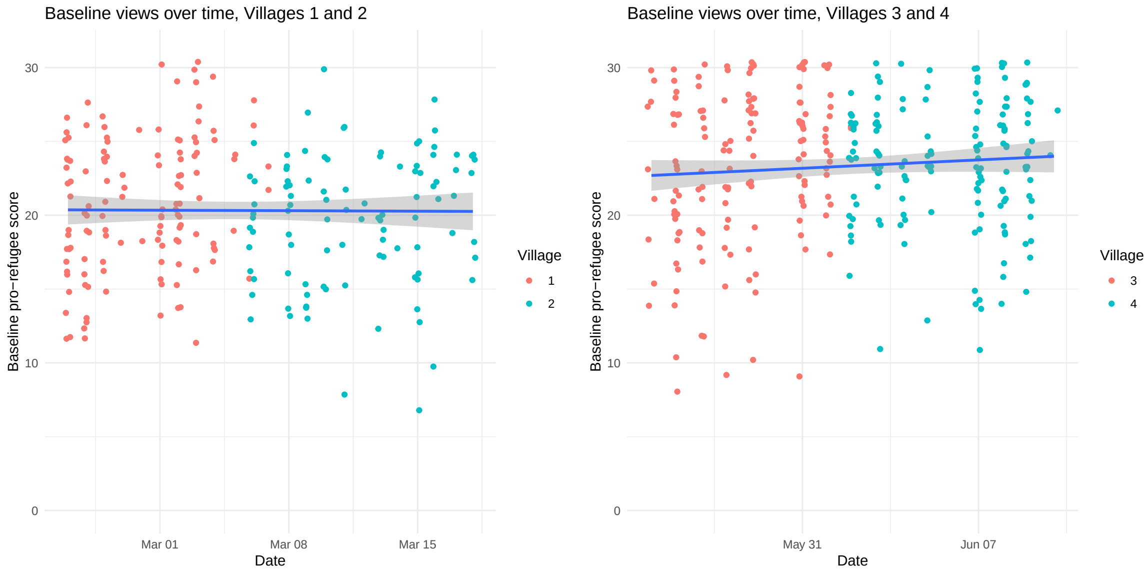

Figure 6: Baseline views in both pairs of villages on the date the views were surveyed with linear trend line. Neither pair shows signs of a significant upward trajectory occurring outside of our study.

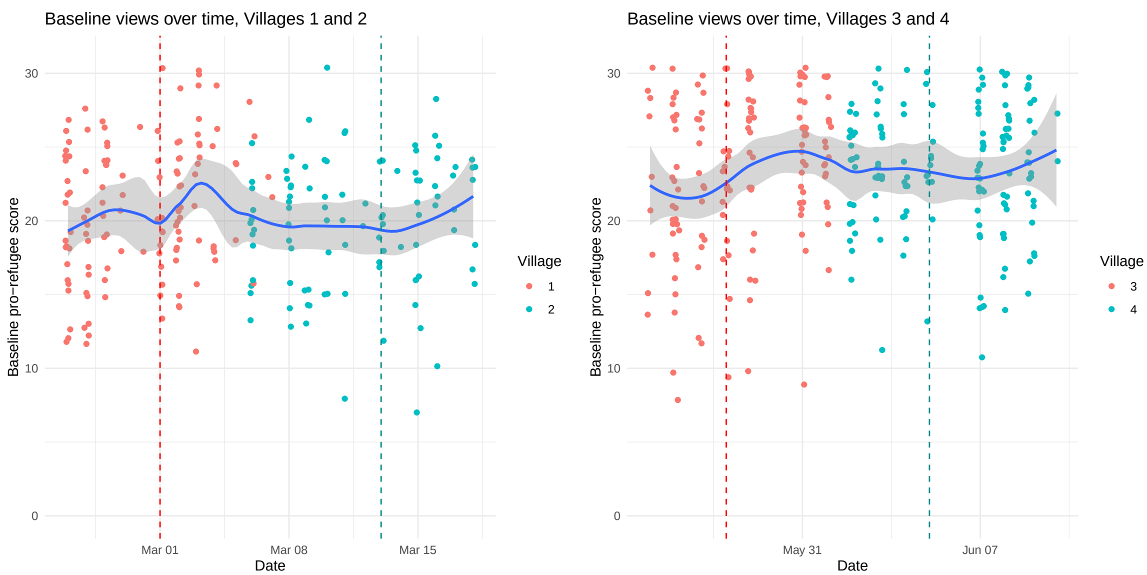

Figure 7: Baseline views in both pairs of villages on the date the views were surveyed, summarized with a flexible Loess smoother. Dashed vertical lines indicate date that treatment began in the village. To the extent that the trend in baseline attitudes shows positive jumps, these track the start of treatment in the village. This combined with the qualitative evidence support the claim that attitudes– even possibly baseline ones for individuals who experienced spillovers from the treated before they were themselves surveyed– were responding to the study and not to some external event.