THE LANCET

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Kim DA, Hwong AR, Stafford D, et al. Social network targeting to maximise population behaviour change: a cluster randomised controlled trial. *Lancet* 2015; published online May 5. http://dx.doi.org/10.1016/S0140-6736(15)60095-2.

WEB EXTRA MATERIAL

Social network targeting to maximise population behaviour change: a cluster randomised controlled trial

David A. Kim, M.D. Candidate

Alison R. Hwong, M.D. Candidate

Derek Stafford, Ph.D. Candidate

D. Alex Hughes, Ph.D. Candidate

A. James O'Malley, Ph.D.

James H. Fowler, Ph.D.

Nicholas A. Christakis, M.D.

Name Generators Used to Identify Social Contacts

We created a photographic census of all adults in each village and used it, with survey software we developed (publicly available in 2015), to map the social networks. Each adult was asked to identify their friends, siblings, and spouse, using the name generator questions below, and to confirm the tie by viewing a photo of the person they named. The name generators we used were:

1. Who are your brothers and sisters that you are friends with?
2. Who are your best friends that are not your brothers and sisters?
3. Who are you married to, or whom are you living with as a husband or wife?

Due to the (deliberate) phrasing with which we elicited sibling ties, not all sets of siblings are fully inter-connected (either in the graphs shown in Figure 1 or in the statistical analyses performed).

Baseline Survey Results

Prior to our intervention, we surveyed all participants regarding their practices with respect to water purification and multivitamin use (see Table S1). The mean village-level rate of water purification at baseline was 0·41 (SD=0·26). The mean village-level rate of daily multivitamin use at baseline was 0·11 (SD=0·05). Due to the unreliability of income and educational data in the region, we also asked six community health workers familiar with the villages to independently rate the socioeconomic status of each village on a 1-10 scale, and we averaged their ratings to create a score for each village; we used this average in the village-level blocking procedure described below. The resulting mean socioeconomic status (SES) rating was 4·4 (SD=2·1). Finally, the mean number of villagers with network data in each village was 146 (SD=108).

Given the limited range of multivitamin use at baseline, we chose to use water purification practices, village SES, and network size to block the villages for treatment assignment.

Choice of Interventions

We chose to study the diffusion of both multivitamins for micronutrient deficiencies and chlorine for water purification because:

1. Diarrheal illness and malnutrition account for a sizeable burden of disease in rural Honduras;
2. We conducted a needs assessment in advance of our trial in which community leaders identified multivitamins for micronutrient deficiencies and chlorine for water purification as useful to their communities, and,
3. We anticipated, based on our pre-intervention assessments and on relevant sociological theory (e.g., on simple vs. complex contagion),[1,2] that the two interventions might diffuse differently owing to their different behavioural and normative demands, thus allowing us to test for differential effects of network targeting methods on two types of interventions, rather than one.

We account for the presence of simultaneous interventions in certain villages in both the design and analysis of the study, as described in several sections below (*Blocking of Villages and Randomisation, Ticket Redemption Models, Models of Knowledge Diffusion)*.

Multivitamin Composition

The multivitamin tablets used in the study (as well as the additional multivitamins donated to all villages after the study's completion) were formulated by Tishcon Corporation as follows, and distributed in 60-count bottles:

Vitamin A 5000IU

Vitamin C 120mg
Vitamin D_3 400IU
Vitamin E 10IU
Thiamin (B_1) 1.5mg
Riboflavin (B_2) 1.7mg
Niacin 20mg
Vitamin B_6 2mg
Folic Acid 400mcg
Vitamin B_{12} 50mcg
Biotin 300mcg
Pantothenic Acid 5mg
Calcium 250mg
Iron 18mg

Blocking of Villages and Randomisation

We used the results from a factor analysis with varimax rotation to form a composite score that explained most of the variance in network size, village-level socioeconomic status, and baseline rates of water purification. The composite score can be considered a weighted average of the three constituent variables with weights reflecting the extent to which each variable accounts for variation in the composite score.

We then assigned each of the 32 villages to eight blocks that minimised the ratio of within-block to between-block variance in the distribution of the composite score. After blocking, we randomised each village to one of four targeting methods (random, indegree, nominated, or none) for the multivitamin and chlorine interventions according to the overlapping fractional-factorial designs outlined below and illustrated in Table S2. By randomising villages to targeting strategies within each block, the between-block component of variation is removed from the error term, yielding more precise results. Even under our multilevel analytic models with village random effects, the randomised block design is advantageous, since adjusting for block (as we do in the models reported in Tables S4-S6) reduces the between-village variance component.

The best way to guard against possible interference of one intervention (multivitamins or chlorine) on the other would be to perform a factorial design in which at least one village (and ideally more) received each combination of targeting methods for the two interventions (random targeting, indegree targeting, nomination targeting, or no intervention). Because we had two interventions (multivitamins and chlorine), each with three different non-null targeting mechanisms, there are $2^6 = 64$ possible arrangements of targeting mechanisms. With 32 villages, we could use only 32 of these; hence, our use of the term "fractional factorial." When including null levels of interventions (to enable estimation of interference effects or interactions of the two products), there are $2^8 = 256$ possible levels. Furthermore, we wanted to prioritize the non-null interventions as well as block on the known, observed village level-factors we thought would be related to outcomes. Therefore, as a compromise, we assigned null treatment to one village for *one* of the interventions in six blocks, and, in two additional blocks, we assigned one village to null treatment for *both* interventions (Table S2). The last two villages (i.e., those receiving no interventions) provide a potential baseline against which to evaluate the effect of a single intervention in the other six villages. The six villages receiving only one intervention, meanwhile, were assigned to targeting methods (random, indegree or nomination) at the same rates (i.e., 1/3 each) as villages receiving both interventions, and thus yield information about treatment effects that is free of any potential interference between interventions, allowing a straightforward village-level interference test to be performed.

For the multivitamin (MVI) intervention (for which we report positive effects of nomination targeting on ticket redemption and knowledge scores), the mean village-level MVI ticket redemption rate is statistically indistinguishable between those villages that received both MVI and chlorine interventions (on average, 73.1% of these villages' available MVI tickets were redeemed), and those villages that received the MVI intervention alone (on average, 73.8% of these villages' available MVI tickets were redeemed; note that these are *village*-level redemption rates, as distinct from the pooled proportions depicted in manuscript Figure 3).

In our individual-level models, we account for potential interference in much more detail, as described below.

In Table S2, Villages 1 to 4 refer to different villages across blocks but the same villages across interventions (multivitamins and chlorine). Thus, for each of the two interventions separately, nine of the 32 villages were targeted randomly, nine by highest indegree, and nine by the nominated friends technique. Six villages received only one intervention, and two villages received neither intervention.

Selection of 5% Target Groups

Table S3 shows each village's assignment to one of the eight blocks, and to one of the four targeting methods for each intervention.

In each village, 5% of the villagers for whom network data were available were selected as targets. Based on past experience, we estimated *a priori* that a 5% targeting rate provided the right balance between sufficient exposure of the villages to the tickets (i.e., the novel stimulus) on the one hand, and avoidance of immediate saturation on the other. And, empirically, our 5% targeting rate produced neither failure of the tickets to diffuse nor complete saturation in any of the study villages, resulting instead in considerable temporal and between-village variation in ticket redemption (Figure 3, main text). The observed variation allowed us to evaluate the differential effects of the three targeting methods, which we do both at the village-level (as described in the *Results* section of the main text) as well as in mixed-effects Cox models controlling for both individual-level and village-level characteristics, and testing for possible interference between the two interventions in villages receiving both (models reported in Tables S4 and S5).

In all cases, targets were drawn from the pool of adult villagers on whom network data had been collected (that is, we did a full census of all adults in every village, mapped the whole network of each village as completely as possible, and then chose subjects from within this population, based on the algorithms below).

1. For randomly targeted villages, we drew a simple 5% random sample of these villagers.

2. For indegree-targeted villages, we targeted the 5% of villagers who had been named as a friend, sibling or spouse most often by others in their village. If tied values of indegree yielded more than 5% of the villagers, we randomly selected individuals who were tied, up to the 5% threshold. For instance, in a village with 100 eligible targets, the highest indegree individuals might have had indegree values of (i.e., may have received friendship/siblinghood/spouse nominations numbering): *15, 13, 12, 10, 10, 10, 10, 9*, etc. Since a 5% target group requires 5 individuals (of which three will be those with indegree values of 15, 13 and 12), how did we select among the four individuals with indegree values of 10? In this case, we would randomly draw two of the four individuals with indegree values of 10 to compose the target group (along with those individuals with values of 15, 13 and 12).

3. In the nominated-friends-targeted villages, we selected 5% of the population *at random*, using the same strategy used for the randomly chosen targets. However, instead of targeting these randomly selected individuals, we selected (again, *at random*) *one friend named by each of these individuals*, and used those *friends* as targets. Notably, because all participants in all villages were asked to nominate friends, targeted individuals in nomination-targeted villages were unaware they had been nominated, such that the effect of the targeting method reflects the structural positions of nominated friends alone, and not any additional psychosocial or "priming" effects of knowing that one had been nominated.

For villages receiving both interventions, we used a different targeting method for each intervention, and target groups were generated independently, such that 26 villagers (across all villages) were selected as targets for both interventions, by chance.

Choice of Outcome Measures

We chose ticket redemption as our *primary measure* of behaviour because it was the most accurately and comprehensively recorded measure of product uptake (i.e., we know the identity without exception of every

individual who redeemed a ticket, and the exact date on which he or she did so), and because it allowed us to trace with the greatest social and temporal resolution the rate and extent of product diffusion through the village networks, without relying on participants' recollection or self-report.

We supplemented this "hard" behavioural measure (ticket redemption) with self-reports of knowledge and practice as well, conducting an extensive follow-up survey in all villages in which we asked villagers (whether or not they had redeemed a ticket) about their use of the products, their attitudes concerning the products' utility and effectiveness, and a series of factual questions about their correct usage and benefits (from which the knowledge scores, our *secondary outcomes*, were derived, as reported in *Models of Knowledge Diffusion,* below).

Conditional on ticket redemption, there was no significant variation in self-reported product use or belief in the products' effectiveness by targeting method (random, indegree, or friend nomination). Self-reported continued product use among confirmed ticket redeemers was uniformly high (> 90%). And high multivitamin knowledge scores, as we report, were more common in friendship nomination-targeted villages.

These results suggest that ticket redemption (in which we *did* observe significant variation by targeting method, both at the population level, as well as in mixed-effects models accounting for individual-level and village-level covariates and for possible interference between interventions) is indeed a *valid omnibus measure of product adoption and continued use*, while the knowledge scores we report provide additional information about the differential spread of *health information* not predicted by product use alone.

Population-level Comparisons

As described in the main text, all pairwise differences in multivitamin redemption rates are statistically significant ($p < 0{\cdot}01$) after correction for multiple comparisons. The specific, pairwise, Holm-corrected p-values are as follows:

1. Nomination (951/1280 = 74·3% redeemed) vs. Random (940/1420 = 66·2% redeemed): $p < 0.0001$
2. Nomination (951/1280 = 74·3% redeemed) vs. Indegree (744/1220 = 61·0% redeemed): $p < 0.0001$
3. Random (940/1420 = 66·2% redeemed) vs. Indegree (744/1220 = 61·0% redeemed): $p = 0{\cdot}0062$

Ticket Redemption Models

For the multivitamin (MVI) intervention, the ticket-level model (Table S4) was estimated as follows. Of the 32 villages, 27 received MVI (24 of these received chlorine as well). 200 MVI targets were chosen across these 27 villages, resulting in 4,000 possible MVI tickets which could be redeemed (200 x 4 = 800 first-wave tickets, 800 x 4 = 3,200 second-wave tickets). Of these 4,000 tickets, 66% were redeemed by eligible participants. 52% of MVI ticket redeemers were also recorded as having redeemed a chlorine ticket. All dates are absolute (rather than zeroed at day of intervention) to control for day-of-week and other secular temporal effects.

Since MVI ticket redemption extended to 15 days after the introduction of the interventions in a given village, each unredeemed MVI ticket received 15 observations (ticket-days), for a total of 20,445 ticket-days for the 1,363 MVI tickets that remained unredeemed by the end of the study. For redeemed tickets, the number of ticket-days per ticket is equal to the number of days from initial targeting through redemption of that ticket.

We controlled for interference between the two interventions by including a ticket redeemer's redemption of a ticket for the opposite intervention as a time-varying covariate: an indicator for chlorine ticket redemption was set to 1 for any ticket-days on or after which the ultimate redeemer of the MVI ticket had redeemed a chlorine ticket. We include as ticket-level covariates the demographics (sex, age, household size, and marital status) of the targets associated with each ticket.

We used separate Cox proportional hazards models for first- and second-wave tickets to estimate the effect of targeting method on time-to-MVI-redemption, controlling for a subject's redemption of a ticket for the other intervention, as well as for basic demographics of the initial targets.

We did not include first-wave redeemer random effects because this would entail conditioning on post-treatment variables (i.e., the identity of the first-wave redeemers) and thus induce post-treatment bias.[3] Instead, we control for the characteristics of the *targets* associated with each ticket, since target selection (through random, indegree, or friendship nomination targeting) can be considered contemporaneous with or preceding "treatment" (i.e., introduction of the interventions to the villages), thus allowing us to control for potential confounders while preserving experimental causal inference.

We also controlled for basic village properties and blocking variables, and for block assignment. We accounted for clustering of observations within villages by estimating frailty models with random village intercepts (using the *coxme* package in R).

The models for chlorine ticket redemption (Table S5) were estimated in analogous fashion.

Models using generalised estimating equations (adjusting variance estimates for clusters of within-village-correlated observations based on a grouped jackknife, using the *geepack* package in R) produce qualitatively similar results.

Tables S4 and S5 show the beta coefficient estimates, hazard ratios (measured as the exponentiated coefficients), 95% confidence intervals for the hazard ratios, and *p* values.

Models of Knowledge Diffusion

Over the course of a single day for each village, we delivered to each targeted individual an intervention consisting of a health product (multivitamins or chlorine), instructions for use, and an educational component. Basic usage and safety information was repeated to ticket-redeemers when they received their products. However, we also gave the initially targeted individuals supplementary information about the interventions that was *not* generally known at baseline or circulated by other means, and asked them to relay it to those to whom they gave tickets, which allowed us to track the diffusion of knowledge as well as of product adoption by the study's completion.

For multivitamins, we taught targets to take 1 tablet per day. The supplementary usage information was to take the pill with food if the vitamin upsets an empty stomach. The educational information was that calcium strengthens bones, that iron prevents anemia, and that vitamin A aids vision.

For the chlorine intervention, recipients were taught to add 3 drops of bleach to 1 liter of tap water, stir, and leave sitting for 20 minutes before drinking. The supplementary usage information was to add 2 additional drops of chlorine if the water was cloudy. The educational information was that the correct use of chlorine kills the germs that cause diarrhea, which is especially harmful to young children.

Upon completion of the interventions, we returned to the 32 villages and administered a follow-up survey to all villagers, including targets, first- and second-wave ticket redeemers, and those who had neither been targeted nor had redeemed tickets. To assess the diffusion of knowledge about the interventions, we asked all non-targeted ticket recipients a series of questions about the use and benefits of the products.

For multivitamins, these questions included: how often to take the vitamins, what to do differently if the vitamin caused stomach upset, what vitamins were included in the multivitamin, and what health benefits were associated with the component vitamins. For chlorine, we asked ticket redeemers how many drops to use per liter of water, how long to let the water rest before consumption, what to do with cloudy water, what are the health consequences of contaminated water, and why purifying water is particularly important for children.

We formed composite 0-10 knowledge scores for each intervention, using the first component from a principal components analysis of the knowledge and usage questions posed during the follow-up survey. Since some of the information tested was introduced only to the original target individuals, the attainment of a high composite knowledge score (which we defined as a score in the top quartile) in an untargeted individual suggested that the usage and health information had diffused through the village network, starting from the target individuals. Because

the composite scores assumed a limited number of discrete values, the proportion of respondents in the top "quartile" is greater than 25%.

We estimated logistic regression models with random village intercepts, fit by maximum likelihood, for high knowledge score attainment among untargeted ticket recipients. We include the same individual-level and village-level covariates as in the ticket-redemption models, with the modification that, here, the individual-level covariates pertain to the ticket recipient herself, rather than to the original target individual. We control for possible interference between the two interventions by estimating the effects of the targeting methods of *both* products on the attainment of a high knowledge score for either product. Results for both interventions are presented in Table S6. Models using other measures of knowledge (including the raw composite score, rather than the PCA-derived score) produce qualitatively similar results (not shown).

Robustness of Results to Omitted Data

We omitted from village-level analysis (Figure 3 in main text) partial data from one village in which the ticket-redemption process was not accurately documented by the participating shopkeeper. This did not affect the main results: in the data as presented, multivitamin ticket redemption rates were 66·2% (940/1420) under random targeting, 61·0% (744/1220) under indegree targeting, and 74·3% (951/1280) under nomination targeting. With no data omitted, the respective rates are 65·9% (988/1500), 61·0% (744/1220) and 74·3% (951/1280).

For the chlorine intervention, aggregate ticket redemption rates as presented were statistically indistinguishable ($p > 0{\cdot}05$): 55·6% (690/1240) under random targeting, 55·9% (995/1780) under indegree, and 54·5% (512/940) under nomination targeting. With no data omitted, the respective rates are again statistically equivalent: 55·6% (690/1240), 55·9% (995/1780) and 53·4% (545/1020).

Tables

Table S1. Baseline survey results used for block assignment

Village number	Proportion of respondents drinking purified water	Proportion of respondents taking multivitamins daily	Village SES rating	Villagers with network data
1	0·38	0·11	5·9	120
2	0·15	0·02	3·3	43
3	0·10	0·06	4·3	39
4	0·10	0·05	7·4	108
5	0·33	0·15	7·4	254
6	0·49	0·10	1·4	141
7	0·41	0·12	6·0	180
8	0·49	0·10	6·4	90
9	0·07	0·03	2·4	130
10	0·77	0·22	6·1	240
11	0·60	0·10	4·4	368
12	0·21	0·10	1·8	112
13	0·21	0·17	5·6	369
14	0·72	0·08	1·3	88
15	0·33	0·11	2·0	72
16	0·19	0·09	4·1	112
17	0·18	0·10	3·4	38
18	0·62	0·21	6·0	25
19	0·13	0·04	2·8	88
20	0·70	0·12	8·0	35
21	0·29	0·07	1·3	60
22	0·98	0·13	3·4	174
23	0·51	0·11	5·0	301
24	0·67	0·22	8·4	348
25	0·17	0·14	6·6	101
26	0·34	0·14	3·9	206
27	0·56	0·10	2·4	151
28	0·14	0·07	4·0	79
29	0·95	0·04	1·0	64
30	0·33	0·10	6·4	390
31	0·15	0·10	4·7	72
32	0·78	0·07	4·1	64

Table S2. Randomised block experimental design

	Multivitamins				Chlorine			
Block	**Village 1**	**Village 2**	**Village 3**	**Village 4**	**Village 1**	**Village 2**	**Village 3**	**Village 4**
1	Random	Indegree	Nominated	Random	Indegree	Nominated	Random	None
2	Random	Indegree	Nominated	Indegree	Nominated	Random	Indegree	None
3	Random	Indegree	Nominated	Nominated	Indegree	Nominated	Random	None
4	Random	Indegree	Nominated	None	Nominated	Random	Indegree	Random
5	Random	Indegree	Nominated	None	Indegree	Nominated	Random	Indegree
6	Random	Indegree	Nominated	None	Nominated	Random	Indegree	Nominated
7	Random	Indegree	Nominated	None	Indegree	Nominated	Random	None
8	Random	Indegree	Nominated	None	Nominated	Random	Indegree	None

Table S3. Assignment of villages to block and targeting method

Village	Block	Multivit. Targeting	Chlorine targeting	Targets	Network size	Mean indegree of random targets (SD)	Mean indegree of nominated targets (SD)	Mean indegree of indegree targets (SD)
1	2	Rand	Nom	6	120	2·50 (2·26)	9·67 (4·59)	
2	5	Indeg	Nom	2	43		6·50 (2·12)	7·50 (3·54)
3	5	Nom	Rand	2	39	5·50 (4·95)	3·00 (0·00)	
4	6	Rand	Nom	5	108	4·20 (0·45)	3·80 (1·64)	
5	1	Indeg	Nom	13	254		5·54 (3·20)	8·58 (1·31)
6	2	Nom	Indeg	7	141		5·29 (3·25)	11·29 (2·29)
7	1	Nom	Rand	9	180	5·11 (3·33)	7·00 (4·87)	
8	7			5	90			
9	5		Indeg	7	130			10·86 (1·95)
10	4		Rand	12	240	4·83 (3·16)		
11	4	Indeg	Rand	18	368	4·67 (4·38)		13·17 (2·53)
12	6	Nom	Indeg	6	112		8·17 (4·71)	10·17 (2·79)
13	1	Rand	Indeg	18	369	5·94 (3·56)		11·50 (2·09)
14	7	Indeg	Nom	4	88		5·67 (1·53)	9·25 (3·20)
15	6	Indeg	Rand	4	72	4·00 (2·16)		8·00 (0·82)
16	6		Nom	6	112		9·67 (5·65)	
17	5	Rand	Indeg	2	38	1·50 (0·71)		7·50 (2·12)
18	7	Nom	Rand	1	25	2·00 (0·00)	3·00 (0·00)	
19	8	Rand	Nom	4	88	2·50 (2·38)	4·50 (1·73)	
20	3	Indeg	Nom	2	35		7·50 (6·36)	9·50 (3·54)
21	8	Indeg	Rand	3	60	5·33 (4·51)		9·67 (0·58)
22	4	Rand	Nom	9	174	3·11 (2·42)	6·89 (5·49)	
23	3	Nom		15	301		7·80 (5·52)	
24	4	Nom	Indeg	17	348		3·76 (2·36)	6·71 (1·76)
25	2	Indeg		5	101			9·80 (3·56)
26	2	Indeg	Rand	10	206	3·00 (2·36)		9·70 (2·11)
27	7	Rand	Indeg	8	151	4·12 (1·13)		12·75 (7·44)
28	8	Nom	Indeg	4	79		5·25 (1·50)	7·00 (0·82)
29	3	Nom	Rand	3	64	7·00 (1·00)	6·33 (3·06)	
30	3	Rand	Indeg	20	390	5·80 (3·76)		9·85 (2·16)
31	8			4	72			
32	1	Rand		3	64	2·33 (1·53)		

Table S4. Cox regression models with random village intercepts for multivitamin ticket redemption

	First wave tickets				*Second wave tickets*			
	Coef	**Hazard Ratio**	**95% CI (HR)**	***p***	**Coef**	**Hazard Ratio**	**95% CI (HR)**	***p***
Nomination targeting	**0·50**	**1·65**	**1·10, 2·47**	**0·02**	**0·14**	**1·15**	**0·74, 1·76**	**0·54**
Indegree targeting	**0·26**	**1·29**	**0·87, 1·93**	**0·20**	**0·07**	**1·07**	**0·70, 1·65**	**0·76**
Chlorine ticket redeemed	2·21	9·07	7·46, 11·03	0·00	2·50	12·23	10·87, 13·75	0·00
Target male	-0·34	0·71	0·59, 0·85	0·00	-0·12	0·89	0·80, 0·98	0·02
Target age	0·01	1·01	1·00, 1·02	0·01	0·00	1·00	1·00, 1·00	0·77
Target persons in house	0·02	1·02	0·98, 1·07	0·34	0·03	1·04	1·01, 1·06	0·01
Target married	-0·03	0·97	0·81, 1·16	0·75	-0·02	0·98	0·88, 1·08	0·63
Village mean indegree	0·11	1·12	0·80, 1·57	0·52	-0·09	0·91	0·63, 1·31	0·62
Village percent male	-0·03	0·97	0·93, 1·01	0·14	-0·04	0·96	0·92, 1·00	0·04
Village mean age	-0·03	0·97	0·88, 1·07	0·54	-0·01	0·99	0·89, 1·09	0·79
Village SES	-0·02	0·98	0·83, 1·17	0·86	-0·09	0·92	0·77, 1·10	0·34
Village population (100s)	-0·44	0·64	0·41, 1·01	0·06	-0·36	0·70	0·43, 1·14	0·15
Village baseline purification	-0·02	0·98	0·95, 1·02	0·35	-0·01	0·99	0·96, 1·02	0·50
Block 2	-0·53	0·59	0·26, 1·35	0·21	-0·19	0·83	0·32, 2·14	0·69
Block 3	0·71	2·04	0·96, 4·34	0·06	0·61	1·84	0·83, 4·04	0·13
Block 4	0·02	1·02	0·24, 4·44	0·98	0·10	1·10	0·23, 5·28	0·90
Block 5	-1·12	0·33	0·04, 2·60	0·29	-0·89	0·41	0·04, 3·90	0·44
Block 6	-1·83	0·16	0·03, 0·83	0·03	-1·39	0·25	0·04, 1·49	0·13
Block 7	-0·25	0·78	0·38, 1·58	0·49	0·08	1·08	0·52, 2·28	0·83
Block 8	-1·06	0·35	0·05, 2·57	0·30	-0·86	0·42	0·05, 3·63	0·43
Tickets redeemed	645				2000			
Random village intercept variance	0·08				0·15			
Deviance (null)	6758				28268			
Deviance (fitted)	6186				25996			

Table S5. Cox regression models with random village intercepts for chlorine ticket redemption

	First wave tickets				*Second wave tickets*			
	Coef	**Hazard Ratio**	**95% CI (HR)**	***p***	**Coef**	**Hazard Ratio**	**95% CI (HR)**	***p***
Nomination targeting	**0·06**	**1·06**	**0·60, 1·89**	**0·84**	**0·26**	**1·30**	**0·79, 2·14**	**0·31**
Indegree targeting	**-0·01**	**0·99**	**0·56, 1·75**	**0·97**	**-0·07**	**0·94**	**0·57, 1·53**	**0·79**
Multivitamin ticket redeemed	2·85	17·29	14·27, 20·95	0·00	2·99	19·98	17·70, 22·56	0·00
Target male	-0·01	0·99	0·82, 1·21	0·94	-0·08	0·92	0·82, 1·04	0·19
Target age	-0·01	0·99	0·99, 1·00	0·13	0·00	1·00	0·99, 1·00	0·21
Target persons in house	0·04	1·04	0·99, 1·10	0·12	-0·02	0·98	0·95, 1·01	0·20
Target married	-0·09	0·91	0·75, 1·10	0·32	-0·25	0·78	0·70, 0·87	0·00
Village mean indegree	0·55	1·74	1·13, 2·68	0·01	0·33	1·39	0·96, 2·03	0·08
Village percent male	0·02	1·02	0·96, 1·07	0·59	-0·03	0·97	0·93, 1·02	0·23
Village mean age	0·01	1·01	0·90, 1·14	0·82	0·01	1·01	0·92, 1·12	0·79
Village SES	0·21	1·24	0·94, 1·63	0·13	0·22	1·24	0·97, 1·59	0·08
Village population (100s)	0·25	1·28	0·49, 3·36	0·61	0·90	2·46	1·05, 5·75	0·04
Village baseline purification	0·02	1·02	0·95, 1·09	0·57	0·07	1·07	1·00, 1·14	0·04
Block 2	0·18	1·20	0·38, 3·82	0·76	0·65	1·91	0·67, 5·44	0·22
Block 3	0·05	1·05	0·26, 4·31	0·94	-1·32	0·27	0·08, 0·94	0·04
Block 4	-1·29	0·28	0·01, 6·13	0·42	-3·25	0·04	0·00, 0·62	0·02
Block 5	2·67	14·46	0·24, 875·70	0·20	4·32	74·83	1·90, 2948·67	0·02
Block 6	1·28	3·59	0·16, 82·45	0·42	2·92	18·52	1·13, 304·30	0·04
Block 7	0·68	1·98	0·73, 5·35	0·18	0·57	1·76	0·77, 4·04	0·18
Block 8	1·91	6·74	0·15, 307·47	0·33	4·01	55·20	1·85, 1646·42	0·02
Tickets redeemed	577				1625			
Random village intercept variance	0·11				0·11			
Deviance (null)	6292				23680			
Deviance (fitted)	5294				20348			

Table S6. Logistic regression models with random village intercepts for high knowledge score attainment

	Multivitamin high knowledge score				*Chlorine high knowledge score*			
	Coef	**Odds Ratio**	**95% CI (OR)**	***p***	**Coef**	**Odds Ratio**	**95% CI (OR)**	***p***
Nomination targeting (multivitamin)	**0·51**	**1·66**	**1·02, 2·70**	**0·04**	0·27	1·32	0·87, 1·98	0·19
Indegree targeting (multivitamin)	**0·23**	**1·25**	**0·84, 1·88**	**0·28**	0·10	1·10	0·76, 1·59	0·60
Nomination targeting (chlorine)	0·61	1·85	1·09, 3·14	0·02	**0·20**	**1·23**	**0·80, 1·88**	**0·34**
Indegree targeting (chlorine)	0·42	1·53	0·98, 2·38	0·06	**0·10**	**1·11**	**0·69, 1·79**	**0·67**
Male	0·02	1·02	0·84, 1·24	0·80	0·05	1·05	0·85, 1·30	0·63
Age	0·00	1·00	0·99, 1·01	0·73	0·00	1·00	0·99, 1·01	0·59
Persons in house	-0·02	0·99	0·93, 1·04	0·57	0·01	1·01	0·96, 1·07	0·69
Married	-0·06	0·94	0·77, 1·15	0·57	-0·14	0·87	0·71, 1·07	0·20
Village mean indegree	0·10	1·10	0·86, 1·41	0·44	0·16	1·17	0·91, 1·51	0·21
Village mean age	-0·11	0·90	0·82, 0·99	0·03	-0·04	0·96	0·90, 1·03	0·30
Village percent male	-0·02	0·98	0·94, 1·03	0·49	0·00	1·00	0·97, 1·04	0·84
Village SES	0·00	1·00	0·91, 1·11	0·92	0·05	1·06	0·95, 1·17	0·29
Village population (100s)	-0·06	0·95	0·76, 1·18	0·62	-0·17	0·85	0·66, 1·08	0·18
Village baseline purification	0·00	1·00	0·99, 1·01	0·72	-0·01	0·99	0·98, 1·00	0·05
Block 2	0·27	1·31	0·85, 2·02	0·21	-0·03	0·97	0·65, 1·44	0·87
Block 3	0·40	1·49	0·95, 2·33	0·08	0·25	1·28	0·75, 2·18	0·36
Block 4	0·38	1·46	0·70, 3·04	0·31	0·44	1·56	0·80, 3·01	0·19
Block 5	1·13	3·11	0·93, 10·35	0·06	0·01	1·01	0·45, 2·25	0·98
Block 6	0·09	1·09	0·54, 2·22	0·81	-0·50	0·61	0·34, 1·08	0·09
Block 7	0·02	1·02	0·58, 1·80	0·95	-0·15	0·86	0·45, 1·64	0·66
Block 8	0·44	1·56	0·72, 3·38	0·26	-0·45	0·64	0·31, 1·34	0·23
Intercept	2·56	12·97	0·15, 1152·94	0·26	0·30	1·35	0·05, 35·09	0·86
N (Individuals)	2110				1698			
N (Villages)	27				27			
Random village intercept variance	$1·80 \times 10^{-11}$				$1·75 \times 10^{-11}$			
Deviance (null)	2523				2222			
Deviance (fitted)	2481				2202			

References

1. Centola D. The Spread of Behavior in an Online Social Network Experiment. Science 2010;329:1194-7.
2. Centola D, Macy MW. Complex Contagions and the Weakness of Long Ties. American Journal of Sociology 2007;113:702-34.
3. King G, Zeng L. The Dangers of Extreme Counterfactuals. Political Analysis 2006;14:131-59.