Online Appendix

“Who knows? The effect of information access on social network position”

For Online Publication

Laura Derksen and Pedro Souza

November 12, 2025

A Appendix

A.1 Randomization Inference for Dyadic Regressions

In Section 4.1, we estimate regressions of the following form:

$$100 \times \text{link}_{ij}^{1} = \beta_0 + \beta_1 \cdot TC_{ij} + \beta_2 \cdot TT_{ij} + \alpha \cdot \text{link}_{ij}^{0} + \mathbf{x}_{ij}'\chi + \epsilon_{ij} \tag{8}$$

where $\text{link}_{ij}^{1} = 1$ if a link is formed between i and j at the endline, $TC_{ij} = 1$ if i is treated and j is control, or vice-versa, $TT_{ij} = 1$ is i and j are treated, $\text{link}_{ij}^{0} = 1$ if a link exists between i and j at the baseline, $\mathbf{x}_{ij}$ is a set of controls, and ϵ_{ij} is the error term. For this appendix, we focus on the case of the undirected networks in Table 4. Other specifications follow with minor modifications.

Consider a null hypothesis under which the intervention does not affect link formation at all. In particular, the sharp null hypothesis is that every dyad in the network would have the same relationship (linked or not linked) regardless of the treatment status of the two nodes involved, and regardless of the treatment assignments of the nodes in the wider network. Under this null hypothesis, the variable link_{ij}^{1} is equal to its potential outcome under any treatment assignment. Thus, under the null hypothesis, the estimates of the effects of the intervention (β_1 and β_2) are statistically indistinguishable under various different treatment assignments.

More specifically, the randomization inference procedure in this case recovers the p-values the following way.

Step 1. Estimate Equation (8) under the original treatment allocation, and store $\hat{\beta}_1$ and $\hat{\beta}_2$.

Step 2. Reshuffle the treatment vector respecting the original stratification bins, and recompute the TC_{ij} and TT_{ij} variables that are consistent with the new treatment allocation, referred to as TC_{ij}^{s} and TT_{ij}^{s}.

Step 3. Reestimate (8) with the reshuffled treatment status,

$$100 \times \text{link}_{ij}^{1} = \beta_0^{s} + \beta_1^{s} \cdot TC_{ij}^{s} + \beta_2^{s} \cdot TT_{ij}^{s} + \alpha^{s} \cdot \text{link}_{ij}^{0} + \mathbf{x}_{ij}'\chi + \epsilon_{ij} \tag{9}$$

and save the estimates $\hat{\beta}_1^{s}$ and $\hat{\beta}_2^{s}$.

Step 4. Repeat steps 2 and 3 above B =10,000 times, and compute the randomization inference p-values

$$p_{\beta_1} \quad = \quad \frac{1}{B}\sum_{s=1}^{B} I\left[|\hat{\beta}_1| > |\hat{\beta}_1^{s}|\right]$$

and similarly for β_2. Those estimates are reported in the paper.

Fredrickson and Chen (2019) show that randomization inference can be used to estimate causal effects on both local (e.g. link formation) and global (e.g. centrality measures) network outcomes with individual-level randomization. Blattman et al. (2021) show how randomization inference can be used to measure treatment effects and spillovers on a geographical network. For further applications and uses of randomization inference, see Duflo et al. (2007), Athey and Imbens (2017) and many others.

A.2 Interpreting Reduced-Form Estimates of Centrality Differences

In equation 3, we regress a node's centrality on its treatment status. These estimates, as discussed in Section 4.2, must be interpreted as relative differences between treatment and control nodes, as opposed to treatment effects. The estimates are, nevertheless, causal, in the sense that these relative differences are due to the intervention. They are unbiased estimates of the expected average difference in centrality between treated and control nodes.

Importantly, our estimates cannot be interpreted directly as *average treatment effects*. In fact, in the context of network centrality, it is not straightforward to define an *average treatment effect*, and such effects are often not the parameter of primary interest. Moreover, centrality measures are highly interdependent, and the Stable Unit Treatment Value Assumption (SUTVA, Rubin 1974) will be violated. While treatment effects bounds can be estimated even when SUTVA violations are present (Manski, 2013), the required assumptions are likely too strong for a setting in which the outcomes themselves are non-localized network measures. In Section 5.3, we use a calibrated network formation model to estimate average treatment effects, and find that they are in fact similar to the relative treatment effects we estimate in reduced form.

Let us demonstrate the appropriate interpretation of our estimates with a simple example. Consider a small network of five nodes, one of which was treated at random, with realized treatment vector

$$\mathbf{T} = \{T_1, T_2, T_3, T_4, T_5\}$$

with $T_i \in \{0, 1\}$ and $|\mathbf{T}| = 1$. We observe realized centrality measures for all nodes:

$$\mathbf{c} = \{c_1, c_2, c_3, c_4, c_5\}$$

One important parameter of interest in this context is the expected difference in centrality between treated and control nodes. This parameter allows us to shed light on the determinants of relative (as opposed to absolute) centrality within a given network. This is particularly relevant for network-based targeting, where policies typically target the most central nodes in a network based on their relative positions as opposed to their raw centrality scores. Within the context of our simple example with a single randomly-treated node, we can define this parameter as follows:

$$\rho = \sum_{i=1}^{5} \left(c_i^{T_i=1,T_{-i}=0} - \frac{1}{4} \sum_{j \neq i} c_j^{T_i=1,T_{-i}=0} \right) \mathbb{P}(T_i = 1)$$

where $c_j^{T_i=1,T_{-i}=0}$ is the potential outcome for node j when only node i is treated. We can produce an unbiased estimate of ρ by taking a simple difference of means. Suppose that in our realized sample, node k was treated. Then, we obtain the estimate

$$\hat{\rho} = c_k - \frac{1}{4} \sum_{j \neq k} c_j.$$

Taking the expectation over different realizations of $\mathbf{T}$, this estimator is unbiased:

$$\mathbb{E}(\hat{\rho}) = \frac{1}{5}\sum_{i=1}^{5} c_i^{T_i=1,T_{-i}=0} - \frac{1}{5}\sum_{i=1}^{5}\frac{1}{4}\sum_{j\neq i} c_j^{T_i=1,T_{-i}=0} = \rho \tag{10}$$

While the parameter we estimate is relevant and has a simple causal interpretation, it does not correspond to any parameter that would typically be thought of as an *average treatment effect*. First, because node centrality measures are interdependent, and depend on the entire vector of treatment statuses, defining a "treatment effect" for node i is not straightforward. For example, we could compare the treated node's centrality in a network in which only that node is treated to its potential outcome under no treatment $c_i^{\mathbf{T}=\mathbf{0}}$. Or, we could compare its potential outcome in a fully treated network $c_i^{\mathbf{T}=\mathbf{1}}$ to the untreated network $c_i^{\mathbf{T}=\mathbf{0}}$. Second, even if we settle on a "treatment effect" definition, we cannot produce an unbiased estimate of the average treatment effect across nodes with our data.

For example, if we define the average treatment effect as the expected effect on the treated node's centrality relative to the node's potential outcome in the untreated network, we then seek to estimate the following parameter:

$$\beta^{ATE} = \sum_{i=1}^{5}\left(c_i^{T_i=1,T_{-i}=0} - c_i^{\mathbf{T}=\mathbf{0}}\right)\mathbb{P}(T_i=1) = \frac{1}{5}\sum_{i=1}^{5} c_i^{T_i=1,T_{-i}=0} - \frac{1}{5}\sum_{i=1}^{5} c_i^{\mathbf{T}=\mathbf{0}}.$$

Referring to equation 10, the problem becomes clear. Because $c_j^{T_i=1,T_{-i}=0} \neq c_j^{\mathbf{T}=\mathbf{0}}$, in general this expected value will not be equal to β^{ATE}.

Nevertheless, with our data we are able to test sharp null hypothesis of no treatment effect, for any node, under any treatment vector (Fredrickson and Chen, 2019). Under this null hypothesis, the potential outcomes are equal to the realized outcomes, $c_i^{\mathbf{T}} = c_i$, for any treatment vector $\mathbf{T}$. This null hypothesis implies zero expected centrality difference between treated and control students,

$$\mathbb{E}(\hat{\rho}) = \frac{1}{5}\sum_{i=1}^{5}\left(c_i - \frac{1}{4}\sum_{j\neq i} c_j\right) = 0,$$

and we can test the null hypothesis using randomization inference p-values for the parameter estimate $\hat{\rho}$.

A.3 Proof of Theorem 5.1

Proof. For a node in the treatment group, the expected degree is

$$\begin{aligned}\mathbb{E}(d_i|T_i=1) &= (N_T-1)P_{TT} + N_C P_{TC}\\ &= (N_T-1)P_{TT} + N_C P_{CC} + N_C(P_{TC}-P_{CC}).\end{aligned}$$

For a node in the control group, the expected degree is

$$\begin{aligned}\mathbb{E}(d_i|C_i=1) &= (N_C-1)P_{CC} + N_T P_{TC}\\ &= (N_C-1)P_{CC} + N_T P_{CC} + N_T(P_{TC}-P_{CC})\\ &= (N-1)P_{CC} + N_T(P_{TC}-P_{CC}).\end{aligned}$$

Because $N_C > N_T$ and $P_{TC} > P_{CC}$,

$$\mathbb{E}(d_i|T_i=1) - \mathbb{E}(d_i|C_i=1) > (N_T-1)P_{TT} + N_C P_{CC} - (N-1)P_{CC} \geq 0.$$

□

A.4 Calibration Details

Network formation. To calibrate the model, we match parameters to moments in our empirical information network as follows. First, we note that under this model both the baseline network and endline network are still general random graphs, with unconditional link probabilities represented similarly to those in equation (5), but with with link probabilities that depend on the academic types of the nodes.

Between pairs of control nodes, these link probabilities are the same at baseline and at endline, and are symmetric in θ_1 and θ_2.

$$\mathbb{P}(g_{ij}^0 = 1 | T_i = T_j = 0) = \mathbb{P}(g_{ij}^1 = 1 | T_i = T_j = 0) = P_{CC}^{\theta_1\theta_2} = P_{CC}^{\theta_2\theta_1} \equiv \mathbb{P}(\nu > -\kappa_{CC}^{\theta_1\theta_2})\mathbb{P}(\nu > -\kappa_{CC}^{\theta_2\theta_1})$$

In our data, the probability of an endline information-link between two control students in the same school and form is used to estimate these probabilities as follows:

$$\hat{P}_{CC}^{LL} = 0.08 \qquad \hat{P}_{CC}^{HL} = 0.11 \qquad \hat{P}_{CC}^{HH} = 0.21$$

These estimates allow us to simulate a simple baseline network. To simulate a corresponding endline network, we start by constructing a "shadow" network $\tilde{g}^1$. This is the network of links that would exist at endline absent the intervention, but allowing for residual network changes to occur over time. In order to simulate a shadow network that is suitably correlated with the baseline network, we must estimate the probability of a shadow link (or equivalently, an endline link) between control nodes conditional on a baseline link

$$\mathbb{P}(g_{ij}^1 = 1 | g_{ij}^0 = 1, T_i = T_j = 0, \theta_i = \theta_1, \theta_j = \theta_2) = (1-\delta)^2 + 2\delta(1-\delta)\sqrt{P_{CC}^{\theta_1\theta_2}} + \delta^2 P_{CC}^{\theta_1\theta_2} \equiv P_{CC|CC}^{\theta_1\theta_2} \qquad (11)$$

$$\mathbb{P}(\tilde{g}_{ij}^1 = 1 | g_{ij}^0 = 1, \theta_i = \theta_1, \theta_j = \theta_2) = P_{CC|CC}^{\theta_1\theta_2}$$

Note that this probability depends on the types $\{\theta_i, \theta_j\}$ but is symmetric in these types. If a link exists in the baseline network, the probability it should appear in the shadow network, $\hat{P}_{CC|CC}^{\theta_1\theta_2}$ is estimated directly from the moment (11) in the data. That is, we take the probability that a control-pair with types θ_1 and θ_2 is linked at endline, conditional on a link existing at baseline:

$$\hat{P}_{CC|CC}^{LL} = 0.35 \qquad \hat{P}_{CC|CC}^{HL} = 0.41 \qquad \hat{P}_{CC|CC}^{HH} = 0.48$$

Conversely, if a link does not exist in the baseline network, the probability that it should appear in the shadow network can be calculated using Bayes' rule.

$$\mathbb{P}(\tilde{g}_{ij}^1 = 1 | g_{ij}^0 = 0, \theta_i = \theta_1, \theta_j = \theta_2) = \frac{P_{CC}^{\theta_1\theta_2}}{1 - P_{CC}^{\theta_1\theta_2}} \left(1 - P_{CC|CC}^{\theta_1\theta_2}\right)$$

This probability, that a pair of control students is linked at endline given there is no link at baseline, is also estimated directly from the corresponding moment in the data.

$$\hat{P}_{CC|!CC}^{LL} = 0.05 \qquad \hat{P}_{CC|!CC}^{HL} = 0.07 \qquad \hat{P}_{CC|!CC}^{HH} = 0.13$$

Next, we simulate an endline network by adding links to to the shadow network. We assume that for fixed θ_1, θ_2, $\kappa_{CC}^{\theta_1\theta_2}$ is weakly smaller than $\kappa_{TC}^{\theta_1\theta_2}$, $\kappa_{CT}^{\theta_1\theta_2}$ and $\kappa_{TT}^{\theta_1\theta_2}$. That is, information is valuable and not costly to spread. This implies that $P_{CC}^{\theta_1\theta_2}$ is weakly smaller than $P_{TC}^{\theta_1\theta_2}$ and $P_{TT}^{\theta_1\theta_2}$, consistent with our reduced-form empirical results (see Table 4). Then, conditional on having a link in the shadow network, the probability of having a link in the endline network is one.

$$\mathbb{P}(g_{ij}^1 = 1|\tilde{g}_{ij}^1 = 1, \theta_i = \theta_1, \theta_j = \theta_2) = 1$$

Conditional on having no link in the shadow network, the probability of a link in the endline network depends on the treatment statuses and academic types of the nodes involved.

$$\mathbb{P}(g_{ij}^1 = 1|\tilde{g}_{ij}^1 = 0, T_i = T_j = 0, \theta_i = \theta_1, \theta_j = \theta_2) = 0$$

$$\mathbb{P}(g_{ij}^1 = 1|\tilde{g}_{ij}^1 = 0, T_i = T_j = 1, \theta_i = \theta_1, \theta_j = \theta_2) = 1 - \frac{\mathbb{P}(\nu < -\kappa_{TT}^{\theta_1\theta_2})\mathbb{P}(\nu < -\kappa_{TT}^{\theta_2\theta_1})}{\mathbb{P}(\nu < -\kappa_{CC}^{\theta_1\theta_2})\mathbb{P}(\nu < -\kappa_{CC}^{\theta_2\theta_1})} = \frac{P_{TT}^{\theta_1\theta_2} - P_{CC}^{\theta_1\theta_2}}{1 - P_{CC}^{\theta_1\theta_2}}$$

$$\mathbb{P}(g_{ij}^1 = 1|\tilde{g}_{ij}^1 = 0, T_i = 1, T_j = 0, \theta_i = \theta_1, \theta_j = \theta_2) = 1 - \frac{\mathbb{P}(\nu < -\kappa_{TC}^{\theta_1\theta_2})\mathbb{P}(\nu < -\kappa_{CT}^{\theta_2\theta_1})}{\mathbb{P}(\nu < -\kappa_{CC}^{\theta_1\theta_2})\mathbb{P}(\nu < -\kappa_{CC}^{\theta_2\theta_1})} = \frac{P_{TC}^{\theta_1\theta_2} - P_{CC}^{\theta_1\theta_2}}{1 - P_{CC}^{\theta_1\theta_2}}$$

We estimate the relevant moments from our endline data as follows:

$$\begin{array}{llll} \hat{P}_{TT}^{LL} = 0.09 & \hat{P}_{TT}^{HL} = 0.13 & & \hat{P}_{TT}^{HH} = 0.29 \\ \hat{P}_{TC}^{LL} = 0.08 & \hat{P}_{TC}^{HL} = 0.12 & \hat{P}_{TC}^{LH} = 0.12 & \hat{P}_{TC}^{HH} = 0.25 \end{array} \tag{12}$$

These calculations allow us to simulate an endline network based on the shadow network. We now have all the required ingredients to simulate an baseline network, a shadow network, and an endline network with appropriately correlated links.

Academic performance. Next, we calibrate our model of academic performance. We modeled a student's academic score y_i as follows.

$$y_i = s_i + \tau(a_i)T_i + \tau(a_i) \sum_{j:g_{ij}^1=1} T_j Q_{ij} \tag{13}$$

We begin by taking the conditional expectation. We abuse notation to write $\tau(\theta_i) = \mathbb{E}(\tau(a_i)|\theta_i)$. While we will focus on capturing this average effect, we do not explicitly assume treatment effects to be uniform within ability-types. In the model of link formation we do assume that the expected number of treated links is independent of ability a_i given type θ_i.

$$\mathbb{E}(y_i|T_i, \theta_i) = \mathbb{E}(s_i|\theta_i) + \tau(\theta_i)T_i + q\tau(\theta_i)\mathbb{E}(\sum_{j:g_{ij}^1=1} T_j|T_i, \theta_i) \tag{14}$$

For a particular T_i and θ_i, we can use final exam scores in the year of the intervention to match $\mathbb{E}(y_i|T_i, \theta_i)$.

$$\begin{aligned}\mathbb{E}(y_i|T_i=1) &= 0.38 \\ \mathbb{E}(y_i|T_i=0) &= 0.26 \\ \mathbb{E}(y_i|T_i=1,\theta_i=\theta_H) &= 1.43 \\ \mathbb{E}(y_i|T_i=0,\theta_i=\theta_H) &= 1.38 \\ \mathbb{E}(y_i|T_i=1,\theta_i=\theta_L) &= 0.26 \\ \mathbb{E}(y_i|T_i=0,\theta_i=\theta_L) &= 0.13\end{aligned} \tag{15}$$

The last expectation in equation 14 can also be matched to a corresponding moment in the data, making use of the number of high and lower-ability treated students as well as the link probabilities we computed in equation 12.

$$\begin{aligned}\mathbb{E}(\sum_{j:g^1_{ij}=1} T_j|T_i=1,\theta_i=\theta_H) &= 3.37 \\ \mathbb{E}(\sum_{j:g^1_{ij}=1} T_j|T_i=0,\theta_i=\theta_H) &= 3.36 \\ \mathbb{E}(\sum_{j:g^1_{ij}=1} T_j|T_i=1,\theta_i=\theta_L) &= 2.29 \\ \mathbb{E}(\sum_{j:g^1_{ij}=1} T_j|T_i=0,\theta_i=\theta_L) &= 2.12\end{aligned} \tag{16}$$

Next, we subtract the conditional expectation for the control arm from the conditional expectation for the treated arm, as follows.

$$\begin{aligned}\mathbb{E}(y_i|T_i=1,\theta_i=\theta_H) - \mathbb{E}(y_i|T_i=0,\theta_i=\theta_H) &= \tau(\theta_H) + q\tau(\theta_H)(3.37-3.36) \\ 0.05 &= \tau(\theta_H)(1+0.01q) \\ \mathbb{E}(y_i|T_i=1,\theta_i=\theta_L) - \mathbb{E}(y_i|T_i=0,\theta_i=\theta_L) &= \tau(\theta_L) + q\tau(\theta_L)(3.41-3.28) \\ 0.12 &= \tau(\theta_L)(1+0.16q)\end{aligned}$$

Then, we take the unconditional expectation for the control arm.

$$\begin{aligned}\mathbb{E}(y_i|T_i=0) &= \mathbb{E}(s_i) + q\tau(\theta_H)\mathbb{E}(\sum_{j:g^1_{ij}=1} T_j|T_i=0,\theta_i=\theta_H) + q\tau(\theta_L)\mathbb{E}(\sum_{j:g^1_{ij}=1} T_j|T_i=0,\theta_i=\theta_L) \\ 0.26 &= 0 + q(3.36\tau(\theta_H) + 2.12\tau(\theta_L))\end{aligned} \tag{17}$$

Here, we matched $\mathbb{E}(y_i|T_i=0)$ to the endline control-arm mean, and $\mathbb{E}(s_i)$ to the mean from the previous school year, which is zero due to normalization. We now have three equations with three unknowns, which we can solve to obtain:

$$\begin{aligned}q &= 0.67 \\ \tau(\theta_H) &= 0.05 \\ \tau(\theta_L) &= 0.11\end{aligned}$$

A.5 The Threshold Model and Precise Targeting

Threshold model. Here, following Beaman et al. (2021), we consider a diffusion model in which a node becomes informed when at least two of its neighbors are informed. The process is again repeated for T periods. This is a "threshold model" of diffusion, with threshold $\lambda = 2$ (Granovetter, 1978).

In Figure A5 we plot the extent of information diffusion as simulated under the threshold model. In each panel, we plot information diffusion under centrality-based targeting (in red) versus random targeting (in blue). We consider three different hypothetical settings with respect to network structure. First, we plot information diffusion on networks that change over time, both due to exogenous link changes and endogenous link formation, as estimated in our empirical setting (solid lines). Second, we plot information diffusion in a hypothetical setting where networks remain stable over time (dashed lines). Third, we compare these plots to a hypothetical setting where networks change exogenously over time, but not endogenously in response to the treatment (dotted line). We use $\lambda = 2$ as in Beaman et al. (2021), and two different values for the number of periods: $T = 1$ and $T = 4$ (as in Beaman et al. 2021).

Precise targeting. It is worth noting that while targeting nodes based on standard centrality measures is often feasible (Banerjee et al., 2019), and can be optimal for total diffusion under the SIR model, it may be far from optimal for the threshold model (Jackson and Storms, 2023). Indeed, for information to spread under the threshold model, it is important to choose seeds with neighbours in common. While this may be more likely for central nodes, it is possible to choose even better seeds based on precise network data.

In general, choosing an optimal targeting strategy based on a particular network's structure is computationally intractable (Kempe et al., 2003). Yet, for a very small number of seeds, it is computationally feasible to identify the precise nodes that would maximize information diffusion. This requires a deterministic model of diffusion, such as the threshold model or the SIR model with $q = 1$, and a stable network. In their study of the diffusion of agricultural technology, Beaman et al. (2021) use network data to identify two top seeds based on a threshold model of diffusion. We will use our baseline network data to perform a similar computation, identifying the top two seeds under both the threshold model and the SIR model with $q = 1$.

In Figure A6, we present results from simulations in which we choose only two seeds, but the two seeds are chosen precisely for maximum diffusion, based on a stable baseline network and a deterministic diffusion process. In the case of total diffusion under the SIR model, this is equivalent to choosing seeds based on diffusion centrality, as in Figure 9. However, it is theoretically possible to choose even better seeds to maximize the number of ever-informed nodes, or to maximize diffusion under the threshold model, as in Beaman et al. (2021).

Consistent with Akbarpour et al. (2023) and Jackson and Storms (2023), we find that under a threshold model, precise targeting vastly outperforms other targeting strategies on a stable network (Figure A6). Precise targeting outperforms centrality-based targeting, and is 3 to 7 times more effective than random targeting.

This advantage shrinks or disappears once the network response is taken into account (Figure A6). Precise targeting relies on specific links remaining intact, and is very sensitive to changes in the network. The gains from precise targeting, relative to random targeting, are reduced by one-half to two-thirds across the range of $T \in \{1, 2, 3, 4\}$. Centrality-based targeting appears more robust, and in fact leads to higher information diffusion across both models. A big part of the reduction in gains is likely due to the fact that links change *exogenously*. That is, two nodes that initially share a neighbor may lose that shared neighbor due simply to the fact that links form and break over time. Indeed, this may explain why Beaman et al. (2021) find that even when using baseline network data to target optimal seeds, the true diffusion of technology over the long run is lower than predicted by their threshold model simulations.

A.6 Information Diffusion and Academic Performance

Previous work analyzing data from the same experiment has shown that Wikipedia access had a direct impact on students' English and Biology scores (Derksen et al., 2022). These two particular exam scores are pre-registered as primary outcomes, as almost all students complete the exams. English is compulsory for graduation, and Biology is the most popular subject, as it is required for entry into the most popular post-secondary programs including nursing and other medical programs. Students in the treated group had significantly higher scores than those in the control group, with effects concentrated among students with below-median scores at baseline (Appendix Table A13). For this subgroup, treated students scored 0.2 standard deviations higher in English, and 0.14 standard deviations higher in Biology, compared to below-median students in the control group. There was no significant effect for students with above-above median scores at baseline, for whom point estimates are zero or negative.

Reduced-form analysis cannot, however, identify the total effect of the intervention nor the extent of spillovers. First, we do not have any pure control schools, so we cannot directly estimate effects relative to a counterfactual in which no student was treated. Second, attempts to estimate spillovers using baseline network data will be biased due to the endogenous network response, and will likely understate their importance.

In this section, we calibrate our model to explore how the intervention affected students both directly and indirectly due to information diffusion, and to characterize the total effect of the intervention. In our setting, access to information may affect a student's academic performance and potentially even later life outcomes. As information diffuses through the network, academic outcomes for non-treated students might also be affected, and the effect on treated students might be amplified.

We model a student's academic score y_i as follows.

$$y_i = s_i + \tau(a_i)T_i + \tau(a_i) \sum_{j:g^1_{ij}=1} T_j Q_{ij} \tag{18}$$

Here, s_i is the student's counterfactual score, if the intervention had not taken place in their school. T_i represents treatment status. τ is the effect of becoming informed, which varies by the student's ability $a_i \in [0, 1]$. Q_{ij} are independent Bernoulli random variables with identical probability parameters q, each indicating information transmission between a particular pair of nodes.

In this model, academic scores depend on total information diffusion under an SIR process with one time period, and probability of information transmission q.[1] Here, by measuring total diffusion, we assume that students receive a direct effect τ each time they become informed. This measure is particularly relevant to our setting, where students receive different information from different peers. This also captures the fact that even treated nodes may benefit from links to other treated nodes, in line with our empirical findings.

We can again calibrate this model by matching moments to empirical moments in our data (see Appendix A.4 for details). To match moments, we will make use of final exam scores in both the year prior to and the year of the intervention. We did not collect form 4 exam scores for the year before the intervention, as this cohort did not include our study participants. We therefore restrict our dataset to include only form 2 and 3 exam scores. We again define y_i to be the average of the student's English and Biology scores. We normalize with respect to the distribution of grades from the previous year's cohort, within the same school and form, by subtracting the mean and dividing by the standard deviation.[2]

Our measure of student ability is, as above, their percentile in the year prior to the intervention. Then, we can map ability a_i to ability type θ_i as follows.

[1] Again, we assume $T = 1$ not only for simplicity, but also because control students report obtaining information primarily from treated students directly, and do not appear to form new links amongst themselves (see Sections 2.2 and 5.1).

[2] This normalization differs from that in Derksen et al. (2022), because for the purposes of this exercise we need to capture changes in both the control and treatment arms, relative to a fixed benchmark.

$$\theta_i = \begin{cases} \theta_H & \text{if } a_i >= 0.9, \\ \theta_L & \text{if } a_i < 0.9. \end{cases}$$

Abusing notation to write $\tau(\theta_i) = \mathbb{E}(\tau(a_i)|\theta_i)$, we obtain the following parameter estimates.

$$q = 0.67 \tag{19}$$
$$\tau(\theta_H) = 0.05 \tag{20}$$
$$\tau(\theta_L) = 0.11 \tag{21}$$

This exercise sheds some light on the likely magnitudes of the direct effects, spillovers, and total effect of the intervention. Taking a weighted average of $\tau(\theta_H)$ and $\tau(\theta_L)$, we obtain an average direct treatment effect of $\tau = 0.10$. These direct effects are not very different from the estimates we obtain by simply comparing treatment and control students at endline, as in equation 15 (Appendix A.4). However, this exercise does suggest that the total effect of the intervention, including spillovers, is larger. Indeed, equation 17 (Appendix A.4) indicates that control group students' end-of-year scores are 0.26 standard deviations higher than those of the previous cohort. The total effect of the intervention, based on the average score overall in the year of the intervention, could be as high as 0.29 standard deviations. This conclusion rests on the assumption that the previous year's cohort, in the same form, is a good comparison group for the cohort that received the intervention. Reassuringly, there does not appear to be significant grade inflation between the two cohorts: the average grade in Chichewa, a subject that should not be impacted by Wikipedia, is unchanged.

To explore the implications for academic welfare, we simulate expected total treatment effects at the node level. Here, we define the *total treatment effect* for student i to be the effect on their exam score relative to (and normalized with respect to) a counterfactual in which the intervention did not take place. Importantly, control students may have non-zero total treatment effects.

$$\begin{aligned} \mathbb{E}(y_i - s_i) &= \tau(\theta_i)T_i + q\tau(\theta_i)\sum_{j:g_{ij}^1=1} T_j \\ &= \begin{cases} 0.05T_i + 0.67 * 0.05 \sum_{j:g_{ij}^1=1} T_j & \text{if } \theta_i = \theta_H, \\ 0.11T_i + 0.67 * 0.11 \sum_{j:g_{ij}^1=1} T_j & \text{if } \theta_i = \theta_L. \end{cases} \end{aligned}$$

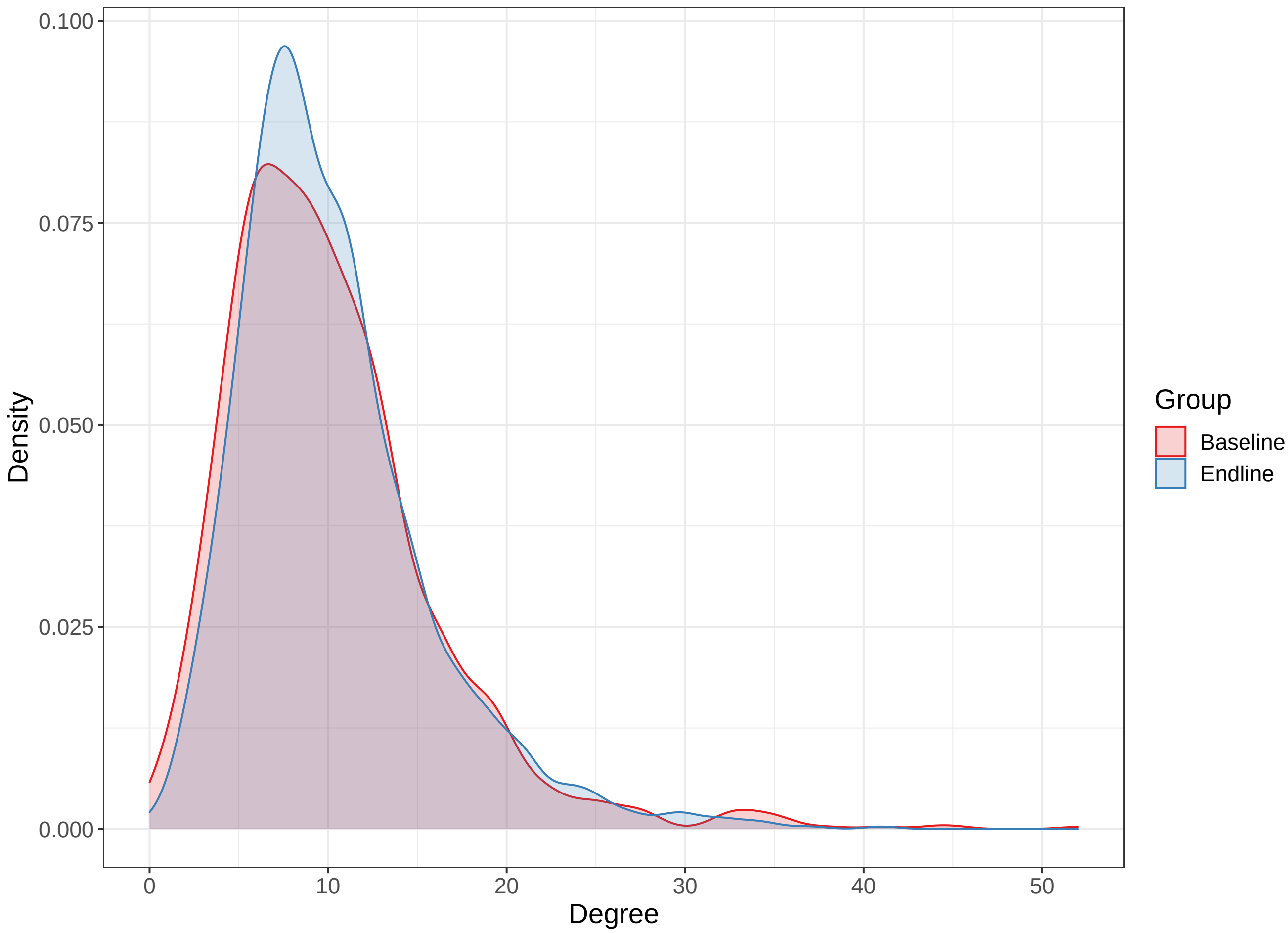
In Figure A13, we plot total treatment effects under network-based targeting (in red) versus random targeting (in blue), again simulating 1000 100-node networks under each set of parameters. We target nodes with top diffusion centralities; because $T = 1$ this is equivalent to degree-centrality targeting. We again consider networks that change over time (solid lines), hypothetical networks that remain stable over time (dashed lines), and hypothetical networks that change exogenously but not endogenously in response to the treatment (dotted line).

Network-based targeting increases academic welfare, as captured by average total treatment effects, with any number of seeds, though the endogenous network response lessens this increase considerably (Panel A of Figure A13). The number of additional random seeds needed to match the effect achieved under network-based targeting is at most 5, compared to 12 under stable networks. The fact that network-based targeting continues to outperform random targeting despite larger direct treatment effects for lower-ability students is due to the fact that spillovers are large in our context. That is, many students benefit indirectly even if treated students are primarily high-ability. Our estimate of $q = 0.67$ relies heavily on the assumption that the previous year's cohort serves as a good counterfactual, yet, even with smaller values of q network-based targeting dominates (Panel B). Random targeting appears to only outperform network-based targeting in networks where information transmission is rare ($q \leq 0.15$).

However, the gains from network-based targeting are attributed disproportionately to high-ability students, and therefore lead to a relatively wider achievement gap (Panels C and D of Figure A13). With 20 seeds, the average total treatment effect for high-ability students jumps from 0.09 standard deviations under random targeting to 0.14 under network-based targeting. This jump is much smaller for lower-ability students, whose total treatment effects are large under both random (0.17 standard deviations) and network-based targeting (0.19 standard deviations). The intervention narrows the achievement gap under both targeting strategies, but this effect is much stronger under random targeting (Panels C and D, solid lines).

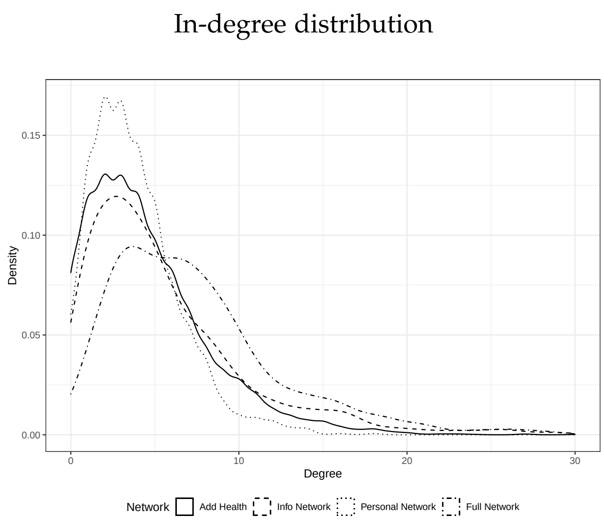
Appendix Figures and Tables

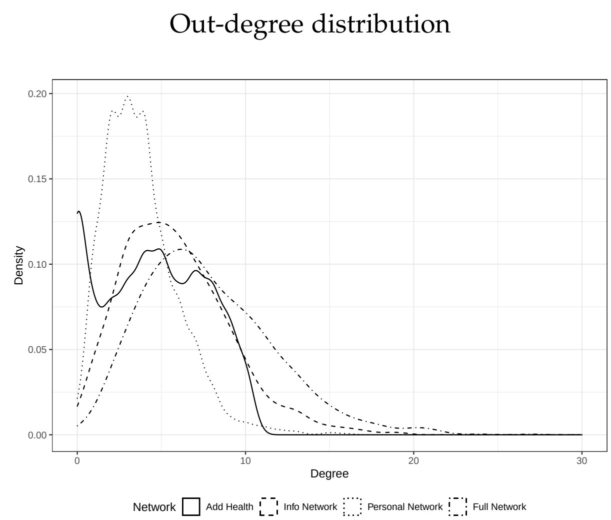
Appendix Figure A1: Degree Distribution at Baseline and Endline, Information Network



Notes: Degree distribution at baseline and endline. Information network. Sample restricted to nodes observed at both times.

Appendix Figure A2: Degree Distribution in Our Data vs AddHealth

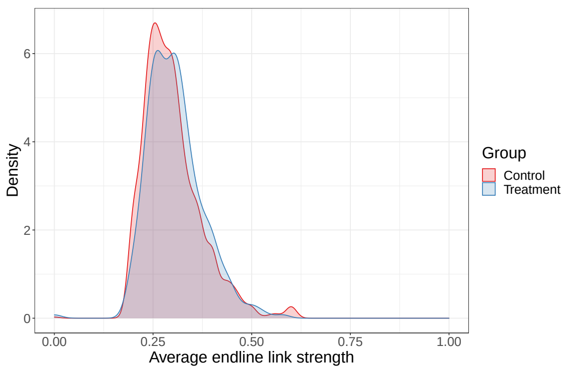




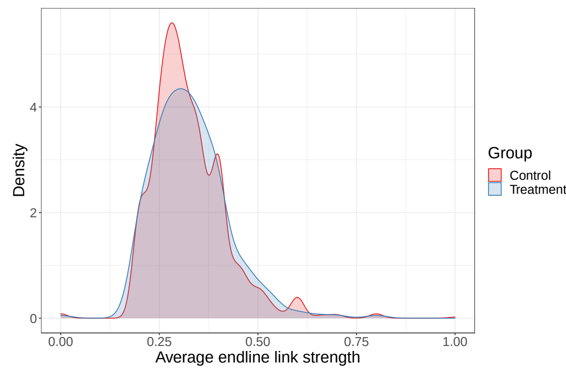
Notes: Degree distributions computed over our data for the information, personal and full networks; and the National Longitudinal Study of Adolescent Health ("AddHealth") obtained from https://www.icpsr.umich.edu/web/ICPSR/studies/21600/datasets/0003/variables/ODGX2?archive=icpsr and https://www.icpsr.umich.edu/web/ICPSR/studies/21600/datasets/0003/variables/IDGX2?archive=icpsr

Appendix Figure A3: Link Strength and Link Dynamics

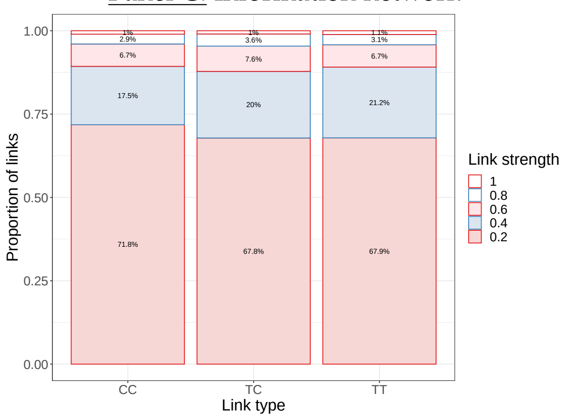
Panel A: Information network



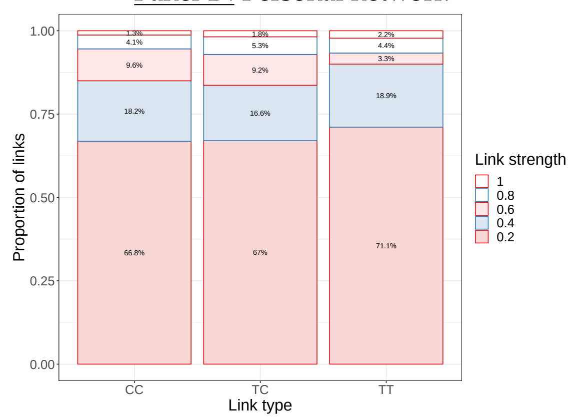
Panel B: Personal network



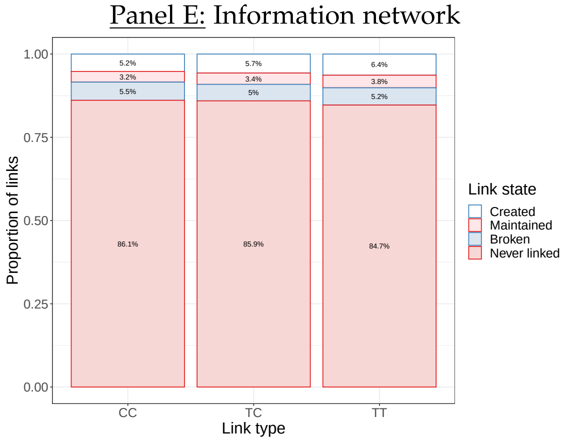
Panel C: Information network



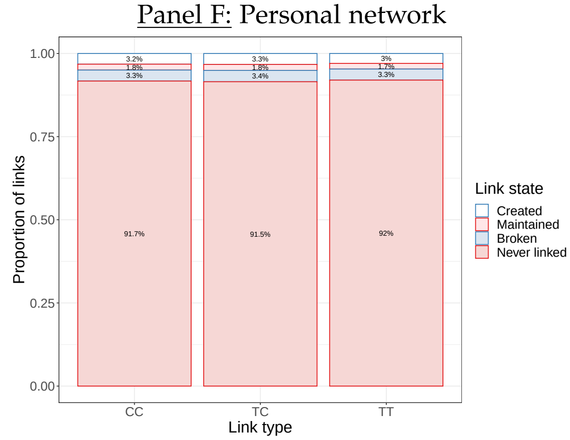
Panel D: Personal network



Panel E: Information network

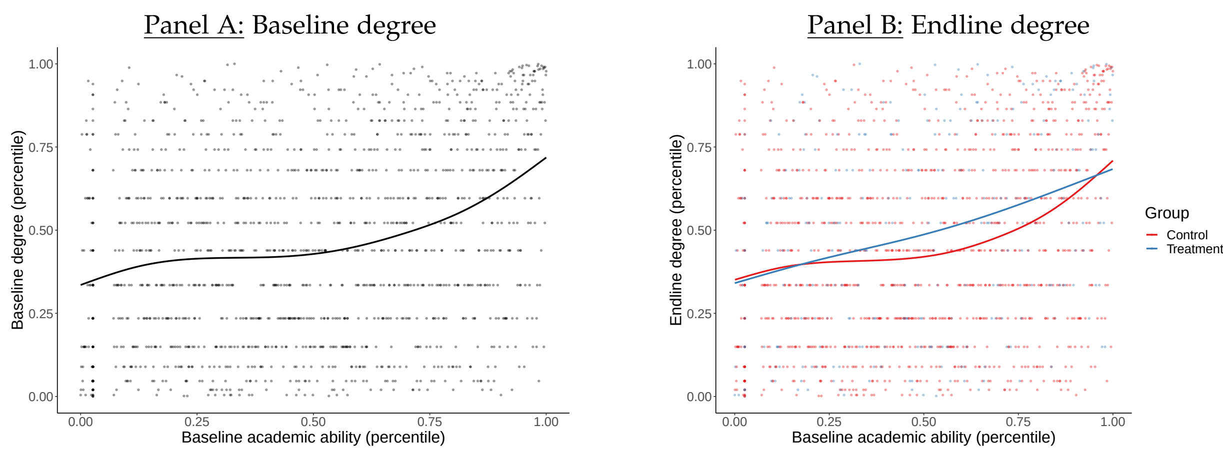


Panel F: Personal network



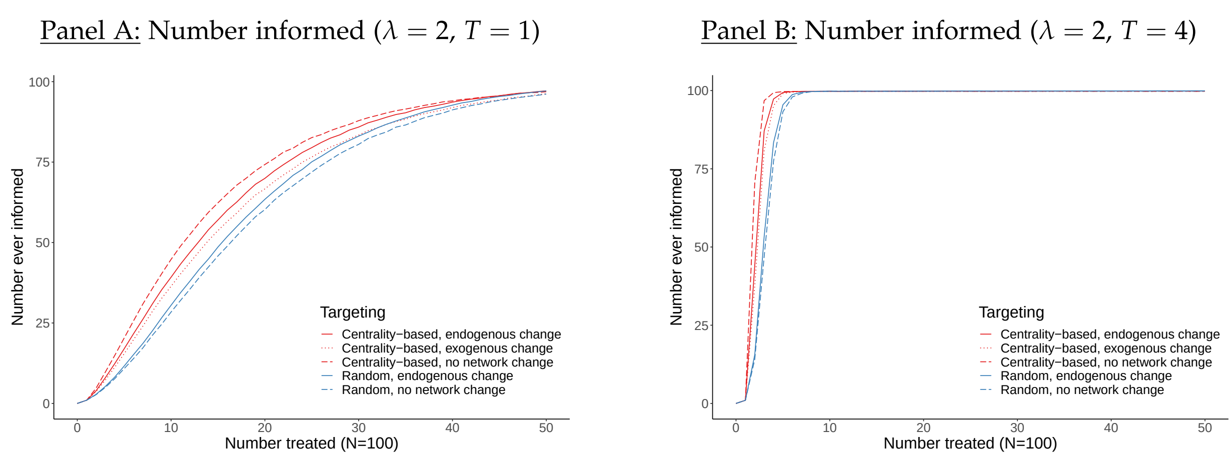
Notes: Panels A-D: link strength between pairs of nodes. Each link consists of five sublinks (see Table 1), strength is defined as the fraction of sublinks present. In all four panels, strength is calculated conditional on the presence of a link. Panels E-F: link dynamics between baseline and endline.

Appendix Figure A4: Baseline Academic Ability and Network Degree



Notes: Correlation between baseline academic ability (percentile) and network degree (percentile) at baseline and endline.

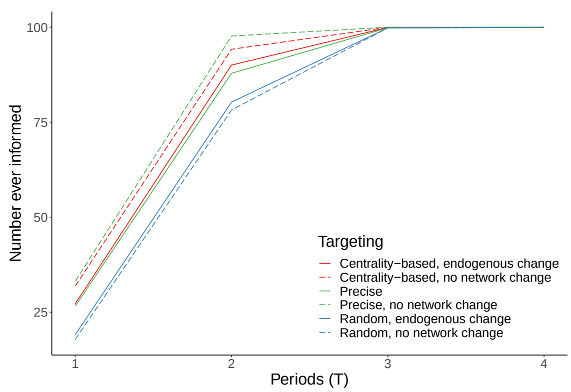
Appendix Figure A5: Threshold Model, Centrality-Based Versus Random Targeting



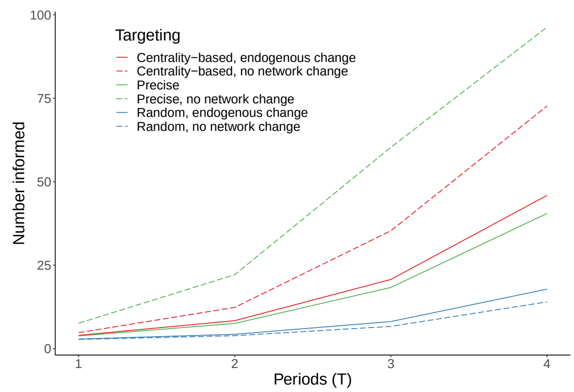
Notes: Simulations of 100-node networks, with 1000 replications for each set of parameter values. Network-based targeting involves targeting the top nodes by diffusion centrality, with parameters $q = 1$ and T matching the parameter of the diffusion model.

Appendix Figure A6: Centrality-Based Targeting Versus Precise Targeting (2 seeds)

Panel A: SIR model, ever informed ($q = 1$)



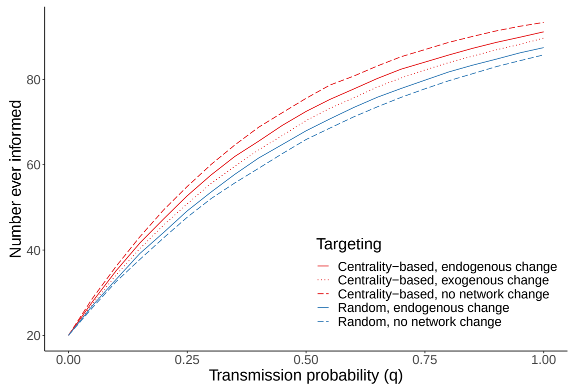
Panel B: Threshold model ($\lambda = 2$)



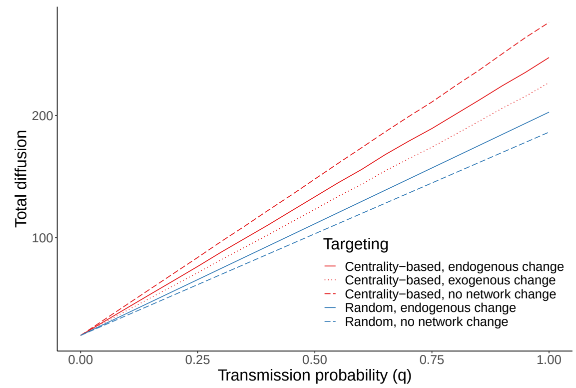
Notes: Simulations of 100-node networks, with 1000 replications for each set of parameter values. Optimal targeting involves targeting the two nodes that maximize diffusion on the baseline network. Network-based targeting involves targeting the top nodes by diffusion centrality, with parameters $q = 1$ and T matching the parameter of the diffusion model.

Appendix Figure A7: SIR Model, Degree Centrality Targeting

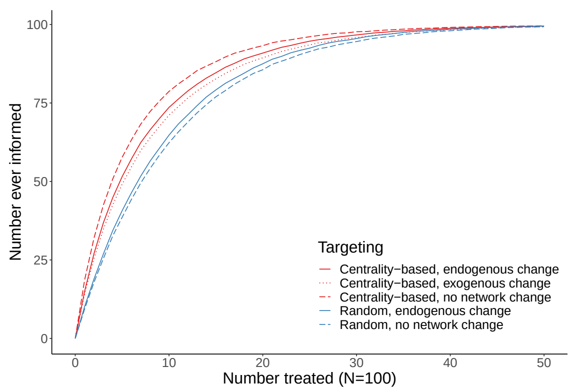
Panel A: Ever informed, 20 seeds and $T = 1$



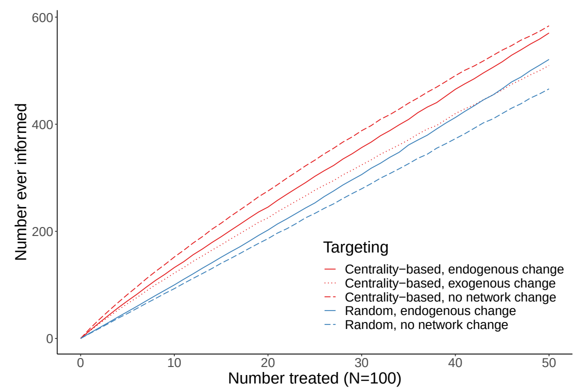
Panel B: Total Diffusion, 20 seeds and $T = 1$



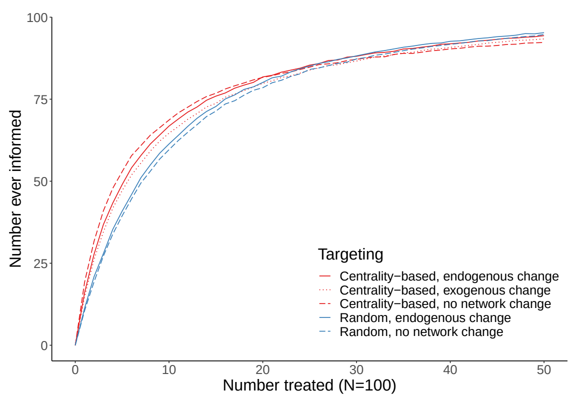
Panel C: Ever informed ($q = 1$, $T = 1$)



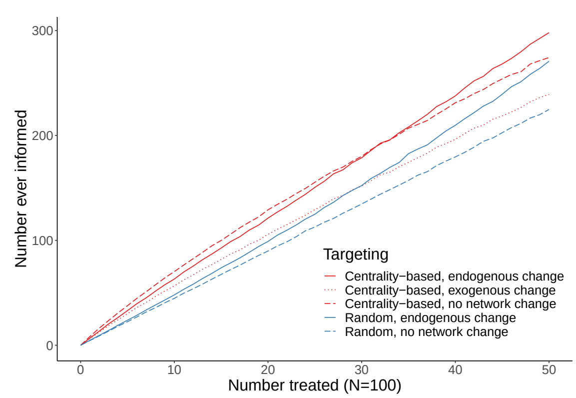
Panel D: Total diffusion ($q = 1$, $T = 1$)



Panel E: Ever informed ($q = 0.1$, $T = 4$)



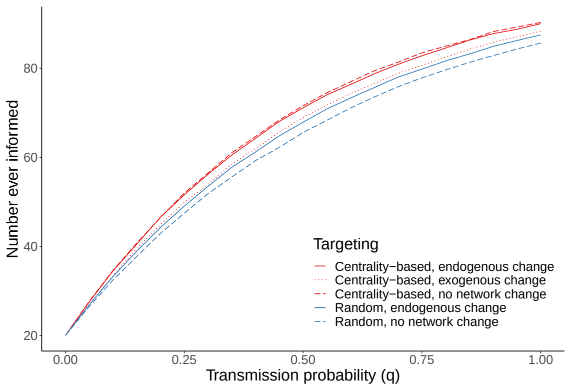
Panel F: Total diffusion ($q = 0.1$, $T = 4$)



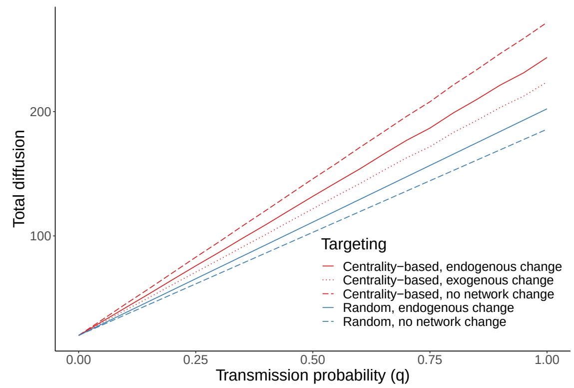
Notes: Simulations of 100-node networks, with 1000 replications for each set of parameter values. Network-based targeting involves targeting the top nodes by degree. $q^* = 0.1$ and $T^* = 4$ are set to equal the reciprocal of the top eigenvalue and diameter of the graph respectively, as in Banerjee et al. (2019).

Appendix Figure A8: SIR Model, Eigenvector Centrality Targeting

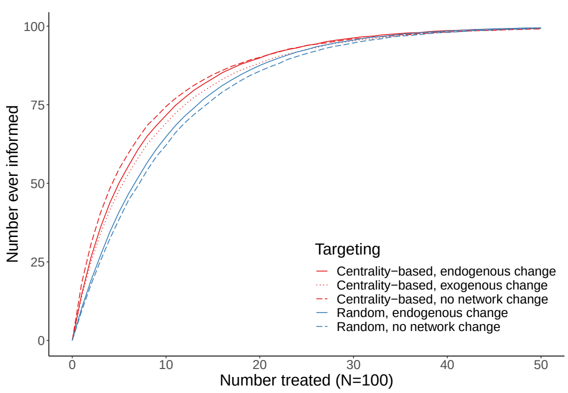
Panel A: Ever informed, 20 seeds and $T = 1$



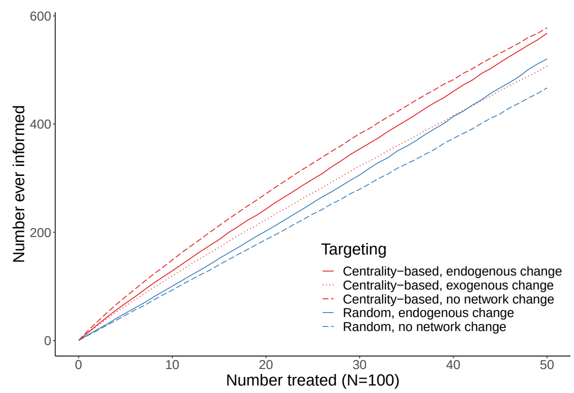
Panel B: Total Diffusion, 20 seeds and $T = 1$



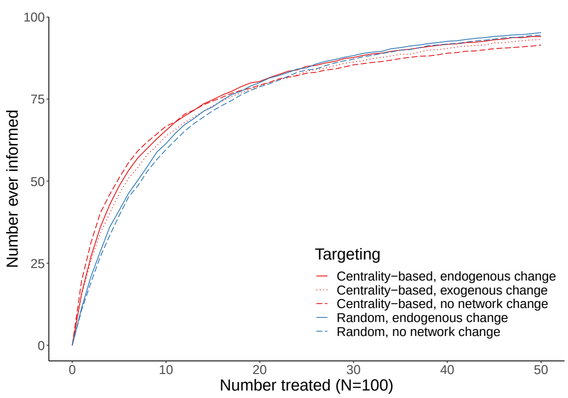
Panel C: Ever informed ($q = 1$, $T = 1$)



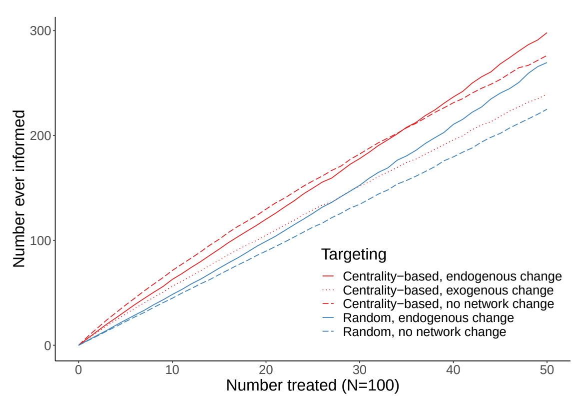
Panel D: Total diffusion ($q = 1$, $T = 1$)



Panel E: Ever informed ($q = 0.1$, $T = 4$)



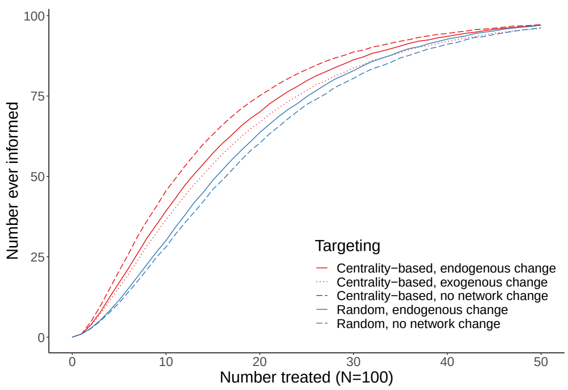
Panel F: Total diffusion ($q = 0.1$, $T = 4$)



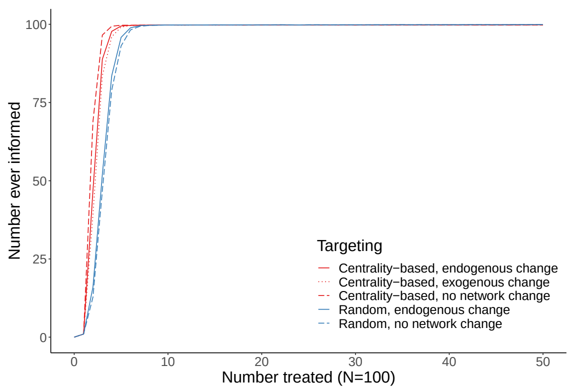
Notes: Simulations of 100-node networks, with 1000 replications for each set of parameter values. Network-based targeting involves targeting the top nodes by eigenvector centrality. $q^* = 0.1$ and $T^* = 4$ are set to equal the reciprocal of the top eigenvalue and diameter of the graph respectively, as in Banerjee et al. (2019).

Appendix Figure A9: Threshold Model, Degree Centrality Targeting

Panel A: Number informed ($\lambda = 2$, $T = 1$)



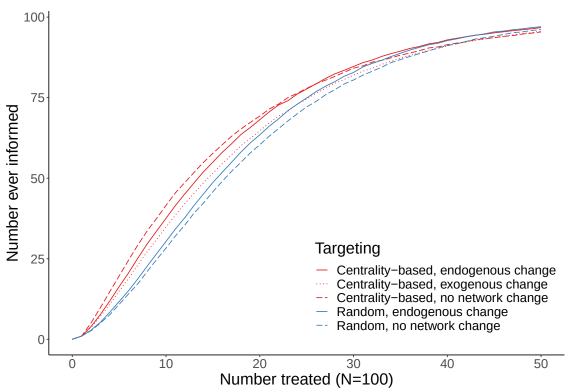
Panel B: Number informed ($\lambda = 2$, $T = 4$)



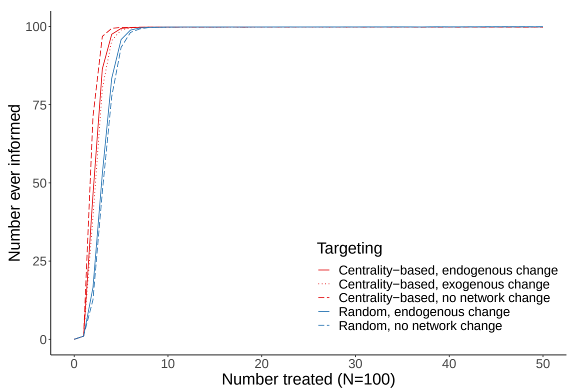
Notes: Simulations of 100-node networks, with 1000 replications for each set of parameter values. Network-based targeting involves targeting the top nodes by degree centrality.

Appendix Figure A10: Threshold Model, Eigenvector Centrality Targeting

Panel A: Number informed ($\lambda = 2$, $T = 1$)



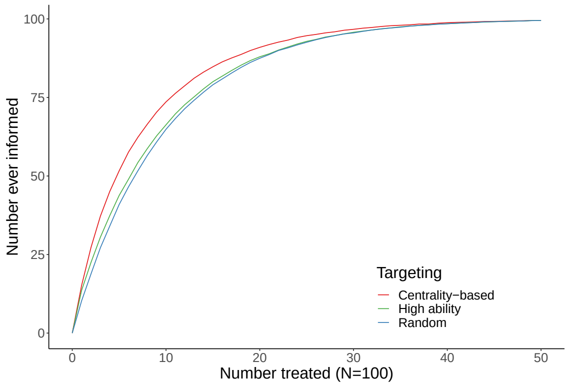
Panel B: Number informed ($\lambda = 2$, $T = 4$)



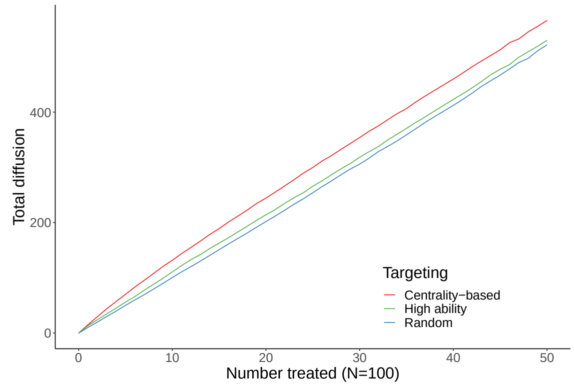
Notes: Simulations of 100-node networks, with 1000 replications for each set of parameter values. Network-based targeting involves targeting the top nodes by eigenvector centrality.

Appendix Figure A11: SIR Model, Centrality-Based Versus High-Ability Targeting

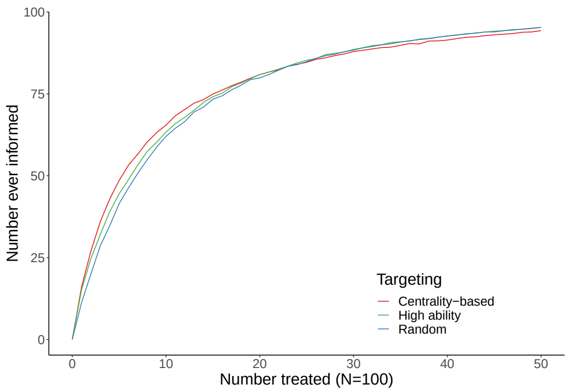
Panel A: SIR, ever informed ($q = 1$, $T = 1$)



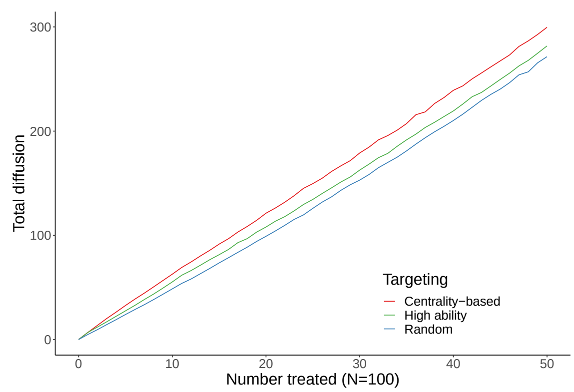
Panel B: SIR, total diffusion ($q = 1$, $T = 1$)



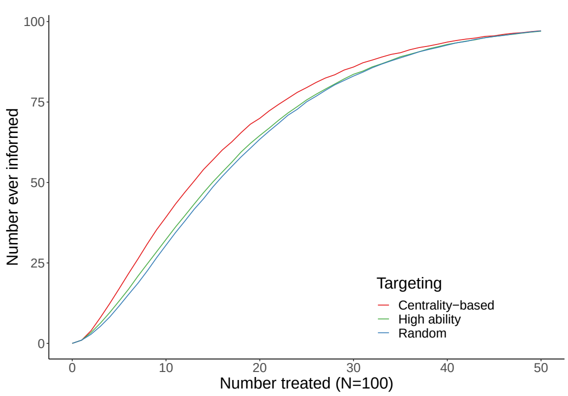
Panel C: SIR, ever informed ($q = 0.1$, $T = 4$)



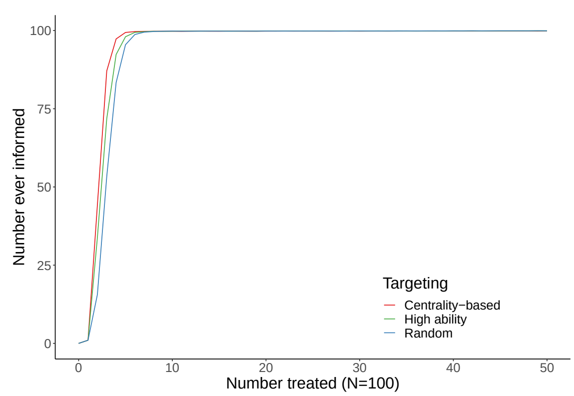
Panel D: SIR, total diffusion ($q = 0.1$, $T = 4$)



Panel E: Threshold model ($\lambda = 2$, $T = 1$)

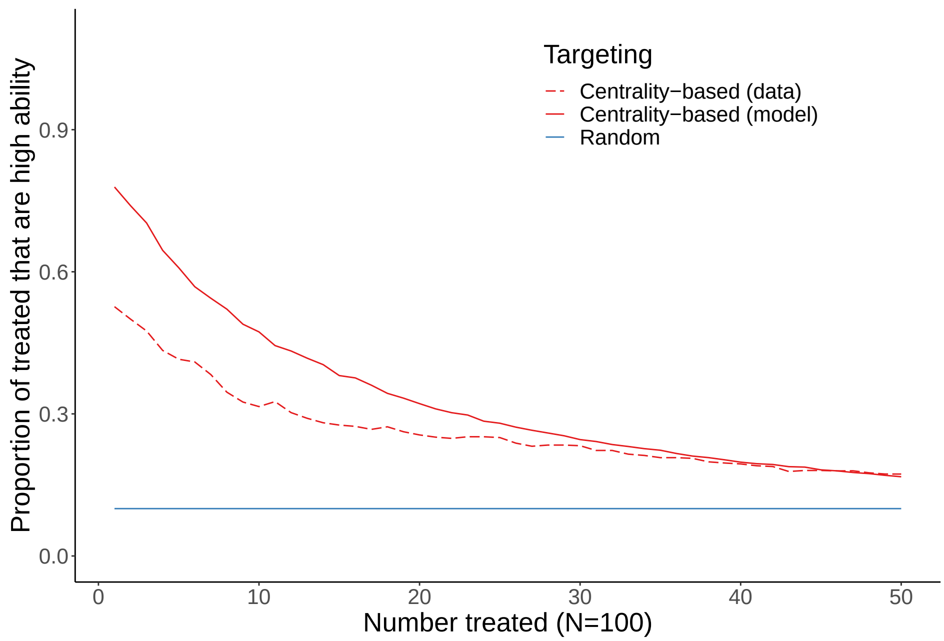


Panel F: Threshold model ($\lambda = 2$, $T = 4$)



Notes: Simulations of 100-node networks, with 1000 replications for each set of parameter values. Network-based targeting involves targeting the top nodes by diffusion centrality, with parameters q and T matching the parameters of the SIR diffusion model.

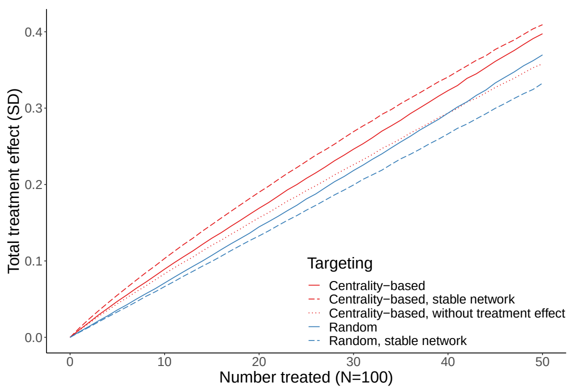
Appendix Figure A12: Centrality-Based Targeting and Baseline Ability



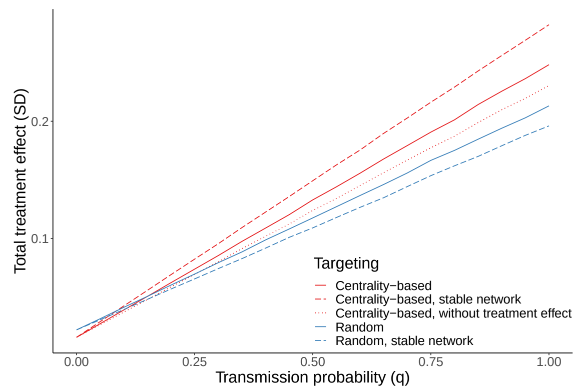
Notes: Simulations of 100-node networks, with 1000 replications for each set of parameter values. Diffusion centrality calculated using $q = 1$ and $T = 1$ (equivalent to degree). The dashed line corresponds to the share of top-centrality nodes that are high-ability in our baseline data.

Appendix Figure A13: Simulated Total Treatment Effects on Academic Performance

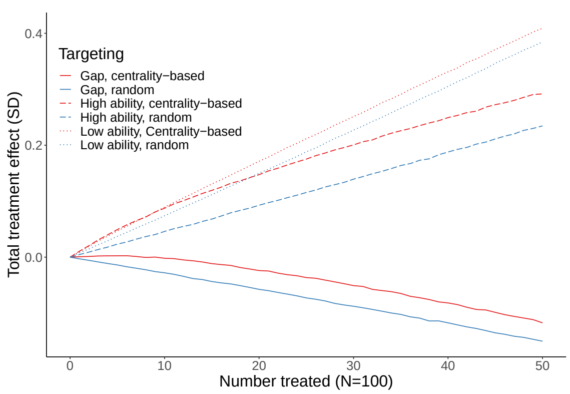
Panel A: Average effect by number of seeds



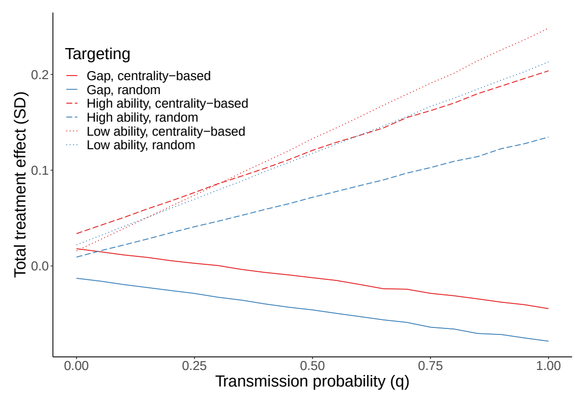
Panel B: Average effect by q



Panel C: Heterogeneous effects by number of seeds



Panel D: Heterogeneous effects by q



Notes: Simulations of 100-node networks, with 1000 replications for each set of parameter values. Network-based targeting involves targeting the top nodes by diffusion centrality with $T = 1$, which is equivalent to targeting by degree.

Appendix Table A1: Comparison to Alternative Networks

	Full Network		Coleman High School		Diffusion of Microfinance	
	Mean	SD	Mean	SD	Mean	SD
Degree	12.71	5.98	7.83	3.43	8.43	5.92
Eigenvector Centrality	0.33	0.18	0.29	0.22	0.07	0.13
Number of length-2 walks	216.45	98.32	80.74	39.62	115.79	113.33
Diffusion	3.52	1.75	4.24	2.40	2.60	3.02
Betweenness	0.01	0.01	0.02	0.03	0.00	0.01

Notes: Comparison to alternative networks. "Coleman High School" is the high-school network from Coleman (1964). "Diffusion of Microfinance" from Banerjee et al. (2013). "SD" refers to the standard deviation across observations.

Appendix Table A2: Dynamics of Link Formation by Baseline Link and Ability (No Controls)

	Link at endline	Link at endline
	(1)	(2)
Treat-Control, No Baseline Link	0.503***	—
	(0.187)	
	p = 0.031	
Treat-Control, Baseline Link	3.70***	—
	(1.26)	
	p = 0.008	
Treat-Treat, No Baseline Link	1.27***	—
	(0.474)	
	p = 0.017	
Treat-Treat, Baseline Link	5.12*	—
	(2.94)	
	p = 0.122	
Treat-Control, High-High	—	4.24
		(3.20)
		p = 0.221
Treat-Control, Low-High(treated)	—	0.812
		(0.738)
		p = 0.549
Treat-Control, Low(treated)-High	—	0.949
		(0.740)
		p = 0.223
Treat-Control, Low-Low	—	0.596***
		(0.226)
		p = 0.043
Treat-Treat, High-High	—	7.34
		(7.83)
		p = 0.391
Treat-Treat, Low-High	—	1.41
		(1.39)
		p = 0.432
Treat-Treat, Low-Low	—	1.69***
		(0.574)
		p = 0.011
Baseline Link	31.3***	—
	(0.725)	
High-High	—	13.7***
		(1.75)
Low-High	—	3.88***
		(0.351)
Constant	5.71***	7.58***
	(0.106)	(0.129)
R^2	0.1045	0.0058
Observations	82,711	82,711

Notes: Dyadic regressions. Unit of observation is a pair of students in the same form and school, *i* and *j*. The outcome is coded as 100 if there is a connection and 0 otherwise. "Treat-Control" is equal to 1 if *i* is treated and *j* is control, or vice-versa. "Treat-Treat" is equal to 1 if both *i* and *j* are in the treatment group. The sample consists of students present at both baseline and endline. Heteroskedasticity-robust standard errors in parentheses and randomization inference p-value with "p = ". Stars represent classical inference p-values with *** p<0.01; ** p<0.05; * p<0.1.

Appendix Table A3: Dynamics of Link Formation, by Baseline Link and Ability (Full Interaction)

	Link at endline	Link at endline
	(1)	(2)
Treat-Control, No Baseline Link, High-High	2.25	1.28
	(3.00)	(2.95)
	p = 0.483	p = 0.694
Treat-Control, No Baseline Link, Low-High(treated)	0.557	0.161
	(0.641)	(0.629)
	p = 0.528	p = 0.860
Treat-Control, No Baseline Link, Low(treated)-High	0.405	0.485
	(0.633)	(0.615)
	p = 0.561	p = 0.471
Treat-Control, No Baseline Link, Low-Low	0.496**	0.488**
	(0.200)	(0.197)
	p = 0.043	p = 0.048
Treat-Control, Baseline Link, High-High	9.07	9.10
	(7.46)	(7.38)
	p = 0.211	p = 0.204
Treat-Control, Baseline Link, Low-High(treated)	2.67	2.33
	(3.19)	(3.15)
	p = 0.517	p = 0.569
Treat-Control, Baseline Link, Low(treated)-High	8.41**	8.33**
	(3.31)	(3.29)
	p = 0.012	p = 0.012
Treat-Control, Baseline Link, Low-Low	2.88*	2.94**
	(1.47)	(1.46)
	p = 0.056	p = 0.048
Treat-Treat, No Baseline Link, High-High	6.58	5.32
	(8.72)	(8.74)
	p = 0.414	p = 0.507
Treat-Treat, No Baseline Link, Low-High	1.04	1.12
	(1.21)	(1.19)
	p = 0.449	p = 0.414
Treat-Treat, No Baseline Link, Low-Low	1.27**	1.10**
	(0.513)	(0.503)
	p = 0.022	p = 0.049
Treat-Treat, Baseline Link, High-High	-5.65	-5.87
	(13.9)	(13.7)
	p = NA	p = NA
Treat-Treat, Baseline Link, Low-High	9.16	8.82
	(6.27)	(6.20)
	p = 0.177	p = 0.189
Treat-Treat, Baseline Link, Low-Low	4.58	4.85
	(3.42)	(3.37)
	p = 0.211	p = 0.182
Baseline Link, High-High	43.2***	38.6***
	(4.32)	(4.26)
Baseline Link, Low-High	36.3***	31.5***
	(1.44)	(1.42)
Baseline Link, Low-Low	29.5***	24.8***
	(0.847)	(0.851)
No Baseline Link, High-High	7.14***	8.10***
	(1.62)	(1.60)
No Baseline Link, Low-High	1.86***	2.28***
	(0.305)	(0.299)
Constant	5.33***	
	(0.114)	
Controls	–	Baseline link, same class, same gender, form FE
R^2	0.1075	0.1375
Observations	82,711	82,711

Notes: Dyadic regressions. Unit of observation is a pair of students in the same form and school, *i* and *j*. The outcome is coded as 100 if there is a connection and 0 otherwise. "Treat-Control" is equal to 1 if *i* is treated and *j* is control, or vice-versa. "Treat-Treat" is equal to 1 if both *i* and *j* are in the treatment group. "High-High" ("Low-Low") means that *i* and *j* are of high (low) ability. "Low-High(treated)" means that one student is high ability and treated, and the other is of low ability. "Low(treated)-High" means that one student is low ability and treated, and the other is of high ability. "Low-High" is a link between one high-ability and a low-ability student. Specification (2) has baseline link, same-class and same-gender controls, and include form fixed effects. The sample consists of students present at both baseline and endline. Heteroskedasticity-robust standard errors in parentheses and randomization inference p-value with "p = ". Stars represent classical inference p-values with *** p<0.01; ** p<0.05; * p<0.1.

Appendix Table A4: Dynamics of Link Formation

	Info link created	Info link broken	Personal link created	Personal link broken	Full network link created	Full network link broken
Treat-Control	0.465***	-0.525***	0.069	0.081	0.423**	-0.456***
	(0.169)	(0.162)	(0.132)	(0.133)	(0.184)	(0.174)
	p = 0.027	p = 0.024	p = 0.655	p = 0.626	p = 0.074	p = 0.064
Treat-Treat	1.05**	-0.378	-0.281	-0.045	0.522	-0.366
	(0.428)	(0.390)	(0.301)	(0.316)	(0.447)	(0.419)
	p = 0.030	p = 0.465	p = 0.426	p = 0.906	p = 0.330	p = 0.506
R^2	0.0273	0.0388	0.0163	0.0235	0.0285	0.0405
Observations	82,711	82,711	82,711	82,711	82,711	82,711

Notes: Dyadic regressions. Unit of observation is a pair of students in the same form and school, *i* and *j*. The outcome is coded as 100 if there is a connection and 0 otherwise. "Treat-Control" is equal to 1 if *i* is treated and *j* is control, or vice-versa. "Treat-Treat" is equal to 1 if both *i* and *j* are in the treatment group. Covariates are interacted with the indicator for presence of link at the baseline. Specifications have baseline link, same-class and same-gender controls, and include form fixed effects. The sample consists of students present at both baseline and endline. Heteroskedasticity-robust standard errors in parentheses and randomization inference p-value with "p = ". Stars represent classical inference p-values with *** $p<0.01$; ** $p<0.05$; * $p<0.1$.

Appendix Table A5: Dynamics of Link Formation by Baseline Link and Ability - F-tests

	F-test p-value (1)	F-test p-value (2)
$\mathcal{H}_0$: Treat-Control, No Baseline Link = Treat-Control, Baseline Link	0.011	—
$\mathcal{H}_0$: Treat-Treat, No Baseline Link = Treat-Treat, Baseline Link	0.161	—
$\mathcal{H}_0$: Treat-Control, High-High = Treat-Control, Low-High(treated)	—	0.361
$\mathcal{H}_0$: Treat-Control, High-High = Treat-Control, Low(treated)-High	—	0.530
$\mathcal{H}_0$: Treat-Control, High-High = Treat-Control, Low-Low	—	0.421
$\mathcal{H}_0$: Treat-Control, Low-High(treated) = Treat-Control, Low(treated)-High	—	0.315
$\mathcal{H}_0$: Treat-Control, Low-High(treated) = Treat-Control, Low-Low	—	0.587
$\mathcal{H}_0$: Treat-Control, Low(treated)-High = Treat-Control, Low-Low	—	0.499
$\mathcal{H}_0$: Treat-Control, High-High = Treat-Control, Low-High(treat) = Treat-Control, Low(treated)-High = Treat-Control, Low-Low	—	0.010
$\mathcal{H}_0$: Treat-Treat, High-High = Treat-Control, Low-High	—	0.683
$\mathcal{H}_0$: Treat-Treat, High-High = Treat-Control, Low-Low	—	0.671
$\mathcal{H}_0$: Treat-Treat, Low-High = Treat-Control, Low-Low	—	0.956
$\mathcal{H}_0$: Treat-Treat, High-High = Treat-Treat, Low-High = Treat-Treat, Low-Low	—	0.913

Notes: Joint F-tests of the hypotheses along rows. Hypoteses in Column (1) and (2) are relative to the specification in Column (1) and (2) of Table 5, respectively. p-values are shown.

Appendix Table A6: Directed Links with Second-order Effects

	Information	Full	Personal
Treat-to-Control Link	0.346** (0.140) p = 0.053	0.275* (0.159) p = 0.207	-0.078 (0.111) p = 0.530
Treat-to-Control Link x Number of Treated Friends of i at Baseline	0.056 (0.098) p = 0.665	-0.029 (0.090) p = 0.823	0.039 (0.087) p = 0.719
Treat-to-Control Link x Number of Friends of i at Baseline	-0.020 (0.031) p = 0.588	-0.017 (0.029) p = 0.632	-0.050 (0.034) p = 0.225
Treat-to-Control Link x Number of Treated Friends of j at Baseline	0.233** (0.102) p = 0.265	0.257*** (0.093) p = 0.323	0.231*** (0.088) p = 0.647
Treat-to-Control Link x Number of Friends of j at Baseline	-0.040 (0.034) p = 0.435	-0.058* (0.030) p = 0.303	-0.041 (0.036) p = 0.632
Control-to-Treat Link	0.738*** (0.146) p = 0.001	0.697*** (0.164) p = 0.003	0.120 (0.115) p = 0.351
Control-to-Treat Link x Number of Treated Friends of i at Baseline	0.233** (0.100) p = 0.271	0.219** (0.093) p = 0.478	0.332*** (0.088) p = 0.256
Control-to-Treat Link x Number of Friends of i at Baseline	-0.047 (0.029) p = 0.331	-0.064** (0.028) p = 0.231	-0.088** (0.035) p = 0.168
Control-to-Treat Link x Number of Treated Friends of j at Baseline	-0.082 (0.110) p = 0.612	-0.020 (0.099) p = 0.893	0.081 (0.094) p = 0.472
Control-to-Treat Link x Number of Friends of j at Baseline	0.065* (0.036) p = 0.239	0.036 (0.033) p = 0.456	-0.050 (0.035) p = 0.243
Treat-to-Treat Link	1.06*** (0.279) p = 0.002	0.742** (0.305) p = 0.067	-0.235 (0.205) p = 0.365
Treat-to-Treat Link x Number of Treated Friends of i at Baseline	0.299 (0.216) p = 0.279	0.123 (0.191) p = 0.716	0.204 (0.174) p = 0.646
Treat-to-Treat Link x Number of Friends of i at Baseline	-0.081 (0.060) p = 0.249	-0.081 (0.055) p = 0.277	-0.072 (0.059) p = 0.426
Treat-to-Treat Link x Number of Treated Friends of j at Baseline	0.185 (0.218) p = 0.532	0.357* (0.197) p = 0.260	0.335* (0.172) p = 0.373
Treat-to-Treat Link x Number of Friends of j at Baseline	0.103 (0.069) p = 0.210	0.069 (0.062) p = 0.378	-0.072 (0.063) p = 0.421
Number of Treated Friends of i at Baseline	-0.051 (0.043) p = 0.474	-0.044 (0.042) p = 0.572	-0.094** (0.040) p = 0.196
Number of Friends of i at Baseline	0.030** (0.013) p = 0.349	0.054*** (0.013) p = 0.260	0.060*** (0.017) p = 0.072
Number of Treated Friends of j at Baseline	-0.105** (0.046) p = 0.177	-0.121*** (0.043) p = 0.107	-0.095** (0.042) p = 0.199
Number of Friends of j at Baseline	0.265*** (0.016) p = 0.392	0.228*** (0.014) p = 0.280	0.104*** (0.017) p = 0.163
R^2	0.1046	0.1272	0.0974
Observations	165,422	165,422	165,422

Notes: Dyadic regressions. Unit of observation is a pair of students in the same form and school, *i* and *j*. The outcome is coded as 100 if *i* named *j* as a contact and 0 otherwise. "Treat-to-Control" is a dummy equal to 1 if *i* is treated and *j* is control, and other covariates are defined similarly. Column "Information" refers to information network, followed by the personal and full networks. Specifications have number of treated friends of *i* and *j* at the baseline, number of friends of *i* and *j* at baseline, baseline link, same-class and same-gender controls, and include form fixed effects. Heteroskedasticity-robust standard errors in parentheses and randomization inference p-value with "p = ". Stars represent classical inference p-values with *** p<0.01; ** p<0.05; * p<0.1.

Appendix Table A7: Information Access and Centrality in the Contact Network

	Degree	Eigenvector	Number of Length-2 Walks	Diffusion	Betweenness	Average Link Strength
Treatment	-0.093	-0.040	-1.41	-0.037	0.013	0.010*
	(0.246)	(0.058)	(3.55)	(0.060)	(0.072)	(0.006)
	p = 0.705	p = 0.497	p = 0.670	p = 0.536	p = 0.849	p = 0.071
Control Mean	10.8	0.000	143.9	0.000	0.000	0.464
R^2	0.285	0.281	0.361	0.237	0.099	0.150
Observations	1,402	1,402	1,402	1,402	1,402	1,402

Notes: Estimated differences between treated and control students in the contact network for five measures of centrality (degree, eigenvector, number of length-2 walks, diffusion, and betweenness centralities) and average link strength (equation 3). Eigenvector, diffusion and betweenness centralities are normalized. Contact links are identified based on the survey question "*[1,2,3] days ago, did you just hang out, have conversations or play with friends?*" Column 6 is calculated based on the fraction of days during which the pair spent time together. Regressions have controls for baseline measure of the outcome, gender, SES, stratification bins and class fixed effects. "Control Mean" represents the mean of the outcome in the control arm. The sample consists of students present at both baseline and endline (N=1,402). Heteroskedasticity-robust standard errors in parentheses and randomization inference p-value with "p = ". Stars represent classical inference p-values with *** p<0.01; ** p<0.05; * p<0.1.

Appendix Table A8: Heterogeneous Effects

	Degree	Eigenvector	Number of Length-2 Walks	Diffusion	Betweenness	Average Link Strength
Panel A. By use of the digital Library						
Treatment	0.426	0.065	7.28	0.059	0.024	0.005
	(0.339)	(0.071)	(4.98)	(0.071)	(0.080)	(0.006)
Treatment x High Browsing	1.03*	0.227**	9.69	0.247**	0.414**	0.005
	(0.550)	(0.115)	(7.59)	(0.116)	(0.165)	(0.008)
Control Mean	10.1	0.000	142.9	0.000	0.000	0.299
R^2	0.512	0.416	0.630	0.416	0.373	0.187
Observations	1,402	1,402	1,402	1,402	1,402	1,402
Panel B. By academic ability						
Treatment	0.759**	0.153**	10.3**	0.152**	0.162	0.008*
	(0.306)	(0.069)	(4.13)	(0.069)	(0.099)	(0.005)
	p = 0.013	p = 0.019	p = 0.020	p = 0.020	p = 0.027	p = 0.089
Treatment x High Ability	1.87*	0.269	18.6	0.316	0.707**	-0.010
	(1.04)	(0.192)	(14.1)	(0.198)	(0.300)	(0.013)
	p = 0.486	p = 0.692	p = 0.745	p = 0.572	p = 0.142	p = 0.354
Mean Outcome Control Arm	10.1	0.000	142.9	0.000	0.000	0.299
R^2	0.512	0.415	0.631	0.415	0.373	0.187
Observations	1,402	1,402	1,402	1,402	1,402	1,402
Panel C. By SES						
Treatment	1.10***	0.256**	14.8***	0.260**	0.282*	0.008
	(0.423)	(0.103)	(5.18)	(0.104)	(0.161)	(0.007)
	p = 0.002	p = 0.001	p = 0.002	p = 0.002	p = 0.007	p = 0.220
Treatment × SES	-0.273	-0.145	-4.96	-0.143	-0.086	-0.002
	(0.608)	(0.132)	(8.16)	(0.133)	(0.195)	(0.009)
	p = 0.606	p = 0.196	p = 0.512	p = 0.205	p = 0.513	p = 0.802
Control Mean	10.1	0.000	142.9	0.000	0.000	0.299
R^2	0.510	0.415	0.630	0.415	0.367	0.187
Observations	1,402	1,402	1,402	1,402	1,402	1,402
Panel D. By gender						
Treatment	0.908**	0.201**	17.1**	0.180**	0.136	0.013**
	(0.444)	(0.082)	(6.63)	(0.083)	(0.106)	(0.006)
	p = 0.023	p = 0.011	p = 0.006	p = 0.022	p = 0.133	p = 0.025
Treatment × Male	0.101	-0.033	-8.69	0.012	0.186	-0.011
	(0.593)	(0.121)	(8.25)	(0.123)	(0.171)	(0.009)
	p = 0.854	p = 0.770	p = 0.260	p = 0.914	p = 0.166	p = 0.225
Control Mean	10.1	0.000	142.9	0.000	0.000	0.299
R^2	0.510	0.414	0.630	0.414	0.368	0.188
Observations	1,402	1,402	1,402	1,402	1,402	1,402
Panel E. By baseline degree						
Treatment	0.751**	0.134*	10.6**	0.142*	0.082	0.007
	(0.349)	(0.071)	(4.75)	(0.073)	(0.078)	(0.007)
	p = 0.025	p = 0.054	p = 0.025	p = 0.045	p = 0.230	p = 0.311
Treatment × High Degree	0.451	0.104	3.67	0.096	0.338*	9.53×10^{-5}
	(0.617)	(0.134)	(8.39)	(0.135)	(0.200)	(0.009)
	p = 0.403	p = 0.360	p = 0.634	p = 0.400	p = 0.014	p = 0.992
Control Mean	10.1	0.000	142.9	0.000	0.000	0.299
R^2	0.510	0.414	0.630	0.414	0.371	0.189
Observations	1,402	1,402	1,402	1,402	1,402	1,402

Notes: Heterogeneous treatment effects on the information network along five measures of centrality (degree, eigenvector, number of length-2 walks, diffusion, and betweenness centralities) and average link strength. Eigenvector, diffusion and betweenness centralities are normalized. Panel A interacts the treatment variable with above-median hours of the digital library use during the experiment ("High Browsing"); Panel B with top decile exam scores at the baseline; Panel C with SES (SES) defined as respondent's house having access to electricity and running water; Panel D with gender; and Panel E with above-median baseline degree. Regressions have controls for the covariate main effect, baseline degree, gender, SES, stratification bins and class fixed effects. "Control Mean" represents the mean of the outcome in the control arm. The sample consists of students present at both baseline and endline (N=1,402). Heteroskedasticity-robust standard errors in parentheses and randomization inference p-value with "p = ". Stars represent classical inference p-values with *** p<0.01; ** p<0.05; * p<0.1.

Appendix Table A9: Alternative Network Definitions

	Links Created	Links Broken	Intersection Degree	In-Degree	Out-Degree	Weighted Degree
Panel A. Information Network						
Treatment	0.647***	-0.316**	0.281***	0.699***	0.247	0.392***
	(0.240)	(0.126)	(0.081)	(0.258)	(0.213)	(0.097)
	p = 0.003	p = 0.009	p = 0.000	p = 0.002	p = 0.218	p = 0.000
Control Mean	6.28	6.33	1.30	5.65	5.77	2.96
R^2	0.247	0.803	0.290	0.599	0.206	0.485
Observations	1,402	1,402	1,402	1,402	1,402	1,402
Panel B. Personal Network						
Treatment	-0.011	0.0002	-0.051	0.043	-0.205	0.009
	(0.147)	(0.078)	(0.068)	(0.134)	(0.129)	(0.051)
	p = 0.948	p = 0.998	p = 0.448	p = 0.751	p = 0.107	p = 0.860
Control Mean	3.81	3.93	1.41	3.62	3.69	1.86
R^2	0.198	0.810	0.264	0.323	0.228	0.358
Observations	1,402	1,402	1,402	1,402	1,402	1,402
Panel C. Full Network						
Treatment	0.497*	-0.325**	0.218**	0.608**	0.121	0.200***
	(0.262)	(0.142)	(0.103)	(0.270)	(0.248)	(0.063)
	p = 0.041	p = 0.015	p = 0.034	p = 0.012	p = 0.618	p = 0.001
Control Mean	7.60	7.33	2.59	7.68	7.82	2.41
R^2	0.250	0.779	0.371	0.586	0.248	0.491
Observations	1,402	1,402	1,402	1,402	1,402	1,402

Notes: Estimated differences in centrality between treated and control students considering alternative definitions of the network. First and second columns decompose the main effects into links that were created and broken, respectively. Third, fourth and fifth columns alternatively use the intersection, in- and out- degrees. Sixth column computes the weighted degree by the number of interactions within the subcomponents of each network. Panel A considers the information network, followed by the personal network (Panel B) and the full network (Panel C). Regressions have controls for baseline degree, gender, SES, stratification bins and class fixed effects. "Control Mean" represents the mean of the outcome in the control arm. The sample consists of students present at both baseline and endline (N=1,402). Heteroskedasticity-robust standard errors in parentheses and randomization inference p-value with "p = ". Stars represent classical inference p-values with *** p<0.01; ** p<0.05; * p<0.1.

Appendix Table A10: Robustness to the Exclusion of Covariates

	Degree	Eigenvector	Number of Length-2 Walks	Diffusion	Betweenness	Average Link Strength
Panel A. Information Networks						
Treatment	0.851**	0.138*	11.4**	0.151**	0.222**	0.007*
	(0.356)	(0.074)	(4.78)	(0.075)	(0.103)	(0.004)
	p = 0.008	p = 0.040	p = 0.011	p = 0.027	p = 0.004	p = 0.116
Control Mean	10.1	0.000	142.9	0.000	0.000	0.299
R^2	0.225	0.097	0.443	0.094	0.071	0.117
Observations	1,402	1,402	1,402	1,402	1,402	1,402
Panel B. Personal Network						
Treatment	-0.018	-0.055	-0.724	-0.043	-0.002	0.004
	(0.181)	(0.061)	(1.58)	(0.063)	(0.066)	(0.006)
	p = 0.923	p = 0.399	p = 0.669	p = 0.509	p = 0.971	p = 0.561
Control Mean	5.91	0.000	49.9	0.000	0.000	0.325
R^2	0.181	0.053	0.363	0.052	0.045	0.105
Observations	1,402	1,402	1,402	1,402	1,402	1,402
Panel C. Full Network						
Treatment	0.734*	0.089	11.0*	0.106	0.183**	0.003
	(0.383)	(0.071)	(6.37)	(0.071)	(0.091)	(0.003)
	p = 0.039	p = 0.176	p = 0.066	p = 0.110	p = 0.014	p = 0.356
Control Mean	12.9	0.000	218.3	0.000	0.000	0.193
R^2	0.257	0.095	0.505	0.094	0.067	0.108
Observations	1,402	1,402	1,402	1,402	1,402	1,402

Notes: Estimated differences between treated and control students for five measures of centrality (degree, eigenvector, number of length-2 walks, diffusion, and betweenness centralities) and average link strength. Eigenvector, diffusion and betweenness centralities are normalized. Regressions include only stratification bins. Panel A considers the information network, followed by the personal network (Panel B) and the full network (Panel C). "Control Mean" represents the mean of the outcome in the control arm. The sample consists of students present at both baseline and endline (N=1,402). Heteroskedasticity-robust standard errors in parentheses and randomization inference p-value with "p = ". Stars represent classical inference p-values with *** $p<0.01$; ** $p<0.05$; * $p<0.1$.

Appendix Table A11: Information Access and Centrality in Low-Attrition Schools

	Degree	Eigenvector	Number of Length- 2 Walks	Diffusion	Betweenness	Average Link Strength
Panel A. Information Network						
Treatment	1.12***	0.155**	14.6**	0.174**	0.228**	0.007
	(0.421)	(0.075)	(5.87)	(0.079)	(0.102)	(0.006)
	p = 0.003	p = 0.029	p = 0.008	p = 0.017	p = 0.004	p = 0.245
Control Mean	10.9	0.000	164.7	0.000	0.000	0.296
R^2	0.515	0.444	0.622	0.428	0.402	0.204
Observations	791	791	791	791	791	791
Panel B. Information Network: probability of being in top 5%						
Treatment	0.024	0.021	0.023	0.022	0.023	0.014
	(0.020)	(0.019)	(0.019)	(0.019)	(0.020)	(0.021)
	p = 0.186	p = 0.243	p = 0.203	p = 0.199	p = 0.189	p = 0.476
Control Mean	0.052	0.047	0.051	0.047	0.047	0.049
R^2	0.305	0.256	0.286	0.304	0.258	0.062
Observations	791	791	791	791	791	791
Panel C. Personal Network						
Treatment	0.192	0.042	1.16	0.033	0.058	-0.004
	(0.236)	(0.071)	(2.05)	(0.077)	(0.085)	(0.007)
	p = 0.417	p = 0.573	p = 0.598	p = 0.677	p = 0.493	p = 0.602
Control Mean	6.27	0.000	56.1	0.000	0.000	0.317
R^2	0.366	0.346	0.527	0.277	0.199	0.160
Observations	791	791	791	791	791	791
Panel D. Full Network						
Treatment	0.984**	0.117	15.4**	0.130*	0.191**	0.004
	(0.458)	(0.076)	(7.77)	(0.079)	(0.095)	(0.004)
	p = 0.020	p = 0.107	p = 0.042	p = 0.079	p = 0.017	p = 0.307
Control Mean	13.9	0.000	253.0	0.000	0.000	0.189
R^2	0.522	0.423	0.679	0.405	0.367	0.211
Observations	791	791	791	791	791	791

Notes: Regression restricting the sample to the two national schools, which have very low attrition. Panel A shows the treatment effects on five measures of centrality (degree, eigenvector, number of length-2 walks, diffusion, and betweenness centralities) and average link strength on the information network (equation 3). Eigenvector, diffusion and betweenness centralities are normalized. Regressions have controls for baseline measure of outcome (and, in Panel B, baseline centrality measure), SES, stratification bins and class fixed effects. Panel B shows the probability of being in the top 5% central within forms on the information network. Panel C observes the effect on personal networks, and Panel D on the full network. "Control Mean" represents the mean of the outcome in the control arm. The sample consists of students present at both baseline and endline. Heteroskedasticity-robust standard errors in parentheses and randomization inference p-value with "p = ". Stars represent classical inference p-values with *** p<0.01; ** p<0.05; * p<0.1.

Appendix Table A12: Predictors of Centrality at Endline

	Degree			Eigenvector Centrality			Diffusion Centrality		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Information Network									
Baseline Degree	0.536***		0.510***						
	(0.026)		(0.028)						
Baseline Eigenvector Centrality				0.607***		0.584***			
				(0.031)		(0.034)			
Baseline Diffusion Centrality							0.614***		0.586***
							(0.030)		(0.034)
Academic Ability		2.38***	1.11***		0.456***	0.233***		0.478***	0.232***
		(0.270)	(0.231)		(0.055)	(0.050)		(0.056)	(0.050)
SES		1.53***	0.867***		0.338***	0.146***		0.334***	0.154***
		(0.273)	(0.229)		(0.056)	(0.049)		(0.056)	(0.049)
Male		-0.568	0.114		-0.129	0.154**		-0.138	0.087
		(0.356)	(0.305)		(0.087)	(0.078)		(0.086)	(0.075)
R^2	0.471	0.227	0.485	0.364	0.127	0.381	0.368	0.112	0.384
Observations	1,402	1,402	1,402	1,402	1,402	1,402	1,402	1,402	1,402
Panel B. Personal Network									
Baseline Degree	0.337***		0.308***						
	(0.027)		(0.027)						
Baseline Eigenvector Centrality				0.364***		0.287***			
				(0.027)		(0.029)			
Baseline Diffusion Centrality							0.366***		0.314***
							(0.028)		(0.030)
Academic Ability		0.810***	0.485***		0.214***	0.138***		0.263***	0.165***
		(0.150)	(0.141)		(0.048)	(0.046)		(0.051)	(0.048)
SES		1.02***	0.822***		0.368***	0.289***		0.374***	0.294***
		(0.156)	(0.148)		(0.051)	(0.050)		(0.054)	(0.052)
Male		-0.768***	-0.284		-0.702***	-0.414***		-0.519***	-0.268***
		(0.224)	(0.216)		(0.077)	(0.080)		(0.080)	(0.081)
R^2	0.287	0.221	0.308	0.286	0.258	0.322	0.226	0.172	0.255
Observations	1,402	1,402	1,402	1,402	1,402	1,402	1,402	1,402	1,402
Panel C. Full Network									
Baseline Degree	0.532***		0.504***						
	(0.025)		(0.026)						
Baseline Eigenvector Centrality				0.599***		0.567***			
				(0.029)		(0.032)			
Baseline Diffusion Centrality							0.594***		0.563***
							(0.029)		(0.032)
Academic Ability		2.58***	1.27***		0.447***	0.233***		0.467***	0.236***
		(0.298)	(0.255)		(0.054)	(0.048)		(0.055)	(0.049)
SES		1.94***	1.28***		0.369***	0.204***		0.368***	0.216***
		(0.301)	(0.254)		(0.055)	(0.047)		(0.055)	(0.048)
Male		-0.749*	0.244		-0.283***	0.034		-0.202**	0.060
		(0.405)	(0.343)		(0.085)	(0.075)		(0.085)	(0.074)
R^2	0.476	0.264	0.493	0.373	0.143	0.391	0.361	0.121	0.381
Observations	1,402	1,402	1,402	1,402	1,402	1,402	1,402	1,402	1,402

Notes: Regressions of endline degree, eigenvector centrality, diffusion centrality on their baseline values (Columns 1, 4, and 7), on academic ability, high SES (SES), and male (Columns 2, 5 and 8) and all (Columns 3, 6 and 9). Eigenvector and diffusion centralities are normalized with respect to the control arm mean and standard deviation. Diffusion centrality parameters follow Banerjee et al. (2019) with q equal to the reciprocal of the top eigenvalue, and T equal to the diameter of the graph. Academic Ability is defined as above-median exam score at baseline. SES is equal to 1 if respondent's house has electricity and running water. All regressions have class fixed effects. Heteroskedasticity-robust standard errors in parentheses. *** $p<0.01$; ** $p<0.05$; * $p<0.1$.

Appendix Table A13: Treatment Effects on Academic Scores

	English	Biology
Panel A. Overall effects		
Treatment	.103**	.063
	(.050)	(.047)
	p = .048	p = .188
Panel B. Heterogeneous treatment effects		
Treatment x Below Median Ability	.195**	.143**
	(.076)	(.067)
	p = .016	p = .041
Treatment x Above Median Ability	.003	-.025
	(.062)	(.064)
	p = .964	p = .718
Control Mean	.000	.000
Observations	1412	1406

Notes: Table reproduced from Derksen et al. (2022). Treatment effects on final exam scores. Ability defined as above (below) median exam scores (average of English and Biology) at the baseline. We include a control for baseline exam score, an indicator for missing baseline score, and strata fixed effects. Randomization was stratified by school, form, above median achievement and past internet use. Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Randomization inference p-values based on 10,000 replications denoted as "p =". P-values slightly different from Derksen et al. (2022) due to changes how the ritest command sets seed.