sociological science

Supplement to:

Lee, Jaemin, David Lazer, Christoph Riedl. 2025. "Complex Contagion in Social Networks: Causal Evidence from a Country-Scale Field Experiment" Sociological Science 12: 685-714.

Appendix for
Complex Contagion in Social Networks: Causal Evidence from a Country-Scale Field Experiment

CONTENTS

1. Conceptual Model

To illustrate the infection force equation given in Eq. (1) we show an example using $\beta = 0.1$ (i.e., 10% infection likelihood) in Figure A1.

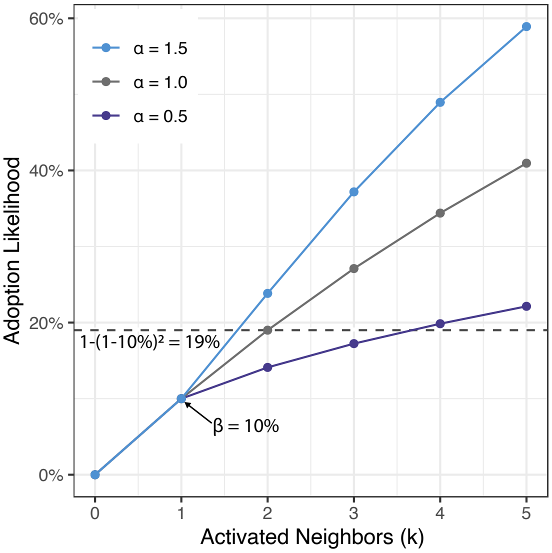

Figure A1. Example illustration of contagion at $\beta = 0.1$.

We leverage the tunable social reinforcement parameter (α) in our infection force equation given in Eq. (1) and show that in most realistic settings, “complex” contagion processes (those that feature positive social reinforcement) do not necessarily exhibit superexponential spreading (i.e., they look quite similar to “simple” contagion processes; Figure A2).

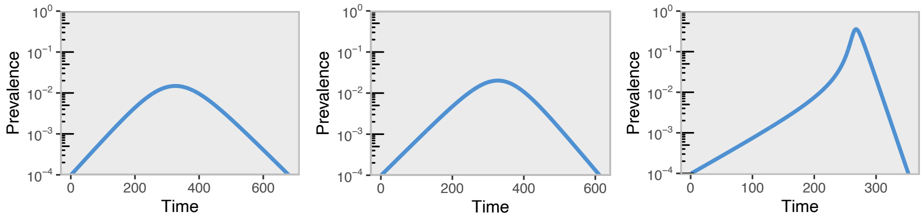

Figure A2. Across a wide range of realistic parameters (R_0 close to 1) the diffusion processes lead to standard exponential spreading (and not superexponential) despite the presence of social reinforcement ($\alpha > 1$). Left: standard exponential spreading of a simple contagion process without social reinforcement ($R_0 = 1.2$ and $\alpha = 1.0$); Middle: visually similar standard exponential spreading despite positive social reinforcement ($R_0 = 1.2$ and $\alpha = 1.2$). Right: superexponential spreading arises only when contagiousness and/or social reinforcement are very high ($R_0 = 1.2$ and $\alpha = 2$).

We simulate the spread of a behavior in a susceptible population using a standard Susceptible-Infected-Recovered (SIR) epidemic process (Barrat et al., 2008) on a random network (random regular graph with $\langle k = 20 \rangle$). Under the assumptions of our model, it is almost impossible to observe superexponential growth in most realistic settings (near criticality when the basic reproduction number (R_0) describing contagiousness or transmissibility of infections is close to 1) even when social reinforcement is strong. The intuition behind this result is that exposure to multiple infected neighbors is rare when the prevalence of the infection in the population is low and there are only few instances in which social reinforcement can unfold. Conversely, when prevalence in the population is high due to high contagiousness, even moderate levels of social reinforcement make little difference on the diffusion process and are not sufficient to push it from the standard exponential regime to a superexponential regime. Taken together, our simulation analysis shows that it is possible that both aspects can hold true at the same time: positive social reinforcement acts as a behavioral mechanism on the individual level (i.e., $\alpha > 1$), while the population-level diffusion process exhibits none of the characteristics that have been theorized—and variably been used as assessment criteria—of complex contagion such as superexponential spreading. This illustrates that individual-level mechanistic insights do not directly translate into population-level empirical criteria nor do population-level observations of the standard exponential spread necessarily rule out positive social reinforcement.

2. Variable Description

Our outcome variable is whether or not a focal node adopted (redeemed) a data voucher. The key predictor variables are Treated Neighbors (K_{ITT}) and Activated Neighbors (K), the number of a focal node's immediate neighbors that were seeded and adopted the voucher code. A core moderator variable in our study is structural embeddedness—the number of friends that two individuals have in common normalized by the degree of the two individuals. Counting common neighbors in a large network like ours with 48M nodes poses a significant computational challenge and is only feasible with the use of advanced graph algorithms. We built on the work by (An et al., 2019) to efficiently compute edge-embeddedness in our large network. We also consider individual and dyadic network characteristics that might account for heterogeneity in peer effects and the propensity of adoption. Degree is the number of the connected neighbors via calls and SMS during the three-month period. Provided that peer influence is more likely and stronger when individuals are close to each other, we control for dyadic tie weights in terms of duration and frequency of their interactions. We turned them into normalized individual-level variables in ways that their value indicates tie strength to the treated seed, relative to other network neighbors connected to that seed.

Variable	Description
Adoption (Y)	Outcome variable, 1 if individual adopts mobile internet by redeeming a voucher code, 0 otherwise. Unless noted otherwise, we consider the entire experiment period of 7 days.
Treated Neighbors (K_{ITT})	Number of treated neighbors that a focal node has.
Activated Peers (K)	Number of immediate neighbors that were seeded and adopted the voucher code. A focal node's seed peer is only considered "activated" if the seed adopted before the focal node adopted.
Degree	Individuals' number of friends in the baseline social network, based on calls and SMS, in a three-month period
Duration	Tie strength between the focal node and the encouraged seed node in the baseline social network measured in voice minutes that an individual had with a treated seed, relative to other network neighbors (i.e., row-normalized to the unit interval [0, 1]). For focal nodes connected to two seeds, we average the tie strength.
Frequency	Tie strength between the focal node and the encouraged seed node in the baseline social network measured as contact frequency, relative to other network neighbors (i.e., row-normalized to the unit interval [0, 1]). We use a composite measurement of (voice+SMS) - (voice×SMS), which showed a better fit than when they are included separately. For focal nodes connected to two seeds, we average the tie strength.
Embeddedness	The number of common friends that an individual had with a treated seed (M_{ij}) normalized by the degree of the focal node (D_i) and the seed (D_j): $\frac{M_{ij}}{D_i+D_j-M_{ij}-2}$. Averaged for focal nodes who are connected to both seeds.

Table A1. Variable Description.

	Mean	SD	Min	Max	(1)	(2)	(3)	(4)	(5)	(6)
Adopted Product (1)	0.00	0.05	0.00	1.00						
k_{ITT} (2)	0.71	0.47	0.00	2.00	0.01***					
k (3)	0.08	0.28	0.00	2.00	0.05***	0.17***				
Degree (4)	64.66	68.46	1.00	4513.00	0.00***	0.00***	-0.03***			
Duration (row normalized) (5)	0.05	0.13	0.00	1.00	-0.01***	-0.01***	-0.01***	-0.22***		
Frequency (row normalized) (6)	0.07	0.17	0.00	1.00	0.00	0.00***	0.02***	-0.23***	0.66***	
Embeddedness (7)	0.02	0.02	0.00	0.47	0.03***	0.03***	0.05***	-0.04***	-0.02***	0.00***

$^{***}p < 0.001$

Table A2. Descriptive statistics of study variables.

3. Contamination Issue

For the data analysis, we removed “contaminated” nodes for which we could not cleanly attribute observed effects to assigned treatment conditions. That is, we removed seed nodes who were themselves connected to other seed nodes, focal nodes who were connected to contaminated seed nodes, and focal nodes who were connected to more than two seed nodes (about 4% of nodes were affected from such contamination). One concern may be that this results in imbalanced observables (e.g., with regard to node degree). We believe this does not constitute a major problem at least with regard to estimating LATE as the majority of nodes are removed because they are connected to contaminated seeds. Since seeds are randomly assigned to treatment, this removal should not result in imbalanced observables.

The biggest challenge in implementing this marketing campaign was online-based diffusion (see [citation removed for blind peer review]). That is, a few customers posted their voucher codes to an online community and those were accessed individually across the population. While this online sharing further helped the product go viral, it certainly raised a concern for our inference of peer-to-peer diffusion through the mobile social network. However, because all seeded vouchers contained unique codes, we were able to identify the four “contaminated” codes reported with an atypically high number of adoptions (10~50 times higher activation counts than other codes). Note that this study does not aim to investigate diffusion processes in the whole population; we are interested in a specific subpopulation that comprises our experimental treatment groups—only the seeds and their one-hop neighbors who would be early spreaders and therefore unaffected by contamination. To provide conservative estimates and avoid counting false-positive product adoptions (i.e., a seed and his neighbor independently adopted the voucher from the online source, but they happened to have a tie to each other), we excluded the cases where any of the two nodes in the randomly assigned edge had adopted a contaminated voucher code. This measure reduced our sample size, but not dramatically (2,640,711 → 2,502,187, a 5% decrease). In a similar vein, we only counted non-contaminated adoption codes.

4. Intent-to-Treat Analysis

Here we consider treatment effect heterogeneity (Table A3).

Dependent Variable:	**Adopted Product within first ...**	
	6h	7 days
	(1)	(2)
Intercept	0.0001***	0.0013***
	(0.0000)	(0.0001)
$k_{ITT} = 1$	−0.0000†	0.0000
	(0.0000)	(0.0001)
$k_{ITT} = 2$	0.0005	0.0023*
	(0.0003)	(0.0009)
Embeddedness		
× $k_{ITT} = 1$	0.0134***	0.0371***
	(0.0015)	(0.0036)
× $k_{ITT} = 2$	0.0132†	0.0495**
	(0.0076)	(0.0190)
Controls		
Embeddedness	0.0027***	0.0248***
	(0.0008)	(0.0026)
Degree	−0.0000**	−0.0001***
	(0.0000)	(0.0000)
Duration	−0.0005***	−0.0035***
	(0.0001)	(0.0002)
Frequency	0.0005***	0.0015***
	(0.0001)	(0.0002)
Adj. R^2	0.0005	0.0010
Num. obs.	2,502,187	

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$

Table A3. Regression results of ITT treatment effect heterogeneity.

Notice that spreading along indirect paths cannot explain these results. To see this, it is useful to recall that two-hop spreading along an indirect path happens only at the squared likelihood of the direct path as both nodes along the indirect paths would have to activate (i.e., $1 - (1 - \beta^2)^N$ where N is the number of available indirect paths). Empirically, we observe seed nodes have about 2.8 friends in common with focal nodes so that this adoption likelihood is about two orders of magnitude *smaller* than the main effect (see the conceptual Figure A1 above). However, empirically, we find a coefficient that is instead about one order of magnitude *larger* than the main effect. This supports the conclusion that our results are inconsistent with the hypothesis that spreading is driven by indirect paths and instead much more consistent with the hypothesis that the tie strength of structurally embedded relationships explains the effect. Finally, notice also that in the model using adoption within the first 6h indirect spreading is effectively ruled out as shown by the near-zero the intercept which indicates that there simply is virtually no indirect spreading at all.

We illustrate the potential effect that indirect spreading along two-hop paths may have relative to the effect of direct spreading in Figure A3. Indirect spreading has only a very small effect relative to the direct effect, especially given the small number of 2.8 indirect paths that we observe in our data between focal nodes and seed nodes (so a total of 5.6 indirect paths between the focal node and the two seeds).

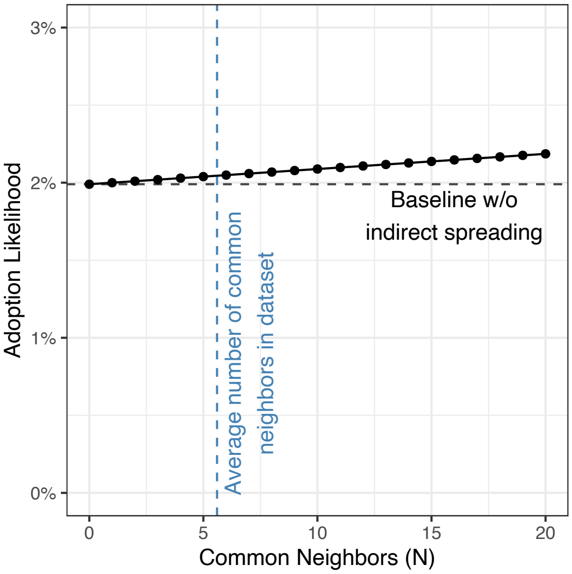

Figure A3. Example illustration of effect of indirect spreading at $\beta = 0.01; k = 2$ and $\alpha = 1$ (simple contagion).

5. Analysis of Encouragement Design with 2SLS

Conditions Here we consider several conditions in applying the 2SLS IV analysis for the peer encouragement design. The potential outcome formulation of the encouragement design requires two assumptions (Eckles et al., 2016). First, it requires an exclusion restriction for instrumental variables (IVs). The exclusion restriction seems plausible in our case as participants are blinded to their assignment. It seems safe to assume that participants did not realize they were randomly assigned to treatments—especially given the minimal nature of the encouragement—or if they were it was just a normal kind of marketing campaign that would not otherwise affect their behavior. The minimal nature of the encouragement makes it plausible that it affects individuals only by causing them to have contact with an activated neighbor and their neighbors own use of mobile internet; however, there may be effects of the encouragement not captured by the adoption status of seed nodes (e.g., changes in content of communication). Next, is the no interference assumption and direct-effect-bounded interference assumption. We expect this assumption to be violated in this setting, even in our finite population. To limit the threat of interference we focus on only one-hop neighbors, exclude individuals who are connected to multiple seed pairs, and the fact that observations in our experiment make up only a relatively small portion of the entire network (about 3M nodes in 48M network). Overall, we therefore work under the assumption that the remaining interference is small compared with the effects of interest (see Eckles et al. 2016 for more details and why this assumption may be justified). To address remaining concerns, we present a complementary analysis using a Fisher exact p-value test that is specifically designed to address the issue of network interference (see the method description in 3.2.3 and its findings in 4.3.). The 2SLS estimate using binary treatment indicators formed from the levels of the encouragement Z_i, estimates a weighted average of effects of changes in increments of D_i using a single binary indicator (Imbens & Angrist, 1994), each of which estimates an average causal response (ACR).

With the parameter estimates for discrete levels of K_i from our 2SLS analysis of the encouragement design in hand, we can then proceed to test our hypothesis of positive reinforcement from redundant exposure. Intuitively, we take the parameter estimates of having zero, one, or two activated neighbors and then apply our criterion for complex contagion to test if the increase in adoption likelihood from exposure to two activated neighbors is larger than the expected counterfactual from two activated neighbors under simple contagion. Finally, we then use the same estimation framework to test for a moderating effect of embeddedness. Specifically, we will add embeddedness as a covariate to the first-stage regression and an interaction term between K_i and embeddedness to the second-stage regression.

First-stage Results

Dependent Variable:	**One Activated Seed Neighbor**		**Two Activated Seed Neighbor**	
	W/O Controls	W/ Controls	W/O Controls	W/ Controls
Intercept	0.064^{***}	0.056^{***}	-0.000	-0.000^{**}
	(0.000)	(0.000)	(0.000)	(0.000)
Single exposure: tie with both seeds	0.050^{***}	0.039^{***}	0.017^{***}	0.017^{***}
	(0.002)	(0.002)	(0.000)	(0.000)
Multiple exposure: tie with one seed	0.047^{***}	0.047^{***}	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Multiple exposure: tie with both seeds	0.119^{***}	0.108^{***}	0.050^{***}	0.049^{***}
	(0.002)	(0.002)	(0.000)	(0.000)
Degree		-0.009^{***}		-0.000^{***}
		(0.000)		(0.000)
Duration		-0.082^{***}		-0.001^{***}
		(0.002)		(0.000)
Frequency		0.063^{***}		0.000^{***}
		(0.001)		(0.000)
Embeddedness		0.371^{***}		0.004^{***}
		(0.007)		(0.001)
R^2	0.008	0.011	0.037	0.037
Adj. R^2	0.008	0.011	0.037	0.037
Num. obs.	2502187	2502187	2502187	2502187
F statistic	6488.094	3884.501	31677.524	13593.571

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table A4. First-stage model results. Omitted category is "Single exposure: tie with one seed."

Dependent Variable:	Adopted Product within first ...					
	6h	12h	24h	36h	48h	7 days
	(1)	(2)	(3)	(4)	(5)	(6)
Peer Effects						
Intercept	-0.000^{**}	-0.000	0.000	0.000	0.000	0.001^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$k = 1$	0.002^{***}	0.003^{***}	0.004^{***}	0.005^{***}	0.005^{***}	0.010^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$k = 2$	0.011^{*}	0.015^{*}	0.016^{*}	0.023^{**}	0.026^{**}	0.060^{***}
	(0.005)	(0.006)	(0.007)	(0.008)	(0.008)	(0.012)
Controls						
Degree	0.000	0.000	0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Duration	-0.000^{***}	-0.001^{***}	-0.001^{***}	-0.001^{***}	-0.001^{***}	-0.003^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Frequency	0.000^{***}	0.000^{**}	0.000^{**}	0.001^{***}	0.001^{***}	0.001^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Embeddedness	0.012^{***}	0.016^{***}	0.019^{***}	0.026^{***}	0.028^{***}	0.047^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
α	2.622	2.556	2.297	2.551	2.449	3.084
Wald $P_{k=2}^{Obs} = P_{k=2}^{Simple}$	8.290	8.694	6.244	10.151	10.455	36.185
Adj. R^2	0.001	0.001	0.001	0.002	0.002	0.003
Num. obs.	2,502,187					

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table A5. ATE effects of activated neighbors on product adoption, replicated in different time windows (Model 4 in Table 2).

Comparison of OLS and 2SLS Estimates. When analyzing peer effects through an instrumental variable framework, it is not uncommon to obtain different effect sizes between OLS and 2SLS estimates (Eckles et al., 2016; Aral and Nicolaides 2017). Eckles et al. (2016) suggests that it cannot be assumed that OLS will necessarily be an upper bound the true effect. This could be due to confounding, simultaneity, or the fact that these different models identify different causal quantities. In our analysis, the effect size was larger in the 2SLS model. This suggests that OLS underestimated the peer effects of *k*, accompanied by an unidentified process where the adoption likelihood *Y decreased* as the endogenous variable *k* increased.

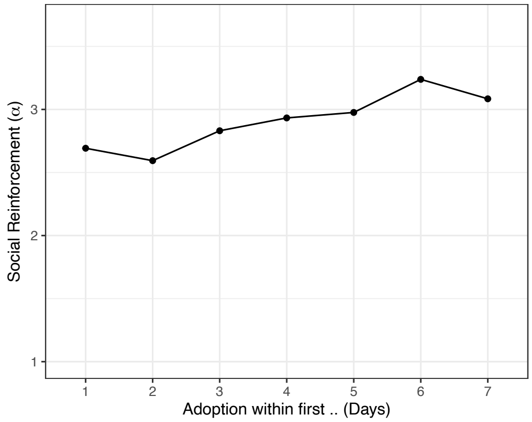

Figure A4. Comparison of the strength of social reinforcement across different time windows (ATE effects from Model 4 in Table 2 estimated on different time windows).

6. Exact P-Values for Network Interference

Network interference is a well-documented problem for networked experiments (Aral, 2016; Aral & Walker, 2014; Aronow, 2012). Randomized controlled experiments on networked populations often violate the stable unit treatment value assumption (SUTVA) because treated individuals are usually connected to other non-treated individuals (Gerber & Green, 2012; Rubin, 1980). SUTVA requires that the outcomes for any given unit be unaffected by other units' treatment assignment. This requirement serves as a condition for the unbiased inference of average treatment effects. Since our experiment was conducted on a single network, even treatments that are randomly assigned do not completely rule out a biased inference. This poses a threat to the inference of treatment effects despite the randomized design, potentially leading to inflated Type 1 errors.

In the Methods section, we have provided an overview of the exact *p*-value approach to estimating spillovers in Athey, Eckles, Imbens (2018, AEI henceforth). Here, we offer details of hypothesis testing that we have reported in the Results section 4.3. The null model and the form of the peer effect term differ by the specific hypotheses to test. We explain these, given N individuals i (nodes), a vector of random treatment assignments W_i, and a network adjacency matrix G with $G_{ij} = 1$ if individual i and j are connected and 0 otherwise.

Hypothesis 0 (No Spillover): We consider the null hypothesis that for the outcome Y_i, only the treatment assignment W_i is relevant, and the treatment assignment for other individuals $W_{j \neq i}$, and consequently exposure to treated neighbors, does not matter. The null hypothesis of no spillover effects is that outcome of Y_i is only affected by the treatment of w_i but not neighbor's treatment (this is H2 in the original AEI). Formally, this can be expressed as: $Y_i(w) = Y_i(w')$ for all i, and for all pairs of assignment vectors $w, w' \in W$ such that $w_i = w_i'$. The null model includes only the own treatment:

$$Y_i^{obs} = \alpha_0 + \tau_{treat} \cdot W_i + \epsilon_i \quad \text{(A1)}$$

Test statistic. To assess the degree to which the observed effects are distinct from the artificial experiment, we calculate a score test statistic, T_{score}. T_{score} for H0 is the covariance between the residual of the null model and i's number of treated neighbors (this is adapted from Eq 7 in AEI). This, and all other T_{score} are calculated for all nodes that are not isolates, which is true for all nodes by design in our setting:

$$T_{score} = cov(\epsilon_i, \textstyle\sum_{j=1}^{N} W_j \cdot G_{ij}) \quad \text{(A2)}$$

Substantially, T_{score} allows us to identify the degree to which the peer effects variable explains the leftover unexplained variance in the treatment-only model. A high T_{score} means that the error term of the no-spillover model substantially covaries with the (observed) peer effect—the number of treated neighbors in this case—and hence substantially explains adoption behavior.

Hypothesis 1 (Threshold Peer Effects): We consider the null hypothesis that for the outcome of Y_i it only matters whether i has any treated neighbors. Here we test the alternative hypothesis that the number of treated neighbors matters. The null hypothesis of threshold peer effects is that the effect of having two or more treated neighbors is not different from the effect of having one treated neighbor on i's adoption (H8 in AEI). Formally: $Y_i(w) = Y_i(w')$ for all i, and for all pairs of assignment vectors $w, w' \in W$ such that $1\{\sum_{j=1}^{N} w_j * G_{ij} > 0\} = 1\{\sum_{j=1}^{N} w'_j * G_{ij} > 0\}$. The null model includes the own treatment term and the discrete variable of $k \geq 1$:

$$Y_i^{obs} = \alpha_0 + \tau_{treat} \cdot W_i + \tau_{peer} \cdot \mathbf{1}\left\{\sum_{j=1}^{N} W_j \cdot G_{ij} \geq 1\right\} + \epsilon_i \quad \text{(A3)}$$

Test statistic. T_{score} is the covariance between the residual of the null model and i's number of treated neighbors:

$$T_{score} = cov(\epsilon_i, \sum_{j=1}^{N} W_j \cdot G_{ij}) \qquad \text{(A4)}$$

Hypothesis 2 (Peer Effect Heterogeneity): Under the null hypothesis of no peer effect heterogeneity it does not matter which specific peers of individual i received the treatment. Here we test specifically the alternative hypothesis that embeddedness matters for the effect of individual i on j. A null hypothesis of peer effect heterogeneity is that only the number of treated peers matter, not which neighbor is treated. In our case, we are specifically interested in the role of embeddedness, so we test the null hypotheses that the embeddedness—the number of common friends—that i has with a treated peer does not matter (H7 in AIE). Formally: $Y_i(w) = Y_i(w')$ for all i, and for all pairs of assignment vectors $w, w' \in W$ such that $\sum_{j=1}^{N} w_j * G_{ij} = \sum_{j=1}^{N} w'_j * G_{ij}$. The null model includes the own treatment term and the continuous peer effects variable (the number of treated neighbors of i):

$$Y_i^{obs} = \alpha_0 + \tau_{treat} \cdot W_i + \tau_{peer} \cdot \sum_{j=1}^{N} W_j \cdot G_{ij} + \epsilon_i \qquad \text{(A5)}$$

Test statistic. T_{score} is the covariance between the residual of the null model and i's number of common friends with the treated neighbor, derived from a weighted matrix G^{emb} where $G^{emb}_{i,j}$ equals to the number of common friends between i and j (adapted from Eq. 12 in AEI):

$$T_{score} = cov(\epsilon_i, \sum_{j=1}^{N} G_{ij} \cdot (G^{emb}_{i,j} \cdot W_j)) \qquad \text{(A6)}$$

For all three hypotheses, the exact p-values (2-tailed test) can be expressed as the chance that T^{Sim}_{score} exceeds the observed T^{Obs}_{score}:

$$p - \widehat{value} = 2 * \frac{1}{R} \sum_{r=1}^{R} 1\{|T^{Sim}_{score,r}| > |T^{Obs}_{score}|\}$$

where R is the total number of simulation trials (Taylor & Eckles, 2018).

Permutation Procedures: We first constructed $\mathbb{P}_F$—a subset of nodes that are fixed with their own treatment status and used to produce score test statistics—by selecting the half of nodes in the following steps: (i) randomly choose one of the seeds in the seed pair; and (ii) choose all focal nodes connected to the randomly chosen seed. The p-values are valid irrespective of the choice of $\mathbb{P}_F$, but maximizing edge comparisons by selecting one seed per the seed pair increases the power of the testing procedure (AEI: p. 235).

In each simulation trial, we permuted treatment status of those not in the fixed subset $\mathbb{P}_F$ and calculated the peer-effects term (e.g., the number of treated neighbors in **H0** and **H1**). We shuffled treatment status *only among seed nodes* to prevent the cases where focal nodes—they are by design "untreated" neighbors of seeds—become treated in the simulation. Substantively, in our simulation, a treated node in the single exposure condition may become an un-treated node, and a pair of seed nodes in the multiple exposure condition may become a single exposure pair. Focal nodes cannot become seed nodes. The distribution of simulated test statistics is shown in Figure A5.

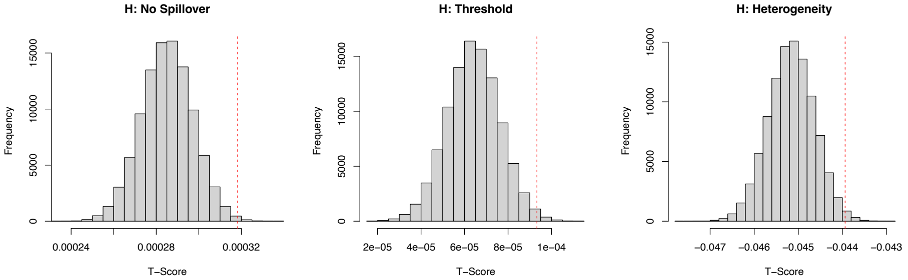

Figure A5. The distribution of test statistics from the exact *p*-value method testing H0: Spillover, H1: Threshold Peer Effects, and H2: Peer Effect Heterogeneity (by embeddedness).

7. Related Literature

Methodologically, our experiment stands as the first large-scale field experiment outside online social media that randomizes the essential theoretical element in the discussion of complex contagions. We put the features of our experiment in perspective by surveying what other studies have contributed. Table A6 is a summary of how our study differs from closely related prior work. Contrary to prior studies that focused on social media (Bakshy et al., 2012), our experiment involves a single, novel product with trackable unique codes so that we can identify precisely which neighbor was more influential. This design prevents social influence from being confounded with the nature of the product which might be the case when studying the spread of such as memes on Twitter or apps on Facebook. Our study context is realistic to the extent that awareness of peer adoption was entirely up to consumers and their interactions—with no automatic notifications helped by an online system.

In terms of the causal effects of the number of exposure to activated peers, which is the central focus of our study, only two studies (Bakshy et al., 2012; Centola, 2010) provide estimates from a randomized experiment. But Bakshy et al.'s estimate was based on the observational association, as their experiment only randomized exposure or non-exposure. Centola's experiment was based on a randomized network structure but in an artificial network that eliminated existing social ties. In Figure A6, we plotted the relative rates of adoptions when there are k and k-1 adopting neighbors, reported in the past studies and our study.

	Studies Complex Contagion?[a]	Randomized Experiment?[b]	Randomized Number of Exposures?[c]	Testing Structural Heterogeneity?[d]	Study Outcome[e]	Network Type[f]
This Study	*Yes*	*Yes*	*Yes*	*Yes*	*Product Adoption*	*Offline social network*
Centola (2010)	Yes	Yes	No	No	Product Adoption	Artificial network
Bakshy et al. (2012)	Yes	Yes	No	No	Content Sharing	Online social network
Ferrali et al. (2020)	Yes	No	No	No	Product Adoption	Offline social network
Karsai et al. (2014)	Yes	No	No	No	Product Adoption	Contact network
Ugander et al. (2012)	Yes	No	No	Yes	Product Adoption	Online social network
Aral et al. (2009)	Yes	No	No	No	Product Adoption	Online social network
Iyengar & Berger (2014)	Yes	No	No	No	Product Adoption	Online social network
Aral & Nicolaides (2017)	Yes	No	No	Yes	Behavioral Change	Online social network
State & Adamic (2015)	Yes	No	No	No	Content Sharing	Online social network
Mønsted et al. (2017)	Yes	No	No	No	Content Sharing	Online social network
Fink et al., (2016)	Yes	No	No	No	Content Sharing	Online social network
Weng et al. (2013)	Yes	No	No	Yes	Content Sharing	Online social network

Table A6. Literature summary.

Note: [a] We only included empirical studies that explicitly addressed the theory of complex contagions and provided at least a partial test of the theory.

[b] We classify as a randomized experiment when researchers imposed a treatment on randomly assigned individuals/groups to estimate peer effects.

[c] This column focuses on whether the effects of an additional exposure were found as an observational association or an exogenous experimental manipulation.

[d] This column indicates whether or not they additionally investigated network-structural variation of peer effects. For example, this study tests the moderating role of network embeddedness in inducing the effect of additional exposure. Ugander et al. (2012) found strong effects of structural diversity (number of components), less the number of exposures. Aral & Nicolaides (2017) similarly tested network embeddedness as a peer influence moderator but do not link it to the heterogeneity of the effectiveness of complex contagions.

[e] Content sharing includes the diffusion of memes and hashtags (Fink et al. 2016; Mønsted et al. 2017; Weng et al. 2013) and URL sharing (Bakshy et al. 2012).

[f] It refers to the context in which network nodes and ties were defined and interactions among peers occurred. Our study uses a whole network of mobile phone users unlike other studies based on Twitter, Facebook, or other created communities.

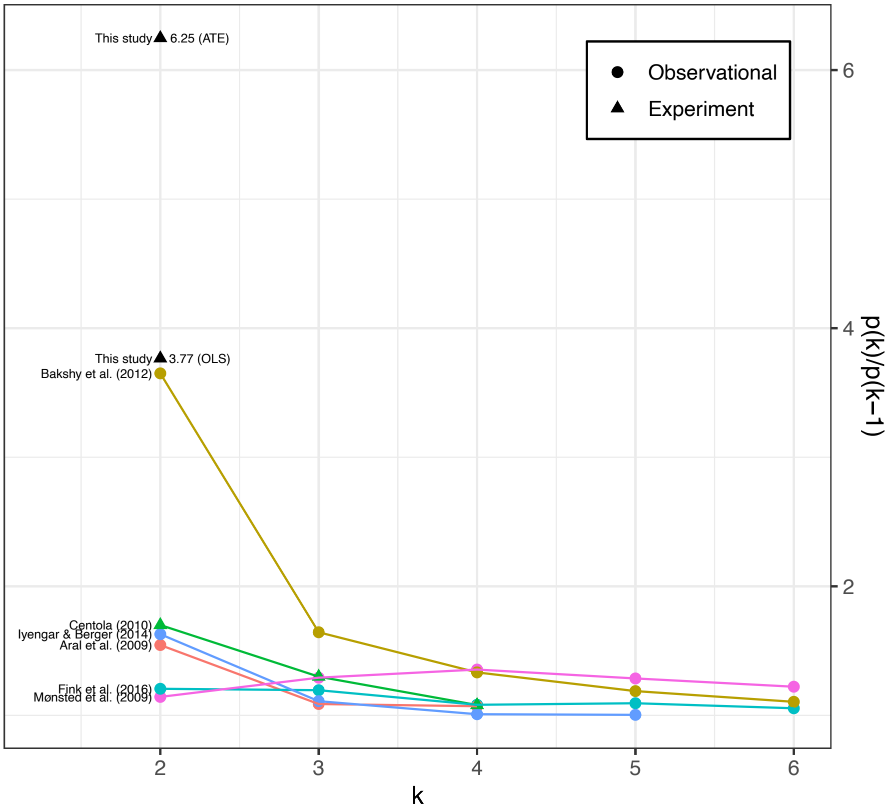

Figure A6. Adoption ratios by *k* reported in past empirical studies.
Note: This plot adapts and extends Figure S1 of Eckles et al. (2019). We selected the empirical studies that show peer effects by specific number of adopted neighbors. Note that other than Bakshy et al. (2012), there is no study showing a multiplied likelihood comparing k=2 and k=1, even when evidence of complex contagions was claimed. Bakshy et al.'s estimate was based on the observational association, as their experiment only randomized exposure or non-exposure. Values for our study are based on estimates from Table 2 Model 4 (ATE) and Model 5 (OLS) (i.e., based on the same values as shown in Figure 3). The characteristics of these studies are summarized in Table A5.

8. The Benefits of Clustered Targeting

Can a manager leverage the benefits of social reinforcement to design a marketing campaign? We examine the implications of our findings for the design of a seeding strategy in viral marketing, a topic of significant interest. The idea is that, if product adoption is substantially improved by social reinforcement engineered by multiple contacts (i.e., clustered targeting: seeding "pair" of individuals), we can leverage it to maximize early spillovers from the seeding stage.

Intuitively the effectiveness of clustered targeting will depend on (a) how likely it is that two clustered seed individuals "activate" together to expose their neighbors to social reinforcement in the first place and (b) the strength of the social reinforcement effect.[1] We construct a stylized model to illustrate how the potential benefits of clustered targeting vary with the adoption probability of seeds and the amplification of the social reinforcement effect. In the model, given adoption probability β, seeds activate to create exposure for one-hop neighbors, and these neighbors may subsequently adopt. We vary the effect of social reinforcement by multiplying the $k = 2$ effect by an amplification factor α. In turn, one-hop neighbors' resulting adoption probability is calculated by:

$$[p(exposure\ to\ one\ seed)\ \times \beta]\ +\ [p(exposure\ to\ both\ seeds) \times \beta\ \times\ \alpha]]$$

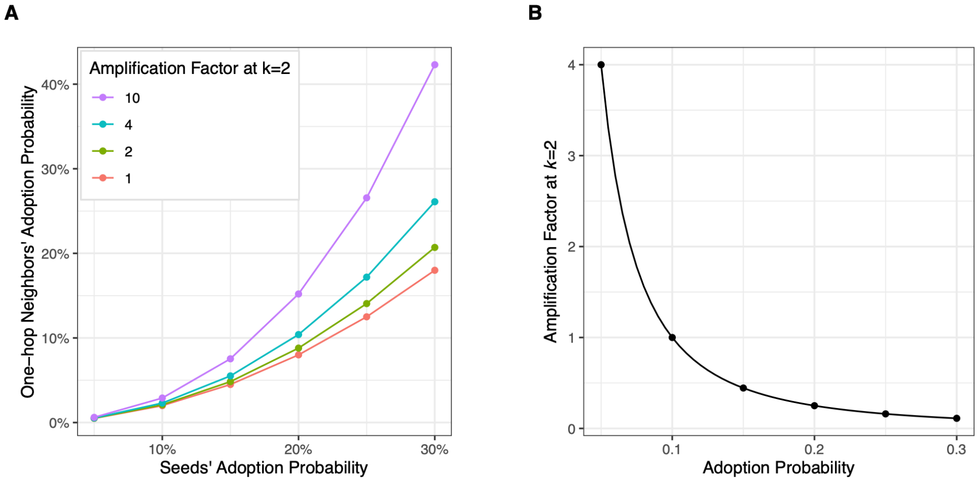

Figure A7. The potential benefits of clustered targeting as a function of seed's adoption probability and amplification of social reinforcement

The results are summarized in Figure A7-A. The key finding is that there are additional benefits generated by amplification at $k = 2$ but to a limited degree. The primary reason is that exposure to two activated peers is necessarily rare. For example, say a customer is tied to two seeds and these seeds adopt with 10% probability. Then the probability for this customer to have two activated peers is only 1% ($0.1^2 = 0.01$). This means that the proportion of adoption generated by $k = 2$ exposure would comprise a very small portion of the total adoption created by clustered targeting. We thus see little difference made by social reinforcement at $\beta = 0.1$.

The simulation result shows that the effect of social reinforcement materializes at a higher rate when a high adoption probability creates more $k = 2$ exposure cases. Yet we note other realisms that were not parameterized in this stylized model but could be detrimental to the benefits of clustered seeding. One such realism is the degree to which an individual happens to have ties to both seeds—in other words, the probability that one forms a triangle to two seeds. Among our focal nodes from the single and multiple exposure treatments, only 2.2% had ties to both seeds. This would further lower the adoption generated by $k = 2$ exposure

---

[1] Naturally, it will also depend on how many friends the two activating individuals have in common but since we are trying to optimize the marketing campaign given the network this is merely a moderating factor.

against the total adoption among clustered seeds' neighbors. Then, clustered targeting with a rare $k = 2$ exposure is unlikely to perform better than random seeding primarily relying on a $k = 1$ exposure.

Next, when a marketer has a given cost of targeting individuals and hence a target adoption rate at which the campaign is profitable, how much weight should be placed on the adoption probability of a seeded customer and social reinforcement at k=2? Let us denote Θ as the target adoption rate, α as the amplification factor at $k = 2$, and β as seed's adoption rate, then $\Theta = \alpha \times \beta^2$. We can infer that regardless of the size of the target adoption rate, the desired social reinforcement effect α is inversely related to the squared term of the desired seeds' adoption probability β. Figure A7-B illustrates this point. Note that this simplified numerical example considers only the effect of clustered seeding on one-hop neighbors. In order to more fully assess whether clustered seeding is better able to start cascades beyond one-hop neighbors it becomes critically important to consider the role of network topology (Eckles et al., 2019; Pastor-Satorras & Vespignani, 2001).

9. Consistency in Linear Probability Models

We consider a concern over using linear probability models (LPM). Horrace and Oaxacana (2006) point out that "[c]onsistency seems to be an exceedingly rare occurrence as one would have to accept extraordinary restrictions on the joint distribution of the regressors. Therefore, OLS is frequently a biased estimator and almost always an inconsistent estimator of the LPM." That is, given outcomes $Y \in \{0,1\}$ and fixed covariate vectors X, the OLS estimator is consistent and unbiased under the linear probability model if $X_i^T \beta \in [0,1]$ for all i, otherwise the OLS estimator is biased and inconsistent.

As a solution to this problem, they suggest a 'trimming' estimator approach: "[t]rimming observations whose predictions lie outside the unit interval and re-estimating the OLS model (based on the trimmed sample) may reduce finite sample bias." The procedures of this approach are as follows: (i) estimate the LPM; (ii) drop observations whose $\hat{y}$ lies outside $[0, 1]$ until $X_i^T \beta \in [0,1]$ for all i; and (iii) re-estimate the model with the remaining observations.

We adopt the trimming approach and examine whether the OLS estimates in our ITT models are consistent even after we trim observations (see Table A7). There are varying rates of observations being trimmed. The trimmed cases are more prevalent in earlier time windows (42% at 6 hours to 6% at 7 days). But, even after trimming out such cases, our findings regarding the difference of adoption probabilities between $k_{ITT} = 1$ and $k_{ITT} = 2$ generally hold, indicated by Wald test statistics large enough to reject the null hypothesis of $p_{k=2}^{Obs} = p_{k=2}^{Simple}$ and $\alpha > 1$ indicative of positive contagion effects across all the time windows.

Dependent Variable:	Adopted Product within first ...					
	6h	12h	24h	36h	48h	7 days
	(1)	(2)	(3)	(4)	(5)	(6)
Peer Effects						
Intercept	-0.0017^{***}	-0.0004^{***}	-0.0000	0.0002^{***}	0.0002^{***}	0.0010^{***}
	(0.0003)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0001)
$k_{ITT} = 1$	0.0017^{***}	0.0006^{***}	0.0004^{***}	0.0004^{***}	0.0005^{***}	0.0007^{***}
	(0.0003)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0001)
$k_{ITT} = 2$	0.0023^{***}	0.0015^{***}	0.0013^{***}	0.0017^{***}	0.0019^{***}	0.0040^{***}
	(0.0004)	(0.0003)	(0.0004)	(0.0004)	(0.0004)	(0.0006)
Controls						
Degree	-0.0001^{***}	-0.0000^{**}	-0.0001^{***}	-0.0001^{***}	-0.0001^{***}	-0.0003^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Duration	-0.0029^{***}	-0.0023^{***}	-0.0045^{***}	-0.0061^{***}	-0.0072^{***}	-0.0151^{***}
	(0.0006)	(0.0005)	(0.0006)	(0.0006)	(0.0007)	(0.0010)
Frequency	0.0007^{***}	0.0007^{***}	0.0008^{***}	0.0009^{***}	0.0010^{***}	0.0016^{***}
	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)
Embeddedness	0.0172^{***}	0.0199^{***}	0.0228^{***}	0.0297^{***}	0.0318^{***}	0.0553^{***}
	(0.0015)	(0.0014)	(0.0014)	(0.0015)	(0.0016)	(0.0020)
Prop. Trimmed	0.4248	0.2409	0.1502	0.0670	0.0691	0.0606
Wald $P^{Obs}_{k=2} = P^{Simple}_{k=2}$	4.7657	11.4399	6.9084	7.9298	9.0378	19.1972
α	5.4056	2.3022	1.7789	1.6266	1.6198	1.5740
Adj. R^2	0.0005	0.0005	0.0005	0.0006	0.0006	0.0010
Num. obs.	1439279	1899399	2126250	2334494	2329409	2350439

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table A7. Robustness check after repeatedly trimming observations whose $\hat{y} \notin [0,1]$

10. Algorithm to find a solution for infection force equation

```
Algorithm 1: Binary-Search Based Computation
Input
  $k$: number of peer exposure
  $\hat{\beta}_k$: estimated infection probability at $k$
Output
  $\alpha$: strength of social reinforcement between $k$ and $k-1$

Function Error_Computation($\alpha$, $\hat{\beta}_k$, $\widehat{\beta_{k-1}}$, $k$):
    $error \leftarrow (1 - (1 - \widehat{\beta_{k-1}}^{\frac{1}{\alpha}})^{k})^{\alpha} - \hat{\beta}_k$
    return $error$

Define
  $l$: numerical solution's lower bound
  $h$: numerical solution's upper bound
  $\varepsilon$: allowed relative error

Function BinarySearch($\hat{\beta}_k$, $\widehat{\beta_{k-1}}$, $k$,$l$,$h$,$\varepsilon$):
    while Error_Computation($l$, $\hat{\beta}_k$, $\widehat{\beta_{k-1}}$, $k$) $> 0$ do
        $l \leftarrow l * 0.1$
    end
    while Error_Computation($h$, $\hat{\beta}_k$, $\widehat{\beta_{k-1}}$, $k$) $< 0$ do
        $h \leftarrow h * 10$
    end
    $solution \leftarrow \frac{l+h}{2}$
    while |Error_Computation($solution$, $\hat{\beta}_k$, $\widehat{\beta_{k-1}}$, $k$)| $\geq \hat{\beta}_k * \varepsilon$ do
        if Error_Computation($solution$, $\hat{\beta}_k$, $\widehat{\beta_{k-1}}$, $k$)$> 0$ then
            $h \leftarrow solution$
        else
            $l \leftarrow solution$
        end
        $solution \leftarrow \frac{l+h}{2}$
    end
    return $solution$
```

Algorithm A1. Pseudo-code for a binary search to solve Eq. 1 and find α.

REFERENCES

An, X., Gabert, K., Fox, J., Green, O., & Bader, D. A. (2019). Skip the intersection: Quickly counting common neighbors on shared-memory systems. *2019 IEEE High Performance Extreme Computing Conference, HPEC 2019*.

Aral, S. (2016). Networked Experiments. In Y. Bramoullé, A. Galeotti, & B. Rogers (Eds.), *The Oxford Handbook on the Economics of Networks*. Oxford University Press.

Aral, S., Muchnik, L., & Sundararajan, A. (2009). Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(51), 21544–21549.

Aral, S., & Nicolaides, C. (2017). Exercise contagion in a global social network. *Nature Communications*, *8*, 14753.

Aral, S., & Walker, D. (2014). Tie Strength, Embeddedness, and Social Influence: A Large-Scale Networked Experiment. *Management Science*, *60*(6), 1–19.

Aronow, P. M. (2012). A General Method for Detecting Interference Between Units in Randomized Experiments. *Sociological Methods and Research*, *40*(1), 3–16.

Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. (2012). The Role of Social Networks in Information Diffusion. *Proceedings of the 21st International Conference on World Wide Web*, *1201.4145*, 519–528.

Barrat, A., Barthelemy, M., & Vespignani, A. (2008). *Dynamical processes on complex networks*. Cambridge university press.

Centola, D. (2010). The spread of behavior in an online social network experiment. *Science (New York, N.Y.)*, *329*(5996), 1194–1197.

Eckles, D., Kizilcec, R. F., & Bakshy, E. (2016). Estimating peer effects in networks with peer encouragement designs. *Proceedings of the National Academy of Sciences*, *113*(27), 7316–7322.

Eckles, D., Mossel, E., Rahimian, M. A., & Sen, S. (2019). Long Ties Accelerate Noisy Threshold-based Contagions. *SSRN Electronic Journal*.

Ferrali, R., Grossman, G., Platas, M. R., & Rodden, J. (2020). It takes a village: Peer effects and externalities in technology adoption. *American Journal of Political Science*, *64*(3), 536–553.

Fink, C., Schmidt, A., Barash, V., Cameron, C., & Macy, M. (2016). Complex contagions and the diffusion of popular Twitter hashtags in Nigeria. *Social Network Analysis and Mining*, *6*(1), 1–19.

Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*. W.W. Norton.

Imbens, G. W., & Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, *62*(2), 467–475. https://doi.org/10.2307/2951620

Iyengar, R., & Berger, J. (2014). *How the Quantity and Timing of Social Influence Impact Product Adoption*.

Karsai, M., Iñiguez, G., Kaski, K., & Kertész, J. (2014). Complex contagion process in spreading of online innovation. *Journal of the Royal Society Interface*, *11*(101), 20140694.

Mønsted, B., Sapieżyński, P., Ferrara, E., & Lehmann, S. (2017). Evidence of complex contagion of information in social media: An experiment using Twitter bots. *PLOS ONE*, *12*(9), e0184148.

Pastor-Satorras, R., & Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Physical Review Letters*, *86*(14), 3200–3203.

Rubin, D. B. (1980). Randomization Analysis of Experimental Data: The Fisher Randomization Test Comment. *Journal of the American Statistical Association*, *75*, 591–593.

State, B., & Adamic, L. (2015). The diffusion of support in an online social movement: Evidence from the adoption of equal-sign profile pictures. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 1741–1750.

Taylor, S. J., & Eckles, D. (2018). Randomized experiments to detect and estimate social influence in networks. In S. Lehmann & Y. Ahn (Eds.), *Complex Spreading Phenomena in Social Systems* (pp. 1–24). Springer.

Ugander, J., Backstrom, L., Marlow, C., & Kleinberg, J. (2012). Structural diversity in social contagion. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(16), 5962–5966.

Weng, L., Menczer, F., & Ahn, Y.-Y. (2013). Virality Prediction and Community Structure in Social Networks. *Nature Scientific Reports*, *3*(2522), 1–6.