Estimating Peer-Influence Effects Under Homophily: Randomized Treatments and Insights

Niloy Biswas and Edoardo M. Airoldi

Abstract When doing causal inference on networks, there is interference among the units. In a social network setting, such interference among individuals is known as peer-influence. Estimating the causal effect of peer-influence under the presence of homophily presents various challenges. In this paper, we present results quantifying the error incurred from ignoring homophily when estimating peer-influence on networks. We then present randomized treatment strategies on networks which can help disentangle homophily from the estimation of peer-influence.

1 Introduction

When doing causal inference there is often *interference* among the units of interest. Interference is when the response to treatment of a unit is affected by the treatments assigned to its neighbors. In a social network setting, where a unit corresponds to an individual and an individual's neighborhood corresponds to their peers, such interference among individuals is known as peer-influence. With the increased usage of social media and availability of network data, understanding the casual effects of peer-influence has garnered much interest. The research area yields a wide range of applications. For example, in advertising Bakshy [5] examined the impact of friends' product affiliation on advertisements via randomized experiments on Facebook users; Aral [3] used randomized experiments on Facebook to examine how firms can design social media marketing campaigns to create peer-influence. In politics, Bond [6] assessed voting behavior results from randomized experiments on Facebook (where political mobilization messages were delivered to Facebook users via a randomized control trial during the 2010 US congressional elections) to find that effect of peer-influence on voting turnout was greater than the effect of the direct messages themselves.

N. Biswas (✉) · E. M. Airoldi
Department of Statistics, Harvard University, Cambridge, USA
e-mail: niloy_biswas@g.harvard.edu

E. M. Airoldi
e-mail: airoldi@fas.harvard.edu

S. Cornelius et al. (eds.), *Complex Networks IX*, Springer Proceedings
in Complexity, https://doi.org/10.1007/978-3-319-73198-8_28

There is recent work on methodology for estimation of causal peer-influence effects (e.g., Toulis and Kao [10], Athey et al. [4]). However, identifying and estimating peer-influence under the presence of homophily have long remained a challenging problem [2, 8, 9]. This is the problem of identifying to what extent the response of an individual is attributable to the treatments given to its neighbors (peer-influence) or attributable to a latent, intrinsic similarity between peers (homophily). This paper makes several contributions toward tackling this problem. In particular, we: (i) Introduce a framework for modeling peer-influence and homophily; (ii) under different models of peer-influence and homophily, quantify the error incurred from ignoring homophily in the estimation of the peer-influence effect; (iii) under a stochastic block model framework for the network, devise randomized treatment strategies which can help disentangle latent homophily from the estimation of peer-influence. Our randomized treatment strategies can also be applied in a more general setting, for general inference of network features in the presence of latent homophily.

2 Peer-Influence Under Homophily: Results and Inference Strategies

2.1 A General Framework for Modeling Peer-Influence and Homophily

Peer-influence is used to denote when the response of one individual is affected by the treatments assigned to its neighbors (e.g., friends in a social network). For individual i, this can be represented by $\mathbf{peer}((Z_j)_{j\in\mathcal{N}_i})$, where $(Z_j)_{j\in\mathcal{N}_i}$ are the treatments assigned to the neighbors of i and $\mathbf{peer}(\cdot)$ is a function taking values in the space of responses.

Homophily represents the latent, intrinsic similarity between close individuals in a network. For $j = 1, \ldots, N$, let X_j be independent and identically distributed random variables corresponding to the latent variable associated with individual j in the network. Then, for individual i, homophily can be represented by $\mathbf{hom}((X_j)_{j\in\mathcal{N}_i})$, where $(X_j)_{j\in\mathcal{N}_i}$ are the latent variables in the neighborhood of i and $\mathbf{hom}(\cdot)$ is a function taking values in the space of responses.

We now introduce the general framework used in our analysis of peer-influence and homophily. Suppose we are interested in the responses of N units in a network. This is represented by the random variables Y_i for $i = 1, \ldots, n$. The response Y_i of the ith unit depends on its treatment, peer-influence, and latent homophily. Our full model is given by:

$$Y_i(Z_i = 0, (Z_j)_{j\in\mathcal{N}_i}, (X_j)_{j\in\mathcal{N}_i}) = \alpha + \beta_0 \mathbf{peer}((Z_j)_{j\in\mathcal{N}_i}) + h_0 \mathbf{hom}((X_j)_{j\in\mathcal{N}_i}) + \epsilon_i(0, \sigma_Y^2) \quad (1)$$

$$Y_i(Z_i = 1, (Z_j)_{j\in\mathcal{N}_i}, (X_j)_{j\in\mathcal{N}_i}) = \tau + Y_i(Z_i = 0, (Z_j)_{j\in\mathcal{N}_i}, (X_j)_{j\in\mathcal{N}_i}) + \beta_1 \mathbf{peer}((Z_j)_{j\in\mathcal{N}_i}) + h_1 \mathbf{hom}((X_j)_{j\in\mathcal{N}_i}) \quad (2)$$

$\epsilon_i(0, \sigma_Y^2)$ for $i = 1, \dots, N$ are the noise terms in the network, indepedent and identically distributed according to an unknown distribution with zero mean and variance σ_Y^2. β_0, β_1 are the unknown peer-influence parameters, and h_0, h_1 are the unknown homophily parameters. Latent effects due to homophily in the model are represented by indepedent and identically distributed random variables $(X_i)_{i=1}^N$ with mean 1 and variance σ_X^2. $\mathbf{Z}$ are the assigned treatments.

Under different models of the peer-influence $\mathbf{peer}(\cdot)$ and homophily $\mathbf{hom}(\cdot)$, we will focus on estimating peer-influence and homophily parameters β_0 and h_0, respectively, assuming the variances are known. Note that our analysis here is focussed on inference concerning the untreated individuals (1 above), but all the methodology can be easily applied to the treated individuals in the network (2 above). In our analysis, we consider the significance of the following factors in the inference of peer-influence under the presence of homophily, their consequences for the design of experiments:

1. Modeling of peer-influence: as a binary ($\mathbf{peer}((Z_j)_{j\in\mathcal{N}_i}) = \mathbf{1}_{\sum_{j\in\mathcal{N}_i} Z_j>0}$) or a linear ($\mathbf{peer}((Z_j)_{j\in\mathcal{N}_i}) = \sum_{j\in\mathcal{N}_i} Z_j$) effect.
2. Modeling of homophily: as an unnormalized ($\mathbf{hom}((X_j)_{j\in\mathcal{N}_i}) = \sum_{j\in\mathcal{N}_i} X_j$) or normalized ($\mathbf{hom}((X_j)_{j\in\mathcal{N}_i}) = \sum_{j\in\mathcal{N}_i} X_j/|\mathcal{N}_i|$) latent factor. Unnormalized homophily corresponds to when dense regions of the network have a stronger homophily effect compared to more sparse regions. Normalized homophily corresponds to when the homophily effect is not affected by the density of different regions in the network.
3. Choice of peer-influence estimate: as a difference of means estimate (for binary peer-influence) or as the average of stratified estimates (for linear peer-influence).
4. Allocation of treatments: fixed optimal treatment allocation or randomized treatment.

In this short paper, we only discuss results and strategies for the case of binary peer-influence under unnormalized homophily. Discussion of the other cases is included in the appendix.

Binary peer-influence effect with unnormalized homophily. Consider the binary peer-influence model with unnormalized homophily. For the untreated individuals, we have

$$Y_i(Z_i = 0, (Z_j)_{j\in\mathcal{N}_i}) = \alpha + \beta_0 \mathbf{1}_{\sum_{j\in\mathcal{N}_i} Z_j>0} + h_0 \sum_{j\in\mathcal{N}_i} X_j + \epsilon_i(0, \sigma_Y^2) \quad (3)$$

where $\epsilon_i(0, \sigma_Y^2)$ are independent and identically distributed with zero mean and σ_Y^2 variance.

Consider estimating the peer-influence parameter β_o using a difference in means estimator. Partition the set of untreated individuals into sets $M_0^{(0)} := \{i : Z_i = 0, \sum_{j \in \mathcal{N}_i} Z_j = 0\}$ (the set of untreated individuals with no treated neighbors) and $M_0^{(1)} := \{i : Z_i = 0, \sum_{j \in \mathcal{N}_i} Z_j > 0\}$ (the set of untreated individuals with at least one treated neighbors). Then, the difference in means estimator for β_0 is given by:

$$\hat{\beta}_0 = \underset{i \in M_0^{(1)}}{avg}\, Y_i - \underset{i \in M_0^{(0)}}{avg}\, Y_i \tag{4}$$

Under the negligence of latent homophily in the model, this difference of means estimator for peer-influence would appear unbiased. However, the presence of latent homophily actually interferes and introduces bias to the estimation of peer-influence, as highlighted in Theorem 1 below.

Theorem 1 *Consider the difference in means estimator $\hat{\beta}_0$ for binary peer-influence effect β_0. Under the presence of unnormalized homophily in our model (3), the mean squared error of $\hat{\beta}_0$ (conditional on the treatment $\mathbf{Z}$) is:*

$$\begin{aligned}\mathbb{E}[(\hat{\beta}_0 - \beta_0)^2 | \mathbf{Z}] = &\left(h_0 \left(\underset{i \in M_0^{(1)}}{avg} |\mathcal{N}_i| - \underset{i \in M_0^{(0)}}{avg} |\mathcal{N}_i| \right) \right)^2 \\ &+ h_0^2 \sigma_X^2 \left(\underset{i,j \in M_0^{(0)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j| + \underset{i,j \in M_0^{(1)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j| - 2 \underset{i \in M_0^{(0)}, j \in M_0^{(1)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j| \right) \\ &+ \sigma_Y^2 \left(\frac{1}{|M_0^{(0)}|} + \frac{1}{|M_0^{(1)}|} \right)\end{aligned} \tag{5}$$

We can interpret (5) to understand the optimal treatment allocation with respect to minimizing the bias and variance. For binary peer-influence effect with unnormalized homophily, the bias of $\hat{\beta}_0$ is minimized through an assignment of treatments $\mathbf{Z}$ which manages to balance the average homophily effect (corresponding to average vertex degrees) between individuals in $M_0^{(1)}$ and $M_0^{(0)}$. Under such balanced treatment assignment, unbiasedness is achieved when

$$\underset{i \in M_0^{(1)}}{avg} |\mathcal{N}_i| = \underset{i \in M_1^{(1)}}{avg} |\mathcal{N}_i| = \underset{i \in M_0^{(0)} \cup M_0^{(1)}}{avg} |\mathcal{N}_i|,$$

where $M_0^{(0)} \cup M_0^{(1)}$ is the set of all (untreated) individuals. For binary peer-influence effect with unnormalized homophily, the variance of $\hat{\beta}_0$ is minimized through treatments $\mathbf{Z}$ which:

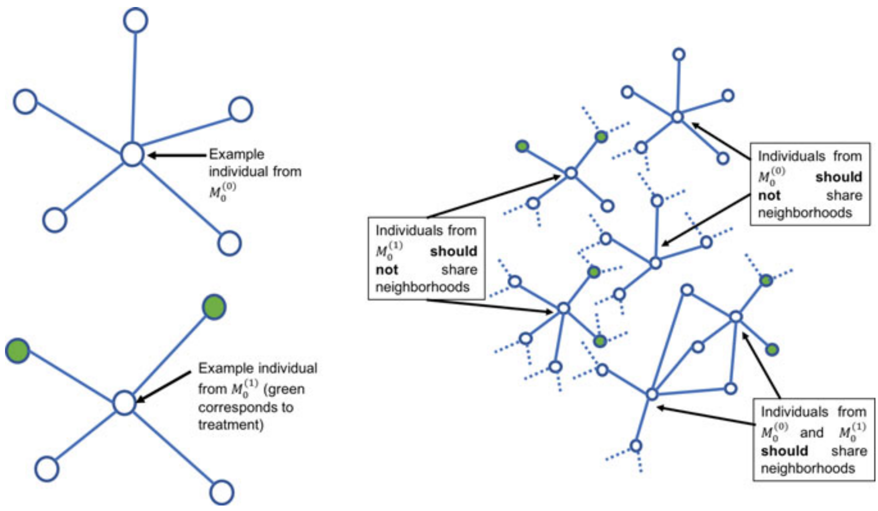


Fig. 1 Optimal treatment allocation

1. Ensure $|M_0^{(1)}| = |M_0^{(0)}|$, such that there is balance between the number of individuals which are affected and not affected by peer-influence (in set $M_0^{(1)}$ and $M_0^{(0)}$, respectively).
2. Ensure that the individuals in $M_0^{(1)}$ are mixed with individuals in $M_0^{(0)}$ as well as possible. In particular, this corresponds to choosing $\mathbf{Z}$ such that elements in $M_0^{(0)}$ have minimal shared neighborhoods between themselves (minimizing $\underset{i,j\in M_0^{(0)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j|$), elements in $M_0^{(1)}$ have minimal shared neighborhoods between themselves (minimizing $\underset{i,j\in M_0^{(1)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j|$), and shared neighborhoods between elements of $M_0^{(0)}$ and $M_0^{(1)}$ are maximal, respectively (maximizing $\underset{i\in M_0^{(0)},j\in M_0^{(1)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j|$). This is illustrated through Fig. 1.

Having developed conceptual insights into what treatment assignments are optimal for inferring peer-influence in the presence of homophily, we can further extend our analysis to consider the cases of randomized treatment. In particular, by considering randomized treatment under a stochastic block model framework (e.g., see Holland [7], Airoldi [1]), we can take advantage of symmetries and exchangeability to gain insight into the optimal design of randomized experiments under such framework. This is explored in Sect. 2.2 where randomized treatment designs which disentangle homophily from the estimation of peer-influence are considered.

2.2 Disentangling Homophily from Estimation of Peer-Influence: Randomized Treatment Strategies

We now propose a general strategy for reducing bias in the inference of binary peer-influence under the presence of homophily. It is applicable to weighted, directed graphs which are clustered. Furthermore, our strategy does not assume any model for homophily $\mathbf{hom}(\cdot)$, which can remain unknown.

Suppose we have a graph G of N vertices which is clustered into r clusters. Assume that given clustering of the graph captures the covariates of the individuals in the network, such that individuals with same or similar covariates are members of the same cluster. Under such a clustering, we fit a corresponding stochastic block model onto the network of N individuals in r communities. Note that by fitting such a stochastic block model, we are implicitly assuming that individuals in the same cluster are exchangeable (hence the need to have a good cluster for this assumption to be justified). Denote the communities of the fitted stochastic block model by the sets $B_1, \ldots, B_r$, which are of respective sizes $A_1, \ldots, A_r$ (where $A_1 + \cdots + A_r = N$). Let $\mathbf{P}$ be the $r \times r$ adjacency probability matrix between the r communities. Values $A_1, \ldots, A_r$ directly are obtained from the cluster sizes, and the entries of the matrix $\mathbf{P}$ can be estimated using MLEs (e.g., in the unweighted graph case, we can choose $[\mathbf{P}]_{i,j} = \frac{\text{number of edges from cluster i to j}}{|A_i||A_j|}$). Within each community B_s, different individuals are affected by different levels of peer-influence. For example, in the binary peer-influence case, untreated individuals are either in set $M_0^{(0)}$ (no treated neighbors—not affected by peer-influence) or in set $M_0^{(1)}$ (at least one treated neighbors—affected by peer-influence); in the linear peer-influence case, untreated individuals in $M_0^{(k)}$ are affected by the k-levels of peer-influence.

When estimating peer-influence, the bias due to homophily arises from imbalances in the homophily effect between the sets of individuals with different levels of peer-influence. This motivates the key idea in our design of randomized treatments to remove bias from homophily: We want to design experiments such that in every community B_s, an equal number of individuals are affected and not affected by peer-influence. By the construction of the cluster, every community B_s has a similar effect due to latent homophily. Therefore by designing randomized experiments which ensure that every such cluster B_s has an equal (expected) number of individuals with different levels of peer-influence, we reduce the bias in the estimation of peer-influence arising from latent homophily. For randomized treatments where individuals in cluster s are treated independently with probability θ_s, our strategy described leads to constrained optimization problems for θ_s. This can then be solved to obtain optimal θ_s^{opt} values as required for reducing bias in the estimation of peer-influence under the presence of latent, unknown homophily. We now highlight our strategy in detail for the estimation of binary peer-influence under the presence of homophily (linear peer-influence case in appendix).

An algorithm for inference of binary peer-influence. For a weighted, directed graph G of N vertices which is clustered into r clusters, consider a corresponding stochastic block model of N individuals in r communities. Denote the communities of the SBM (clusters of G) by the sets $B_1, \ldots, B_r$, which are of respective sizes $A_1, \ldots, A_r$ (where $A_1 + \cdots + A_r = N$). Let $\mathbf{P}$ be the $r \times r$ adjacency probability matrix between the r communities. We assign treatments independently to individuals such that individuals in B_s are treated with probability θ_s for $s = 1, \ldots, r$. We want to choose θ_s with the aim of reducing bias, such that homophily does not interfere with the estimation of peer-influence. Note that the general bias of the binary peer-influence estimator $\hat{\beta}_0$ is

$$\underset{i \in M_0^{(1)}}{avg}\, \mathbb{E}_X[\mathbf{hom}((X)_{j \in \mathcal{N}_i})] - \underset{i \in M_0^{(0)}}{avg}\, \mathbb{E}_X[\mathbf{hom}((X)_{j \in \mathcal{N}_i})].$$

This highlights that the bias in our estimation arises from an imbalance in the average homophily effect between the sets $M_0^{(1)}$ and $M_0^{(0)}$ (the individuals who are and are not affected by peer-influence, respectively). This observation motivates the key idea in our design of randomized treatments to remove bias from homophily: We want to design experiments such that in every cluster B_s, an equal number of individuals are affected and not affected by peer-influence. For randomized treatment assignment, this means we want

$$\forall s = 1, \ldots, r, \quad \mathbb{E}[|M_0^{(1)} \cap B_s|] = \mathbb{E}[|M_0^{(0)} \cap B_s|]. \tag{6}$$

Let us now derive a result about $M_0^{(0)}$ and $M_0^{(1)}$ under our framework to proceed further with (6).

Proposition 1 *Consider a stochastic block model (SBM) of N individuals in r communities. Denote the communities of the SBM by the sets $B_1, \ldots, B_r$, which are of respective sizes $A_1, \ldots, A_r$ (where $A_1 + \cdots + A_r = N$). Let $\mathbf{P}$ be the $r \times r$ adjacency probability matrix between the r communities. We assign treatments independently to individuals such that individuals in B_s are treated with probability θ_s for $s = 1, \ldots, r$. Under such setup, let $M_0^{(0)}$ denote the set of untreated individuals which have no treated neighbors and let $M_0^{(1)}$ denote the set of untreated individuals which have at least one treated neighbor. For ease of notation, let $\{s \in M_0^{(0)}\}$, $\{s \in M_0^{(1)}\}$ denote the event that a fixed vertex in community s is in the set $M_0^{(0)}$, $M_0^{(1)}$, respectively. Then,*

$$\mathbb{P}(s \in M_0^{(0)}) = (1 - \theta_s) \prod_{v=1}^{r} (1 - P_{s,v}\theta_v)^{A_v - \mathbf{1}_{v=s}}, \text{ and} \tag{7}$$

$$\mathbb{P}(s \in M_0^{(1)}) = (1 - \theta_s)\left(1 - \prod_{v=1}^{r} (1 - P_{s,v}\theta_v)^{A_v - \mathbf{1}_{v=s}}\right). \tag{8}$$

Using Proposition (1), we can now directly derive an algorithm to reduce the effect of homophily during inference. Let s denote any vertex in the graph which is in community B_s. Note that $\mathbb{E}[|M_0^{(0)} \cap B_s|] = A_s \mathbb{P}(s \in M_0^{(0)})$ and $\mathbb{E}[|M_0^{(1)} \cap B_s|] = A_s \mathbb{P}(s \in M_0^{(1)}) = A_s(1 - \mathbb{P}(s \in M_0^{(0)}))$, as all untreated individuals are in either $M_0^{(0)}$ or $M_0^{(1)}$. This gives,

$$\begin{aligned}\mathbb{E}[|M_0^{(1)} \cap B_s|] = \mathbb{E}[|M_0^{(0)} \cap B_s|] &\iff \mathbb{P}(s \in M_0^{(0)}) = \frac{1}{2}\mathbb{P}(Z_s = 0) \\ &\iff (1 - \theta_s)\prod_{v=1}^{r}(1 - P_{s,v}\theta_v)^{A_v - \mathbf{1}_{v=s}} = \frac{1}{2}(1 - \theta_s) \\ &\iff \sum_{v=1}^{r}(A_v - \mathbf{1}_{v=s})log(1 - P_{s,v}\theta_v) + log(2) = 0\end{aligned}$$

For $|P_{s,v}| \approx 0$, $log(1 - P_{s,v}\theta_v) \approx -P_{s,v}\theta_v$. This allows us to approximate the optimal θ_s values by simply solving a set of linear equations. Our algorithm is given below.

Algorithm 1: Randomized treatment design for more accurate inference of peer-influence

1 function optimal_treatment_values $(\mathbf{A}(G), \mathbf{B}(G))$;
Input : Adjacency matrix $\mathbf{A}$ and clustering $\mathbf{B}$ (with r clusters) of some graph G
Output: Treatment probabilities $\theta \in [0, 1]^r$ for Bernoulli assignment on each cluster
2 Fit an SBM, giving an adjacency matrix $\mathbf{P}$ for clusters $B_1, \ldots, B_r$ of sizes $A_1, \ldots, A_r$.
3 Choose treatment probabilities for the clusters $\theta \in [0, 1]^r$ as the solution to:

$$\begin{pmatrix} P_{1,1}(A_1-1) & P_{1,2}A_2 & \ldots & \ldots & P_{1,r}A_r \\ P_{2,1}A_1 & P_{2,2}(A_2-1) & \ldots & \ldots & P_{2,r}A_r \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ P_{r-1,1}A_1 & P_{r-1,2}A_2 & \ldots & P_{r-1,r-1}(A_{r-1}-1) & P_{r-1,r}A_r \\ P_{r,1}A_1 & P_{r,2}A_2 & \ldots & P_{r,r-1}A_{r-1} & P_{r,r}(A_r-1) \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_{r-1} \\ \theta_r \end{pmatrix} = log(2) \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix} \quad (9)$$

In practice, if (9) does not have a solution in $[0, 1]^r$, we can solve the constrained optimization problem of minimizing

$$\left\| \begin{pmatrix} P_{1,1}(A_1-1) & P_{1,2}A_2 & \ldots & \ldots & P_{1,r}A_r \\ P_{2,1}A_1 & P_{2,2}(A_2-1) & \ldots & \ldots & P_{2,r}A_r \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ P_{r-1,1}A_1 & P_{r-1,2}A_2 & \ldots & P_{r-1,r-1}(A_{r-1}-1) & P_{r-1,r}A_r \\ P_{r,1}A_1 & P_{r,2}A_2 & \ldots & P_{r,r-1}A_{r-1} & P_{r,r}(A_r-1) \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_{r-1} \\ \theta_r \end{pmatrix} - log(2) \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix} \right\| \quad (10)$$

for $\theta \in [0, 1]^r$ and for some chosen norm $\| \ \|$ on $\mathbb{R}^d$ (e.g., L^2). Note that under the optimal treatment probabilities θ^{opt} obtained from (9), the total expected number of treated individuals is $\sum_{s=1}^{r} A_s \theta_s^{opt}$. In practice, often it is desirable to control the expected number of individuals treated under randomized treatment. This is can be done under our framework by considering the constrained optimization problem of minimizing the norm in (9) subject to $\theta \in [0, 1]^r$ and $\sum_{s=1}^{r} A_s \theta_s = Nx$, where x is our chosen percentage of individuals treated.

Analysis of treatment strategies via simulations. We highlight the performance of our randomized treatment strategy compared to alternatives via numerical results from Monte Carlo simulations under a stochastic block model. We consider the bias and mean squared error of our optimal randomized treatment compared to other common randomized treatment strategies. Unsuccessful treatment occurs when either one of the sets $M_0^{(0)}$ or $M_0^{(1)}$ is empty, and the difference in means estimator for binary peer-influence (4) is ill-defined.

We consider the unnormalized sum of latent variables ($\mathbf{hom}((X)_{j \in \mathcal{N}_i}) = \sum_{j \in \mathcal{N}_i} X_j$ for X_j i.i.d. latent random variables with mean 1, variance σ_X) as our homophily function. The baseline simulation model (with binary peer-influence, unnormalized homophily) here is:

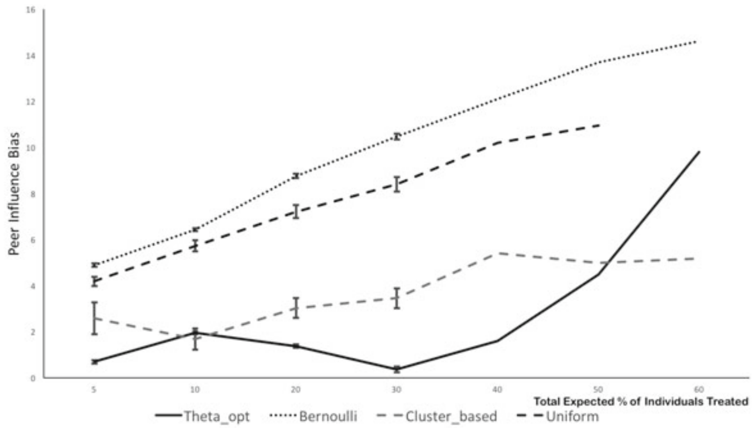
$$Y_i(Z_i = 0, (Z_j)_{j \in \mathcal{N}_i}, (X_j)_{j \in \mathcal{N}_i}) = \alpha + \beta_0 \mathbf{1}_{\sum_{j \in \mathcal{N}_i} Z_j > 0} + h_0 \sum_{j \in \mathcal{N}_i} X_j + \epsilon_i(0, \sigma_Y^2)$$

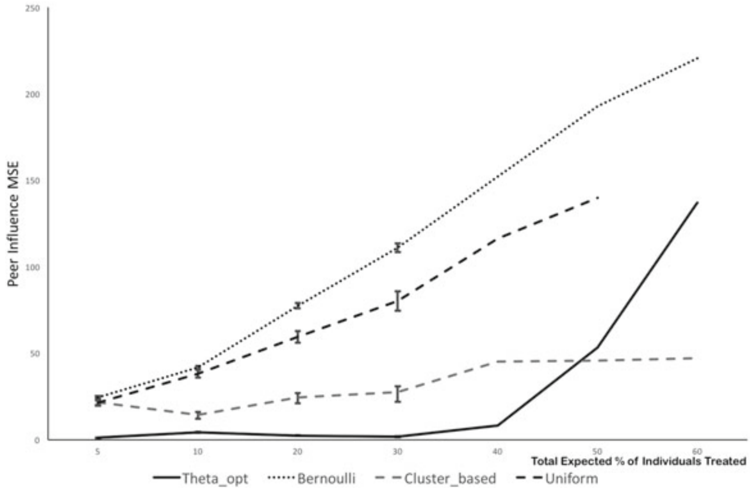
$$Y_i(Z_i = 1, (Z_j)_{j \in \mathcal{N}_i}, (X_j)_{j \in \mathcal{N}_i}) = \tau + Y_i(Z_i = 0, (Z_j)_{j \in \mathcal{N}_i}, (X_j)_{j \in \mathcal{N}_i}) + \beta_1 \mathbf{1}_{\sum_{j \in \mathcal{N}_i} Z_j > 0} + h_1 \sum_{j \in \mathcal{N}_i} X_j$$

for $\alpha = 3, \beta_0 = 0.1, h_0 = 1, \tau = 0.2, \beta_1 = 0.05, h_1 = 0.5, \sigma_Y = 1.5^2, \sigma_X = 1^2$.

SBM Graph Simulation. The following figures display the bias and variance of the difference in means estimator (y-axis) against the controlled expected percentage of treated individuals (x-axis) for our strategy in comparison with other common randomized treatment strategies. The figures highlight that our randomized treatment strategy leads to improved estimation of peer-influence. We are working with a directed SBM of 1530 vertices and 7 clusters with

$$(A_1, A_2, A_3, A_4, A_5, A_6, A_7) = (600, 340, 200, 150, 100, 90, 50)\ , \ \mathbf{P} = \begin{pmatrix} 50 & 10 & 20 & 5 & 5 & 15 & 3 \\ 10 & 30 & 5 & 15 & 15 & 10 & 5 \\ 10 & 5 & 40 & 5 & 10 & 13 & 12 \\ 4 & 5 & 10 & 25 & 15 & 14 & 12 \\ 14 & 15 & 10 & 5 & 20 & 10 & 10 \\ 13 & 14 & 5 & 2 & 10 & 35 & 15 \\ 10 & 14 & 14 & 5 & 5 & 10 & 45 \end{pmatrix} \Big/ 1530.$$





2.3 Concluding Remarks

When doing causal inference in networks, neglecting latent homophily can lead to inaccurate inference of peer-influence. In this paper, we have introduced a general framework for modeling peer-influence and homophily, quantified the error incurred from ignoring homophily, and devised randomised treatment strategies which allow the estimation of peer-influence in the presence of homophily. Simulations highlight our method's performance relative to other randomized treatment strategies. This work is a preliminary insight into a forthcoming project. Our future extensions will involve further statistical analysis and theoretical guarantees on the performance of the randomized treatment strategies, and results from experimentation on large real-world social networks.

A Appendices

A.1 Peer-Influence Under Homophily: Results and Inference Strategies

Binary peer-influence effect with normalized homophily: Consider now binary peer-influence effect with normalized homophily. For the untreated individuals, we have

$$Y_i(Z_i = 0, (Z_j)_{j\in\mathcal{N}_i}) = \alpha + \beta_0 \mathbf{1}_{\sum_{j\in\mathcal{N}_i} Z_j > 0} + h_0 \sum_{j\in\mathcal{N}_i} \frac{X_j}{|\mathcal{N}_i|} + \epsilon_i(0, \sigma_Y^2) \tag{11}$$

where $\epsilon_i(0, \sigma_Y^2)$ are idependent and identically distributed with zero mean and σ_Y^2 variance.

As before, consider estimating the peer-influence parameter β_o using a difference in means estimator. Partition the set of untreated individuals into sets $M_0^{(0)} := \{i : Z_i = 0, \sum_{j\in\mathcal{N}_i} Z_j = 0\}$ (the set of untreated individuals with no treated neighbors) and $M_0^{(1)} := \{i : Z_i = 0, \sum_{j\in\mathcal{N}_i} Z_j > 0\}$ (the set of untreated individuals with at least one treated neighbors). Then, the difference in means estimator for β_0 is given by:

$$\hat{\beta}_0 = \underset{i\in M_0^{(1)}}{avg}\, Y_i - \underset{i\in M_0^{(0)}}{avg}\, Y_i \tag{12}$$

Unlike in the case with unnormalized homophily, the difference of means estimator for peer-influence remains unbiased in the presence of normalized homophily. This is further highlighted in Theorem 2 below. Furthermore, for most sparse and dense models for the underlying graph, Theorem 2 can be used to show that $\hat{\beta}_0$ is a consistent estimator of peer-influence under normalized homophily.

Theorem 2 *Consider the difference in means estimator $\hat{\beta}_0$ for binary peer-influence effect β_0. Under the presence of normalized homophily in our model (11), the mean squared error of $\hat{\beta}_0$ (conditional on the treatment $\mathbf{Z}$) is:*

$$\mathbb{E}[(\hat{\beta}_0 - \beta_0)^2|\mathbf{Z}] =$$
$$h_0^2\sigma_X^2\left(\underset{i,j\in M_0^{(0)}}{avg} \frac{|\mathcal{N}_i \cap \mathcal{N}_j|}{|\mathcal{N}_i||\mathcal{N}_j|} + \underset{i,j\in M_0^{(1)}}{avg} \frac{|\mathcal{N}_i \cap \mathcal{N}_j|}{|\mathcal{N}_i||\mathcal{N}_j|} - 2 \underset{i\in M_0^{(0)}, j\in M_0^{(1)}}{avg} \frac{|\mathcal{N}_i \cap \mathcal{N}_j|}{|\mathcal{N}_i||\mathcal{N}_j|}\right) + \sigma_Y^2\left(\frac{1}{|M_0^{(0)}|} + \frac{1}{|M_0^{(1)}|}\right) \tag{13}$$

Linear peer-influence effect with unnormalized homophily: We now consider modeling peer-influence as a linear function of the number of treated neighbors

$\mathbf{peer}((Z_j)_{j\in\mathcal{N}_i}) = \sum_{j\in\mathcal{N}_i} Z_j$. For the untreated individuals under unnormalized homophily, this gives:

$$Y_i(Z_i = 0, (Z_j)_{j\in\mathcal{N}_i}) = \alpha + \beta_0 \sum_{j\in\mathcal{N}_i} Z_j + h_0 \sum_{j\in\mathcal{N}_i} X_j + \epsilon_i(0, \sigma_Y^2) \tag{14}$$

where $\epsilon_i(0, \sigma_Y^2)$ are idependent and identically distributed with zero mean and σ_Y^2 variance.

Consider estimating the peer-influence parameter β_0. Generalizing our methodology from the binary peer-influence case, we now develop a stratified estimator for β_0. Let

$$M_0^{(k)} := \{i : Z_i = 0, \sum_{j\in\mathcal{N}_i} Z_j = k\}$$

be the set of untreated individuals which have k treated neighbors. Then, an average of difference in means estimator for peer-influence is:

$$\hat{\beta}_0 = \frac{\sum_k \hat{\beta}_0^{(k)}}{\sum_k 1} \; for \; \hat{\beta}_0^{(k)} = \frac{1}{k}\left(\frac{\sum_{i\in M_0^{(k)}} Y_i}{|M_0^{(k)}|} - \frac{\sum_{i\in M_0^{(0)}} Y_i}{|M_0^{(0)}|}\right) = \frac{1}{k}\left(\underset{i\in M_0^{(k)}}{avg}\, Y_i - \underset{i\in M_0^{(0)}}{avg}\, Y_i\right). \tag{15}$$

where we average over all k such that $|M_0^{(k)}| > 0$ (so that $\hat{\beta}_0^{(k)}$ is well-defined). Note that here we are averaging over the class of estimators $\hat{\beta}_0^{(k)}$ under the assumption of linear peer-influence. In the case of nonlinearity, we can also consider each $\hat{\beta}_0^{(k)}$ separately to understand the kth-level peer-influence effect in the network.

The presence of latent unnormalized homophily interferes and introduces bias to the estimation of linear peer-influence, as highlighted in Theorem 3 below.

Theorem 3 *Consider the estimator $\hat{\beta}_0$ for linear peer-influence effect β_0. Under the presence of unnormalized homophily in our model (3), the mean squared error of $\hat{\beta}_0$ (conditional on the treatment $\mathbf{Z}$) is:*

$$\begin{aligned}
&\mathbb{E}[(\hat{\beta}_0 - \beta_0)^2|\mathbf{Z}] = \\
&\left(\frac{h_0}{\sum_{k>0} 1} \sum_{k>0} \frac{1}{k}\left(\underset{i\in M_0^{(k)}}{avg} |\mathcal{N}_i| - \underset{i\in M_0^{(0)}}{avg} |\mathcal{N}_i|\right)\right)^2 \\
&+ \frac{1}{(\sum_{k>0} 1)^2} \sum_{k,l>0} \frac{1}{kl}\Bigg[h_0^2\sigma_X^2\left(\underset{i\in M_0^{(k)}, j\in M_0^{(l)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j| + \underset{i,j\in M_0^{(0)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j| - 2 \underset{i\in M_0^{(0)}, j\in M_0^{(k)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j|\right) \\
&\qquad + \sigma_Y^2\left(\frac{1}{|M_0^{(0)}|} + \frac{\mathbf{1}_{k=l}}{|M_0^{(k)}|}\right)\Bigg]
\end{aligned} \tag{16}$$

Equation (16) highlights that unbiasedness estimation via optimal treatment allocation may be difficult computationally, as now we need to ensure balance across all the strata $(M_0^{(k)})_{k\geq 0}$. This motivates an alternative approach of unbiased estimation.

Linear peer-influence effect with normalized homophily: For the peer-influence effect on untreated individuals under normalized homophily, we obtain:

$$Y_i(Z_i = 0, (Z_j)_{j\in\mathcal{N}_i}) = \alpha + \beta_0 \sum_{j\in\mathcal{N}_i} Z_j + h_0 \sum_{j\in\mathcal{N}_i} \frac{X_j}{|\mathcal{N}_i|} + \epsilon_i(0, \sigma_Y^2) \tag{17}$$

where $\epsilon_i(0, \sigma_Y^2)$ are idependent and identically distributed with zero mean and σ_Y^2 variance.

To estimate the peer-influence parameter β_0, the same stratified estimator as in the linear peer-influence with unnormalized homophily case can be applied:

$$\hat{\beta}_0 = \frac{\sum_k \hat{\beta}_0^{(k)}}{\sum_k 1} \; for \; \hat{\beta}_0^{(k)} = \frac{1}{k}\left(\frac{\sum_{i\in M_0^{(k)}} Y_i}{|M_0^{(k)}|} - \frac{\sum_{i\in M_0^{(0)}} Y_i}{|M_0^{(0)}|}\right) = \frac{1}{k}\left(\underset{i\in M_0^{(k)}}{avg}\, Y_i - \underset{i\in M_0^{(0)}}{avg}\, Y_i\right). \tag{18}$$

where $M_0^{(k)} := \{i : Z_i = 0, \sum_{j\in\mathcal{N}_i} Z_j = k\}$ and we are averaging over all k such that $|M_0^{(k)}| > 0$.

In the presence of normalized homophily, $\hat{\beta}_0$ remains an unbiased estimator of peer-influence. This is highlighted in Theorem 4 below.

Theorem 4 *Consider the estimator $\hat{\beta}_0$ for linear peer-influence effect β_0. Under the presence of normalized homophily in our model (11), $\hat{\beta}_0$ is unbiased and the mean squared error of $\hat{\beta}_0$ (conditional on the treatment $\mathbf{Z}$) is:*

$$\mathbb{E}[(\hat{\beta}_0 - \beta_0)^2|\mathbf{Z}] = \frac{1}{(\sum_{k>0} 1)^2} \sum_{k,l>0} \frac{1}{kl}\Bigg[h_0^2\sigma_X^2\Bigg(\underset{i\in M_0^{(k)}, j\in M_0^{(l)}}{avg} \frac{|\mathcal{N}_i \cap \mathcal{N}_j|}{|\mathcal{N}_i||\mathcal{N}_j|} + \underset{i,j\in M_0^{(0)}}{avg} \frac{|\mathcal{N}_i \cap \mathcal{N}_j|}{|\mathcal{N}_i||\mathcal{N}_j|} - 2 \underset{i\in M_0^{(0)}, j\in M_0^{(k)}}{avg} \frac{|\mathcal{N}_i \cap \mathcal{N}_j|}{|\mathcal{N}_i||\mathcal{N}_j|}\Bigg) + \sigma_Y^2\left(\frac{1}{|M_0^{(0)}|} + \frac{\mathbf{1}_{k=l}}{|M_0^{(k)}|}\right)\Bigg] \tag{19}$$

The difference of means estimator for linear peer-influence remains unbiased in the presence of normalized homophily. Furthermore, for most sparse and dense models for the underlying graph, Theorem 2 can be used to show that $\hat{\beta}_0$ is a consistent estimator of linear peer-influence under normalized homophily.

A.2 Disentangling Homophily from Estimation of Peer-Influence: Randomized Treatment Strategies

An algorithm for inference of linear peer-influence. We now use our general framework to design randomized treatments for the inference of linear peer-influence effects under homophily. We proceed to find the optimal treatment probabilities θ_s for $s = 1, \ldots, r$ under a stochastic block model with r communities as before.

Let $M_0^{(k)}$ denote the set of untreated individuals which have k neighbors (note that we are abusing notation here: Now, $M_0^{(1)}$ represents untreated individuals which have exactly 1 neighbor, rather than at least 1 neighbor as before in the binary peer-influence case). First, we derive a proposition about $M_0^{(k)}$ under our framework.

Proposition 2 *Consider a stochastic block model (SBM) of N individuals in r communities. Denote the communities of the SBM by the sets $B_1, \ldots, B_r$, which are of respective sizes $A_1, \ldots, A_r$ (where $A_1 + \cdots + A_r = N$). Let $\boldsymbol{P}$ be the $r \times r$ adjacency probability matrix between the r communities. We assign treatments independently to individuals such that individuals in B_s are treated with probability θ_s for $s = 1, \ldots, r$. Under such setup, let $M_0^{(k)}$ denote the set of untreated individuals which have k treated neighbors. For ease of notation, let $\{s \in M_0^{(k)}\}$ denote the event that a fixed vertex in community s is in the set $M_0^{(k)}$. Then,*

$$\mathbb{P}(s \in M_0^{(k)}) = (1 - \theta_s) \sum_{\substack{t_1,\ldots,t_r: \\ \forall v=1,\ldots,r \ 0 \le t_v \le A_v - \boldsymbol{1}_{\{v=s\}}, \\ t_1+\cdots+t_r=k}} \left(\prod_{v=1}^{r} \boldsymbol{Bin}(t_v; A_v - \boldsymbol{1}_{\{v=s\}}, \theta_v P_{s,v}) \right) \tag{20}$$

where $\boldsymbol{Bin}(t_v; A_v - \boldsymbol{1}_{\{v=s\}}, \theta_v P_{s,v}) = \binom{A_v - \boldsymbol{1}_{\{v=s\}}}{t_v} \left(\theta_v P_{s,v}\right)^{t_v} \left(1 - \theta_v P_{s,v}\right)^{A_v - \boldsymbol{1}_{\{v=s\}} - t_v}$.

The main idea behind the homophily disentangling strategy is to ensure that in every community B_s on our stochastic block model, there are equal (expected) numbers of individuals being affected by different levels of peer-influence. In the case of linear peer-influence, this means choosing treatment values such that inside every community s, each individual has an equal probability of being in sets $M_0^{(k)}$ for different peer-influence levels k. Under a stochastic block model, values of k range from 0 to $N - 1$ (as one individual can have at most $N - 1$ treated neighbors). However, in practice, we can choose to consider $k = 0, 1, \ldots, K$ where K is the maximum degree of the actual observed network. Therefore, through an optimal assignment of treatments, we wish to satisfy

$$\forall s = 1, \ldots, r, \quad \mathbb{P}(s \in M_0^{(0)}) = \mathbb{P}(s \in M_0^{(1)}) = \cdots = \mathbb{P}(s \in M_0^{(K-1)}) = \mathbb{P}(s \in M_0^{(K)}),$$

where expressions for each $\mathbb{P}(s \in M_0^{(k)})$ as functions of θ_s for $s = 1, \ldots, r$ are obtained from Proposition 2 above. This gives Kr conditions to satisfy for r variables $\theta_s \in [0, 1]$ (for $s = 1, \ldots, r$), so we can approach this as a constrained optimization problem as considered in the binary peer-influence case before.

B Tables of Main Results

B.1 Analytical Results

Model	Peer-Influence Estimate	Results
Binary peer-influence with unnormalized homophily : $\mathbf{peer}((Z_j)_{j\in\mathcal{N}_i}) = \mathbf{1}_{\sum_{j\in\mathcal{N}_i} Z_j>0}$ $\mathbf{hom}((X_j)_{j\in\mathcal{N}_i}) = \sum_{j\in\mathcal{N}_i} X_j$	$\hat\beta_0 = \underset{i\in M_0^{(1)}}{avg}\ Y_i - \underset{i\in M_0^{(0)}}{avg}\ Y_i$ for $M_0^{(0)} := \{i : Z_i = 0, \sum_{j\in\mathcal{N}_i} Z_j = 0\}$ and $M_0^{(1)} := \{i : Z_i = 0, \sum_{j\in\mathcal{N}_i} Z_j > 0\}$	Bias of $\hat\beta_0$ (conditional on $\mathbf{Z}$): $h_0\Big(\underset{i\in M_0^{(1)}}{avg}\lvert\mathcal{N}_i\rvert - \underset{i\in M_0^{(0)}}{avg}\lvert\mathcal{N}_i\rvert\Big)$. Variance of $\hat\beta_0$ (conditional on $\mathbf{Z}$): $h_0^2\sigma_X^2\Big(\underset{i,j\in M_0^{(0)}}{avg}\lvert\mathcal{N}_i\cap\mathcal{N}_j\rvert + \underset{i,j\in M_0^{(1)}}{avg}\lvert\mathcal{N}_i\cap\mathcal{N}_j\rvert - 2\underset{\substack{i\in M_0^{(0)}\\ j\in M_0^{(1)}}}{avg}\lvert\mathcal{N}_i\cap\mathcal{N}_j\rvert\Big) + \sigma_Y^2\Big(\frac{1}{\lvert M_0^{(0)}\rvert} + \frac{1}{\lvert M_0^{(1)}\rvert}\Big)$
Binary peer-influence with normalized homophily : $\mathbf{peer}((Z_j)_{j\in\mathcal{N}_i}) = \mathbf{1}_{\sum_{j\in\mathcal{N}_i} Z_j>0}$ $\mathbf{hom}((X_j)_{j\in\mathcal{N}_i}) = \sum_{j\in\mathcal{N}_i} \frac{X_j}{\lvert\mathcal{N}_i\rvert}$	$\hat\beta_0 = \underset{i\in M_0^{(1)}}{avg}\ Y_i - \underset{i\in M_0^{(0)}}{avg}\ Y_i$ for $M_0^{(0)} := \{i : Z_i = 0, \sum_{j\in\mathcal{N}_i} Z_j = 0\}$ and $M_0^{(1)} := \{i : Z_i = 0, \sum_{j\in\mathcal{N}_i} Z_j > 0\}$	Bias of $\hat\beta_0$ (conditional on $\mathbf{Z}$): $\mathbf{0}$. Variance of $\hat\beta_0$ (conditional on $\mathbf{Z}$): $h_0^2\sigma_X^2\Big(\underset{i,j\in M_0^{(0)}}{avg}\frac{\lvert\mathcal{N}_i\cap\mathcal{N}_j\rvert}{\lvert\mathcal{N}_i\rvert\lvert\mathcal{N}_j\rvert} + \underset{i,j\in M_0^{(1)}}{avg}\frac{\lvert\mathcal{N}_i\cap\mathcal{N}_j\rvert}{\lvert\mathcal{N}_i\rvert\lvert\mathcal{N}_j\rvert} - 2\underset{\substack{i\in M_0^{(0)}\\ j\in M_0^{(1)}}}{avg}\frac{\lvert\mathcal{N}_i\cap\mathcal{N}_j\rvert}{\lvert\mathcal{N}_i\rvert\lvert\mathcal{N}_j\rvert}\Big) + \sigma_Y^2\Big(\frac{1}{\lvert M_0^{(0)}\rvert} + \frac{1}{\lvert M_0^{(1)}\rvert}\Big)$
Linear peer-influence with unnormalized homophily : $\mathbf{peer}((Z_j)_{j\in\mathcal{N}_i}) = \sum_{j\in\mathcal{N}_i} Z_j$ $\mathbf{hom}((X_j)_{j\in\mathcal{N}_i}) = \sum_{j\in\mathcal{N}_i} X_j$	$\hat\beta_0 = \frac{\sum_k \hat\beta_0^{(k)}}{\sum_k 1}$ for $\hat\beta_0^{(k)} := \frac{1}{k}\Big(\underset{i\in M_0^{(k)}}{avg}\ Y_i - \underset{i\in M_0^{(0)}}{avg}\ Y_i\Big)$, $M_0^{(k)} := \{i : Z_i = 0, \sum_{j\in\mathcal{N}_i} Z_j = k\}$.	Bias of $\hat\beta_0$ (conditional on $\mathbf{Z}$): $\frac{h_0}{\sum_{k>0}1}\sum_{k>0}\frac{1}{k}\Big(\underset{i\in M_0^{(k)}}{avg}\lvert\mathcal{N}_i\rvert - \underset{i\in M_0^{(0)}}{avg}\lvert\mathcal{N}_i\rvert\Big)$ Variance of $\hat\beta_0$ (conditional on $\mathbf{Z}$): $\frac{1}{(\sum_{k>0}1)^2}\sum_{k,l>0}\frac{1}{kl}\Big[h_0^2\sigma_X^2\Big(\underset{i\in M_0^{(k)},j\in M_0^{(l)}}{avg}\lvert\mathcal{N}_i\cap\mathcal{N}_j\rvert + \underset{i,j\in M_0^{(0)}}{avg}\lvert\mathcal{N}_i\cap\mathcal{N}_j\rvert - 2\underset{i\in M_0^{(0)},j\in M_0^{(k)}}{avg}\lvert\mathcal{N}_i\cap\mathcal{N}_j\rvert\Big) + \sigma_Y^2\Big(\frac{1}{\lvert M_0^{(0)}\rvert} + \frac{\mathbf{1}_{k=l}}{\lvert M_0^{(k)}\rvert}\Big)\Big]$
Linear peer-influence with normalized homophily : $\mathbf{peer}((Z_j)_{j\in\mathcal{N}_i}) = \sum_{j\in\mathcal{N}_i} Z_j$ $\mathbf{hom}((X_j)_{j\in\mathcal{N}_i}) = \sum_{j\in\mathcal{N}_i} \frac{X_j}{\lvert\mathcal{N}_i\rvert}$	$\hat\beta_0 = \frac{\sum_k \hat\beta_0^{(k)}}{\sum_k 1}$ for $\hat\beta_0^{(k)} := \frac{1}{k}\Big(\underset{i\in M_0^{(k)}}{avg}\ Y_i - \underset{i\in M_0^{(0)}}{avg}\ Y_i\Big)$, $M_0^{(k)} := \{i : Z_i = 0, \sum_{j\in\mathcal{N}_i} Z_j = k\}$.	Bias of $\hat\beta_0$ (conditional on $\mathbf{Z}$): $\mathbf{0}$ Variance of $\hat\beta_0$ (conditional on $\mathbf{Z}$): $\frac{1}{(\sum_{k>0}1)^2}\sum_{k,l>0}\frac{1}{kl}\Big[h_0^2\sigma_X^2\Big(\underset{i\in M_0^{(k)},j\in M_0^{(l)}}{avg}\frac{\lvert\mathcal{N}_i\cap\mathcal{N}_j\rvert}{\lvert\mathcal{N}_i\rvert\lvert\mathcal{N}_j\rvert} + \underset{i,j\in M_0^{(0)}}{avg}\frac{\lvert\mathcal{N}_i\cap\mathcal{N}_j\rvert}{\lvert\mathcal{N}_i\rvert\lvert\mathcal{N}_j\rvert} - 2\underset{i\in M_0^{(0)},j\in M_0^{(k)}}{avg}\frac{\lvert\mathcal{N}_i\cap\mathcal{N}_j\rvert}{\lvert\mathcal{N}_i\rvert\lvert\mathcal{N}_j\rvert}\Big) + \sigma_Y^2\Big(\frac{1}{\lvert M_0^{(0)}\rvert} + \frac{\mathbf{1}_{k=l}}{\lvert M_0^{(k)}\rvert}\Big)\Big]$

B.2 Randomized Treatment Strategies to Disentangle Homophily

Model	Peer-Influence Estimate	Randomised treatment strategy
Binary peer-influence, $\mathbf{peer}((Z_j)_{j\in\mathcal{N}_i}) = \mathbf{1}_{\sum_{j\in\mathcal{N}_i} Z_j>0}$, with an *unknown* homophily function $\mathbf{hom}((X_j)_{j\in\mathcal{N}_i})$.	$\hat{\beta}_0 = \underset{i\in M_0^{(1)}}{avg}\ Y_i - \underset{i\in M_0^{(0)}}{avg}\ Y_i$ for $M_0^{(0)} := \{i : Z_i = 0, \sum_{j\in\mathcal{N}_i} Z_j = 0\}$ (the set of untreated individuals with no treated neighbours) and $M_0^{(1)} := \{i : Z_i = 0, \sum_{j\in\mathcal{N}_i} Z_j > 0\}$ (the set of untreated individuals with at least one treated neighbour).	For graph G of N vertices which is clustered into r clusters, consider a corresponding Stochastic Block Model of N individuals in r communities $B_1, ..., B_r$. We assign treatments independently such that individuals in B_s are treated with probability θ_s for $s = 1, .., r$. We choose θ_s such that homophily does not interfere with the estimation of binary peer-influence. Let $\{s \in M_0^{(0)}\}$, $\{s \in M_0^{(1)}\}$ denote the event that a fixed vertex in community s is in the set $M_0^{(0)}$, $M_0^{(1)}$ respectively. In our optimal assignment of treatments we wish to satisfy $\mathbb{P}(s \in M_0^{(0)}) = \mathbb{P}(s \in M_0^{(1)})$, where $\mathbb{P}(s \in M_0^{(0)}) = (1-\theta_s)\prod_{v=1}^{r}(1 - P_{s,v}\theta_v)^{A_v - \mathbf{1}_{v=s}}$.. This can be approximated by solving the contrained optimisation problem of minimising: $$\left\lVert \begin{pmatrix} P_{1,1}(A_1-1) & P_{1,2}A_2 & \dots & \dots & P_{1,r}A_r \\ P_{2,1}A_1 & P_{2,2}(A_2-1) & \dots & \dots & P_{2,r}A_r \\ \vdots & \vdots & \vdots & \dots & \vdots \\ P_{r-1,1}A_1 & P_{r-1,2}A_2 & \dots & P_{r-1,r-1}(A_{r-1}-1) & P_{r-1,r}A_r \\ P_{r,1}A_1 & P_{r,2}A_2 & \dots & P_{r,r-1}A_{r-1} & P_{r,r}(A_r-1) \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_{r-1} \\ \theta_r \end{pmatrix} - log(2) \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix} \right\rVert$$ for $\theta \in [0,1]^r$ and some chosen norm $\lVert\ \rVert$ on $\mathbb{R}^d$ (e.g. L^2).
Linear peer-influence, $\mathbf{peer}((Z_j)_{j\in\mathcal{N}_i}) = \sum_{j\in\mathcal{N}_i} Z_j$, with an *unknown* homophily function $\mathbf{hom}((X_j)_{j\in\mathcal{N}_i})$.	$\hat{\beta}_0 = \frac{\sum_k \hat{\beta}_0^{(k)}}{\sum_k 1}$ for $\hat{\beta}_0^{(k)} := \frac{1}{k}\left(\underset{i\in M_0^{(k)}}{avg}\ Y_i - \underset{i\in M_0^{(0)}}{avg}\ Y_i\right)$, $M_0^{(k)} := \{i : Z_i = 0, \sum_{j\in\mathcal{N}_i} Z_j = k\}$ (the set of untreated individuals which have k treated neighbours).	For graph G of N vertices which is clustered into r clusters, consider a corresponding Stochastic Block Model of N individuals in r communities $B_1, ..., B_r$. We assign treatments independently such that individuals in B_s are treated with probability θ_s for $s = 1, .., r$. We choose θ_s such that homophily does not interfere with the estimation of binary peer-influence. Let $\{s \in M_0^{(k)}\}$ denote the event that a fixed vertex in community s is in the set $M_0^{(k)}$. In our optimal assignment of treatments we wish to satisfy $$\forall s = 1, ..., r, \quad \mathbb{P}(s \in M_0^{(0)}) = \mathbb{P}(s \in M_0^{(1)}) = ... = \mathbb{P}(s \in M_0^{(K-1)}) = \mathbb{P}(s \in M_0^{(K)}),$$ where $$\mathbb{P}(s \in M_0^{(k)}) = (1-\theta_s) \sum_{\substack{t_1,\dots,t_r: \\ \forall v=1,\dots,r\ 0\le t_v \le A_v - \mathbf{1}_{\{v=s\}}, \\ t_1+\dots+t_r=k}} \left(\prod_{v=1}^{r} \mathbf{Bin}(t_v; A_v - \mathbf{1}_{\{v=s\}}, \theta_v P_{s,v})\right)$$ for $\mathbf{Bin}(t_v; A_v - \mathbf{1}_{\{v=s\}}, \theta_v P_{s,v}) = \binom{A_v - \mathbf{1}_{\{v=s\}}}{t_v}(\theta_v P_{s,v})^{t_v}(1-\theta_v P_{s,v})^{A_v - \mathbf{1}_{\{v=s\}} - t_v}$. We approach this as a constrained optimisation problem as in the binary peer-influence case.

C Proofs

C.1 Proof of Theorem 1 (See p. xxx)

Theorem 1 *Consider the difference in means estimator $\hat{\beta}_0$ for binary peer-influence effect β_0. Under the presence of unnormalized homophily in our model (3), the mean squared error of $\hat{\beta}_0$ (conditional on the treatment $\mathbf{Z}$) is:*

$$\mathbb{E}[(\hat{\beta}_0-\beta_0)^2|\mathbf{Z}] = \left(h_0\left(\underset{i\in M_0^{(1)}}{avg}|\mathcal{N}_i| - \underset{i\in M_0^{(0)}}{avg}|\mathcal{N}_i|\right)\right)^2 + h_0^2\sigma_X^2\left(\underset{i,j\in M_0^{(0)}}{avg}|\mathcal{N}_i\cap\mathcal{N}_j| + \underset{i,j\in M_0^{(1)}}{avg}|\mathcal{N}_i\cap\mathcal{N}_j| - 2\underset{i\in M_0^{(0)},j\in M_0^{(1)}}{avg}|\mathcal{N}_i\cap\mathcal{N}_j|\right) + \sigma_Y^2\left(\frac{1}{|M_0^{(0)}|}+\frac{1}{|M_0^{(1)}|}\right) \tag{5}$$

Proof Recall the definition of the difference in means estimator for binary peer-influence (4).

$$\hat{\beta}_0 = \underset{i\in M_0^{(1)}}{avg} Y_i - \underset{i\in M_0^{(0)}}{avg} Y_i$$

where $M_0^{(0)} := \{i : Z_i = 0, \sum_{j\in\mathcal{N}_i} Z_j = 0\}$ (the set of untreated individuals with no treated neighbors) and $M_0^{(1)} := \{i : Z_i = 0, \sum_{j\in\mathcal{N}_i} Z_j > 0\}$ (the set of untreated individuals with at least one treated neighbors). The response variables $(Y_i)_{i=1,\ldots,N}$ are defined by:

$$Y_i(Z_i = 0, (Z_j)_{j\in\mathcal{N}_i}) = \alpha + \beta_0\mathbf{1}_{\sum_{j\in\mathcal{N}_i} Z_j>0} + h_0\sum_{j\in\mathcal{N}_i} X_j + \epsilon_i(0,\sigma_Y^2)$$

$$Y_i(Z_i = 1, (Z_j)_{j\in\mathcal{N}_i}, (X_j)_{j\in\mathcal{N}_i}) = \tau + Y_i(Z_i = 0, (Z_j)_{j\in\mathcal{N}_i}, (X_j)_{j\in\mathcal{N}_i}) + \beta_1\mathbf{1}_{\sum_{j\in\mathcal{N}_i} Z_j>0} + h_1\sum_{j\in\mathcal{N}_i} X_j$$

$\epsilon_i(0,\sigma_Y^2)$ for $i = 1,\ldots,N$ are the noise terms in the network, indepedent and identically distributed with zero mean and variance σ_Y^2. Note that the sets $M_0^{(0)}$ and $M_0^{(1)}$ are $\mathbf{Z}$ measurable and that latent homophily variables $\mathbf{X} = (X_j)_{j=1,\ldots,N}$ are independent of $\mathbf{Z} = (Z_j)_{j=1,\ldots,N}$. Therefore,

$$\begin{aligned}\mathbb{E}[\hat{\beta}_0|\mathbf{Z}] &= \frac{\sum_{i\in M_0^{(1)}}\mathbb{E}\Big[Y_i|\mathbf{Z}\Big]}{|M_0^{(1)}|} - \frac{\sum_{i\in M_0^{(0)}}\mathbb{E}\Big[Y_i|\mathbf{Z}\Big]}{|M_0^{(0)}|} \\ &= \frac{\sum_{i\in M_0^{(1)}}\mathbb{E}\Big[\beta_0 + \sum_{j\in\mathcal{N}_i} X_j\Big]}{|M_0^{(1)}|} - \frac{\sum_{i\in M_0^{(0)}}\mathbb{E}\Big[\sum_{j\in\mathcal{N}_i} X_j\Big]}{|M_0^{(0)}|}\end{aligned}$$

$$= \beta_0 + \frac{\sum_{i \in M_0^{(1)}} h_0 |\mathcal{N}_i|}{|M_0^{(1)}|} - \frac{\sum_{i \in M_0^{(0)}} h_0 |\mathcal{N}_i|}{|M_0^{(0)}|}$$
$$= \beta_0 + h_0 \Big(\underset{i \in M_0^{(1)}}{avg} |\mathcal{N}_i| - \underset{i \in M_0^{(1)}}{avg} |\mathcal{N}_i| \Big).$$

This gives the bias of $\hat{\beta}_0$: $\mathbb{E}\Big[\hat{\beta}_0 - \beta_0 | \mathbf{Z}\Big] = h_0 \Big(\underset{i \in M_0^{(1)}}{avg} |\mathcal{N}_i| - \underset{i \in M_0^{(1)}}{avg} |\mathcal{N}_i| \Big)$. Similarly,

$$var[\hat{\beta}_0 | \mathbf{Z}] = var\Big(\frac{\sum_{i \in M_0^{(1)}} Y_i}{|M_0^{(1)}|} - \frac{\sum_{j \in M_0^{(0)}} Y_j}{|M_0^{(0)}|} \Big| \mathbf{Z} \Big)$$
$$= \frac{var(\sum_{i \in M_0^{(1)}} Y_i \mid \mathbf{Z})}{|M_0^{(1)}|^2} + \frac{var(\sum_{j \in M_0^{(0)}} Y_j \mid \mathbf{Z})}{|M_0^{(0)}|^2} - \frac{2cov(\sum_{i \in M_0^{(1)}} Y_i, \sum_{j \in M_0^{(0)}} Y_j \mid \mathbf{Z})}{|M_0^{(0)}||M_0^{(1)}|}$$
$$= \frac{\sum_{i \in M_0^{(1)}} \sum_{k \in M_0^{(1)}} cov(Y_i, Y_k \mid \mathbf{Z})}{|M_0^{(1)}|^2} + \frac{\sum_{j \in M_0^{(0)}} \sum_{l \in M_0^{(0)}} cov(Y_j, Y_l \mid \mathbf{Z})}{|M_0^{(0)}|^2}$$
$$- \frac{2 \sum_{i \in M_0^{(1)}} \sum_{j \in M_0^{(0)}} cov(Y_i, Y_j \mid \mathbf{Z})}{|M_0^{(0)}||M_0^{(1)}|}$$
$$= \underset{i,k \in M_0^{(1)}}{avg} cov(Y_i, Y_k \mid \mathbf{Z}) + \underset{j,l \in M_0^{(0)}}{avg} cov(Y_i, Y_k \mid \mathbf{Z}) - 2 \underset{i \in M_0^{(0)}, j \in M_0^{(1)}}{avg} cov(Y_i, Y_j \mid \mathbf{Z}).$$

For $i \in M_0^{(1)}$ and $k \in M_0^{(1)}$, by the law of total covariance and as $\mathbf{X}$ are i.i.d.,

$$cov(Y_i, Y_k \mid \mathbf{Z}) = \mathbb{E}[cov(Y_i, Y_k \mid \mathbf{X}, \mathbf{Z}) \mid \mathbf{Z}] + cov\Big(\mathbb{E}[Y_i | \mathbf{X}, \mathbf{Z}], \mathbb{E}[Y_k | \mathbf{X}, \mathbf{Z}] \Big| \mathbf{Z} \Big)$$
$$= \sigma_Y^2 \mathbb{1}_{\{i=k\}} + cov\Big(\alpha + \beta_0 + h_0 \sum_{a \in \mathcal{N}_i} X_a, \alpha + \beta_0 + h_0 \sum_{b \in \mathcal{N}_k} X_b \Big)$$
$$= \sigma_Y^2 \mathbb{1}_{\{i=k\}} + h_0^2 \, cov\Big(\sum_{a \in \mathcal{N}_i} X_a, \sum_{b \in \mathcal{N}_k} X_b \Big)$$
$$= \sigma_Y^2 \mathbb{1}_{\{i=k\}} + h_0^2 \sigma_X^2 |\mathcal{N}_i \cap \mathcal{N}_j|$$

Similarly for $j \in M_0^{(0)}$ and $l \in M_0^{(0)}$,

$$cov(Y_j, Y_l \mid \mathbf{Z}) = \sigma_Y^2 \mathbb{1}_{\{j=l\}} + cov\Big(\alpha + h_0 \sum_{a \in \mathcal{N}_j} X_a, \alpha + h_0 \sum_{b \in \mathcal{N}_l} X_b \Big| \mathbf{Z} \Big)$$
$$= \sigma_Y^2 \mathbb{1}_{\{j=l\}} + h_0^2 \sigma_X^2 |\mathcal{N}_j \cap \mathcal{N}_l|$$

For $i \in M_0^{(1)}$ and $j \in M_0^{(0)}$, by the law of total covariance and as $\mathbf{X}$ are i.i.d.,

$$
\begin{aligned}
cov(Y_i, Y_j \mid \mathbf{Z}) &= \mathbb{E}[cov(Y_i, Y_j \mid \mathbf{X}, \mathbf{Z}) \mid \mathbf{Z}] + cov\Big(\mathbb{E}[Y_i \big| \mathbf{X}, \mathbf{Z}], \mathbb{E}[Y_k \big| \mathbf{X}, \mathbf{Z}] \,\Big|\, \mathbf{Z}\Big) \\
&= 0 + \ cov\Big(\alpha + \beta_0 + h_0 \sum_{a \in \mathcal{N}_i} X_a, \alpha + h_0 \sum_{b \in \mathcal{N}_k} X_b\Big) \\
&= h_0^2 \, cov\Big(\sum_{a \in \mathcal{N}_i} X_a, \sum_{b \in \mathcal{N}_k} X_b\Big) \\
&= h_0^2 \sigma_X^2 |\mathcal{N}_i \cap \mathcal{N}_j|.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
var[\hat{\beta}_0|\mathbf{Z}] = &\mathop{avg}_{i,k \in M_0^{(1)}} cov(Y_i, Y_k \mid \mathbf{Z}) + \mathop{avg}_{j,l \in M_0^{(0)}} cov(Y_i, Y_k \mid \mathbf{Z}) - 2 \mathop{avg}_{i \in M_0^{(0)}, j \in M_0^{(1)}} cov(Y_i, Y_j \mid \mathbf{Z}) \\
= &\mathop{avg}_{i,k \in M_0^{(1)}} \Big(\sigma_Y^2 \mathbb{1}_{\{i=k\}} + h_0^2 \sigma_X^2 |\mathcal{N}_i \cap \mathcal{N}_k|\Big) + \mathop{avg}_{j,l \in M_0^{(0)}} \Big(\sigma_Y^2 \mathbb{1}_{\{j=l\}} + h_0^2 \sigma_X^2 |\mathcal{N}_j \cap \mathcal{N}_l|\Big) \\
&- 2 \mathop{avg}_{i \in M_0^{(0)}, j \in M_0^{(1)}} \Big(h_0^2 \sigma_X^2 |\mathcal{N}_j \cap \mathcal{N}_l|\Big) \\
= &h_0^2 \sigma_X^2 \Big(\mathop{avg}_{i,j \in M_0^{(0)}} |\mathcal{N}_i \cap \mathcal{N}_j| + \mathop{avg}_{i,j \in M_0^{(1)}} |\mathcal{N}_i \cap \mathcal{N}_j| - 2 \mathop{avg}_{i \in M_0^{(0)}, j \in M_0^{(1)}} |\mathcal{N}_i \cap \mathcal{N}_j|\Big) \\
&+ \sigma_Y^2 \Big(\frac{1}{|M_0^{(0)}|} + \frac{1}{|M_0^{(1)}|}\Big)
\end{aligned}
$$

Now we can recall the bias–variance decomposition of the MSE to obtain

$$
\begin{aligned}
&\mathbb{E}[(\hat{\beta}_0 - \beta_0)^2|\mathbf{Z}] = \Big(\mathbb{E}[\hat{\beta}_0 - \beta_0|\mathbf{Z}]\Big)^2 + var[\hat{\beta}_0|\mathbf{Z}] \\
&\quad = \Big(h_0\Big(\mathop{avg}_{i \in M_0^{(1)}} |\mathcal{N}_i| - \mathop{avg}_{i \in M_0^{(0)}} |\mathcal{N}_i|\Big)\Big)^2 \\
&\qquad + h_0^2 \sigma_X^2 \Big(\mathop{avg}_{i,j \in M_0^{(0)}} |\mathcal{N}_i \cap \mathcal{N}_j| + \mathop{avg}_{i,j \in M_0^{(1)}} |\mathcal{N}_i \cap \mathcal{N}_j| - 2 \mathop{avg}_{i \in M_0^{(0)}, j \in M_0^{(1)}} |\mathcal{N}_i \cap \mathcal{N}_j|\Big) \\
&\qquad + \sigma_Y^2 \Big(\frac{1}{|M_0^{(0)}|} + \frac{1}{|M_0^{(1)}|}\Big)
\end{aligned}
$$

as required. □

C.2 Proof of Theorem 3 (See p. xxx)

Theorem 3 *Consider the difference in means estimator $\hat{\beta}_0$ for binary peer-influence effect β_0. Under the presence of unnormalized homophily in our model (3), the mean* *squared error of $\hat{\beta}_0$ (conditional on the treatment $\mathbf{Z}$) is:*

$$\mathbb{E}[(\hat{\beta}_0-\beta_0)^2|\mathbf{Z}] = \Big(h_0\Big(\underset{i\in M_0^{(1)}}{avg}|\mathcal{N}_i| - \underset{i\in M_0^{(0)}}{avg}|\mathcal{N}_i|\Big)\Big)^2 + h_0^2\sigma_X^2\Big(\underset{i,j\in M_0^{(0)}}{avg}|\mathcal{N}_i\cap\mathcal{N}_j| + \underset{i,j\in M_0^{(1)}}{avg}|\mathcal{N}_i\cap\mathcal{N}_j| - 2\underset{i\in M_0^{(0)},j\in M_0^{(1)}}{avg}|\mathcal{N}_i\cap\mathcal{N}_j|\Big) + \sigma_Y^2\Big(\frac{1}{|M_0^{(0)}|}+\frac{1}{|M_0^{(1)}|}\Big) \tag{16}$$

Proof We proceed as similar to the binary peer-influence estimator case. Recall the definition of the estimator for linear peer-influence (21):

$$\hat{\beta}_0 = \frac{\sum_k \hat{\beta}_0^{(k)}}{\sum_k 1} \; for \; \hat{\beta}_0^{(k)} = \frac{1}{k}\Big(\frac{\sum_{i\in M_0^{(k)}} Y_i}{|M_0^{(k)}|} - \frac{\sum_{i\in M_0^{(0)}} Y_i}{|M_0^{(0)}|}\Big) = \frac{1}{k}\Big(\underset{i\in M_0^{(k)}}{avg} Y_i - \underset{i\in M_0^{(0)}}{avg} Y_i\Big), \tag{21}$$

where $M_0^{(k)} := \{i : Z_i = 0, \sum_{j\in\mathcal{N}_i} Z_j = k\}$ (the set of untreated individuals with k treated neighbors). The response variables $(Y_i)_{i=1,\dots,N}$ are defined by:

$$Y_i(Z_i=0,(Z_j)_{j\in\mathcal{N}_i}) = \alpha + \beta_0\sum_{j\in\mathcal{N}_i} Z_j + h_0\sum_{j\in\mathcal{N}_i} X_j + \epsilon_i(0,\sigma_Y^2)$$

$$Y_i(Z_i=1,(Z_j)_{j\in\mathcal{N}_i},(X_j)_{j\in\mathcal{N}_i}) = \tau + Y_i(Z_i=0,(Z_j)_{j\in\mathcal{N}_i},(X_j)_{j\in\mathcal{N}_i}) + \beta_1\sum_{j\in\mathcal{N}_i} Z_j + h_1\sum_{j\in\mathcal{N}_i} X_j$$

$\epsilon_i(0,\sigma_Y^2)$ for $i = 1,\dots,N$ are the noise terms in the network, indepedent and identically distributed with zero mean and variance σ_Y^2. Note that sets $M_0^{(k)}$ are $\mathbf{Z}$ measurable and that latent homophily variables $\mathbf{X} = (X_j)_{j=1,\dots,N}$ are independent of $\mathbf{Z} = (Z_j)_{j=1,\dots,N}$. Therefore,

$$\begin{aligned}
\mathbb{E}[\hat{\beta}_0^{(k)}|\mathbf{Z}] &= \frac{1}{k}\Big(\frac{\sum_{i\in M_0^{(k)}}\mathbb{E}\big[Y_i|\mathbf{Z}\big]}{|M_0^{(k)}|} - \frac{\sum_{i\in M_0^{(0)}}\mathbb{E}\big[Y_i|\mathbf{Z}\big]}{|M_0^{(0)}|}\Big)\\
&= \frac{1}{k}\Big(\frac{\sum_{i\in M_0^{(k)}}\mathbb{E}\big[k\beta_0 + \sum_{j\in\mathcal{N}_i} X_j|\mathbf{Z}\big]}{|M_0^{(k)}|} - \frac{\sum_{i\in M_0^{(0)}}\mathbb{E}\big[\sum_{j\in\mathcal{N}_i} X_j|\mathbf{Z}\big]}{|M_0^{(0)}|}\Big)\\
&= \beta_0 + \frac{1}{k}\Big(\frac{\sum_{i\in M_0^{(k)}}\mathbb{E}\big[\sum_{j\in\mathcal{N}_i} X_j|\mathbf{Z}\big]}{|M_0^{(k)}|} - \frac{\sum_{i\in M_0^{(0)}}\mathbb{E}\big[\sum_{j\in\mathcal{N}_i} X_j|\mathbf{Z}\big]}{|M_0^{(0)}|}\Big)\\
&= \beta_0 + \frac{1}{k}\Big(\frac{\sum_{i\in M_0^{(k)}} h_0|\mathcal{N}_i|}{|M_0^{(k)}|} - \frac{\sum_{i\in M_0^{(0)}} h_0|\mathcal{N}_i|}{|M_0^{(0)}|}\Big)
\end{aligned}$$

$$= \beta_0 + \frac{h_0}{k}\left(\underset{i \in M_0^{(k)}}{avg}\, |\mathcal{N}_i| - \underset{i \in M_0^{(1)}}{avg}\, |\mathcal{N}_i| \right),$$

which gives the bias of the estimator $\hat{\beta}_0 = \frac{\sum_k \hat{\beta}_0^{(k)}}{\sum_k 1}$ to be:

$$\mathbb{E}[\hat{\beta}_0 - \beta_0 | \mathbf{Z}] = \frac{h_0}{\sum_{k>0} 1} \sum_{k>0} \frac{1}{k} \left(\underset{i \in M_0^{(k)}}{avg}\, |\mathcal{N}_i| - \underset{i \in M_0^{(0)}}{avg}\, |\mathcal{N}_i| \right).$$

Similarly, $var[\hat{\beta}_0 | \mathbf{Z}] = \frac{1}{(\sum_k 1)^2} \sum_{k>0} \sum_{l>0} cov(\hat{\beta}_0^{(k)}, \hat{\beta}_0^{(l)})$, where

$$cov(\hat{\beta}_0^{(k)}, \hat{\beta}_0^{(l)} \Big| \mathbf{Z}) = \frac{1}{kl} cov\Big(\frac{\sum_{i \in M_0^{(k)}} Y_i}{|M_0^{(k)}|} - \frac{\sum_{j \in M_0^{(0)}} Y_j}{|M_0^{(0)}|}, \frac{\sum_{i \in M_0^{(l)}} Y_i}{|M_0^{(l)}|} - \frac{\sum_{j \in M_0^{(0)}} Y_j}{|M_0^{(0)}|} \ \Big|\ \mathbf{Z} \Big)$$
$$= \frac{1}{kl}\Big(\frac{\sum_{i \in M_0^{(k)}, j \in M_0^{(l)}} cov(Y_i, Y_j | \mathbf{Z})}{|M_0^{(k)}||M_0^{(l)}|} + \frac{\sum_{i \in M_0^{(0)}, j \in M_0^{(0)}} cov(Y_i, Y_j | \mathbf{Z})}{|M_0^{(0)}|^2}$$
$$- \frac{\sum_{i \in M_0^{(k)}, j \in M_0^{(0)}} cov(Y_i, Y_j | \mathbf{Z})}{|M_0^{(k)}||M_0^{(0)}|} - \frac{\sum_{i \in M_0^{(0)}, j \in M_0^{(l)}} cov(Y_i, Y_j | \mathbf{Z})}{|M_0^{(0)}||M_0^{(l)}|} \Big).$$

For $i \in M_0^{(k)}$ and $j \in M_0^{(l)}$, by the law of total covariance and as $\mathbf{X}$ are i.i.d.,

$$cov(Y_i, Y_j \mid \mathbf{Z}) = \mathbb{E}[cov(Y_i, Y_j \mid \mathbf{X}, \mathbf{Z}) \Big| \mathbf{Z}] + cov\Big(\mathbb{E}[Y_i \big| \mathbf{X}, \mathbf{Z}], \mathbb{E}[Y_j \big| \mathbf{X}, \mathbf{Z}] \ \Big|\ \mathbf{Z} \Big)$$
$$= \sigma_Y^2 \mathbb{1}_{\{i=j\}} \ + \ cov\Big(\alpha + k\beta_0 + h_0 \sum_{a \in \mathcal{N}_i} X_a, \alpha + l\beta_0 + h_0 \sum_{b \in \mathcal{N}_j} X_b \Big)$$
$$= \sigma_Y^2 \mathbb{1}_{\{i=j\}} \ + h_0^2\, cov\Big(\sum_{a \in \mathcal{N}_i} X_a, \sum_{b \in \mathcal{N}_j} X_b \Big)$$
$$= \sigma_Y^2 \mathbb{1}_{\{i=j\}} \ + h_0^2 \sigma_X^2 |\mathcal{N}_i \cap \mathcal{N}_j|.$$

This gives

$$cov(\hat{\beta}_0^{(k)}, \hat{\beta}_0^{(l)} \Big| \mathbf{Z}) = \frac{1}{kl}\Big(\frac{\sum_{i \in M_0^{(k)}, j \in M_0^{(l)}} cov(Y_i, Y_j | \mathbf{Z})}{|M_0^{(k)}||M_0^{(l)}|} + \frac{\sum_{i \in M_0^{(0)}, j \in M_0^{(0)}} cov(Y_i, Y_j | \mathbf{Z})}{|M_0^{(0)}|^2}$$
$$- \frac{\sum_{i \in M_0^{(k)}, j \in M_0^{(0)}} cov(Y_i, Y_j | \mathbf{Z})}{|M_0^{(k)}||M_0^{(0)}|} - \frac{\sum_{i \in M_0^{(0)}, j \in M_0^{(l)}} cov(Y_i, Y_j | \mathbf{Z})}{|M_0^{(0)}||M_0^{(l)}|} \Big)$$
$$= \frac{1}{kl}\Big(\frac{\sum_{i \in M_0^{(k)}, j \in M_0^{(l)}} \sigma_Y^2 \mathbb{1}_{\{i=j\}} \ + h_0^2 \sigma_X^2 |\mathcal{N}_i \cap \mathcal{N}_j|}{|M_0^{(k)}||M_0^{(l)}|}$$
$$+ \frac{\sum_{i \in M_0^{(0)}, j \in M_0^{(0)}} \sigma_Y^2 \mathbb{1}_{\{i=j\}} \ + h_0^2 \sigma_X^2 |\mathcal{N}_i \cap \mathcal{N}_j|}{|M_0^{(0)}|^2}$$
$$- \frac{\sum_{i \in M_0^{(k)}, j \in M_0^{(0)}} h_0^2 \sigma_X^2 |\mathcal{N}_i \cap \mathcal{N}_j|}{|M_0^{(k)}||M_0^{(0)}|} - \frac{\sum_{i \in M_0^{(0)}, j \in M_0^{(l)}} h_0^2 \sigma_X^2 |\mathcal{N}_i \cap \mathcal{N}_j|}{|M_0^{(0)}||M_0^{(l)}|} \Big)$$

$$
\begin{aligned}
&= \frac{1}{kl}\Bigg(\sigma_Y^2 \frac{\mathbb{1}_{\{l=k\}}}{|M^{(k)}|} + h_0^2\sigma_X^2 \frac{\sum_{i\in M_0^{(k)}, j\in M_0^{(l)}} |\mathcal{N}_i \cap \mathcal{N}_j|}{|M_0^{(k)}||M_0^{(l)}|} \\
&\quad + \sigma_Y^2 \frac{1}{|M^{(0)}|} + h_0^2\sigma_X^2 \frac{\sum_{i\in M_0^{(0)}, j\in M_0^{(0)}} |\mathcal{N}_i \cap \mathcal{N}_j|}{|M_0^{(0)}||M_0^{(0)}|} \\
&\quad - h_0^2\sigma_X^2 \frac{\sum_{i\in M_0^{(k)}, j\in M_0^{(0)}} |\mathcal{N}_i \cap \mathcal{N}_j|}{|M_0^{(k)}||M_0^{(0)}|} - h_0^2\sigma_X^2 \frac{\sum_{i\in M_0^{(0)}, j\in M_0^{(l)}} |\mathcal{N}_i \cap \mathcal{N}_j|}{|M_0^{(0)}||M_0^{(l)}|}\Bigg) \\
&= \frac{1}{kl}\Bigg[h_0^2\sigma_X^2\Bigg(\underset{i\in M_0^{(k)}, j\in M_0^{(l)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j| + \underset{i,j\in M_0^{(0)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j| - \underset{i\in M_0^{(k)}, j\in M_0^{(0)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j| \\
&\quad - \underset{i\in M_0^{(0)}, j\in M_0^{(l)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j|\Bigg) + \sigma_Y^2\Bigg(\frac{1}{|M_0^{(0)}|} + \frac{\mathbf{1}_{k=l}}{|M_0^{(k)}|}\Bigg)\Bigg].
\end{aligned}
$$

Now, we can recall the bias–variance decomposition of the MSE to obtain

$$
\begin{aligned}
\mathbb{E}[(\hat{\beta}_0 - \beta_0)^2|\mathbf{Z}] &= \Big(\mathbb{E}[\hat{\beta}_0 - \beta_0|\mathbf{Z}]\Big)^2 + var[\hat{\beta}_0|\mathbf{Z}] \\
&= \Big(\mathbb{E}[\hat{\beta}_0 - \beta_0|\mathbf{Z}]\Big)^2 + \frac{1}{(\sum_k 1)^2}\sum_{k>0}\sum_{l>0} cov(\hat{\beta}_0^{(k)}, \hat{\beta}_0^{(l)}) \\
&= \Bigg(\frac{h_0}{\sum_{k>0} 1}\sum_{k>0}\frac{1}{k}\Bigg(\underset{i\in M_0^{(k)}}{avg} |\mathcal{N}_i| - \underset{i\in M_0^{(0)}}{avg} |\mathcal{N}_i|\Bigg)\Bigg)^2 \\
&\quad + \frac{1}{(\sum_{k>0} 1)^2}\sum_{k,l>0}\frac{1}{kl}\Bigg[h_0^2\sigma_X^2\Bigg(\underset{i\in M_0^{(k)}, j\in M_0^{(l)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j| + \underset{i,j\in M_0^{(0)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j| \\
&\quad - \underset{i\in M_0^{(0)}, j\in M_0^{(k)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j|\Bigg) - \underset{i\in M_0^{(l)}, j\in M_0^{(0)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j| + \sigma_Y^2\Bigg(\frac{1}{|M_0^{(0)}|} + \frac{\mathbf{1}_{k=l}}{|M_0^{(k)}|}\Bigg)\Bigg] \\
&= \Bigg(\frac{h_0}{\sum_{k>0} 1}\sum_{k>0}\frac{1}{k}\Bigg(\underset{i\in M_0^{(k)}}{avg} |\mathcal{N}_i| - \underset{i\in M_0^{(0)}}{avg} |\mathcal{N}_i|\Bigg)\Bigg)^2 \\
&\quad + \frac{1}{(\sum_{k>0} 1)^2}\sum_{k,l>0}\frac{1}{kl}\Bigg[h_0^2\sigma_X^2\Bigg(\underset{i\in M_0^{(k)}, j\in M_0^{(l)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j| + \underset{i,j\in M_0^{(0)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j| \\
&\quad - 2\underset{i\in M_0^{(0)}, j\in M_0^{(k)}}{avg} |\mathcal{N}_i \cap \mathcal{N}_j|\Bigg) + \sigma_Y^2\Bigg(\frac{1}{|M_0^{(0)}|} + \frac{\mathbf{1}_{k=l}}{|M_0^{(k)}|}\Bigg)\Bigg]
\end{aligned}
$$

as required. □

C.3 Proof of Theorem 1 (See p. xxx)

Proposition 1 *Consider a stochastic block model (SBM) of N individuals in r communities. Denote the communities of the SBM by the sets $B_1, \ldots, B_r$, which are of respective sizes $A_1, \ldots, A_r$ (where $A_1 + \cdots + A_r = N$). Let $\boldsymbol{P}$ be the $r \times r$ adjacency probability matrix between the r communities. We assign treatments independently to individuals such that individuals in B_s are treated with probability θ_s for $s = 1, \ldots, r$. Under such setup, let $M_0^{(0)}$ denote the set of untreated individuals* *which have no treated neighbors and let* $M_0^{(1)}$ *denote the set of untreated individuals which have at least one treated neighbor. For ease of notation, let* $\{s \in M_0^{(0)}\}$, $\{s \in M_0^{(1)}\}$ *denote the event that a fixed vertex in community* s *is in the sets* $M_0^{(0)}$, $M_0^{(1)}$ *respectively. Then,*

$$\mathbb{P}(s \in M_0^{(0)}) = (1-\theta_s)\prod_{v=1}^{r}(1-P_{s,v}\theta_v)^{A_v - \mathbf{1}_{v=s}}, \text{ and} \tag{7}$$

$$\mathbb{P}(s \in M_0^{(1)}) = (1-\theta_s)\left(1-\prod_{v=1}^{r}(1-P_{s,v}\theta_v)^{A_v - \mathbf{1}_{v=s}}\right). \tag{8}$$

Proof Note that each vertex in the graph is assigned treatment independently and that under the stochastic block model the events of any pair of vertices being adjacent are independent. Therefore,

$$\mathbb{P}(s \text{ has 0 treated neighbors} \mid s \text{ is untreated}) = \mathbb{P}(s \text{ has 0 treated neighbors})$$

for all k and $s = 1, \ldots, r$. This gives

$$\begin{aligned}
\mathbb{P}(s \in M_0^{(0)}) &= \mathbb{P}(s \text{ is untreated})\mathbb{P}(s \text{ has 0 treated neighbors}) \\
&= (1-\theta_s)\mathbb{P}(s \text{ has 0 treated neighbors}) \\
&= (1-\theta_s)\mathbb{P}\Big(\bigcap_{v=1}^{r}\{s \text{ has 0 treated neighbors in } B_v\}\Big) \\
&= (1-\theta_s)\prod_{v=1}^{r}\mathbb{P}(s \text{ has 0 treated neighbors in } B_v) \\
&= (1-\theta_s)\prod_{v=1}^{r}(1-P_{s,v}\theta_v)^{A_v - \mathbf{1}_{v=s}}.
\end{aligned}$$

where the $A_v - \mathbf{1}_{v=s}$ arises from noting that s can have at most $A_s - 1$ neighbors in B_s (it cannot connect to itself). Note that sets $M_0^{(0)}$ and $M_0^{(1)}$ partition the set of untreated individuals. Therefore,

$$\begin{aligned}
\mathbb{P}(s \in M_0^{(1)}) &= \mathbb{P}(Z_s = 0) - \mathbb{P}(s \in M_0^{(0)}) \\
&= (1-\theta_s) - (1-\theta_s)\prod_{v=1}^{r}(1-P_{s,v}\theta_v)^{A_v - \mathbf{1}_{v=s}} \\
&= (1-\theta_s)\Big(1-\prod_{v=1}^{r}(1-P_{s,v}\theta_v)^{A_v - \mathbf{1}_{v=s}}\Big).
\end{aligned}$$

□

C.4 Proof of Proposition 2 (See p. xxx)

Proposition 2 *Consider a stochastic block model (SBM) of N individuals in r communities. Denote the communities of the SBM by the sets $B_1, \ldots, B_r$, which are of respective sizes $A_1, \ldots, A_r$ (where $A_1 + \cdots + A_r = N$). Let $\boldsymbol{P}$ be the $r \times r$ adjacency probability matrix between the r communities. We assign treatments independently to individuals such that individuals in B_s are treated with probability θ_s for $s = 1, \ldots, r$. Under such setup, let $M_0^{(k)}$ denote the set of untreated individuals which have k treated neighbors. For ease of notation, let $\{s \in M_0^{(k)}\}$ denote the event that a fixed vertex in community s is in the set $M_0^{(k)}$. Then,*

$$\mathbb{P}(s \in M_0^{(k)}) = (1-\theta_s) \sum_{\substack{t_1,\ldots,t_r: \\ \forall v=1,\ldots,r\ 0 \le t_v \le A_v - \boldsymbol{1}_{\{v=s\}}, \\ t_1+\cdots+t_r=k}} \left(\prod_{v=1}^{r} \boldsymbol{Bin}(t_v;\, A_v - \boldsymbol{1}_{\{v=s\}}, \theta_v P_{s,v}) \right) \tag{20}$$

$$\textit{where } \boldsymbol{Bin}(t_v;\, A_v - \boldsymbol{1}_{\{v=s\}}, \theta_v P_{s,v}) = \binom{A_v - \boldsymbol{1}_{\{v=s\}}}{t_v} \big(\theta_v P_{s,v}\big)^{t_v} \big(1 - \theta_v P_{s,v}\big)^{A_v - \boldsymbol{1}_{\{v=s\}} - t_v}.$$

Proof Note that each vertex in the graph is assigned treatment indpendently. Therefore,

$$\mathbb{P}(s \text{ has k treated neighbors} \mid s \text{ is untreated}) = \mathbb{P}(s \text{ has k treated neighbors})$$

for all k and $s = 1, \ldots, r$. This gives

$$\begin{aligned}
\mathbb{P}(s \in M_0^{(k)}) &= \mathbb{P}(s \text{ is untreated})\mathbb{P}(s \text{ has k treated neighbors}) \\
&= (1-\theta_s)\mathbb{P}(s \text{ has k treated neighbors}) \\
&= (1-\theta_s) \sum_{\substack{t_1,\ldots,t_r: \\ \forall v=1,\ldots,r\ 0 \le t_v \le A_v - \mathbf{1}_{\{v=s\}}, \\ t_1+\cdots+t_r=k}} \mathbb{P}\Big(\bigcap_{v=1}^{r} \{s \text{ has } t_v \text{ treated neighbors in } B_v\} \Big) \\
&= (1-\theta_s) \sum_{\substack{t_1,\ldots,t_r: \\ \forall v=1,\ldots,r\ 0 \le t_v \le A_v - \mathbf{1}_{\{v=s\}}, \\ t_1+\cdots+t_r=k}} \Big(\prod_{v=1}^{r} \mathbb{P}(s \text{ has } t_v \text{ treated neighbors in } B_v) \Big).
\end{aligned}$$

We now wish to evaluate $\mathbb{P}(s$ has t_k treated neighbors in $B_v)$. Let n_v be the number of neighbors s (denoting a fixed individual in community B_s) has in B_v. Under a stochastic block model setup,

$$n_v \sim Bin(A_v - \mathbf{1}_{v=s}, P_{s,v})$$

$$t_v | n_v \sim Bin(n_v, \theta_v)$$

where the $A_v - \mathbf{1}_{v=s}$ arises from noting that s can have at most $A_s - 1$ neighbors in B_s (it cannot connect to itself). We want the unconditional distribution of t_v. Recall that moment generating function of $X \sim Bin(N, p)$ is $\mathbb{E}(z^X) = ((1-p) + pz)^N$. Therefore,

$$\mathbb{E}[z^{t_v}] = \mathbb{E}[\mathbb{E}[z^{t_v}|n_v]] = \mathbb{E}\Big[\Big((1-\theta_s) + \theta_s z\Big)^{n_v}\Big] = \Big((1-P_{s,v}) + P_{s,v}\Big((1-\theta_s)+\theta_s z\Big)\Big)^{A_v - \mathbf{1}_{v=s}}$$
$$= \Big((1-\theta_s P_{s,v}) + P_{s,v}\theta_s z\Big)^{A_v},$$

giving $t_v \sim Bin(A_v - \mathbf{1}_{v=s}, \theta_s P_{s,v})$. This gives

$$\mathbb{P}(s \text{ has } t_v \text{ treated neighbors in } B_v) = Bin(t_v; A_v - \mathbf{1}_{\{v=s\}}, \theta_v P_{s,v})$$

from which (20) directly follows. □

References

1. E.M. Airoldi, D.M. Blei, S.E. Fienberg, E.P. Xing, Mixed membership stochastic block- models. J. Mach. Learn. Res. (2008)
2. J.D. Angrist, The perils of peer effects. Labour Econ. (2014)
3. S.Aral, D. Walker, Creating social contagion through viral product design: a randomized trial of peer influence in networks. Manag. Sci. (2011)
4. S. Athey, D. Eckles, G.W. Imbens, Exact P-values for network interference. J. Am. Stat. Assoc. (2016)
5. E. Bakshy, D. Eckles, R. Yan, I. Rosenn, Social influence in social advertising: Evidence from field experiments, in *Proceedings of the 13th ACM Conference on Electronic Commerce*, 2012
6. R.M. Bond, C.J. Fariss, J.J. Jones, A.D.I. Kramer, C. Marlow, J.E. Settle, J.H. Fowler, A 61-million-person experiment in social influence and political mobilization. Nature (2012)
7. P. Holland, K. Laskey, S. Leinhardt, Stochastic block models: first steps. Soc. Netw. (1983)
8. C.F. Manski, Identification of endogenous social effects: the reflection problem. Rev. Econ. Stud. (1993)
9. C.R. Shalizi, A.C. Thomas, Homophily and contagion are generically confounded in observational social network studies. Sociol. Methods Res. (2011)
10. P. Toulis, E. Kao, Estimation of causal peer influence effects. J. Mach. Learn. Res. (2013)