Social Networks 54 (2018) 101–117

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

Alternative estimation methods for identifying contagion effects in dynamic social networks: A latent-space adjusted approach

Ran Xu*

Grado Department of Industrial and System Engineering, Virginia Polytechnic Institute and State University, United States

ARTICLE INFO

Article history:

Keywords:
Contagion effects
Social influence
Estimation methods
Latent-space approach

ABSTRACT

Contagion effects, also known as peer effects or social influence process, have become more and more central to social science, especially with the availability of longitudinal social network data. However, contagion effects are usually difficult to identify, as they are often entangled with other factors, such as homophily in the selection process, the individual's preference for the same social settings, etc. Methods currently available either do not solve these problems or require strong assumptions. Following Shalizi and Thomas (2011), I frame this difficulty as an omitted variable bias problem, and I propose several alternative estimation methods that have potentials to correctly identify contagion effects when there is an unobserved trait that co-determines the influence and the selection. The Monte-Carlo simulation results suggest that a latent-space adjusted estimator is especially promising. It outperforms other estimators that are traditionally used to deal with the unobserved variables, including a structural equation based estimator and an instrumental variable estimator.

1. Introduction

Endogenous social effects, which have long been central to the field of social science (Asch, 1952; Merton, 1957; Erbring and Young, 1979; Bandura, 1986), are defined as the propensity for the behavior of an individual to vary along with the prevalence of that behavior in some reference group containing the individual (Manski, 1993). Within the framework of social network analysis, the endogenous social effects are also known as "contagion" or "social influence", and the reference group can be one's network neighborhood. Contagion effects have also received much attention and have been widely studied (Kandel, 1978; Marsden and Friedkin, 1993; Doreian, 2001; An, 2011) as they have various implications for issues such as health behavior (e.g. obesity and smoking), information diffusion, or change in teacher practices, among others (Christakis and Fowler, 2007, 2008; Valente, 1995, 1996; Frank et al., 2004).

However, these types of contagion effects are usually difficult to identify, as it is difficult to separate such influences from other processes when there is network autocorrelation in the data, i.e. when we observe that people who are closely related to each other tend to be similar in some salient individual behavior and attitude dimensions, it is difficult to tell which is the underlying mechanism that generates these patterns. It could be influence and contagion (Friedkin and Johnsen, 1999; Friedkin, 2001; Oetting and Donnermeyer, 1998) whereby actors assimilate the behavior of their network members; or selection mechanisms, more specifically homophily (Lazarsfeld and Merton, 1954; McPherson and Smith-Lovin, 1987; McPherson et al., 2001), where actors seek to interact with similar others; or it could be due to different social contexts where people with previous similarities can select themselves into the same social setting, and actual friendship formation just reflects the opportunities of meeting in this social setting (Feld, 1981, 1982; Kalmijn and Flap, 2001).[1] Several notable attempts that try to identify contagion effects include modeling the co-evolution of selection and influence (Snijders et al., 2007; Steglich et al., 2010), using indirect ties from third parties as instrumental variables (Bramoullé et al., 2009; An, 2011), or Propensity Score Matching (Aral et al., 2009). But there is still considerable misconception about when it is problematic to identify contagion effects, and why these methods would need to be applied. Furthermore, all the methods mentioned above require some form of strong

[1] There are also structural constraints such as transitivity, preferential attachment etc. which could cause people to become friends. However these mechanisms in themselves do not entangle with influence (e.g. one befriends with another having high popularity but different behavior). In these cases another mechanism must be present to induce similarity between these friends (e.g. selection of common friends based on similarity in attributes), and thus the entanglement goes back to the original three mechanisms, namely influence, selection based on homophily, and social-environmental factors.

* Correspondence to: 250 Durham Hall, Virginia Tech, Blacksburg, VA, 24061, United States.
E-mail addresses: ranxu@vt.edu, ranxu@msu.edu

https://doi.org/10.1016/j.socnet.2018.01.002
0378-8733/

assumptions such as the exponential-family parametric assumption, the standard IV assumption, the assumption that all of the dependence is captured by observable covariates, and so on, each of which imposes substantial limits on the forms of data where these methods can actually be applied.

The difficulty of identification caused by entanglement between contagion effects and other confounding variables (social-environmental factors, or the attributes of egos and alters, for example) can be easily framed as an omitted variable bias problem. What is less obvious is that the dilemma caused by co-evolution of the influence and selection processes can essentially be framed as an omitted variable bias problem as well. As pointed out by Steglich et al. (2010), one of the important concerns is the "possibility that there may be non-observed variables co-determining the probabilities of change in network and/or behavior". Shalizi and Thomas (2011) have shown that when there is a latent trait that co-determines both influence and selection in network data, contagion effects are generally unidentifiable, mainly due to the fact that contagion and homophily (selection) are generically confounded through this latent trait.

In this paper, first I will clarify why contagion effects are difficult to identify; specifically, I will frame identification as an omitted variable bias problem. Then I will give some simulation examples exploring the magnitude of bias of the estimates for the contagion effects when there is an unobserved variable that codetermines influence and selection. Next, I will propose several estimation methods that have potentials to correctly identify contagion effects, and the focus of which is a latent-space based estimation method. Finally I will carry out simulation studies to examine the performance of these estimation methods.

2. Theoretical framework

2.1. Identifying contagion effects: where does the bias come from?

To understand where the bias comes from when identifying contagion effects, first we need to specify our "causal" models in terms of influence and selection. After specifying our model we then show how the estimation of contagion effects can suffer from bias.

A network behavioral (influence) model can be represented as

$$Y_{it} = f\left(Z_{ij}, Y_j, X_i, c_i\right) \tag{1}$$

where the behavior of node *i* at time t is a function of the behavior of network members Y_j, other variables X specific to node i, network relations Z and unobserved effects c_i. For example, adolescents' alcohol use (Y_{it}) can be a function of their previous alcohol use (Y_{it-1}), their close friends' alcohol use (Y_{jt-1}), their own cigarette use (X_{it-1}) and some latent disposition for substance abuse (c_i). Specifically, one form of such behavior models can be in a dynamic linear form (Friedkin and Johnsen, 1990)[2]:

$$Y_{it} = \beta_0 + \beta_1 Y_{it-1} + \beta_2 \frac{\Sigma Z_{ijt-1} Y_{jt-1}}{\Sigma Z_{ijt-1}} + \beta_3 X_{it-1} + c_i + e_{it} \tag{2}$$

where Y_{it-1} is the previous behavior of *i*, Z_{ijt-1} is a dummy variable indicating if there is a link from *i* to *j* at time $t-1$, i.e. 1 if yes and 0 otherwise, and $\frac{\Sigma Z_{ijt-1} Y_{jt-1}}{\Sigma Z_{ijt-1}}$ represents the weighted average behavior among the network neighbors of *i*, which is the exposure term (contagion) of interest,[3] and X_{it-1} represents other variables that might affect the behavioral outcome Y. We choose this form of behavioral model for several reasons: (1) we choose linear models as they have greater flexibility when compared with models like SIENA (Steglich et al., 2010), and because of the availability of methods that they may produce to deal with unobserved time-constant actor differences (Steglich et al., 2010; Mouw, 2006); (2) we only use lagged endogenous variables (if X is exogenous, i.e. if X is not "caused" by Y then X can be contemporaneous), which to us is a more realistic assumption that there is some lag in the transmission of social effects. In addition, such formulations require less strict conditions for identification of social effects (Manski, 1993). One might argue that there are also contemporaneous social effects that should be included, which constitute the true "structural model" (Sims, 1980; Bramoullé et al., 2009). Even if this is true, however, the identification of contemporaneous effects often requires strong structural restrictions or valid instrumental variables (Sims, 1980; Manski, 1993; Wooldridge, 2010), and including both contemporaneous and lagged effects when identifying contagion effects can cause problems both in estimation and interpretation (Lyons, 2011; VanderWeele and An, 2013; VanderWeele et al., 2012). As Sims (1980) has argued for vector-autoregressive models, the type of "reduced form" models in model 2 do not require "too many incredible restrictions" for identification, and are still very useful in forecasting and analysis.

For the selection process, let $Z_{ijt} = 1$ if there is a connection from node i to node j at time t, and let Z^*_{ijt} be a latent variable defined as

$$Z^*_{ijt} = \alpha_0 + \alpha_1 Z_{ijt-1} + \alpha_2 |c_i - c_j| + \alpha_3 |Y_{it-1} - Y_{jt-1}| + \alpha_4 |X_{it-1} - X_{jt-1}| + \varepsilon_{ijt} \tag{3}$$

where c represents a time invariant unobserved trait for i and j, Y represents the behavior of interest, X represents the observed variables and $\varepsilon_{ijt} \sim N(0, 1)$. By defining Z_{ijt} as

$$Z_{ijt} = \begin{cases} 1 \ \ if \ \ Z^*_{ijt} > 0 \\ 0 \ \ \ \ \ \ \ \ otherwise \end{cases} \tag{4}$$

we know that Z_{ijt} follows a standard probit model (Wooldridge, 2010) where

$$P\left(Z_{ijt} = 1\right) = \Phi(\alpha_0 + \alpha_1 Z_{ijt-1} + \alpha_2 |c_i - c_j| + \alpha_3 |Y_{it-1} - Y_{jt-1}| + \alpha_4 |X_{it-1} - X_{jt-1}| + \varepsilon_{ijt}) \tag{5}$$

The models described in Eqs. (2) and (5) are now called simply models 2 and 5. Model 5 represents the selection model. Through models 2 and 5 we now have described the co-evolution of the influence and selection processes, which operate through the same sets of observed and unobserved variables. And the magnitude of contagion effects is represented by the parameter β_2 in model 2.

To understand where any bias comes from, we need to know that in order to get consistent estimates in model 2 using conventional approaches such as OLS, one key assumption is that unobserved errors have to be uncorrelated with observed variables. In this case, if either the idiosyncratic error e_{it} or a latent trait c_i is correlated with observed variables, we will have biased estimates. For now we only focus on the latent trait c_i and assume that e_{it} does not correlate with observed variables in model 2. (Different exogeneity assumptions must hold for different estimation methods, for more details see Wooldridge (2010)). We already know that c_i correlates with Y_{it-1}, and if $\alpha_2 < 0$ in model 5, such that there is a homophily based selection which operates through an unobserved trait, we know (i) person i will select person j with a similar latent trait, and

[2] Note that although we choose a linear form of the model with mean exposure term, these assumptions can be easily relaxed and our arguments can be applied to other types of models.

[3] From now on I will use the term "contagion effects" to represent β_2, "network exposure term" to represent $\frac{\Sigma Z_{ijt-1} Y_{jt-1}}{\Sigma Z_{ijt-1}}$.

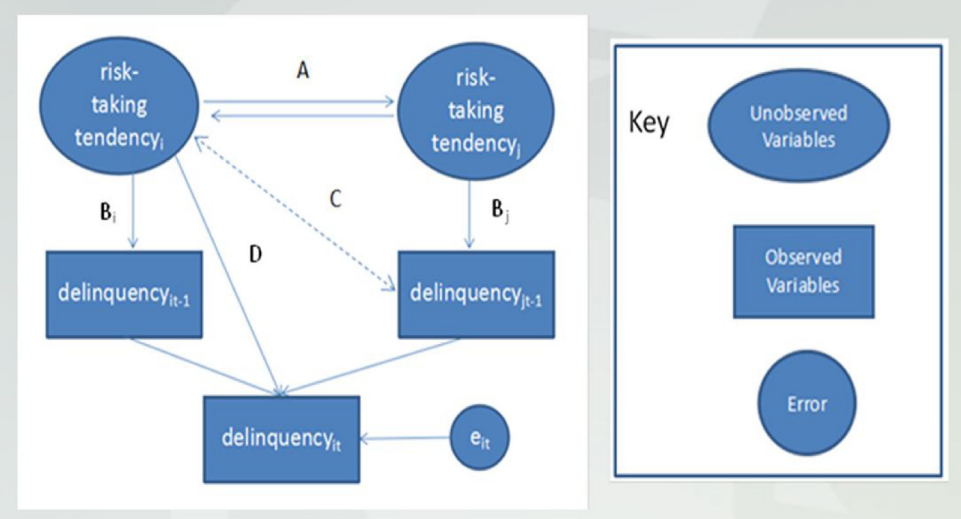

Fig. 1. Omitted variable bias.

(ii) person j's behavior is a function of person j's latent trait c_j, which is similar to c_i through selection; and (iii) together c_i will be correlated with person j's behavior, which is analogous to the exposure term in model 2. As c is unobserved, this violates the key assumption of OLS (as well as many other methods), so that estimates in model 2 will be inconsistent, and the contagion (exposure) effect is unidentifiable.

To give an example, assuming that delinquency is a function of an unobserved risk-taking tendency (arrow B_i in the Fig. 1), and when there is homophily based selection which operates through this unobserved variable, (i) person *i* will select person *j* who is similar on the unobserved risk-taking tendency (arrow A in the figure); (ii) person j's delinquency behavior is a function of person j's risk-taking tendency (arrow B_j), which is similar to person i's risk-taking tendency through selection; and (iii) because of (i) and (ii) the risk-taking tendency for person *i* will be correlated with person *j*'s delinquency behavior (arrow C in the figure). As the risk-taking tendency is unobserved, this violates the key assumption of OLS, so that estimates may be inconsistent, and the contagion (exposure) effect is unidentifiable. For an analogous algebraic argument see Appendix c in Supplementary material.

Hence through a regression framework we have explained how selection (homophily) is confounded with influence, and as we can see, this can directly translate into an "omitted variable bias" problem, under which there are some omitted variables that we do not control for, but which affect both selection and behavioral outcomes.[4] So instead of stating that contagion effects are unidentifiable because selection operates at the same time, a more meaningful question to ask might be "what factors in the selection process might also affect behavioral outcomes?" And as long as we have that variable controlled in the influence model, we should not be worried about selection (homophily) being confounded with influence any more.

2.2. Magnitude of bias when the unobserved trait is ignored

In an ideal world we would correctly measure all variables that can affect influence, selection, and social context, so that there will be no remaining omitted variable problem, and conventional approaches such as OLS can be applied.[5] But given the limited richness of the data, an omitted variable problem is almost inevitable in any empirical study, in which we know OLS estimates for contagion effects will be biased. Furthermore, most methods that are currently widely used also suffer from this omitted variable bias problem. For example, one might attempt to use models that can model influence and selection at the same time (SIENA for example) to separate influence from selection, however as Steglich et al. (2010) pointed out, such models still will not work when "non-observed variables co-determine the probabilities of change in network and/or behavior". One might borrow from the causal inference literature (Rosenbaum and Rubin, 1983) and use methods such as propensity matching (Aral et al., 2009), but that still does not deal with unobserved variable problems as strong ignorability assumes observed variables carry all the dependency between outcomes and treatment assignments (contagion or network exposure).

To better understand how estimates of the contagion effects will be biased when selection is confounded with influence, we provide some intuitions of the magnitude of the bias through some simulation examples. Specifically, we are interested in the magnitude of bias of the estimates of the contagion effects when we ignore the unobserved trait that codetermines influence and selection. For simplicity we do not include other observed variable Xs, and we let the "true" influence model be:

$$Y_{it} = \beta_1 Y_{it-1} + \beta_2 \frac{\Sigma Z_{ijt-1} Y_{jt-1}}{\Sigma Z_{ijt-1}} + \beta_3 c_i + e_{it} \quad (6)$$

And correspondingly, the "true" selection model can be represented as

$$P\left(Z_{ijt} = 1\right) = \Phi(\alpha_0 + \alpha_1 |c_i - c_j|) \quad (7)$$

where c is an individual specific time-invariant unobserved trait that is present in both the influence and the selection process.

In the simulations we start from a random network and simulate a network data based on models in (6) and (7), and estimate the influence model while ignoring the unobserved trait:

$$Y_{it} = \beta_1 Y_{it-1} + \beta_2 \frac{\sum Z_{ijt-1} Y_{jt-1}}{\sum Z_{ijt-1}} + e_{it} \quad (8)$$

[4] Note that unobserved variables, that only affect behavioral outcomes but not selection, may cause estimation problems as well, but that is not the focus here.

[5] Note that in this case a reflection problem still exists if contagion effect is contemporaneous. For a formal derivation of the problem see Appendix b in Supplementary material.

Table 1
OLS estimates of the influence model when unobserved trait is ignored.

		Lagged Dependent Variable		Network Exposure	
	True Coefficient	Mean Bias	SD	Mean Bias	SD
High Homophily	0	0.176	0.048	0.062	0.032
	0.1	0.184	0.046	0.082	0.046
	0.2	0.185	0.042	0.121	0.059
	0.3	0.185	0.039	0.161	0.071
	0.4	0.183	0.036	0.195	0.078
	0.5	0.173	0.031	0.227	0.089
	0.6	0.156	0.028	0.255	0.097
	0.7	0.136	0.022	0.276	0.103
	0.8	0.112	0.018	0.282	0.106
	0.9	0.076	0.013	0.293	0.113
Low Homophily	0	0.180	0.052	0.018	0.032
	0.1	0.197	0.051	0.027	0.050
	0.2	0.204	0.051	0.043	0.067
	0.3	0.211	0.046	0.059	0.082
	0.4	0.209	0.042	0.061	0.092
	0.5	0.203	0.036	0.073	0.108
	0.6	0.190	0.029	0.074	0.114
	0.7	0.167	0.023	0.077	0.117
	0.8	0.133	0.017	0.068	0.128
	0.9	0.089	0.011	0.068	0.134
No Homophily	0	0.172	0.054	−0.003	0.029
	0.1	0.185	0.053	−0.059	0.045
	0.2	0.198	0.051	−0.094	0.062
	0.3	0.203	0.047	−0.116	0.078
	0.4	0.207	0.043	−0.116	0.096
	0.5	0.200	0.039	−0.114	0.107
	0.6	0.191	0.031	−0.113	0.115
	0.7	0.169	0.025	−0.097	0.126
	0.8	0.137	0.016	−0.083	0.119
	0.9	0.090	0.011	−0.068	0.131

and find OLS estimates for β_1 and β_2.

Simulation configuration. While there are many factors that could affect the magnitude of the bias, such as density, the magnitude of the latent trait (β_3), variance of idiosyncratic error, etc, we focus on the following:

(1) level of homophily. As homophily operates through the unobserved trait in our simulated data, higher homophily means higher correlation between network exposure and the unobserved trait (also higher correlation between network exposure and lagged dependent variable). We are interested in how homophily affects the magnitude of bias.
(2) magnitude of true coefficients for lagged dependent variable and network exposure, as different levels of influence might affect the magnitude of bias in estimation.

Specifically, let a simulation configuration be as follows: (1) we fix the number of nodes as 80 and the number of time points to be 5; (2) we vary the homophily level to be (i) no homophily, where we start the simulation from a random network that does not change over time, (ii) low homophily ($\alpha_0 = -0.75$, $\alpha_1 = -0.1$. Note that α_1 affects the level of homophily based on the unobserved trait and α_0 affects the overall density of the network); (iii) high homophily ($\alpha_0 = -0.55$, $\alpha_1 = -0.3$); and (3) we vary both β_1 and β_2 from 0.1 to 0.9, while keeping $\beta_1 + \beta_2 = 0.9$. And we set $\beta_3 = 0.1$. This only changes the dynamics of the relationship, and keeps the consensus within the initial range for c_i (Friedkin and Johnsen, 1999). Other model configurations include $e_{it} \sim N(0, 0.2^2)$, initial density of the network is 0.2. And in all three configurations α_0 and α_1 are chosen to deliver the overall density similar to that of the initial network.

The Mean biases for the lagged dependent variable and the network exposure term under various conditions are shown in Table 1. There are several things to note concerning the lagged dependent variable term: (1) the bias for lagged dependent variable is generally smaller when the true coefficient is larger; and (2) the OLS estimates for the lagged dependent variable term are consistently upwardly biased, and the magnitude of bias does not change much with different homophilous selection processes.

The bias for network exposure presents a more interesting pattern: (1) contrary to estimates of the lagged dependent variable term, the bias for network exposure is generally smaller when the true coefficient is smaller; (2) when homophily is present in the selection process, the network exposure is upwardly biased, and the magnitude of bias is much smaller for lower levels of homophily; (3) when there is no homophily in the selection process and the network is static, estimates for network exposure are downwardly biased. Note that as true coefficient of network exposure increases, standard error of the estimate also increases. This is because as influence becomes stronger in the network, nodes' behaviors converge and exhibit smaller variance, resulting in larger standard error.

One possible reason for the lagged dependent variable term to be consistently upwardly biased is that it has a consistently high correlation with the unobserved trait (>0.6). The direction of bias for the contagion effects shows an interesting pattern: it is upward when there is latent homophily in the selection process, but downward when the network is static. We provide some explanations in the technical Appendix a in Supplementary material.

From the results above we can see that an uncontrolled unobserved trait that either only affects influence or codetermines selection and influence will indeed create biased OLS estimates (as well as many other methods) for the contagion effects, possibly leading to invalid inference.[6] And given the limited richness of the data, an omitted variable problem is almost inevitable in any empirical study. So in the next section, we turn to some alternative estimation methods that have the potential to correctly identify

[6] We did not fully explore the magnitude of bias under various scenarios. And simulation examples here merely show the existence of the bias.

contagion effects, assuming that there are unobserved variables that co-determine the influence and the selection.

3. Alternative estimation methods

In this section we will propose three estimation methods that have potentials to correctly identify contagion effects when there are unobserved variables which co-determine influence and selection. We will introduce each, and explain how they can be applied in a social network context to identify contagion effects under weaker assumptions. After that we will perform Monte-Carlo simulations to test the performance of each estimator we propose. The main method that we will focus on is what we called a "latent-space adjusted" approach, which utilize the latent social positions from the latent space models (Hoff et al., 2002). And we will compare this new method with two existing methods which are traditionally used to deal with omitted variable bias problem, yet less commonly used by social network researchers. One of them is Structural-equation model (SEM) based and the other is a variation of the instrumental variable (IV) methods.

3.1. Latent-space adjusted approach

In a case where there is an unobserved trait that co-determines the influence and the selection process (homophily based on the unobserved trait), in theory if there is any information about this unobserved trait from the selection process, it can be borrowed and used in the estimation of the influence model, and this will reduce the bias in estimating the contagion effects. However, most selection models (e.g. SIENA) will likely not contribute to the influence estimation in this case, as their estimations are based on observed variables and thus do not attend to those variables that are not observed. And here we will propose an estimation procedure that borrows information beyond what is observed from the selection process, and use that to help estimating the contagion effects in the influence model.

Our approach builds on the theoretical logic of latent space models as applied to social-network data (Hoff et al., 2002). Latent space models assume that each individual has a "latent position" that lies in an unobserved n-dimensional social space, and the probability of interaction between any two actors depends on the latent positions of these two actors. Specifically, they take a logistic form and specify the selection model as

$$log\ \ odds\left(Z_{ij}=1|c_i,c_j,x_{ij},\alpha,\beta\right)=\alpha+\beta' x_{ij}-|c_i-c_j| \tag{9}$$

Here, Z_{ij} indicates whether there is an interaction from i to j, x_{ij} is a vector of observed covariates (at dyadic level or node level), c indicates the latent social position of i and j, and $|c_i - c_j|$ represents the Euclidean distance between i and j's latent position. A smaller distance between i and j's latent position indicates a larger probability of having a tie. And these latent social positions can be regarded as determinants of interactions that have not been accounted for by the observed variables in the selection process. The parameters α and β are estimated using either Maximum-Likelihood Estimation (MLE) or Markov Chain Monte Carlo (MCMC) methods, and the latent position c can estimated by Minimum Kullback-Leibler (MKL) estimates (Shortreed et al., 2006). Note that if there are no covariates in model 9, this model is similar in principle to multidimensional scaling (Kruskal, 1964), which put nodes at positions in n-dimensional space based on their network relations. As described in Hoff et al. (2002), crude estimates of individual positions from multidimensional scaling are actually used as starting values in their estimation procedure.

It is not difficult to see that the latent space model described as model 9 is very similar to the selection process as we defined in model 3, except that c represents latent position in the latent space model, while c represents individual's unobserved trait in model 3.[7] Therefore, if two individuals are close to each other in terms of the latent social positions, they should also be close to each other in terms of the unobserved trait. And if these latent positions are estimated accurately enough, the estimates of these latent positions can be used as proxies for the unobserved trait that determines the homophily in the selection process. Furthermore, if these unobserved traits are also present in the influence process, the estimated latent social positions from the latent space model can also be included when estimating an influence model such as in model 2, and this will in-principle reduce the bias in estimation of contagion effects that are due to the omitted variable problem (Wooldridge, 2010).[8] For example, to model adolescents' delinquency behavior, we can first use a latent space model to model the friendship network of adolescents and acquire an estimated "latent social position" for each individual, and then use these estimates as proxies for the unobserved risk-taking tendency in the influence model, and thus achieve a better estimation of the true contagion effects.

There are two things to note here: (1) in principle this method can apply to any functional form of the selection model, because it accounts for the homophily determinants that are unobserved in the selection process, which are also present in the influence process. This is different from many traditional methods that only depend on the observed variables; (2) as the scale and the actual position of the estimated latent social positions are essentially arbitrary (Hoff et al., 2002), the actual value of the latent social position might be very different from the actual value of the unobserved trait that codetermines influence and selection. However, as long as the estimated latent social positions are highly correlated with the unobserved trait (actors who are close to each other on the latent social positions are also close to each other in terms of the unobserved trait), the contagion effects can still be consistently estimated.

However, if the social network data is longitudinal, the latent space model as described in model 9 cannot produce consistent latent position estimates across different time points, as it is static in nature. Extensions of latent space models that apply to dynamic social network data have been proposed (Sarkar and Moore, 2005), but they do not assume constant latent positions across time, which violates one of our key assumptions, that individuals possess time-invariant latent-traits. In addition, there is the difficulty of implementation in software. So instead we propose a two-step estimator for contagion effects: (1) we estimate a latent space model and acquire the latent position estimates for each time point, and (2) we include estimated values of latent positions for all available time points as proxies for the unobserved trait, and estimate the influence model using OLS. Specifically, let c be the unobserved trait that codetermines influence and selection, q_t be the latent position estimates for time t, assume that

$$c=\gamma_0+\sum_{1}^{T}\gamma_t q_t+r \tag{10}$$

Under the assumption that E(y|x,c,q) = E(y|x,c) and $Cov(x,r)=0$ (X represents all independent variables in the influence model), q can be valid proxies for the unobserved trait c and thus the con-

[7] Here we only choose one-dimensional latent social positions to mimic the unobserved trait that drives the homophily in the selection process. The arguments can easily be extended to multi-dimensional latent positions.

[8] In principle this approach is similar to Heckman's control function approach to deal with selection bias. However, one important difference is that Heckman's selection model requires instrumental variables to be uncorrelated with unobserved variables, while in our approach latent social positions should be proxies for the unobserved variables.

tagion effect is identified. Essentially we assume that each q is an imperfect measure of c, and by including all q's we will have a better approximation to c, and thus better estimation of contagion effects.[9]

3.2. Instrumental variable methods

Instrumental variable (IV) methods are often used in situations where explanatory variables are correlated with the error terms, which can be caused by simultaneity, omitted variables, measurement error, etc. These type of methods work through identifying a set of new variables that only correlate with endogenous explanatory variables, but not with the unobserved error terms, and thus achieve consistent estimation by "blocking out" the correlation between the endogenous variable and unobserved errors (An, 2011; Wooldridge, 2010). There have been a few studies that used IV methods to identify contagion effects. For example Duncan et al. (1968) used a friend's intelligence as an IV for the friend's occupational and educational aspirations. Angrist and Lang (2004) used the predicted number of transferred-in disadvantaged students to study their effects on the academic performance of students in the receiving schools. O'Malley et al. (2014) used genetic alleles as IVs to estimate peer effects on weight status. An (2015) used friends' family smoking status to estimate peer effects on smoking. However, all these IV methods require a strong theoretical argument of validity for the instrumental variables, which thus are essentially untestable. And we will also encounter inconsistency problems and large standard error of estimates when we have weak instruments or data with small sample sizes (Bound et al., 1995; Wooldridge, 2010).

There are also studies exploiting structural properties of networks to identify instrumental variables. For example, Bramoullé et al. (2009) argued that if there are intransitive triads, for example i−>j−>k but i and k are not connected, then i's outcome can be used as instruments for j to estimate contagion effects for k's outcome, since k is not directly influenced by i. However, the identification of the model would require the validity of the instrument such that i does not influence k through any alternative path, and simulations by Bramoullé et al. (2009) have shown that the quality of IV estimates depend on specific network structural properties as well (precision decreases with denser networks and complex functions of intransitivity).

Given the strong assumptions required by various instrumental variables above, alternatively we propose to exploit the dynamic nature of our data and our model. Specifically, as pointed out by Anderson and Hsiao (1982), under specific assumptions, past values of one's own outcomes can be used as instruments for endogenous variables in a dynamic model. To see this, we first-difference our original influence model to remove the time invariant unobserved trait c_i (we exclude X in model for simplicity):

$$\begin{aligned} &\text{Originalinfluencemodel}: \quad Y_{it} = \beta_1 Y_{it-1} + \beta_2 \tilde{Y}_{it-1} + c_i + e_{it} \\ &\textit{Afterfirst-difference}: \quad \Delta Y_{it} = \beta_1 \Delta Y_{it-1} + \beta_2 \Delta \tilde{Y}_{it-1} + \Delta e_{it} \end{aligned} \tag{11}$$

Where $\tilde{Y}$ represents the original network exposure term, and $\Delta Y_{it} = Y_{it} - Y_{it-1}$, $\Delta Y_{it-1} = Y_{it-1} - Y_{it-2}$, $\Delta \tilde{Y}_{it-1} = \tilde{Y}_{it-1} - \tilde{Y}_{it-2}$ etc. As in the fixed effects approach, this transformation will induce a correlation between $\Delta Y_{it}$and Δe_{it}, thus biasing the estimates (Nickell, 1981). However, under a sequential exogeneity assumption, which states that errors (shocks) in the future are independent of past values of y (which seems to be a very reasonable assumption if errors do not contain omitted variables and are structural/idiosyncratic), plus an assumption that errors are serial-independent, a natural instrument would be the past values of Y for each time period, which will correlate with ΔY_{it-1} but not with Δe_{it}, and hence satisfy the IV assumption. For example, in a panel data with 3 time points, for $Y_{i2} - Y_{i1}$ the instrumental variable can be Y_{i1}, as it does not correlate with e_{i3}-e_{i2}, and this will generate consistent estimates of β_1. To be more concrete, using the delinquency example, to model the change score of delinquency from time 2 to time 3, we can use adolescents' delinquency score at time 1 as an instrumental variable for the change score of delinquency from time 1 to time 2. Note that in our setup the transformed network exposure (contagion) term does not correlate with the transformed error term, as the effect of network exposure on outcomes in the influence model is not simultaneous, but lagged. To see this more clearly write $\Delta \tilde{Y}_{it-1} = \tilde{Y}_{it-1} - \tilde{Y}_{it-2}$and $\Delta e_{it} = e_{it} - e_{it-1}$, showing that these two terms are independent, since any change in e_{it-1} will be reflected in $\tilde{Y}_{it}$ but not in $\tilde{Y}_{it-1}$ or $\tilde{Y}_{it-2}$. In this sense, the exposure term is "exogenous", so that it can be identified without extra instrumental variables.[10]

However, since all past values of Y can potentially be instruments, Arellano and Bond (1991) proposed using the entire set of instruments in a generalized method of moments (GMM) procedure to improve efficiency. Specifically, let the matrix of instrumental variables for individual i to be Z_i, as follows[11]:

$$Z_i = \begin{bmatrix} Y_{i1} & 0 & 0 & 0 & 0 & 0 & . & . & 0 & 0 & 0 & 0 & \Delta\tilde{Y}_{i2} \\ 0 & Y_{i1} & Y_{i2} & 0 & 0 & 0 & . & . & 0 & 0 & 0 & 0 & \Delta\tilde{Y}_{i3} \\ 0 & 0 & 0 & Y_{i1} & Y_{i2} & Y_{i3} & . & . & 0 & 0 & 0 & 0 & \Delta\tilde{Y}_{i4} \\ . & . & . & . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & . & . & . & Y_{i1} & Y_{i2} & . & . & Y_{iT-2} & \Delta\tilde{Y}_{iT-1} \end{bmatrix} \tag{12}$$

Each column in Z represents an instrumental variable z. As we can see, each variable z is uncorrelated with the error term in the model in (11), such that E[z'e] = 0. Writing the model in (11) as Y = XB + E, we can use 2-stage least squares (2SLS) estimation to achieve consistent estimates: in the first stage regress X on Z; in the second stage regress Y on the predicted value of X from the first stage regression. Together 2SLS can be written as

$$\hat{B}_{2SLS} = \left(\left(\sum_i^N X_i' Z_i \right) \left(\sum_i^N Z_i' Z_i \right)^{-1} \left(\sum_i^N Z_i' X_i \right) \right)^{-1} \left(\left(\sum_i^N X_i' Z_i \right) \left(\sum_i^N Z_i' Z_i \right)^{-1} \left(\sum_i^N Z_i' Y_i \right) \right)$$

Alternatively, we can use a GMM-IV estimator, which can be represented as$\hat{B}_{GMM} = \left((\Sigma^N X_i' Z_i) \hat{W} \left(\Sigma^N Z_i' X_i \right) \right)^{-1} \left(\left(\Sigma^N X_i' Z_i \right) (\hat{W}) \left(\Sigma^N Z_i' Y_i \right) \right)$, where W is a weighting matrix which is the inverse of the variance-covariance matrix of $Z_i' E_i$. And as we can see, the only difference between the 2SLS and GMM-IV estimators is that they use a different weighting matrix.

[9] As all q's are probably highly correlated, this will possibly create a multicollinearity issue when estimating the influence model. But this will not affect estimation of lagged dependent variable and contagion effects, as all these q's are explaining unique variances represented by c under the assumption E(y|x,c,q) = E(y|x,c)

[10] Note that if influence is contemporaneous instead of lagged, it is also possible to use past values of exposure terms as instrumental variables and thus achieve identification.

[11] The instrumental variable matrix was constructed this way to create as many moment condition as possible, zeroes were added instead of missing values, and thus still keeping orthogonality.

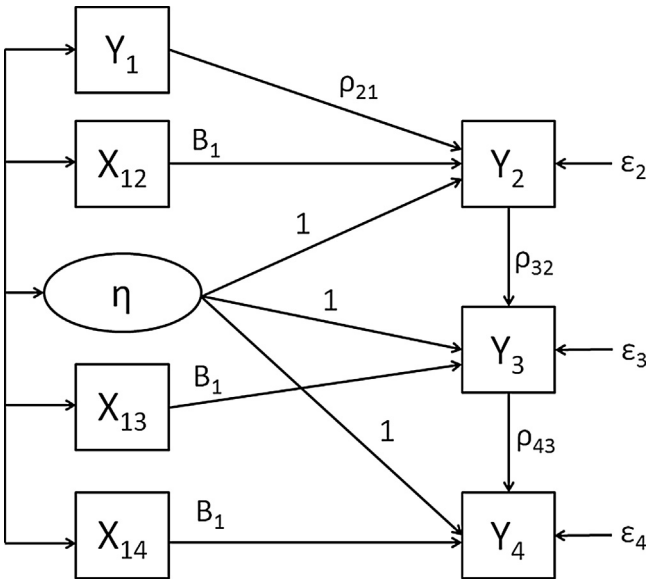

Fig. 2. Dynamic model with unobserved term.

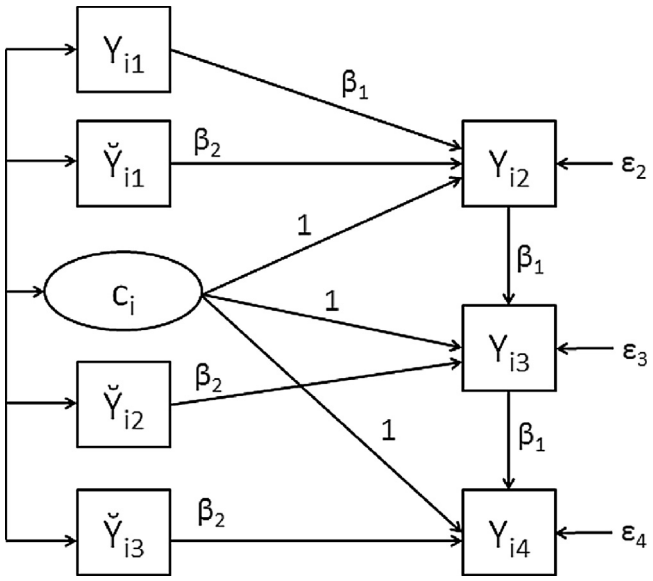

Fig. 3. Influence model in structural equation model.

This GMM IV approach is shown to be generally consistent and efficient as $N \to \infty$, but in empirical work the optimal number of moment conditions that should be used for estimation is not that clear (Judson and Owen, 1999; Kiviet, 1995; and Wansbeek and Bekker, 1996). And simulation by Ziliak (1997) also has shown that there could be a downward bias in GMM estimates as the number of moment conditions expands. Furthermore, it is shown that this method will also suffer from the weak-instrument problem when β_1 approaches 1 (Wooldridge 2010).[12] As none of these similar methods have been applied to social network data, the performance of such estimators remains largely unknown. So here we incorporate this estimator of contagion effects and examine how well it performs.

3.3. Structural equation modeling approach

Structural equation modeling (SEM) is also known as an alternative approach to deal with latent variables (Kaplan, 2007; Kline, 2011). SEM is widely used in the social sciences mainly due to its ability to isolate observational error from measurement using latent constructs (Hancock, 2003). But it can also be used to model unobserved variables in the estimation procedure. For example Barnes et al. (2000) use latent growth modeling to study the alcohol use of adolescents, with latent variables representing adolescents' initial drinking behavior and rates of increase in alcohol use. In a social network context, as described above, if we treat an unobserved trait that codetermines influence and selection as a latent variable, borrowing from the SEM framework we can estimate the unobserved trait as a latent variable, and thus correctly identify the contagion effects.

In a paper by Bollen and Brand (2010), a structural equation modeling based approach is discussed to estimate parameters in dynamic models with unobserved heterogeneity. Fig. 3 provides a graphic depiction of their model, where Y represents the outcome of interest and X is the contemporaneous exogenous variable (subscripts indicate different time points). As can be seen in Fig. 2, error variances and the coefficients for the time-varying variables across different time points are set to be equal. The latent time-invariant variables, η_i, representing unobserved heterogeneity, are allowed to correlate with both the exogenous variables and the lagged values of the outcome variable, Y. In principle, the Bollen and Brand (2010) model should provide accurate estimates of the ARDL (auto-regressive distributed lag) model with unobserved heterogeneity, since it models the unobserved heterogeneity without running into any incidental parameter problem (Lancaster, 2000). It allows correlations between unobserved effects and exogenous variables and a lagged dependent variable. Further, it models a dependent variable conditioned on an initial observation, y_1, thereby avoiding the initial condition problem (Anderson and Hsiao, 1981; Wooldridge, 2005). Unfortunately, few simulation studies have been performed to evaluate the performance of this SEM approach. Therefore, we will incorporate this method with our influence model, and use simulation to examine the performance of this method. Specifically, we represent the influence model as in Fig. 3.

Here, Y represents a behavioral outcome, $\tilde{Y}$ represents network exposure, and c represents the latent trait that codetermines influence and selection. For example, Y can represent the delinquency behavior of a focal adolescent, and $\tilde{Y}$ can represent the delinquency behavior of his/her friends, while c represents the unobserved risk-taking tendency. By setting up as in Fig. 3, we follow Bollen and Brand's (2010) framework and allow c to be estimated as a latent variable, which at the same time correlates with a lagged dependent variable as well as the network exposure term. We will obtain model estimates by maximum likelihood estimation.

After the description of all the proposed methods, our main research question is whether estimation methods proposed above can correctly identify contagion effects when there are unobserved traits that co-determine the influence and selection. So in the next section we perform Monte-Carlo simulation to test how well these methods can recover the true contagion effects.

4. Monte-Carlo simulation

In this section, we use Monte-Carlo simulations to examine the performance of each estimator for contagion effects proposed above: the latent space adjusted estimator, the GMM-IV estimator,

[12] Note empirically that an important diagnostic test uses auto-correlation of the error terms in model 11. By construction, errors should exhibit AR(1) behavior but not AR(2).

Table 2a
Simulation results for lagged dependent variable.

Lagged Dependent Variable											
		LSDV		Random Effects		GMM		SEM		Latent Space	
		Mean Bias	SD	Mean Bias	SD	Mean Bias	SD	Mean Bias	SD	Mean Bias	SD
N = 40,T = 6	β1 = 0	−0.174	0.066	0.102	0.100	−0.022	0.097	0.011	0.090	−0.023	0.080
	β1 = 0.1	−0.201	0.070	0.124	0.102	−0.037	0.111	0.007	0.096	0.000	0.057
	β1 = 0.2	−0.227	0.071	0.151	0.090	−0.034	0.120	0.010	0.102	−0.023	0.070
	β1 = 0.3	−0.254	0.073	0.162	0.080	−0.039	0.130	0.020	0.119	−0.019	0.062
	β1 = 0.4	−0.289	0.072	0.165	0.065	−0.069	0.147	0.016	0.129	−0.006	0.059
	β1 = 0.5	−0.315	0.074	0.167	0.053	−0.086	0.161	0.046	0.161	−0.006	0.062
	β1 = 0.6	−0.355	0.073	0.154	0.040	−0.114	0.182	0.043	0.178	0.003	0.045
	β1 = 0.7	−0.388	0.076	0.134	0.033	−0.194	0.222	0.041	0.188	−0.002	0.051
	β1 = 0.8	−0.423	0.079	0.107	0.027	−0.329	0.293	0.038	0.202	0.013	0.045
	β1 = 0.9	−0.423	0.082	0.065	0.017	−0.486	0.400	0.010	0.189	0.011	0.027
N = 80,T = 6	β1 = 0	−0.185	0.049	0.114	0.083	−0.009	0.072	0.011	0.064	0.011	0.045
	β1 = 0.1	−0.213	0.050	0.136	0.079	−0.013	0.076	0.006	0.065	0.008	0.046
	β1 = 0.2	−0.240	0.051	0.160	0.073	−0.016	0.085	0.005	0.069	0.012	0.052
	β1 = 0.3	−0.264	0.052	0.176	0.061	−0.028	0.095	0.010	0.075	0.008	0.054
	β1 = 0.4	−0.296	0.051	0.180	0.040	−0.032	0.100	0.008	0.081	0.014	0.059
	β1 = 0.5	−0.326	0.055	0.174	0.033	−0.045	0.112	0.010	0.094	0.007	0.051
	β1 = 0.6	−0.356	0.054	0.159	0.027	−0.057	0.133	0.023	0.107	0.017	0.035
	β1 = 0.7	−0.392	0.052	0.136	0.024	−0.090	0.167	0.030	0.133	0.021	0.042
	β1 = 0.8	−0.424	0.056	0.112	0.019	−0.192	0.218	0.029	0.156	0.022	0.030
	β1 = 0.9	−0.413	0.056	0.065	0.012	−0.362	0.315	0.026	0.136	0.009	0.017
N = 40,T = 3	β1 = 0	−0.441	0.138	0.148	0.122	0.037	0.270	0.042	0.237	0.045	0.101
	β1 = 0.1	−0.506	0.142	0.168	0.109	0.055	0.355	0.073	0.355	0.030	0.111
	β1 = 0.2	−0.556	0.152	0.172	0.101	0.061	0.379	0.078	0.328	0.063	0.115
	β1 = 0.3	−0.609	0.155	0.176	0.097	0.095	0.511	0.120	0.539	0.069	0.119
	β1 = 0.4	−0.661	0.153	0.174	0.086	0.131	0.633	0.207	1.035	0.060	0.102
	β1 = 0.5	−0.724	0.156	0.164	0.073	0.113	0.767	0.184	0.871	0.056	0.108
	β1 = 0.6	−0.766	0.172	0.149	0.063	0.103	0.978	0.287	1.267	0.041	0.087
	β1 = 0.7	−0.830	0.164	0.131	0.059	−0.013	1.520	0.280	1.829	0.069	0.070
	β1 = 0.8	−0.884	0.177	0.108	0.043	−0.521	1.872	0.559	2.492	0.054	0.053
	β1 = 0.9	−0.929	0.161	0.068	0.029	−0.594	2.175	0.706	2.595	0.027	0.036
N = 80,T = 3	β1 = 0	−0.478	0.096	0.163	0.092	0.016	0.190	0.019	0.176	0.051	0.067
	β1 = 0.1	−0.526	0.101	0.183	0.076	0.028	0.225	0.034	0.193	0.048	0.078
	β1 = 0.2	−0.580	0.104	0.183	0.072	0.033	0.267	0.042	0.247	0.062	0.087
	β1 = 0.3	−0.642	0.106	0.182	0.066	0.042	0.344	0.044	0.277	0.047	0.081
	β1 = 0.4	−0.678	0.099	0.178	0.059	0.031	0.348	0.047	0.292	0.076	0.074
	β1 = 0.5	−0.740	0.108	0.166	0.054	0.090	0.516	0.068	0.353	0.060	0.070
	β1 = 0.6	−0.790	0.112	0.156	0.044	0.155	0.740	0.185	0.812	0.060	0.065
	β1 = 0.7	−0.836	0.109	0.138	0.038	0.153	0.897	0.285	1.127	0.064	0.055
	β1 = 0.8	−0.895	0.114	0.112	0.029	−0.097	1.760	0.498	1.632	0.062	0.041
	β1 = 0.9	−0.916	0.117	0.067	0.019	−0.509	1.769	0.568	2.042	0.033	0.025

and the SEM estimator. For simplicity, we do not include exogenous observed variable Xs, and we let the "true" influence model be:

$$Y_{it} = \beta_1 Y_{it-1} + \beta_2 \frac{\sum Z_{ijt-1} Y_{jt-1}}{\sum Z_{ijt-1}} + \beta_3 c_i + e_{it} \quad (13)$$

Correspondingly let the "true" selection model be represented as

$$P\left(Z_{ijt} = 1\right) = \Phi(\alpha_0 + \alpha_1 |c_i - c_j|) \quad (14)$$

All the notations are same as before. The influence process and the selection process are linked through c – the individual specific time-invariant unobserved trait.

Simulation configuration. While there are many factors which could affect the performance of estimators, such as density, the magnitude of the latent trait (β_3), the variance of idiosyncratic error, etc., we focus on the performance of estimators under the following cases. (1) High homophilous selection determined by the latent trait; (2) sample size: As N and T are important in panel data, which usually have large N and small T, we will focus on the number of nodes and the number of time points; (3) the magnitude of the true coefficient for lagged dependent variable and network exposure, as different levels of influence might affect the performance of estimators.

Specifically, let a simulation configuration be as follow: (i) in each simulation we fix each actor's unobserved trait c to be a constant drawn from a normal distribution N(0,1); (ii) we vary number of time points to be 3 or 6. And the simulated data used for estimation is collected at time 18–20 and 15–20 respectively, which is less dependent on the initial condition; (iii) we vary number of nodes to be 40 or 80; (iv) we keep the homophily level to be high, and the average correlation between the lagged dependent variable term and the network exposure term across simulations to be around 0.4 (Penuel et al., 2012; Venkatesh and Morris, 2000) ($\alpha_0 = -0.4$, $\alpha_1 = -0.45$ for N = 40; $\alpha_0 = -0.55$, $\alpha_1 = -0.3$ for N = 80); (v) same as before we vary both β_1 and β_2 from 0.1 to 0.9, while keeping $\beta_1 + \beta_2 = 0.9$.

In each configuration, we start from a random network and simulate a network data based on model 13 and 14, and then use the proposed estimation methods to get estimates for β_1 and β_2.[13] Furthermore, we include fixed and random effects estima-

[13] In T = 3 cases, for IV estimation we use 2SLS with behavioral outcomes from first time point as instrumental, which is shown to be the same as GMM-IV estimation when the model is just-identified (Wooldridge, 2010, 2002).

Table 2b
Simulation results for network exposure.

Network Exposure											
		LSDV		Random Effects		GMM		SEM		Latent Space	
		Mean Bias	SD	Mean Bias	SD	Mean Bias	SD	Mean Bias	SD	Mean Bias	SD
N = 40,T = 6	β2 = 0	0.003	0.042	0.065	0.042	0.002	0.049	0.006	0.052	−0.001	0.049
	β2 = 0.1	−0.019	0.073	0.084	0.069	−0.012	0.085	0.011	0.089	0.010	0.082
	β2 = 0.2	−0.037	0.101	0.113	0.086	−0.011	0.113	0.013	0.114	−0.036	0.094
	β2 = 0.3	−0.034	0.100	0.144	0.095	0.020	0.133	0.034	0.124	−0.017	0.108
	β2 = 0.4	−0.043	0.130	0.179	0.115	0.023	0.163	0.035	0.152	−0.013	0.116
	β2 = 0.5	−0.049	0.137	0.207	0.125	0.031	0.177	0.034	0.153	0.001	0.127
	β2 = 0.6	−0.059	0.150	0.217	0.145	0.039	0.188	0.028	0.168	−0.047	0.147
	β2 = 0.7	−0.053	0.156	0.230	0.151	0.054	0.185	0.030	0.172	−0.034	0.178
	β2 = 0.8	−0.056	0.166	0.233	0.168	0.053	0.212	0.028	0.186	−0.041	0.155
	β2 = 0.9	−0.057	0.171	0.218	0.170	0.054	0.214	0.032	0.191	−0.045	0.148
N = 80,T = 6	β2 = 0	0.001	0.046	0.096	0.044	0.004	0.052	0.005	0.053	0.008	0.049
	β2 = 0.1	−0.019	0.071	0.133	0.064	−0.002	0.085	0.013	0.082	0.013	0.063
	β2 = 0.2	−0.019	0.098	0.204	0.086	0.022	0.119	0.033	0.112	0.042	0.088
	β2 = 0.3	−0.035	0.112	0.253	0.100	0.026	0.139	0.027	0.130	0.009	0.096
	β2 = 0.4	−0.039	0.130	0.310	0.116	0.036	0.169	0.036	0.148	0.005	0.131
	β2 = 0.5	−0.050	0.155	0.357	0.144	0.042	0.199	0.035	0.177	0.021	0.157
	β2 = 0.6	−0.059	0.154	0.375	0.151	0.050	0.207	0.031	0.181	−0.038	0.142
	β2 = 0.7	−0.066	0.170	0.382	0.173	0.053	0.217	0.020	0.192	−0.056	0.144
	β2 = 0.8	−0.065	0.177	0.387	0.191	0.055	0.229	0.025	0.202	−0.051	0.183
	β2 = 0.9	−0.055	0.178	0.382	0.205	0.072	0.235	0.047	0.207	−0.016	0.151
N = 40,T = 3	β2 = 0	0.004	0.085	0.060	0.072	0.002	0.186	0.006	0.238	0.004	0.067
	β2 = 0.1	−0.047	0.128	0.092	0.116	−0.026	0.299	0.058	0.408	0.027	0.123
	β2 = 0.2	−0.086	0.167	0.125	0.141	0.004	0.364	0.030	0.415	0.046	0.156
	β2 = 0.3	−0.102	0.190	0.179	0.160	0.079	0.403	0.087	0.392	0.039	0.196
	β2 = 0.4	−0.139	0.241	0.209	0.180	0.091	0.422	0.079	0.386	0.022	0.184
	β2 = 0.5	−0.172	0.261	0.238	0.211	0.072	0.444	0.069	0.435	−0.011	0.222
	β2 = 0.6	−0.179	0.288	0.255	0.227	0.085	0.460	0.093	0.418	−0.023	0.212
	β2 = 0.7	−0.214	0.297	0.283	0.248	0.086	0.426	0.102	0.430	0.052	0.244
	β2 = 0.8	−0.202	0.301	0.297	0.260	0.096	0.437	0.104	0.429	−0.045	0.231
	β2 = 0.9	−0.210	0.339	0.293	0.280	0.078	0.423	0.084	0.418	−0.041	0.283
N = 80,T = 3	β2 = 0	0.000	0.075	0.089	0.067	0.000	0.158	0.001	0.207	0.009	0.075
	β2 = 0.1	−0.044	0.122	0.129	0.103	0.013	0.276	0.047	0.283	0.035	0.118
	β2 = 0.2	−0.085	0.162	0.196	0.140	0.028	0.277	0.041	0.325	0.067	0.159
	β2 = 0.3	−0.109	0.190	0.288	0.171	0.053	0.370	0.051	0.362	0.037	0.178
	β2 = 0.4	−0.140	0.226	0.338	0.196	0.060	0.354	0.060	0.324	0.060	0.196
	β2 = 0.5	−0.200	0.259	0.370	0.216	0.041	0.390	0.046	0.378	0.059	0.269
	β2 = 0.6	−0.179	0.296	0.436	0.231	0.078	0.383	0.083	0.382	−0.039	0.245
	β2 = 0.7	−0.204	0.318	0.466	0.276	0.086	0.429	0.087	0.421	0.080	0.288
	β2 = 0.8	−0.210	0.323	0.488	0.271	0.121	0.423	0.127	0.422	0.024	0.325
	β2 = 0.9	−0.238	0.353	0.465	0.296	0.070	0.437	0.072	0.433	0.049	0.298

tor as baseline comparisons.[14] That is, we model the individual specific time-invariant trait c as either random or fixed effect, as estimates from these two models most likely will provide upper or lower bound for the bias (Nickell, 1981). Note that we use a least square dummy variable (LSDV) approach to estimate the fixed effects model (Wooldridge, 2010), which we refer to as "LSDV" below. Other model configurations include $\beta_3 = 0.1$ (to keep consensus within the initial range of c_i (Friedkin and Johnsen, 1999)), $e_{it} \sim N(0, 0.2^2)$, and initial network density = 0.2. In each configuration, each model is simulated 500 times, except for latent-space adjusted approach, which are simulated 100 times each due to their algorithm complexity and longer running time.

We perform our simulation experiments in R 3.3.2. LSDV model is estimated using "lm" functions; random-effects model is estimated using "lmer" function in "lme4" package (Bates et al., 2015); GMM is estimated using "pgmm" function in "plm" package (Croissant and Millo, 2008); SEM is estimated using "sem" function in "lavaan" package (Rosseel, 2012); latent space model is estimated using "ergmm" function in "latentnet" package (Krivitsky and Handcock, 2008; Krivitsky and Handcock, 2015).

The mean biases and standard deviation of the estimates for lagged dependent variable and network exposure are shown in Tables 2a and 2b respectively (See also Fig. S1a and Fig. S1b in the Appendix e in Supplementary material). For the estimates of the lagged dependent variable: (1) the magnitude of bias is smaller for all estimation methods when we include more time points (bigger T), while it is not affected much by increasing node sizes (bigger N); (2) random effects estimates are always positively biased and fixed effects estimates are always negatively biased, which is as expected and consistent with Nickell (1981); and (3) GMM-IV estimates exhibit small bias when the true coefficient is small, but as the true coefficient of the lagged dependent variable increases, we will be more likely to encounter weak instrument problems, and the magnitude of bias thus increases (Wooldridge, 2010; Arellano and Bond, 1991). Also when T = 3, GMM-IV estimates are more unstable with larger variance (we exclude several outliers in this case). (4) SEM estimates are one of the least biased estimates when T is large; however the bias is much larger when T is small and true coefficient is large. (5) Latent space estimates are the least biased estimates in all cases, out-performing other estimates. The standard deviation of the latent space-estimates is also much smaller than SEM and GMM estimates when T is small.

[14] We did not include results from SIENA models here, because the coefficients from SIENA models are not directly comparable with the methods studied in this paper, and the treatment of the time unit is different. That is, while we assume time is discrete, SIENA assumes time is continuous.

Table 3a
Simulation results for lagged dependent variable with inclusion of lagged transitivity.

Lagged Dependent Variable											
		LSDV		Random Effects		GMM		SEM		Latent Space	
		Mean Bias	SD	Mean Bias	SD	Mean Bias	SD	Mean Bias	SD	Mean Bias	SD
N = 40,T = 6	β1 = 0	−0.171	0.07	0.09	0.102	−0.021	0.1	0.011	0.092	0.034	0.078
	β1 = 0.1	−0.197	0.071	0.116	0.106	−0.025	0.109	0.013	0.098	0.024	0.068
	β1 = 0.2	−0.227	0.069	0.135	0.097	−0.046	0.109	0.006	0.098	0.021	0.074
	β1 = 0.3	−0.253	0.071	0.163	0.084	−0.047	0.119	0.016	0.112	0.041	0.071
	β1 = 0.4	−0.287	0.072	0.17	0.073	−0.069	0.144	0.017	0.138	0.028	0.076
	β1 = 0.5	−0.312	0.074	0.173	0.061	−0.076	0.165	0.043	0.161	0.044	0.065
	β1 = 0.6	−0.353	0.077	0.155	0.044	−0.128	0.175	0.038	0.172	0.041	0.068
	β1 = 0.7	−0.388	0.077	0.141	0.032	−0.175	0.22	0.051	0.184	0.038	0.054
	β1 = 0.8	−0.42	0.08	0.113	0.025	−0.318	0.276	0.046	0.206	0.041	0.048
	β1 = 0.9	−0.416	0.076	0.068	0.017	−0.458	0.345	0.012	0.186	0.034	0.022
N = 80,T = 6	β1 = 0	−0.182	0.048	0.113	0.077	−0.009	0.074	0.006	0.063	0.027	0.059
	β1 = 0.1	−0.208	0.047	0.139	0.075	−0.02	0.075	0.004	0.064	0.025	0.051
	β1 = 0.2	−0.234	0.051	0.161	0.065	−0.02	0.08	0.005	0.069	0.023	0.052
	β1 = 0.3	−0.261	0.053	0.176	0.052	−0.023	0.096	0.007	0.076	0.026	0.054
	β1 = 0.4	−0.292	0.051	0.176	0.038	−0.032	0.1	0.011	0.082	0.026	0.044
	β1 = 0.5	−0.318	0.055	0.17	0.033	−0.047	0.113	0.018	0.094	0.04	0.046
	β1 = 0.6	−0.356	0.05	0.152	0.031	−0.076	0.128	0.013	0.109	0.028	0.043
	β1 = 0.7	−0.39	0.057	0.134	0.023	−0.105	0.156	0.028	0.128	0.035	0.04
	β1 = 0.8	−0.42	0.057	0.107	0.019	−0.188	0.222	0.029	0.147	0.032	0.028
	β1 = 0.9	−0.409	0.054	0.064	0.013	−0.337	0.313	0.044	0.135	0.017	0.018
N = 40,T = 3	β1 = 0	−0.446	0.137	0.149	0.128	0.034	0.293	0.039	0.263	0.092	0.119
	β1 = 0.1	−0.508	0.14	0.154	0.122	0.016	0.321	0.027	0.27	0.086	0.117
	β1 = 0.2	−0.552	0.145	0.175	0.099	0.061	0.409	0.069	0.369	0.096	0.099
	β1 = 0.3	−0.616	0.15	0.177	0.1	0.05	0.467	0.114	0.642	0.094	0.116
	β1 = 0.4	−0.666	0.149	0.179	0.09	0.066	0.553	0.155	0.922	0.121	0.107
	β1 = 0.5	−0.735	0.166	0.171	0.072	0.129	0.733	0.201	1.082	0.124	0.086
	β1 = 0.6	−0.773	0.165	0.157	0.069	0.136	1.109	0.437	1.839	0.111	0.084
	β1 = 0.7	−0.831	0.162	0.142	0.054	−0.036	1.348	0.535	2.04	0.103	0.074
	β1 = 0.8	−0.89	0.171	0.113	0.039	−0.445	1.724	0.356	2.438	0.095	0.049
	β1 = 0.9	−0.911	0.164	0.07	0.026	−0.797	2.546	0.881	2.647	0.051	0.035
N = 80,T = 3	β1 = 0	−0.463	0.091	0.172	0.081	0.016	0.178	0.026	0.166	0.077	0.093
	β1 = 0.1	−0.515	0.092	0.179	0.08	0.015	0.216	0.021	0.185	0.065	0.092
	β1 = 0.2	−0.564	0.106	0.186	0.072	0.037	0.248	0.045	0.214	0.077	0.09
	β1 = 0.3	−0.624	0.096	0.18	0.066	0.033	0.334	0.044	0.262	0.09	0.074
	β1 = 0.4	−0.691	0.107	0.173	0.06	0.064	0.387	0.064	0.321	0.08	0.074
	β1 = 0.5	−0.725	0.117	0.165	0.056	0.064	0.453	0.07	0.37	0.094	0.063
	β1 = 0.6	−0.789	0.11	0.155	0.045	0.141	0.762	0.136	0.679	0.089	0.069
	β1 = 0.7	−0.835	0.113	0.133	0.039	0.101	0.918	0.207	1.041	0.085	0.054
	β1 = 0.8	−0.886	0.116	0.109	0.029	−0.136	1.442	0.529	1.686	0.074	0.038
	β1 = 0.9	−0.914	0.114	0.066	0.019	−0.469	2.141	0.614	2.005	0.041	0.03

For estimates of the network exposure term: (1) same as above, the magnitude of bias is smaller for all estimation methods when we include more time points (bigger T), while the bias is not decreasing when we increase node size (bigger N); (2) random effects estimates are always positively biased and fixed effects estimates are always negatively biased, although fixed effects estimates for the exposure term produce much smaller bias than fixed effects estimates for the lagged dependent variable term, especially when T is large; (3) GMM-IV estimates overall have small bias, though they are much more unstable when T is small; (4) SEM and latent space estimates out-perform other estimates in terms of producing the smallest bias across all cases. However, the standard deviation for latent space estimates are much smaller than GMM and SEM estimates when T is small.

Overall, random-effects and fixed-effect estimates produce the biggest bias among all estimators; GMM-IV estimates sometimes produce small bias but are generally unstable, and it can be largely biased when T is small and the true coefficient of the lagged dependent variable is large; SEM estimates produce small bias in estimating contagion effects, but SEM estimates for the lagged dependent variable term can be severely biased when T is small and true coefficient is large; latent space adjusted estimators outperform other estimators in terms of producing the smallest bias across most cases, and it is also considerably more efficient than GMM and SEM estimates, especially when T is small. However, one should note that and the latent space method is relatively time consuming, as it uses a simulation based estimation method (MCMC) and requires lots of burn-in time (Hoff et al., 2002).

5. Robustness test

In this section we perform some robustness tests to determine whether the estimation methods we proposed can still recover the true contagion effects under alternative scenarios.

5.1. Including transitivity in network selection

One potential limitation of our main selection model is that it does not involve any network dependence beyond dyads in the selection process. Network dependence involving more than two nodes, such as transitivity, are also shown to be fundamental properties of the network (Simmel and Wolff, 1950; Newman and Park, 2003; Rivera et al., 2010). Although these types of network dependencies are important for the selection process, ignoring them do not necessarily bias the estimation in the influence process. This is because these network dependencies usually do not directly affect behavior. And as long as all the factors that directly affect behav-

Table 3b
Simulation results for network exposure with inclusion of lagged transitivity.

Network Exposure											
		LSDV		Random Effects		GMM		SEM		Latent Space	
		Mean Bias	SD	Mean Bias	SD	Mean Bias	SD	Mean Bias	SD	Mean Bias	SD
N = 40,T = 6	β2 = 0	0.001	0.046	0.059	0.047	0.004	0.048	0.007	0.054	0.015	0.051
	β2 = 0.1	−0.012	0.074	0.078	0.068	0	0.085	0.022	0.087	0.026	0.059
	β2 = 0.2	−0.036	0.091	0.104	0.087	−0.001	0.118	0.019	0.111	0.015	0.098
	β2 = 0.3	−0.028	0.121	0.14	0.112	0.016	0.149	0.039	0.137	0.05	0.121
	β2 = 0.4	−0.044	0.134	0.15	0.132	0.022	0.166	0.029	0.159	0.014	0.129
	β2 = 0.5	−0.054	0.141	0.174	0.134	0.022	0.172	0.024	0.164	0.013	0.141
	β2 = 0.6	−0.048	0.153	0.195	0.148	0.042	0.19	0.035	0.17	0.025	0.15
	β2 = 0.7	−0.054	0.158	0.187	0.166	0.04	0.199	0.023	0.179	0.013	0.147
	β2 = 0.8	−0.067	0.171	0.171	0.174	0.036	0.22	0.012	0.192	0.014	0.173
	β2 = 0.9	−0.061	0.161	0.166	0.163	0.053	0.205	0.016	0.181	−0.027	0.162
N = 80,T = 6	β2 = 0	0.001	0.038	0.087	0.036	0.003	0.043	0.004	0.043	0.019	0.035
	β2 = 0.1	−0.016	0.065	0.118	0.06	−0.005	0.079	0.008	0.076	0.025	0.068
	β2 = 0.2	−0.036	0.088	0.162	0.079	0.003	0.108	0.012	0.097	0.033	0.095
	β2 = 0.3	−0.025	0.105	0.23	0.094	0.025	0.124	0.032	0.115	0.033	0.111
	β2 = 0.4	−0.045	0.119	0.264	0.112	0.02	0.215	0.024	0.181	0.022	0.105
	β2 = 0.5	−0.05	0.123	0.301	0.116	0.033	0.152	0.022	0.135	0.029	0.118
	β2 = 0.6	−0.055	0.131	0.324	0.119	0.028	0.17	0.014	0.143	0.015	0.138
	β2 = 0.7	−0.055	0.139	0.339	0.148	0.036	0.173	0.018	0.149	0.022	0.145
	β2 = 0.8	−0.061	0.143	0.33	0.15	0.035	0.178	0.009	0.162	−0.002	0.14
	β2 = 0.9	−0.048	0.148	0.331	0.158	0.059	0.181	0.026	0.163	0.013	0.154
N = 40,T = 3	β2 = 0	0.006	0.081	0.059	0.073	−0.002	0.197	0.014	0.281	0.04	0.082
	β2 = 0.1	−0.039	0.13	0.082	0.106	−0.007	0.3	0.059	0.443	0.054	0.124
	β2 = 0.2	−0.083	0.164	0.116	0.147	−0.011	0.378	0.078	0.54	0.044	0.152
	β2 = 0.3	−0.118	0.204	0.152	0.185	0.02	0.445	0.118	0.714	0.06	0.21
	β2 = 0.4	−0.138	0.229	0.182	0.197	0.078	0.424	0.081	0.448	0.096	0.207
	β2 = 0.5	−0.177	0.266	0.214	0.229	0.051	0.401	0.103	0.551	0.122	0.26
	β2 = 0.6	−0.184	0.285	0.218	0.249	0.065	0.441	0.074	0.418	0.051	0.238
	β2 = 0.7	−0.195	0.315	0.242	0.257	0.078	0.494	0.098	0.506	0.045	0.266
	β2 = 0.8	−0.219	0.34	0.256	0.279	0.073	0.437	0.074	0.424	0.054	0.287
	β2 = 0.9	−0.224	0.341	0.235	0.278	0.07	0.421	0.07	0.407	0.085	0.274
N = 80,T = 3	β2 = 0	0.004	0.067	0.082	0.064	−0.004	0.161	−0.005	0.191	0.043	0.075
	β2 = 0.1	−0.042	0.105	0.121	0.096	−0.02	0.26	0.039	0.258	0.043	0.108
	β2 = 0.2	−0.082	0.137	0.178	0.122	0.026	0.281	0.063	0.37	0.094	0.138
	β2 = 0.3	−0.128	0.173	0.237	0.139	0.049	0.305	0.035	0.259	0.085	0.168
	β2 = 0.4	−0.146	0.208	0.287	0.183	0.041	0.288	0.045	0.278	0.106	0.2
	β2 = 0.5	−0.177	0.235	0.335	0.193	0.053	0.346	0.047	0.324	0.094	0.233
	β2 = 0.6	−0.187	0.24	0.369	0.203	0.053	0.339	0.056	0.326	0.104	0.239
	β2 = 0.7	−0.204	0.276	0.383	0.228	0.067	0.352	0.07	0.353	0.123	0.241
	β2 = 0.8	−0.218	0.289	0.404	0.247	0.054	0.371	0.056	0.366	0.111	0.217
	β2 = 0.9	−0.231	0.302	0.409	0.251	0.045	0.36	0.049	0.358	0.096	0.269

ior have been accounted for in the influence model, the contagion effects can still be identified.

However, one exception is that ignoring network dependencies such as transitivity will likely to bias the estimation of contagion effects using latent-space adjusted approach. Latent space model is shown to capture transitivity in a natural way (Hoff et al., 2002; Rastelli et al., 2015). When two actors have a neighbor in common, it is to be expected that the three corresponding nodes will be close in the latent space, making triangles more likely. And when transitivity is ignored in the latent space model, it is likely to confound with unobserved trait that co-determines influence and selection, as a result the estimated latent social position will be poor proxies of the true unobserved trait, thus biasing the estimation of contagion effects in the influence model.

In this section we include transitivity in the selection process, and test whether the models we proposed can still recover the true contagion effects. Specifically, let the "true" influence model be:

$$Y_{it} = \beta_1 Y_{it-1} + \beta_2 \frac{\sum Z_{ijt-1} Y_{jt-1}}{\sum Z_{ijt-1}} + \beta_3 c_i + e_{it} \quad (15)$$

Correspondingly let the "true" selection model be represented as

$$P\left(Z_{ijt} = 1\right) = \Phi(\alpha_0 + \alpha_1 |c_i - c_j| + \alpha_2 NCommonfriends_{ijt-1}) \quad (16)$$

All the notations are same as above. In the selection process we represent transitivity using $NCommonfriends_{ijt-1}$–the number of common friends i and j have at time t − 1.[15] And we also include this variable as a covariate when estimating the latent space model. All other simulation configurations are the same as in the main experiment, except that we change the value of α_0, α_1, α_2 to keep the density of the network at the same level as in the main experiment ($\alpha_0 = -1$, $\alpha_1 = -0.4$, $\alpha_2 = 0.35$ for N = 40; $\alpha_0 = -1.15$, $\alpha_1 = -0.35$, $\alpha_2 = 0.35$ for N = 80).

Mean bias and standard error for estimates of lagged dependent variable and network exposure term are shown in Tables 3a and 3b respectively (See also Fig. S2a and Fig. S2b in the Appendix e in Supplementary material). Results are generally consistent with the main experiment, that is: (1) in general fixed and random effects estimates generate largest bias, both in estimating lagged dependent variable and network exposure term; (2) GMM and SEM estimates generate large bias in estimating lagged dependent variable when T is small and true coefficient of lagged dependent variable is close to 1, while latent-space adjusted estimates generate the smallest bias in estimating the lagged dependent variable under most scenarios; (3) All three estimators proposed

[15] The variable is normalized to have mean 0 and standard deviation of 1.

Table 4a
Simulation results for lagged dependent variable under misspecification.

Lagged Dependent Variable											
		LSDV		Random Effects		GMM		SEM		Latent Space	
		Mean Bias	SD	Mean Bias	SD	Mean Bias	SD	Mean Bias	SD	Mean Bias	SD
N = 40,T = 6	β1 = 0	−0.182	0.065	−0.027	0.070	−0.017	0.090	0.008	0.090	−0.037	0.072
	β1 = 0.1	−0.210	0.069	−0.032	0.071	−0.027	0.089	0.005	0.092	−0.022	0.068
	β1 = 0.2	−0.236	0.071	−0.032	0.075	−0.025	0.099	0.011	0.102	−0.041	0.073
	β1 = 0.3	−0.269	0.075	−0.038	0.077	−0.035	0.106	0.010	0.118	−0.034	0.064
	β1 = 0.4	−0.290	0.071	−0.039	0.073	−0.038	0.109	0.022	0.126	−0.027	0.057
	β1 = 0.5	−0.331	0.080	−0.046	0.077	−0.053	0.121	0.017	0.143	−0.038	0.072
	β1 = 0.6	−0.359	0.080	−0.040	0.073	−0.061	0.130	0.030	0.159	−0.035	0.065
	β1 = 0.7	−0.392	0.077	−0.038	0.064	−0.085	0.142	0.031	0.169	−0.035	0.058
	β1 = 0.8	−0.422	0.079	−0.029	0.056	−0.101	0.170	0.030	0.186	−0.025	0.044
	β1 = 0.9	−0.432	0.084	−0.008	0.027	−0.116	0.175	0.007	0.183	−0.009	0.028
N = 80,T = 6	β1 = 0	−0.194	0.050	−0.020	0.054	−0.009	0.067	0.001	0.066	−0.008	0.051
	β1 = 0.1	−0.213	0.048	−0.015	0.054	−0.007	0.069	0.009	0.065	−0.011	0.054
	β1 = 0.2	−0.241	0.050	−0.022	0.054	−0.013	0.074	0.005	0.070	−0.019	0.049
	β1 = 0.3	−0.269	0.051	−0.021	0.052	−0.013	0.076	0.009	0.075	−0.015	0.045
	β1 = 0.4	−0.298	0.051	−0.026	0.052	−0.021	0.078	0.007	0.080	−0.019	0.050
	β1 = 0.5	−0.326	0.053	−0.026	0.054	−0.022	0.086	0.009	0.091	−0.013	0.049
	β1 = 0.6	−0.360	0.052	−0.025	0.050	−0.034	0.092	0.008	0.099	−0.015	0.038
	β1 = 0.7	−0.394	0.054	−0.024	0.047	−0.047	0.103	0.014	0.114	−0.013	0.038
	β1 = 0.8	−0.425	0.053	−0.016	0.036	−0.054	0.121	0.021	0.135	−0.012	0.034
	β1 = 0.9	−0.428	0.054	−0.005	0.018	−0.065	0.115	0.007	0.125	0.000	0.018
N = 40,T = 3	β1 = 0	−0.460	0.142	−0.043	0.121	0.035	0.234	0.044	0.233	−0.007	0.105
	β1 = 0.1	−0.506	0.144	−0.046	0.125	0.051	0.268	0.062	0.279	−0.032	0.108
	β1 = 0.2	−0.572	0.155	−0.051	0.126	0.042	0.311	0.058	0.341	−0.019	0.124
	β1 = 0.3	−0.625	0.152	−0.057	0.131	0.050	0.339	0.087	0.440	−0.024	0.102
	β1 = 0.4	−0.706	0.162	−0.050	0.135	0.043	0.431	0.081	0.766	−0.015	0.110
	β1 = 0.5	−0.741	0.161	−0.045	0.123	0.088	0.491	0.153	0.694	−0.027	0.114
	β1 = 0.6	−0.791	0.160	−0.041	0.114	0.106	0.593	0.225	1.114	−0.015	0.088
	β1 = 0.7	−0.847	0.165	−0.037	0.096	0.128	0.726	0.257	2.015	−0.028	0.078
	β1 = 0.8	−0.896	0.153	−0.023	0.074	0.207	1.035	0.456	2.522	−0.011	0.068
	β1 = 0.9	−0.925	0.163	−0.003	0.046	0.101	1.277	0.474	2.497	−0.003	0.048
N = 80,T = 3	β1 = 0	−0.489	0.100	−0.032	0.083	0.015	0.164	0.019	0.163	−0.009	0.079
	β1 = 0.1	−0.535	0.097	−0.033	0.090	0.022	0.177	0.026	0.179	−0.015	0.071
	β1 = 0.2	−0.589	0.107	−0.036	0.089	0.025	0.214	0.029	0.222	−0.009	0.074
	β1 = 0.3	−0.635	0.105	−0.031	0.088	0.047	0.247	0.050	0.241	−0.002	0.076
	β1 = 0.4	−0.695	0.109	−0.034	0.086	0.039	0.263	0.048	0.265	−0.011	0.071
	β1 = 0.5	−0.742	0.112	−0.034	0.080	0.007	0.294	0.015	0.312	−0.004	0.068
	β1 = 0.6	−0.788	0.112	−0.027	0.070	0.059	0.330	0.067	0.338	−0.012	0.068
	β1 = 0.7	−0.855	0.107	−0.019	0.061	0.045	0.480	0.057	0.514	−0.007	0.047
	β1 = 0.8	−0.891	0.113	−0.012	0.049	0.109	0.578	0.199	0.968	−0.001	0.042
	β1 = 0.9	−0.932	0.110	−0.004	0.029	0.162	0.733	0.420	1.730	−0.003	0.030

(GMM, SEM, Latent-space adjusted) generate small bias in estimating network exposure term across most of the scenarios. While latent-space adjusted estimate of network exposure term is slightly more biased than GMM and SEM when T is small and N is large, the difference is not significant comparing to the size of the standard error; (4) the standard deviation of Latent-space adjusted estimates is much smaller than GMM and SEM estimates for both lagged dependent variable and network exposure, especially when T is small.

Overall, when transitivity is included in the network selection process, the estimates proposed are not more biased than the main experiment, indicating that the methods proposed are robust to network dependencies beyond dyads. Specifically, after including transitivity as covariates, the latent space model can still generate good proxies for the unobserved trait and recover the true contagion effects.

5.2. Misspecification: no unobserved trait

While the methods proposed exhibit good properties in recovering the true coefficients in the influence model, it is also important to test the performance of the proposed methods when the model is misspecified. Here we explore one scenario, where we estimate our proposed models while there is no unobserved trait co-determining influence and selection in the data generating process.

Specifically, let the "true" influence model be

$$Y_{it} = \beta_1 Y_{it-1} + \beta_2 \frac{\sum Z_{ijt-1} Y_{jt-1}}{\sum Z_{ijt-1}} + \beta_3 c_i + e_{it} \quad (17)$$

Correspondingly, let the "true" selection model be represented as

$$P\left(Z_{ijt} = 1\right) = \Phi(\alpha_0 + \alpha_1 NCommonfriends_{ijt-1}) \quad (18)$$

All the notations are same as above. Note that in the true data generating process we exclude the unobserved trait that co-determines influence and selection. The selection process is only driven by transitivity – the number of friends i and j have in common at time t − 1. All other simulation configurations are the same as in the main experiment, except that we change the value of α_0, α_1, to keep the density of the network at the same level as in the main experiment ($\alpha_0 = -1.6$, $\alpha_1 = 0.4$ for N = 40; $\alpha_0 = -1.4$, $\alpha_1 = 0.2$ for N = 80). And we test the performance of the proposed methods by estimating the exact same models in the main experiment as if there were unobserved trait co-determining influence and selection.

Table 4b
Simulation results for network exposure under misspecification.

Network Exposure											
		LSDV		Random Effects		GMM		SEM		Latent Space	
		Mean Bias	SD	Mean Bias	SD	Mean Bias	SD	Mean Bias	SD	Mean Bias	SD
N = 40,T = 6	β2 = 0	0.006	0.067	−0.004	0.068	0.014	0.090	0.017	0.080	−0.010	0.060
	β2 = 0.1	−0.016	0.113	−0.019	0.114	0.023	0.143	0.031	0.135	−0.012	0.116
	β2 = 0.2	−0.029	0.139	−0.032	0.137	0.037	0.173	0.044	0.174	−0.070	0.133
	β2 = 0.3	−0.037	0.158	−0.035	0.145	0.045	0.202	0.060	0.187	−0.034	0.141
	β2 = 0.4	−0.033	0.172	−0.028	0.167	0.057	0.225	0.057	0.204	−0.031	0.144
	β2 = 0.5	−0.052	0.177	−0.051	0.171	0.063	0.230	0.047	0.205	−0.071	0.172
	β2 = 0.6	−0.063	0.181	−0.054	0.170	0.067	0.238	0.037	0.216	−0.042	0.174
	β2 = 0.7	−0.057	0.184	−0.047	0.172	0.081	0.248	0.049	0.220	−0.058	0.174
	β2 = 0.8	−0.062	0.191	−0.060	0.184	0.071	0.239	0.041	0.220	−0.047	0.158
	β2 = 0.9	−0.063	0.195	−0.055	0.177	0.082	0.235	0.036	0.218	−0.069	0.193
N = 80,T = 6	β2 = 0	−0.001	0.073	−0.013	0.077	0.010	0.098	0.012	0.089	−0.018	0.073
	β2 = 0.1	−0.016	0.130	−0.028	0.127	0.036	0.168	0.036	0.152	−0.048	0.117
	β2 = 0.2	−0.024	0.151	−0.031	0.151	0.060	0.199	0.055	0.177	−0.046	0.151
	β2 = 0.3	−0.043	0.170	−0.046	0.162	0.078	0.232	0.061	0.205	−0.056	0.177
	β2 = 0.4	−0.043	0.167	−0.047	0.165	0.098	0.225	0.072	0.201	−0.062	0.168
	β2 = 0.5	−0.048	0.202	−0.054	0.185	0.100	0.262	0.068	0.233	−0.089	0.202
	β2 = 0.6	−0.063	0.192	−0.065	0.178	0.101	0.262	0.058	0.230	−0.086	0.175
	β2 = 0.7	−0.070	0.225	−0.066	0.211	0.094	0.292	0.053	0.266	−0.104	0.227
	β2 = 0.8	−0.071	0.212	−0.063	0.195	0.092	0.280	0.048	0.248	−0.089	0.170
	β2 = 0.9	−0.080	0.210	−0.077	0.194	0.116	0.282	0.043	0.244	−0.084	0.203
N = 40,T = 3	β2 = 0	−0.003	0.117	−0.010	0.112	0.027	0.250	0.019	0.319	−0.012	0.120
	β2 = 0.1	−0.031	0.197	−0.011	0.191	0.055	0.378	0.070	0.822	−0.025	0.190
	β2 = 0.2	−0.090	0.246	−0.028	0.233	0.060	0.400	0.136	0.713	−0.066	0.210
	β2 = 0.3	−0.112	0.274	−0.032	0.257	0.108	0.419	0.139	0.524	−0.054	0.255
	β2 = 0.4	−0.152	0.312	−0.043	0.276	0.085	0.474	0.082	0.634	−0.059	0.296
	β2 = 0.5	−0.164	0.332	−0.055	0.281	0.075	0.491	0.083	0.545	−0.058	0.282
	β2 = 0.6	−0.183	0.337	−0.043	0.285	0.141	0.498	0.147	0.497	−0.043	0.289
	β2 = 0.7	−0.220	0.337	−0.057	0.283	0.084	0.463	0.090	0.469	−0.099	0.290
	β2 = 0.8	−0.224	0.355	−0.030	0.297	0.114	0.486	0.120	0.480	−0.057	0.307
	β2 = 0.9	−0.199	0.385	−0.034	0.315	0.127	0.484	0.130	0.482	−0.042	0.324
N = 80,T = 3	β2 = 0	0.004	0.126	−0.008	0.129	0.021	0.232	0.039	0.344	−0.002	0.128
	β2 = 0.1	−0.045	0.219	−0.017	0.203	0.074	0.359	0.078	0.365	0.022	0.205
	β2 = 0.2	−0.092	0.265	−0.055	0.262	0.047	0.414	0.046	0.417	−0.048	0.250
	β2 = 0.3	−0.122	0.322	−0.031	0.281	0.090	0.436	0.091	0.439	−0.026	0.254
	β2 = 0.4	−0.159	0.330	−0.050	0.313	0.059	0.462	0.060	0.464	−0.076	0.330
	β2 = 0.5	−0.171	0.345	−0.034	0.328	0.131	0.484	0.135	0.487	−0.016	0.350
	β2 = 0.6	−0.176	0.370	−0.051	0.333	0.135	0.507	0.136	0.507	−0.095	0.363
	β2 = 0.7	−0.255	0.404	−0.070	0.345	0.078	0.526	0.078	0.526	−0.097	0.379
	β2 = 0.8	−0.237	0.387	−0.063	0.326	0.115	0.487	0.117	0.489	−0.072	0.342
	β2 = 0.9	−0.240	0.400	−0.047	0.324	0.121	0.485	0.124	0.486	−0.064	0.327

Mean bias and standard error for estimates of lagged dependent variable and network exposure term are shown in Tables 4a and 4b respectively (see also Fig. S3a and Fig. S3b in the Appendix e in Supplementary material). Results exhibit the same pattern as in the main experiment, except that random effects estimates generate smallest bias in estimating both lagged dependent variable and network exposure, comparing with other estimates. This is because all variables affecting behavior have now been accounted for in the random effects model when there is no unobserved trait that co-determines influence and selection. GMM, SEM and latent space adjusted approach estimates for lagged dependent variable did not exhibit more bias than in the main experiments. However, all three estimates for network exposure are slightly more biased than the estimates in the main experiments, especially when T is large. But again the difference is small compared with the size of the standard error. This shows that although modeling unobserved trait when it does not exist will cause some bias in the estimation, the magnitude of the bias is generally small. A possible reason is that although latent space adjusted approach estimates the unobserved trait when it does not exist, the estimated unobserved trait does not strongly correlate with the lagged dependent variable or the network exposure term, as a result the estimates are not affected much by inclusion of the "false" unobserved latent trait in the influence model.

5.3. Including covariates

We also test the robustness of the results when we include covariates in both influence and selection models. Specifically, in the data generating process, let the influence and selection process be

$$Y_{it} = \beta_1 Y_{it-1} + \beta_2 \frac{\sum Z_{ijt-1} Y_{jt-1}}{\sum Z_{ijt-1}} + \beta_3 c_i + \beta_4 X_i + e_{it} \quad (19)$$

Correspondingly, let the "true" selection model be represented as

$$P\left(Z_{ijt} = 1\right) = \Phi(\alpha_0 + \alpha_1 |c_i - c_j| + \alpha_2 |X_i - X_j|) \quad (20)$$

All notations are the same as previously used, and X is an observed time invariant attribute that follows a N(0,1) distribution. All simulation setups are as before, except that we include X when estimating each proposed model.

Mean bias and standard error for estimates of lagged dependent variable and network exposure term are shown in Tables 5a and 5b respectively (See also Fig. S4a and Fig. S4b in the Appendix e in Supplementary material). Results are consistent with the main experiment. GMM and SEM estimates generate large bias in estimating lagged dependent variable when T is small and true coefficient of lagged dependent variable is large. All three estima-

Table 5a
Simulation results for lagged dependent variable with inclusion of covariates.

Lagged Dependent Variable											
		LSDV		Random Effects		GMM		SEM		Latent Space	
		Mean Bias	SD	Mean Bias	SD	Mean Bias	SD	Mean Bias	SD	Mean Bias	SD
N = 40,T = 6	β1 = 0	−0.170	0.064	0.052	0.092	−0.028	0.102	0.008	0.084	0.021	0.067
	β1 = 0.1	−0.198	0.068	0.072	0.111	−0.037	0.109	0.002	0.099	0.035	0.067
	β1 = 0.2	−0.222	0.066	0.096	0.107	−0.042	0.118	0.011	0.096	0.037	0.078
	β1 = 0.3	−0.248	0.066	0.131	0.107	−0.052	0.128	0.012	0.106	0.049	0.069
	β1 = 0.4	−0.275	0.073	0.165	0.086	−0.067	0.145	0.029	0.130	0.052	0.057
	β1 = 0.5	−0.310	0.075	0.173	0.066	−0.095	0.170	0.041	0.161	0.054	0.072
	β1 = 0.6	−0.344	0.076	0.167	0.046	−0.135	0.187	0.037	0.177	0.051	0.062
	β1 = 0.7	−0.383	0.076	0.152	0.034	−0.212	0.243	0.061	0.207	0.070	0.046
	β1 = 0.8	−0.416	0.074	0.122	0.024	−0.411	0.291	0.030	0.194	0.060	0.042
	β1 = 0.9	−0.417	0.077	0.072	0.017	−0.576	0.347	0.016	0.194	0.038	0.025
N = 80,T = 6	β1 = 0	−0.169	0.047	0.078	0.071	−0.015	0.077	−0.002	0.058	0.024	0.052
	β1 = 0.1	−0.192	0.048	0.109	0.080	−0.015	0.078	0.002	0.062	0.036	0.048
	β1 = 0.2	−0.215	0.049	0.145	0.073	−0.021	0.085	0.005	0.066	0.043	0.063
	β1 = 0.3	−0.244	0.050	0.169	0.063	−0.026	0.095	0.006	0.069	0.054	0.040
	β1 = 0.4	−0.277	0.050	0.185	0.047	−0.047	0.107	0.003	0.080	0.060	0.050
	β1 = 0.5	−0.300	0.053	0.184	0.038	−0.048	0.126	0.014	0.089	0.057	0.047
	β1 = 0.6	−0.340	0.053	0.169	0.027	−0.075	0.154	0.016	0.112	0.057	0.046
	β1 = 0.7	−0.381	0.053	0.151	0.021	−0.125	0.187	0.019	0.131	0.067	0.032
	β1 = 0.8	−0.411	0.056	0.121	0.017	−0.270	0.254	0.033	0.152	0.060	0.027
	β1 = 0.9	−0.408	0.051	0.073	0.012	−0.561	0.347	0.018	0.146	0.032	0.020
N = 40,T = 3	β1 = 0	−0.433	0.133	0.142	0.130	0.051	0.320	0.057	0.258	0.110	0.112
	β1 = 0.1	−0.490	0.147	0.153	0.137	0.078	0.414	0.070	0.325	0.131	0.134
	β1 = 0.2	−0.543	0.148	0.169	0.123	0.101	0.460	0.133	0.707	0.096	0.113
	β1 = 0.3	−0.602	0.145	0.179	0.102	0.092	0.579	0.152	1.036	0.128	0.107
	β1 = 0.4	−0.650	0.152	0.185	0.088	0.108	0.714	0.187	0.936	0.135	0.108
	β1 = 0.5	−0.723	0.159	0.179	0.081	0.110	0.927	0.258	1.253	0.134	0.085
	β1 = 0.6	−0.763	0.155	0.171	0.068	0.166	1.438	0.425	1.696	0.098	0.088
	β1 = 0.7	−0.825	0.173	0.154	0.054	−0.203	1.563	0.662	2.290	0.127	0.062
	β1 = 0.8	−0.884	0.167	0.123	0.039	−0.639	1.972	0.588	2.727	0.084	0.057
	β1 = 0.9	−0.933	0.158	0.077	0.024	−0.864	2.589	0.928	2.891	0.053	0.029
N = 80,T = 3	β1 = 0	−0.418	0.098	0.165	0.084	0.021	0.196	0.023	0.157	0.094	0.073
	β1 = 0.1	−0.482	0.099	0.179	0.083	0.017	0.259	0.013	0.186	0.097	0.069
	β1 = 0.2	−0.541	0.100	0.189	0.072	0.054	0.323	0.041	0.220	0.102	0.076
	β1 = 0.3	−0.601	0.106	0.188	0.064	0.039	0.381	0.031	0.231	0.129	0.066
	β1 = 0.4	−0.649	0.110	0.192	0.057	0.087	0.491	0.048	0.277	0.134	0.061
	β1 = 0.5	−0.710	0.109	0.180	0.051	0.110	0.532	0.080	0.374	0.117	0.056
	β1 = 0.6	−0.767	0.114	0.168	0.043	0.106	0.907	0.160	0.647	0.108	0.064
	β1 = 0.7	−0.818	0.114	0.147	0.036	0.069	1.329	0.287	1.137	0.112	0.045
	β1 = 0.8	−0.894	0.113	0.119	0.027	−0.301	1.632	0.465	1.823	0.086	0.035
	β1 = 0.9	−0.910	0.120	0.075	0.018	−0.855	2.218	0.925	2.541	0.056	0.025

tors proposed (GMM, SEM, Latent-space adjusted) generate small bias in estimating network exposure term across most of the scenarios (When N is large and T is small, the bias in latent space adjusted approach is slightly larger than GMM and SEM). The standard deviation of Latent-space adjusted estimates is smaller than GMM and SEM estimates for both lagged dependent variable and network exposure, especially when T is small. Thus we conclude our proposed methods are robust to the inclusion of other observed covariates.

5.4. Other robustness tests

We also performed other robustness tests. Details can be found in Appendix d in Supplementary material. Specifically:

(1) We test the robustness of the latent-space adjusted approach to the scaling of the latent social positions by starting the simulation with a different prior distribution of the intercept in the latent space model. Results show that estimates from the new latent space adjusted approach are almost identical to the original latent space adjusted approach in the main experiment.
(2) We test the robustness of our methods to the non-linear misspecification of the model. Specifically, we assume behavior is a function of the square of the unobserved trait in the data generating process, while we still estimate all models as linear. Results show that GMM and SEM estimates are robust to this non-linear misspecification, while the latent-space adjusted approach generates large bias in estimating lagged dependent variable, and smaller but still considerable bias in estimating network exposure term. This is because GMM and SEM estimates should be consistent as long as the unobserved trait is constant within unit, while as latent-space adjusted approach estimates proxies for the unobserved trait, it is sensitive to the functional form of the unobserved trait in both selection and influence model.
(3) We compare results from latent-space adjusted approach to the one where we use cluster membership as proxies for the unobserved trait. Results show that latent space adjusted approach clearly outperforms the one using cluster membership in terms of producing smaller bias in estimating lagged dependent variable as well as network exposure term.

6. Discussion and conclusion

While contagion effects have important implications for both theoretical and empirical studies, they are generally difficult to identify, as influence processes are often entangled with other processes such as selection and environmental factors. Here we show

Table 5b
Simulation results for network exposure with inclusion of covariates.

Network Exposure											
		LSDV		Random Effects		GMM		SEM		Latent Space	
		Mean Bias	SD	Mean Bias	SD	Mean Bias	SD	Mean Bias	SD	Mean Bias	SD
N = 40,T = 6	β2 = 0	−0.001	0.037	0.026	0.035	0.001	0.035	0.002	0.043	0.008	0.034
	β2 = 0.1	−0.016	0.053	0.037	0.054	−0.012	0.060	0.009	0.066	0.011	0.059
	β2 = 0.2	−0.031	0.075	0.050	0.074	−0.011	0.090	0.012	0.093	0.013	0.086
	β2 = 0.3	−0.035	0.094	0.076	0.090	0.003	0.116	0.020	0.106	0.023	0.105
	β2 = 0.4	−0.041	0.107	0.088	0.108	0.010	0.137	0.025	0.122	0.035	0.096
	β2 = 0.5	−0.036	0.130	0.105	0.124	0.026	0.155	0.035	0.147	0.012	0.121
	β2 = 0.6	−0.047	0.131	0.108	0.135	0.022	0.168	0.020	0.149	0.015	0.124
	β2 = 0.7	−0.044	0.146	0.101	0.150	0.042	0.180	0.029	0.162	0.006	0.120
	β2 = 0.8	−0.060	0.146	0.093	0.149	0.028	0.183	0.011	0.166	0.005	0.127
	β2 = 0.9	−0.057	0.156	0.077	0.165	0.020	0.196	0.016	0.175	−0.018	0.165
N = 80,T = 6	β2 = 0	0.001	0.023	0.031	0.024	−0.001	0.024	0.001	0.027	0.012	0.025
	β2 = 0.1	−0.014	0.036	0.041	0.034	−0.009	0.044	0.005	0.042	0.016	0.041
	β2 = 0.2	−0.027	0.047	0.066	0.046	−0.004	0.057	0.008	0.052	0.028	0.048
	β2 = 0.3	−0.033	0.058	0.091	0.056	−0.002	0.073	0.008	0.064	0.030	0.054
	β2 = 0.4	−0.042	0.072	0.110	0.068	−0.002	0.087	0.007	0.078	0.017	0.076
	β2 = 0.5	−0.045	0.082	0.132	0.077	−0.003	0.100	0.003	0.088	0.029	0.082
	β2 = 0.6	−0.042	0.084	0.151	0.089	0.008	0.109	0.011	0.093	0.017	0.087
	β2 = 0.7	−0.046	0.091	0.154	0.096	0.009	0.110	0.005	0.098	0.023	0.084
	β2 = 0.8	−0.043	0.108	0.155	0.109	0.016	0.133	0.008	0.113	0.030	0.101
	β2 = 0.9	−0.041	0.103	0.147	0.108	0.023	0.126	0.007	0.110	0.010	0.107
N = 40,T = 3	β2 = 0	−0.003	0.059	0.023	0.053	−0.006	0.133	0.004	0.182	0.019	0.059
	β2 = 0.1	−0.041	0.086	0.039	0.090	−0.024	0.228	0.046	0.342	0.042	0.083
	β2 = 0.2	−0.068	0.128	0.067	0.126	−0.007	0.291	0.071	0.411	0.018	0.129
	β2 = 0.3	−0.115	0.154	0.088	0.146	0.027	0.379	0.094	0.501	0.030	0.173
	β2 = 0.4	−0.126	0.199	0.124	0.185	0.074	0.385	0.130	0.600	0.041	0.207
	β2 = 0.5	−0.167	0.215	0.130	0.192	0.055	0.385	0.074	0.401	0.027	0.211
	β2 = 0.6	−0.182	0.245	0.137	0.227	0.065	0.390	0.124	0.708	0.074	0.228
	β2 = 0.7	−0.196	0.264	0.154	0.238	0.073	0.409	0.083	0.510	0.038	0.237
	β2 = 0.8	−0.239	0.300	0.115	0.269	0.042	0.397	0.038	0.382	0.029	0.255
	β2 = 0.9	−0.225	0.289	0.146	0.261	0.072	0.424	0.073	0.388	0.005	0.241
N = 80,T = 3	β2 = 0	−0.002	0.038	0.029	0.036	−0.004	0.078	0.002	0.133	0.019	0.032
	β2 = 0.1	−0.045	0.060	0.045	0.053	−0.010	0.147	0.031	0.201	0.040	0.059
	β2 = 0.2	−0.084	0.084	0.066	0.078	0.004	0.212	0.034	0.165	0.053	0.080
	β2 = 0.3	−0.109	0.107	0.104	0.095	0.015	0.225	0.039	0.182	0.056	0.092
	β2 = 0.4	−0.141	0.124	0.119	0.107	0.038	0.223	0.027	0.176	0.034	0.106
	β2 = 0.5	−0.170	0.141	0.147	0.119	0.025	0.221	0.023	0.189	0.055	0.119
	β2 = 0.6	−0.196	0.160	0.157	0.141	0.018	0.244	0.009	0.201	0.062	0.150
	β2 = 0.7	−0.201	0.171	0.174	0.150	0.027	0.243	0.022	0.221	0.059	0.190
	β2 = 0.8	−0.202	0.197	0.190	0.164	0.017	0.253	0.016	0.232	0.105	0.158
	β2 = 0.9	−0.208	0.197	0.181	0.170	0.015	0.237	0.018	0.232	0.047	0.161

that this entanglement/difficulty can essentially be framed as an omitted variable bias problem, and the methods currently used (e.g. SIENA, propensity score etc.) either do not deal with this problem or require strong assumptions.

In this paper, we propose several alternative estimation methods that have the potential to identify contagion effects when there are omitted variables present, and we use Monte Carlo simulation to test the performance of these estimators. Results show that all three methods proposed generally perform well in terms of recovering the true contagion effects when there is an unobserved variable that codetermines influence and selection. Specifically, a latent space adjusted approach outperforms the other methods that are traditionally used to deal with omitted variable bias problem, including a SEM based approach and an IV based approach: (1) the latent space adjusted approach produces much smaller bias in estimating the lagged dependent variable when T is small and the true coefficient for lagged dependent variable is large; (2) latent space adjusted approach generates slightly smaller bias in estimating the contagion effects than the other two methods in most scenarios; (3) the latent space adjusted approach is generally more efficient than the other methods, especially when T is small.

Although the latent space adjusted approach proposed in this paper hold much promise as an alternative approach to identify the contagion effect, there are also several limitations with this approach: (1) The latent space adjusted approach requires that the same unobserved traits occur in both influence and selection process. It can not account for the unobserved traits that are only present in one of the process but not the other. This is different from the GMM and SEM approach, which can also deal with unobserved trait that only occurs in influence but not selection process; (2) the approach assumes influence is lagged rather than contemporaneous. Although in many cases it is a reasonable assumption that there are some lags in the transmission of social effects, in some cases influence can also be regarded as contemporaneous, especially when the time interval between two observations is long. In this cases the latent space adjusted approach can not identify the true contagion effects due to reflection problem (see formal derivation in Appendix b in supplementary material); (3) Similarly, our approach does not account for concurrent network dependencies in the selection process. Although simulation results show that our approach is robust to the inclusion of lagged network dependencies (such as transitivity) in the selection process, in general it does not deal with concurrent network dependencies as opposed to other models such as ERGM. Note that there are some recent extensions of latent space models hold promise in this realm, such as latent position cluster model (Handcock et al., 2007) and Gaussian latent position cluster model (Rastelli et al., 2015). However, they have not been tested in the problem of iden-

tifying contagion effects, and thus beyond the scope of this paper; (4) the choice of dimensions of latent social space in latent space model is not clear. In our simulation we choose latent social position to be one-dimensional, however, this need not to be the case and there is no clear rule deciding how many dimensions users should use; (5) as our approaches assume time is a discrete variable, the interpretation of the parameters and the validity of some model assumptions (e.g. conditional independence) are critically dependent on the length of the time interval between observations (Lerner et al., 2013). While continuous-time based approaches and process-based models such as SIENA are not constrained by these limitations (Voelkle et al., 2012; Block et al., 2017); (6) finally, the computation of latent social position is very time consuming, and the computation time increases significantly with the increase of data or the number of dimensions in latent social position.

There are also some immediate extensions of the model. Although we choose a specific form of influence model (dynamic model with mean influence), the formulas of our methods can be easily changed and adapt to other forms of influence models. Similarly, the network variable is binary in the current model setup, but the model proposed can also be adapted and applied to data with continuous network variables.

Another possible extension of the proposed latent space adjusted approach is to apply it to multilevel data. For example, if we have students nested within classrooms, most of the networks we observe will be within classrooms and there will be few ties between classrooms. And to identify contagion/peer effects across classrooms, we need to adjust our approach to reflect the network structure in the latent space. If we estimate the latent social position using all networks as one global network, and estimate the influence model using all available data, the estimated latent social positions will not reflect the difference in actors' unobserved trait, because people holding similar unobserved traits might not be aware of the existence of each other due to the structural constraint. For example, if node A from organization 1 hold a similar unobserved trait to node B from organization 2, when we estimate networks from organizations 1 and 2 as a whole, the latent space estimates of persons A and B will deviate greatly from each other as they do not have a tie, even though they are similar in terms of the unobserved trait. And this will give biased results in the subsequent estimation of influence. Note that their lack of relationship is due to structural constraint rather than dissimilarity in the unobserved trait, and by not accounting for the structural constraints in the latent space model the estimated social position will be a poor proxy for the actual unobserved trait. And for the multilevel network, we propose to estimate the contagion effect in each sub graph/network using latent space adjusted approach as proposed. Then we can use a meta-analysis method to estimate overall contagion/peer effects.[16]

Although the methods we proposed in this paper still have considerable limitations, and none of these methods can 100% eliminate the bias in the estimation of contagion effects when influence is confounded with selection, we do believe that the methods proposed in this paper are still useful in many scenarios, especially the latent space adjusted approach. Furthermore, they have laid foundations and pointed directions for future work developing unique estimation methods to deal with omitted variable bias in the context of social network analysis. As we conclude, we have no intention in stating that we have found the cure for identifying contagion effects, since there is no universal cure, and an estimation method is only part of the solution. Nonetheless, we believe that with plausible alternative explanations that come from good theory, carefully measured covariates from longitudinal data, and a set of appropriate estimation models, we can effectively inform the debate about the contagion effects, and move forward scientifically.

[16] For the same reason, a large network should also be divided into smaller subgraphs before applying the latent space model to acquire the latent social positions.

Acknowledgements

Previous version of this paper has been presented at Sunbelt Conference, 2016. I am grateful to Michigan State University for financial support during the work reported here. I would like to thank Kenneth Frank, Spiro Maroulis, Rick Deshon, Spyros Konstantopoulos, Tracy Sweet and the anonymous reviewers for their thoughtful comments.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.socnet.2018.01.002.

References

An, W., 2011. Models and methods to identify peer effects. In: The Sage Handbook of Social Network Analysis. Sage, London, pp. 515–532.
An, W., 2015. Instrumental variables estimates of peer effects in social networks. Soc. Sci. Res. 50, 382–394.
Anderson, T.W., Hsiao, C., 1981. Estimation of dynamic models with error components. J. Am. Stat. Assoc. 76 (375), 598–606.
Anderson, T.W., Hsiao, C., 1982. Formulation and estimation of dynamic models using panel data. J. Econometrics 18 (1), 47–82.
Angrist, J.D., Lang, K., 2004. Does school integration generate peer effects? Evidence Boston's Metco Program. Am. Econ. Rev., 1613–1634.
Aral, S., Muchnik, L., Sundararajan, A., 2009. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. Proc. Natl. Acad. Sci. 106 (51), 21544–21549.
Arellano, M., Bond, S., 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Rev. Econ. Stud. 58 (2), 277–297.
Asch, S.E., 1952. Group Forces in the Modification and Distortion of Judgments.
Bandura, A., 1986. Social Foundations of Thought and Action. Prentice-Hall, Englewood Cliffs, N.J.
Barnes, G.M., Reifman, A.S., Farrell, M.P., Dintcheff, B.A., 2000. The effects of parenting on the development of adolescent alcohol misuse: a Six-Wave latent growth model. J. Marriage Family 62 (1), 175–186.
Bates, D., Maechler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Software 67 (1), 1–48, http://dx.doi.org/10.18637/jss.v067. i01.
Block, P., Koskinen, J., Hollway, J., Steglich, C., Stadtfeld, C., 2017. Change we can believe in: comparing longitudinal network models on consistency, interpretability and predictive power. Soc. Netw.
Bollen, K.A., Brand, J.E., 2010. A general panel model with random and fixed effects: a structural equations approach. Soc. Forces 89 (1), 1–34.
Bound, J., Jaeger, D.A., Baker, R.M., 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. J. Am. Stat. Assoc. 90 (430), 443–450.
Bramoullé, Y., Djebbari, H., Fortin, B., 2009. Identification of peer effects through social networks. J. Econometrics 150 (1), 41–55.
Christakis, N.A., Fowler, J.H., 2007. The spread of obesity in a large social network over 32 years. New Engl. J. Med. 357 (4), 370–379.
Christakis, N.A., Fowler, J.H., 2008. The collective dynamics of smoking in a large social network. New Engl. J. Med. 358 (21), 2249–2258.
Croissant, Y., Millo, G., 2008. Panel data econometrics in R: the plm package. J. Stat. Software 27 (2), 1–43.
Doreian, P., 2001. Causality in social network analysis. Sociol. Methods Res. 30 (1), 81–114.
Duncan, O.D., Haller, A.O., Portes, A., 1968. Peer influences on aspirations: a reinterpretation. Am. J. Sociol., 119–137.
Erbring, L., Young, A.A., 1979. Individuals and social structure contextual effects as endogenous feedback. Sociol. Methods Res. 7 (4), 396–430.
Feld, S.L., 1981. The focused organization of social ties. Am. J. Sociol., 1015.
Feld, S.L., 1982. Social structural determinants of similarity among associates. Am. Sociol. Rev., 797–801.
Frank, K.A., Zhao, Y., Borman, K., 2004. Social capital and the diffusion of innovations within organizations: the case of computer technology in schools. Sociol. Educ. 77 (2), 148–171.
Friedkin, N.E., Johnsen, E.C., 1990. Social influence and opinions. J. Math. Sociol. 15 (3-4), 193–206.
Friedkin, N.E., Johnsen, E.C., 1999. Social influence networks and opinion change. Adv. Group Processes 16 (1), 1–29.

Friedkin, N.E., 2001. Norm formation in social influence networks. Social Netw. 23 (3), 167–189.
Hancock, G.R., 2003. Fortune cookies, measurement error, and experimental design. J. Modern Appl. Stat. Methods 2 (2), 3.
Handcock, M.S., Raftery, A.E., Tantrum, J.M., 2007. Model-based clustering for social networks. J. R. Stat. Soc.: Series A (Stat. Soc.) 170 (2), 301–354.
Hoff, P.D., Raftery, A.E., Handcock, M.S., 2002. Latent space approaches to social network analysis. J. Am. Stat. Assoc. 97 (460), 1090–1098.
Judson, R.A., Owen, A.L., 1999. Estimating dynamic panel data models: a guide for macroeconomists. Econ. Lett. 65 (1), 9–15.
Kalmijn, M., Flap, H., 2001. Assortative meeting and mating: unintended consequences of organized settings for partner choices. Soc. Forces 79 (4), 1289–1312.
Kandel, D.B., 1978. Homophily, selection, and socialization in adolescent friendships. Am. J. Sociol., 427–436.
Kaplan, D., 2007. Structural Equation Modeling. Sage, pp. 1089–1093.
Kiviet, J.F., 1995. On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. J. Econometrics 68 (1), 53–78.
Kline, R., 2011. Principles and Practice of Structural Equation Modeling, third ed. Guilford.
Krivitsky, P.N., Handcock, M.S., 2008. Fitting Latent Cluster Models for Networks with Latentnet.
Krivitsky, P.N., Handcock, M.S., 2015. Latentnet: Latent Position and Cluster Models for Statistical Networks. The Statnet Project. R Package Version 2.7.1, http://www.statnet.org http://CRAN.R-project.org/package=latentnet.
Kruskal, J.B., 1964. Nonmetric multidimensional scaling: a numerical method. Psychometrika 29 (2), 115–129.
Lancaster, T., 2000. The incidental parameter problem since 1948. J. Econometrics 95 (2), 391–413.
Lazarsfeld, P.F., Merton, R.K., 1954. Friendship as a social process: a substantive and methodological analysis. Freedom Control Mod. Soc. 18 (1), 18–66.
Lerner, J., Indlekofer, N., Nick, B., Brandes, U., 2013. Conditional independence in dynamic networks. J. Math. Psychol. 57 (6), 275–283.
Lyons, R., 2011. The spread of evidence-poor medicine via flawed social-network analysis. Stat. Politics Policy 2 (1).
Manski, C.F., 1993. Identification of endogenous social effects: the reflection problem. Rev. Econ. Stud. 60 (3), 531–542.
McPherson, J.M., Smith-Lovin, L., 1987. Homophily in voluntary organizations: status distance and the composition of face-to-face groups. Am. Sociol. Rev., 370–379.
McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: homophily in social networks. Ann. Rev. Sociol., 415–444.
Merton, R., 1957. Social Theory and Social Structure. The Free Press, Glencoe.
Mouw, T., 2006. Estimating the causal effect of social capital: a review of recent research. Ann. Rev. Sociol., 79–102.
Newman, M.E., Park, J., 2003. Why social networks are different from other types of networks. Phys. Rev. E 68 (3), 036122.
Nickell, S., 1981. Biases in dynamic models with fixed effects. Econometrica: J. Econometric Soc., 1417–1426.
O'Malley, A.J., Elwert, F., Rosenquist, J.N., Zaslavsky, A.M., Christakis, N.A., 2014. Estimating peer effects in longitudinal dyadic data using instrumental variables. Biometrics 70 (3), 506–515.
Oetting, E.R., Donnermeyer, J.F., 1998. Primary socialization theory: the etiology of drug use and deviance. I. Subst. Use Misuse 33 (4), 995–1026.
Penuel, W.R., Sun, M., Frank, K.A., Gallagher, H.A., 2012. Using social network analysis to study how collegial interactions can augment teacher learning from external professional development. Am. J. Edu. 119 (1), 103–136.
Rastelli, Riccardo, Friel, Nial, Adrian, E., 2015. Raftery Properties of Latent Variable Network Models, arXiv preprint arXiv:1506.07806.
Rivera, M.T., Soderstrom, S.B., Uzzi, B., 2010. Dynamics of dyads in social networks: assortative, relational, and proximity mechanisms. Ann. Rev. Sociol. 36, 91–115.
Rosseel, Y., 2012. Lavaan: an r package for structural equation modeling. J. Stat. Software 48 (2), 1–36 (http://www.jstatsoft.org/v48/i02/).
Sarkar, P., Moore, A.W., 2005. Dynamic social network analysis using latent space models. ACM SIGKDD Explorations Newslett. 7 (2), 31–40.
Shalizi, C.R., Thomas, A.C., 2011. Homophily and contagion are generically confounded in observational social network studies. Sociol. Methods Res. 40 (2), 211–239.
Shortreed, S., Handcock, M.S., Hoff, P., 2006. Positional estimation within a latent space model for networks. Methodology 2 (1), 24–33.
Simmel, G., Wolff, K.H., 1950. The Sociology of Georg Simmel, vol. 92892. Simon and Schuster.
Sims, C.A., 1980. Macroeconomics and reality. Econometrica: J. Econometric Soc., 1–48.
Snijders, T., Steglich, C., Schweinberger, M., 2007. Modeling the Coevolution of Networks and Behavior., pp. 41–71.
Steglich, C., Snijders, T.A., Pearson, M., 2010. Dynamic networks and behavior: separating selection from influence. Sociol. Methodol. 40 (1), 329–393.
Valente, T.W., 1995. Network Models of the Diffusion of Innovations, vol. 2. Hampton Press, Cresskill, NJ, No. 2.
Valente, T.W., 1996. Social network thresholds in the diffusion of innovations. Social Netw. 18 (1), 69–89.
Vander Weele, T.J., An, W., 2013. Social networks and causal inference. In: Handbook of Causal Analysis for Social Research. Springer, Netherlands, pp. 353–374.
VanderWeele, T.J., Ogburn, E.L., Tchetgen, E.J., 2012. Why and when flawed social network analyses still yield valid tests of no contagion. Stat. Politics Policy 3, 1–11, Article 4.
Venkatesh, V., Morris, M.G., 2000. Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. MIS Q., 115–139.
Voelkle, M.C., Oud, J.H., Davidov, E., Schmidt, P., 2012. An SEM approach to continuous time modeling of panel data: relating authoritarianism and anomia. Psychol. Methods 17 (2), 176.
Wansbeek, T., Bekker, P., 1996. On IV, GMM and ML in a dynamic panel data model. Econ. Lett. 51 (2), 145–152.
Wooldridge, J.M., 2005. Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. J. Appl. Econometrics 20 (1), 39–54.
Wooldridge, J.M., 2010. Econometric Analysis of Cross Section and Panel Data. MIT Press.
Ziliak, J.P., 1997. Efficient estimation with panel data when instruments are predetermined: an empirical comparison of moment-condition estimators. J. Bus. Econ. Stat. 15 (4), 419–431.