Social Networks and the Identification of Peer Effects*

Paul Goldsmith-Pinkham† Guido W. Imbens‡

December 2011

Abstract

JEL Classification: C14
Keywords:

*Financial support for this research was generously provided through NSF grants 0631252, 0820361, and 0961707. We are grateful for discussions with Nicholas Christakis.

†.

‡Department of Economics, Harvard University, M-24 Littauer Center, 1805 Cambridge Street, Cambridge, MA 02138, and NBER. Electronic correspondence: imbens@harvard.edu, http://www.economics.harvard.edu/faculty/imbens/imbens.html.

1 Introduction

There is a large and growing literature on peer effects in economics. In the current paper we focus on the linear-in-means model that has proved to be popular in empirical work. In the context of data on friendship networks from the Add Health data where we are interested in peer effects on high school grades we examine some aspects of the model. Specifically we focus on three aspects.

First, we explore the possible endogeneity of the network. This issue has been raised in the econometric literature, and is often mentioned as a potential concern in the interpretation of estimates in empirical work. This is a particular concern in settings where the peer effects are hypothesized to arise through networks that are formed by individuals making choices to establish links. The specific concern is that individuals have unobserved characteristics that are correlated with their outcomes, and that these characteristics also affect the formation of links. Often the precise mechanism that is articulated is that individuals exhibit homophily in these unobserved characteristics, making it more likely that individuals who have common or similar values for these characteristcs form links. If these characteristics are also correlated with the outcomes, researchers will find that individuals who are connected have correlated outcomes even though there are no peer effects. In this paper we explore evidence for the presence of endogeneity. We develop a specific model for network formation that can incorporate this specific form of endogeneity.

Second, in settings where the peer groups are self-reported, are the result of choices by the individuals and do not form a partitioning of the population, there is likely to be measurement error in the links. Moreover, links may be of different strengths. The linear-in-means model restricts the influence of other individuals to be identical for all peers and zero for all peers, and assumes peer relationships are measured without error. Here we explore the implications of these restrictions. We find that outcomes for individuals who are not reported to be in one's peer group may still have direct effects on an individual's outcome in the context of a generalization of the linear-in-means model. Moreover, within peer groups some individuals may have stronger influences than others.

In the third part of this paper we explore another aspect of the linear-in-means model. This model implies that although individuals who are not part of particular individual's peer group have no direct effect on that individual's outcomes, there are indirect effects through common friends. This implies that changing covariate values for individuals not in one's peer group may still change that individual's outcome. Here we investigate whether there is direct evidence on these effects. This issue can partly be understood in the context of the connection between networks and peer effects and the spatial statistics literature. In the spatial statistics literature the researcher has units that are located in some space, with associated with each pair of units a distance between them. In the peer effects literature this distance between pairs of units is integer-valued, equal to the number of links one needs to travel to find a connection between two units. In the spatial literature researchers have used both autoregressive type models, where correlations decrease slowly with the distance without ever completely disappearing, as well as moving average type models where the correlations vanish at some finite distance. The linear-in-means model fits in with the autoregressive structure. In this paper we explore moving average type correlation structures and compare them to autoregressive type structures.

2 Set Up and Notation

We have data on $N = 534$ individuals in a single high school in the United States from the Add Health data. For these individuals we know their academic performance (grade point average) at two points in time. For each pair of individuals we know at two points in time whether they list each other as friends. Friendships here are interpreted as symmetric relationships: the pair i and j are coded as being friends if either i lists j as a friend, or j lists i as a friend, or both.

The outcome of interest is denoted by Y_i. In our application this is the grade point average (gpa) for a student at the end of the observation period. The vector of outcomes is denoted by $\mathbf{Y}$ with typical element Y_i. For each individual we also observe a $K-$vector of exogenous covariates, X_i, with $\mathbf{X}$ the $N \times K$ matrix with ith row equal to X_i'. In our application X_i is a scalar, the initial grade point average. The network in the final period is captured by the symmetric adjacency matrix $\mathbf{D}$, with typical element D_{ij} equal to one if i and j are friends and zero otherwise. The links are not directed, so $D_{ij} = D_{ji}$. The diagonal elements elements of $\mathbf{D}$ are zero. We also observe the network in the previous period, with adjacency matrix $\mathbf{D}_0$. For individual i the number of friends is

$M_i = \sum_{j=1}^{N} D_{ij}$, with $\mathbf{M}$ the N vector with ith element equal to M_i. It is also convenient to have a notation for the row-normalized adjacency matrix $\mathbf{G} = \text{diag}(\mathbf{M})^{-1}\mathbf{D}$ with

$$G_{ij} = \begin{cases} D_{ij}/M_i & \text{if } M_i > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Note that although $\mathbf{D}$ is symmetric, the row-normalized $\mathbf{G}$ is not in general symmetric. We drop from our analysis all students with no friends in the sample.

We observe the vector of outcomes $\mathbf{Y}$, the covariate matrix $\mathbf{X}$, and the networks $\mathbf{D}$ and $\mathbf{D}_0$ for a given sample size N. We assume these can be viewed as a draw from joint distribution $f(\mathbf{Y}, \mathbf{X}, \mathbf{D}, \mathbf{D}_0)$. For identification questions we assume we know this distribution. This is not as straightforward as it is in other settings where identification questions are studied. In many settings it is assumed that unit level responses can be viewed as independent and identically distributed random variables, so that in a large sample the joint distribution of these variables can be assumed to be known. Here the responses for pairs of individuals are not independent, and so we cannot simply refer to large sample arguments to claim that the joint distribution of the variables is known. Without giving formal arguments here, the motivation for focusing on this question can take two different forms.

The simplest case is one where we have a large number of networks. In that case we can estimate the joint distribution of $(\mathbf{Y}, \mathbf{X}, \mathbf{D})$ from that sample and proceed from there.

The more complicated case is that where we observe data from a single network. We need to impose some structure that limits the dependencies between observations, so that in a large sample we can construct subsamples of size N that are approximately independent. Suppose there is a covariate Z_i such that if Z_i and Z_j are far apart, then the probability of a link between individuals i and j is very small. This could arise if the individuals exhibit homophily in this variable Z_i, and its support is large. Then, if we think of large samples such where the distribution of Z_i is indexed by i, with the location of that distribution increasing in i, it may under some conditions be possible to view blocks of observations as essentially independent. That in turn may allow us to view the data as similar to data based on samples of networks rather than as a single network. This is similar in spirit to domain-increasing asymptotics in timeseries analyses where large sample arguments are based on increasing the sample size by adding observations further and further away in time, as opposed to the infill asymptotics where the large sample arguments are based on increasing the number of observations within the time range of the current sample. The same issues arise in spatial statistics, where this choice between infill and domain-increasing asymptotics is also important. Formal conditions that lead to consistency for estimators in such settings have not been established, although in many cases researchers proceed as if consistency and asymptotic normality holds. See

3 Friendship Networks and High School Grades

To illustrate the methods and models developed in this paper we use data from the Add Health survey. Here we use 534 individuals from a single high school. We observe their grade point average at two points in time, as well as their friendship links at the same two points in time. Table 1 presents some summary statistics for this sample. On average the students have 5 friends in their school.

Table 2 gives the distribution of the degree of separation. There are 142,311 pairs of students. Out of this set there are 1374 pairs of friends, and 5215 pairs of students who are not friends but who do have friends in common. For 8,892 pairs of students there are no links that connect them.

4 The Linear-In-Means Model for Peer Effects

The starting point is a linear-in-means model of the type studied by Manski (1993). Here we focus on the specification

$$Y_i = \beta_0 + \beta_x X_i + \beta_{\overline{y}} \overline{Y}_{(i)} + \beta_{\overline{x}} \overline{X}_{(i)} + \eta_i, \tag{4.1}$$

also used in Bramoullé, Djebbaria and Fortin (2009). Here the peer effects are based on averages over the peer groups, excluding the own outcome or covariate:

$$\overline{Y}_{(i)} = \frac{1}{M_i} \sum_{j=1}^{N} D_{ij} Y_j = \sum_{j=1}^{N} G_{ij} Y_j, \qquad \overline{X}_{(i)} = \frac{1}{M_i} \sum_{j=1}^{N} D_{ij} X_j = \sum_{j=1}^{N} G_{ij} X_j.$$

The main object of interest is the effect of peer's outcomes on own outcomes, $\beta_{\overline{y}}$, the endogenous peer effect in Manski's terminology. Also of interest is the exogenous peer effect $\beta_{\overline{x}}$. Here we interpret the endogenous effect as the average change we would see inan individual's outcome if we changed their peer's outcomes directly. In some cases this may be difficult to envision. Here we will think of this as some along the lines of the direct effect of providing special tutoring to one's peers on one's own outcome. The exogenous effect is interpreted as the causal effect of changing the peer's covariate values. For some covariates this thought experiment may be difficult, but for others, especially lagged values of choices, it may be feasible to consider interventions that would change those values for the peers.

It is useful to write the linear-in-means model in matrix notation. First, using the definition of the row-normalized adjacency matrix $\mathbf{G}$, we have

$$\mathbf{GY} = \begin{pmatrix} \overline{Y}_{(1)} \\ \vdots \\ \overline{Y}_{(N)} \end{pmatrix}, \qquad \mathbf{GX} = \begin{pmatrix} \overline{X}_{(1)} \\ \vdots \\ \overline{X}_{(N)} \end{pmatrix},$$

so that the model can be written as

$$\mathbf{Y} = \beta_0 \iota_N + \beta_x \mathbf{X} + \beta_{\overline{y}} \mathbf{GY} + \beta_{\overline{x}} \mathbf{GX} + \eta.$$

Initially we assume that the η are independent of the exogenous covariates and the peer groups:

$$\eta \perp \mathbf{X}, \mathbf{D}. \tag{4.2}$$

Because $\mathbf{G}$ is a deterministic function of $\mathbf{D}$, it follows that also $\eta \perp \mathbf{G}$. We also assume normality:

$$\eta \big| \mathbf{X}, \mathbf{D} \sim \mathcal{N}(0, \sigma^2 I_N).$$

This is mainly for convenience and can be relaxed. The independence assumption is a critical assumption. Manski (1993) raises the concern that the residual is correlated with the network. In his setting one can define a group indicator C_i associated with each individual, so that having a link between i and j, or $D_{ij} = 1$, is equivalent to the condition that $C_i = C_j$. In that case Manski formulates the concern with the independence assumption as the concern that $\mathbb{E}[\eta_i | C_i = c]$ may depend on c. In the next section we return to this issue, where we will explicity view this as a concern with the exogeneity of the network.

Although under Assumption (4.2) η is independent of $\mathbf{G}$ and $\mathbf{X}$, it is not independent, or even uncorrelated, with $\mathbf{Y}$ and therefore not independent of $\mathbf{GY}$. Hence we cannot simply regress $\mathbf{Y}$ on a constant, $\mathbf{X}$, $\mathbf{GY}$ and $\mathbf{GX}$ to get unbiased estimators for β. Manski (1993) and Bramoullé, Djebbaria and Fortin (2009) study identification in a setting close to this one. Manski focuses on the case where the peer groups partition the population, and highlights the identification problems that arise in that case. Bramoullé, Djebbaria and Fortin (2009) focus on the identifying power of peer groups that do not partition the population, and in particular on the assumption that $\mathbf{GG} \neq \mathbf{G}$.

To study identification of this model in the Bramoullé, Djebbaria and Fortin (2009) case it is useful to look at the conditional distribution of $\mathbf{Y}$ given $\mathbf{X}$ and $\mathbf{D}$:

$$\mathbf{Y} = (\mathbf{I} - \beta_{\overline{y}}\mathbf{G})^{-1}\beta_0\iota_N + (\mathbf{I} - \beta_{\overline{y}}\mathbf{G})^{-1}(\beta_x + \beta_{\overline{x}}\mathbf{G})\mathbf{X} + (\mathbf{I} - \beta_{\overline{y}}\mathbf{G})^{-1}\eta,$$

where ι_N is the N-vector with all elements equal to one. Now, under the normality assumption,

$$\mathbf{Y}|\mathbf{X}, \mathbf{G} \sim \mathcal{N}(\mu_Y, \Sigma_Y)$$

where

$$\mu_Y = (\mathbf{I} - \beta_{\overline{y}}\mathbf{G})^{-1}\beta_0\iota_N + (\mathbf{I} - \beta_{\overline{y}}\mathbf{G})^{-1}(\beta_x + \beta_{\overline{x}}\mathbf{G})\mathbf{X},$$

and

$$\Sigma_Y = \sigma^2(\mathbf{I} - \beta_{\overline{y}}\mathbf{G})^{-1}(\mathbf{I} - \beta_{\overline{y}}\mathbf{G}')^{-1}.$$

Manski(2003) shows that identification of $\beta_{\overline{y}}$ is difficult in settings where the the peer groups partition the population. Under conditions described in Bramoullé, Djebbaria and Fortin (2009) we can identify the parameters β_0, β_x, $\beta_{\overline{x}}$, and $\beta_{\overline{y}}$ from the conditional distribution of $\mathbf{Y}$ given $\mathbf{X}$ and $\mathbf{G}$ even without normality. A key condition is that the network $\mathbf{D}$ does not partition the population, and therefore $\mathbf{GG} \neq \mathbf{G}$. It must be the case, at least for some individuals, that their friends' friends are not their friends.

Exploiting the normality assumption we can estimate this model using maximum likelihood methods. Establishing large sample properties of the maximum likelihood estimator is difficult however. In order to establish large sample properties we need to make assumptions concerning the sequence of adjacency matrices $\mathbf{D}$ as the sample size grows. The most appropriate sequences of adjacency matrices would appear preserve the marginal distribution of the number of friends. This rules out settings with a large, infinite, population and random sampling of nodes and their associated links. Such sequences would lead to all individuals having increasing numbers of links. Instead a more appropriate sequence may shift the distribution of the covariates so that the probability of links between individuals far away in covariate space goes to zero.

We therefore use Bayesian methods for estimating the models. There are two main reasons for this. One is that the posterior distributions given the model and given the prior distributions have clear interpretations. In contrast, maximum likelihood estimates are difficult to interpret. There is no well-developed theory for the properties of maximum likelihood estimates, even in large samples. Although there is evidence that the logarithm of the likelihood is approximately quadratic around its maximum, there have been no formal properties established for these estimators and for confidence intervals based on maximum likelihood estimators and the information matrix. Second, for some of the models we consider here maximum likelihood estimators are difficult to calculate. In contrast, obtaining draws from the posterior distribution is relatively straightforward, although computationally intensive in many cases. These advantages of Bayesian methods have been noted before, and many researchers use Bayesian computational methods, often suggesting classical (frequentist) interpretations of the resulting estimators may be appropriate, without formal justification. Here we follow a fully Bayesian approach and focus on the posterior distributions.

First, we estimate the model assuming exogeneity of the network. We use normal independent prior distributions for the parameters β_0, β_x, $\beta_{\overline{y}}$ and $\beta_{\overline{x}}$, and an inverse chi-squared distribution for σ^2. Some summary statistics for posterior distributions are reported in the first panel of Table 3. The model for the outcomes exhibits evidence of substantial peer effects. The posterior mean for the endogenous peer effect is 0.16, and the posterior mean for the exogenous effect is 0.11. The posterior standard deviations are 0.05 and 0.07 respectively. Having current friends with good past academic performance, or friends with good current academic performance is associated with better performance for the individual.

5 Network Formation

In the setting where the peer groups are non-overlapping, a key issue is how the networks are formed, and how they might have been different. As a result we may need to worry about peer effects not merely through the direct effect of the peer groups characteristics and outcomes on an individual's outcomes, but also on the effects of outcomes and covariates on the peer group itself.

Currently $\beta_{\overline{x}}$ measures the effect of changing the average characteristics of one's peer group on an individual's outcome, keeping fixed the peer group itself. It is possible that changing these covariates may change the peer group, even under the exogeneity condition $\eta \perp (\mathbf{D}, \mathbf{X})$. Suppose that X_i is the grade point average of student i at the beginning of the period. Now suppose that we could have changed this by giving some of the students special tutoring. The effect of such a change would be to raise X_j for the affected students. This in turn would affect the gpa in the final period of their peers/friends. That in turn would affect their friends' friends, and so on. However, it may also affect whom the tutored students form friendships with. In order to find the total effect of the change in the covariates we would need to model the effect of the covariates on the outcome as well as the effect of the covariates on the network.

We start by modelling the network formation. The first assumption we make is that the decision to form a link is the result of two choices. Both individuals need to agree to form the link, and will do so if they view the net utility from the link as positive. Formally,

$$D_{ij} = \mathbf{1}_{U_i(j)>0} \cdot \mathbf{1}_{U_j(i)>0},$$

where $U_i(j)$ is the utility for individual i of forming a link with individual j. Following Jackson (2008) we refer to this type of model as strategic network formation models. In the sociological literature they are also referred to as Network Evolution Models (Toivonen, Kovanen, Kivelä, Onnela, Saramäki, and K. Kaski, 2009), or Actor Based Models (Snijders, 2009; Snijders, Koskinen, and Schweinberger, 2010). These models differ in the utility the agents associate with links given the characteristics of the other agents, and given the current state of the network, and in the opportunities the agents have for establishing or changing the status of their links.

Here we focus on a particular version of these models, exploiting the presence of data on the network at two points in time. At the same point in time each pair of agents evaluates the utility of a link between them. This utility depends on the characteristics of the two agents, and on the status of the network at the beginning of the period. Unlike, for example, Christakis, Fowler, Imbens and Kalyanaraman (2010) this utility does not depend on the decisions of the other pairs of agents. Conditional on the network at the beginning of the period $\mathbf{D}_0$, the utility for i of forming a link with j depends on X_i, X_j, on whether the two were friends in the previous period, $D_{0,ij}$, and whether they had friends in common in the previous period, F_{0ij}:

$$U_i(j) = \alpha_0 + \alpha_x|X_i - X_j| + \alpha_d D_{0ij} + \alpha_f F_{0,ij} + \epsilon_{ij}. \tag{5.1}$$

Initially let us assume that the ϵ_{ij} are independent accross all i and j and have a logistic distribution. The covariates enter in a specific way, reflecting homophily: the utility of a friendship goes down with the distance in covariate space. This implies that the probability of a link between i and j, given the previous version of the network, and given the covariates, is

$$\text{pr}(D_{ij} = 1|\mathbf{D}_0, \mathbf{X}) = p_{ij} \cdot p_{ji},$$

where the probability that i values the link with j is the same as the probability that j values the link with i:

$$p_{ij} = p_{ji} = \frac{\exp\left(\alpha_0 + \alpha_x|X_i - X_j| + \alpha_d D_{0ij} + \alpha_f F_{0,ij}\right)}{1 + \exp\left(\alpha_0 + \alpha_x|X_i - X_j| + \alpha_d D_{0ij} + \alpha_f F_{0,ij}\right)}.$$

More general models are possible here. Such models need not have the implication that $p_{ij} = p_{ji}$. For example, the utility associated with a friendship link may depend on the number of friends the potential friend had in the initial period, or on the level of the attribute X_i for the potential friend rather than solely on the difference $|X_i - X_j|$.

Again we use Bayesian methods for inference. The prior distributions for α_0, α_x, α_d and α_f are also independent normal. The normal prior distributions are centered at zero with prior standard deviation equal to one. Results are not sensitive to these assumptions about the prior distributions. Some summary statistics for posterior distributions are reported in the second panel of Table 3. The network model suggests that there is substantial sorting on academic performance, with the utility of a friendship link decreasing in the difference in past academic performance. The utility increases with the presence of past friendships, and with having in friends in common in the past.

6 Assessing Endogeneity of Networks

A major concern with a causal interpretation of estimates of the parameters in equation (4.1) is that the peer groups themselves may be endogenous. Correlations in outcomes between peers need not be the (causal) effect of peers. Instead, because peers are partly the result of individual choices, these correlations may reflect prior similarities between individuals, what Manski (1993) calls correlated effects. Individuals who are peers may be similar in terms of unobserved characteristics that also affect the outcomes. As a result there is a correlation in outcomes between peers. This is not the endogenous peer effect in Manski's terminology that stems from a simultaneity problem, but it is a correlated effect arising from omitted variable bias. Such omitted variable problems are traditionally in econometrics also referred to as endogenity problems (e.g., the ability bias in regression estimates of the returns to education).

In the setting Manski studies peer groups partition the sample. Thus we can assign each individual a cluster or group indicator C_i, so that $D_{ij} = 1$ if $C_i = C_j$. In this setting we can conceptualize the effect in Manski's terminology as a correlated effect defined as $\mathbb{E}[\eta_i|C_i = c] = \delta_c$, with δ_c varying with the group c. In the case we study here, with the peer groups individual specific, we need a generalization of this notion. What we wish to capture is that the $N-$vector of unobserved components of the outcome, η, is not independent of the $N \times N$ adjacency matrix $\mathbf{D}$ and the matrix of covariates $\mathbf{X}$. We model this potential dependence through the presence of unobserved individual characteristics. Let ξ_i be a individual-specific unobserved component that enters the outcome equation. We generalize the outcome equation to

$$Y_i = \beta_0 + \beta_x X_i + \beta_{\overline{y}} \overline{Y}_{(i)} + \beta_{\overline{x}} \overline{X}_{(i)} + \beta_\xi \xi_i + \eta_i. \tag{6.1}$$

We also modify the network formation process by generalizing the utility associated with equation governing the network formation. The utility associated with a link between individuals i and j now depends also on the distance between these two individuals in terms of the unobserved characteristic ξ_i:

$$U_i(j) = \alpha_0 + \alpha_x|X_i - X_j| + \alpha_\xi|\xi_i - \xi_j| + \alpha_d D_{0ij} + \alpha_f F_{0,ij} + \epsilon_{ij}. \tag{6.2}$$

Individuals with similar values for ξ_i are more likely to form links (if α_ξ is negative), and if β_ξ is differs from zero, this covariate has a direct effect on the outcome.

The question we wish to address in this section is whether the endogeneity is testable given knowledge of the joint distribution of **X**, **D** and **Y**, or, in other words, whether there is a basis for estimating the model with endogeneity. The answer is that endogeneity, within the context of the model with exchangeability of peers within the peer groups has some testable implications, and thus there is some basis for estimating models with network endogeneity. Given the model without endogeneity, (4.1) and (5.1), suppose we have the parameter values for β and α. Then we can calculate for each pair of individuals the probability of being friends,

$$P_{ij}^* = \text{pr}(U_i(j) > 0, U_j(i) > 0|\mathbf{X}) = P_{ij} \cdot P_{ji},$$

where P_{ij} is the probability that i attaches net positive utility to a link with j conditional on characteristics and the past value of the network:

$$P_{ij} = \frac{\exp\left(\alpha_0 + \alpha_x|X_i - X_j| + \alpha_d D_{0ij} + \alpha_f F_{0,ij}\right)}{1 + \exp\left(\alpha_0 + \alpha_x|X_i - X_j| + \alpha_d D_{0ij} + \alpha_f F_{0,ij}\right)},$$

and similarly for P_{ji}. We can also calculate the residuals η_i:

$$\eta_i = Y_i - \beta_0 - \beta_x X_i - \beta_{\overline{y}} \overline{Y}_{(i)} - \beta_{\overline{x}} \overline{X}_{(i)}.$$

Under the exogeneity assumption it follows that

$$\eta \perp P_{ij}^*.$$

However, if the ξ_i are non-degenerate, and β_ξ and α_ξ differ from zero, this no longer holds. For example, if $\alpha_\xi < 0$ and $\beta_\xi > 0$, then the absolute value of the difference in residuals, $|\eta_i - \eta_j|$, is, in expectation, increasing in the *ex ante* probability of the link. Formally,

$$\mathbb{E}\left[|\eta_i - \eta_j| \,\middle|\, P_{ij}^* = p, D_{ij} = d\right]$$

is increasing in p for $d = 0, 1$.

The intuition goes as follows. Among friend pairs a low value of P^*_{ij} implies that $|\xi_i - \xi_j|$ must be relatively close to zero. As a result, the absolute value of the difference $|\eta_i - \eta_j|$ is expected to be relatively low. High values of P^*_{ij} do not contain information regarding $|\xi_i - \xi_j|$, and thus do not predict the value of $|\eta_i - \eta_j|$. A similar argument goes through for non-friend pairs. If P^*_{ij} is high, one would expect for non-friend pairs that $|\xi_i - \xi_j|$ is relatively large, and hence $|\eta_i - \eta_j|$ is relatively large.

We can also look at this directly in terms of covariates.

$$\mathbb{E}\left[|\xi_i - \xi_j| \;\middle|\; |X_i - X_j| = x, D_{ij} = 1\right]$$

is decreasing in x. Hence, under endogeneity,

$$\mathbb{E}\left[|\eta_i - \eta_j| \;\middle|\; |X_i - X_j| = x, D_{ij} = 1\right]$$

is decreasing in x. Hence we can look, among friend pairs, at the correlation between the absolute value of the difference in the covariates and compare that to the absolute value of the difference in residuals in the outcome equations.

One concern with this comparison is that it relies on the influence of all peers being equal. If on the other hand peers have different effects on an individual depending on their proximity, one might find that the correlation in residuals η_i is higher for friends with similar characteristics. Although this is not captured in the linear-in-means model, one might generalize the model by modifying the weights in the row-normalized adjacency matrix $\mathbf{G}$. For example, one might model

$$G_{ij} = \frac{D_{ij} \cdot |X_i - X_j|}{\sum_{k \neq i} D_{ik} \cdot |X_i - X_k|}.$$

A second approach to assessing evidence for or against exogeneity does not rely on the equality of peer effects among peers. Instead we look at correlations in outcomes for non-friends. Again we compare within this set the correlation between the absolute value of the difference in residuals η_i and η_j and the *ex ante* probability of a link. Pairs of individual who are not friends, but who had a high probability of a link, have, in expectation, a larger difference in the absolute value of the difference ξ_i and ξ_j, and thus a larger expected value for the difference in absolute values of the residuals η_i and η_j. Here the key assumption is that under the null model there is no correlation in residuals for pairs of individuals who are not friends.

7 Models with Endogenous Networks

We now turn to an estimable version of the model with network endogeneity. We use the model in equations (6.1) and (6.2), with an unobserved scalar component that appears in the outcome equation and the utility associated with links:

$$Y_i = \beta_0 + \beta_x X_i + \beta_{\overline{y}} \overline{Y}_{(i)} + \beta_{\overline{x}} \overline{X}_{(i)} + \beta_\xi \xi_i + \eta_i,$$

$$U_i(j) = \alpha_0 + \alpha_x |X_i - X_j| + \alpha_\xi |\xi_i - \xi_j| + \alpha_d D_{0ij} + \alpha_f F_{0,ij} + \epsilon_{ij}.$$

We make the distributional assumptions

$$\eta \big| \xi, \mathbf{X}, \mathbf{D} \sim \mathcal{N}(0, \sigma^2 I_N),$$

and a logistic distribution for ϵ_{ij}, independent of η and ξ, and with all ϵ_{ij} independent. Finally, we assume that the unobserved type ξ_i is binary:

$$\text{pr}(\xi_i = 1 | \mathbf{X}, \mathbf{D}_0) = 1 - \text{pr}(\xi_i = 0 | \mathbf{X}, \mathbf{D}_0) = p.$$

In the empirical analysis we fix $p = 1/2$. This leads to a parametric distribution for $(\mathbf{Y}, \mathbf{G})$ given $(\mathbf{X}, \mathbf{D}_0)$. First, define the probability of a link conditional on observed and unobserved covariates:

$$p(x_1, x_2, d_0, f_0, \xi_1, \xi_2; \alpha_0, \alpha_x, \alpha_\xi, \alpha_d, \alpha_f) = \frac{\exp(\alpha_0 + \alpha_x |x_1 - x_2| + \alpha_\xi |\xi_1 - \xi_2| + \alpha_d d_0 + \alpha_f f_0)}{1 + \exp(\alpha_0 + \alpha_x |x_1 - x_2| + \alpha_\xi |\xi_1 - \xi_2| + \alpha_d d_0 + \alpha_f f_0)}.$$

Then, conditional on ξ, $\mathbf{X}$, and $\mathbf{D}_0$, we have the likelihood function for the network,

$$\mathcal{L}_{\text{network}}(\alpha | \mathbf{G}; \xi, \mathbf{X}, \mathbf{D}_0) =$$

$$\prod_{i \neq j} \left(p(X_i, X_j, D_{0,ij}, F_{0,ij}, \xi_i, \xi_j; \alpha_0, \alpha_x, \alpha_\xi, \alpha_d, \alpha_f) \times p(X_j, X_i, D_{0,ji}, F_{0,ji}, \xi_j, \xi_i; \alpha_0, \alpha_x, \alpha_\xi, \alpha_d, \alpha_f) \right)^{D_{ij}}$$

$$\times \left(1 - p(X_i, X_j, D_{0,ij}, F_{0,ij}, \xi_i, \xi_j; \alpha_0, \alpha_x, \alpha_\xi, \alpha_d, \alpha_f) \times p(X_j, X_i, D_{0,ji}, F_{0,ji}, \xi_j, \xi_i; \alpha_0, \alpha_x, \alpha_\xi, \alpha_d, \alpha_f)\right)^{1 - D_{ij}}.$$

The likelihood function for the outcome is

$$\mathcal{L}_{\text{outcome}}(\beta, \sigma^2 | \mathbf{Y}; \mathbf{D}, \mathbf{X}, \xi) = \frac{1}{(2\pi)^{N/2} |\Sigma_Y|} \exp\left(- (\mathbf{Y} - \mu_Y) \Sigma_Y^{-1} (\mathbf{Y} - \mu_Y) / 2 \right),$$

where

$$\Sigma_Y = \sigma^2 (\mathbf{I} - \beta_{\overline{y}} \mathbf{G})^{-1} (\mathbf{I} - \beta_{\overline{y}} \mathbf{G}')^{-1},$$

and

$$\mu_Y = (\mathbf{I} - \beta_{\overline{y}}\mathbf{G})^{-1}\beta_0\iota_N + (\mathbf{I} - \beta_{\overline{y}}\mathbf{G})^{-1}(\beta_x + \beta_{\overline{x}}\mathbf{G}) + (\mathbf{I} - \beta_{\overline{y}}\mathbf{G})^{-1}\beta_\xi\xi.$$

Finally,

$$\mathcal{L}(\beta, \alpha, \sigma^2, p|\mathbf{Y}, \mathbf{G}, \xi; \mathbf{X}, \mathbf{D}_0) = \mathcal{L}_{\text{outcome}}(\beta, \sigma^2|\mathbf{Y}; \mathbf{D}, \mathbf{X}, \xi) \times \mathcal{L}_{\text{network}}(\alpha|\mathbf{G}; \xi, \mathbf{X}, \mathbf{D}_0),$$

is the conditional likelihood function. Integrating out the ξ_i we have

$$\mathcal{L}(\beta, \alpha, \sigma^2, p|\mathbf{Y}, \mathbf{G}; \mathbf{X}, \mathbf{D}_0) =$$

We could use alternative parametric distributions for the unobserved component ξ_i. The main issues are the flexibility of the distribution and the computational tractibility of the resulting model.

In this case maximum likelihood estimation is particularly computationally demanding because of the difficulty of integrating out the unobserved ξ_i. Moreoever, there are no results for the repeated sampling properties of the maximum likelihood estimators for this type of model, even in large samples. We therefore again focus on Bayesian methods. We specify a prior distribution for the parameter $\theta = (\beta, \sigma^2, \alpha, p)$, and use mcmc methods to obtain draws from the posterior distribution of θ given the data $(\mathbf{D}, \mathbf{D}_0, \mathbf{X}, \mathbf{Y})$. This Bayesian approach allows us to treat the ξ_i as unobserved random variables, and exploit the fact that with the ξ_i known, we would have exogeneity of the network and be able to obtain draws from the posterior distribution effectively.

The appendix contains details on the mcmc algorithm for the specific case. We use metropolis-hastings steps separately for α, β, and the ξ, and use exact draws for the conditional posterior for σ^2 given the other parameters and the data. The results are reported in the second part of Table 4. The unobserved component ξ matters substantially for the network estimation. The coefficient on the difference in the unobserved characteristics is large and precisely estimated. This appears to allow the model to fit the observed clustering better than the model without the unobserved component. The unobserved component does not appear to matter much for the outcome. Its coefficient is close to zero with the 95% posterior probability interval comfortably including zero. The posterior distribution for the endogenous peer effect is not much affected by this, with the posterior mean equal to 0.15, and the posterior standard deviation equal to 0.05, compared to 0.16 and 0.05 in the exogenous network model.

8 Who are the Peers?

Obviously in order to estimate peer effects we need to know who one's peers are. In pratice this is not so easy in the current setting. One specific concern that arises in the setting with self reported networks such as the friends networks in Add Health is that the links may be observed with error. This is much less of a problem in settings where the population is partitioned in peer groups, such as the case where the peer groups correspond to classes. In the friends networks we rely heavily on individuals consistently reporting the state of the network where in reality the network is a constantly changing system.

Here it may be useful to think of the connection between networks and general notations of spatial correlations. The network literature typically models the distance between units as integers, one if connected, and two if not connected but with friends in common. In reality units, individuals in our case, may form stronger links with some individuals than with others, but be tied to many individuals with weak links. In the spatial statistics literature this would create complications because the distance measure would be partially unknown.

In this section we look at the implications of this for the estimation and interpretation of peer effects.

First, let us look at the number of friendships that are only reported by one side and not the other. We can also see evidence of the fluidity of the concept by looking at the changes in the network over time. On average each student has 3.5 former friends, individual who were listed as friends in the first period, but not in the second period. On average each student has 2.5 long term friends, individuals listed as friends in both periods.

We study the implications of the uncertainty in the measurement of peers by estimating a generalization of the linear-in-means model where we consider the presence of two networks side by side, each with their associated peer effects. Let $\mathbf{D}_A$ denote first network, and $\mathbf{D}_B$ the second network. Then let the averages be

$$\overline{Y}_{A,(i)} = \frac{1}{M_{A,i}} \sum_{j=1}^{N} D_{A,ij} Y_j, \qquad \overline{X}_{A,(i)} = \frac{1}{M_{A,i}} \sum_{j=1}^{N} D_{A,ij} X_j,$$

and similarly for $\overline{Y}_{B,(i)}$ and $\overline{X}_{B,(i)}$. Then we estimate a linear-in-means model where both networks have exogenous and endogenous peer effects:

$$Y_i = \beta_0 + \beta_x X_i + \beta_{A,\overline{y}}\overline{Y}_{A,(i)} + \beta_{A,\overline{x}}\overline{X}_{A,(i)} + \beta_{B,\overline{y}}\overline{Y}_{B,(i)} + \beta_{B,\overline{x}}\overline{X}_{B,(i)} + \eta_i. \tag{8.1}$$

We implement this by setting the first network to be that of friends in the second period, and the second one to be the network of former friends, where where links imply friendships in the initial period but not in the second period. Formally,

$$D_{A,ij} = D_{ij}, \qquad D_{B,ij} = (1 - D_{ij}) \cdot D_{0,ij}.$$

Table 5 reports posterior means and standard deviations for this model under exogeneity of the network formation. We find that the peer effects of former friends are substantial (posterior mean 0.13, posterior standard deviation 0.05), almost the same as the peer effects for current friends (posterior mean 0.15, posterior standard deviation 0.05). This casts substantial doubt on the notion that relying on self reported friendship links captures all the connections that matter for correlations in outcomes.

9 Indirect Peer Effects

The linear-in-means model has strong implications for the correlations in outcomes between individuals. Peer effects are mediated through friends' outcomes, but ultimately outcomes for non-friends can still affects individual's outcomes indirectly. In this section we investigate some implicastions of this.

Even though outcomes for individuals who are not friends with i but who have friends in common with i do not directly affect the outcome for individual i, they do so indirectly through the common friends. As a result changing the covariates for these second-friends affects the expected value of the outcome for i.

Taking the matrix version of the linear-in-means model,

$$\mathbf{Y} = \beta_0 \iota_N + \beta_x \mathbf{X} + \beta_{\overline{y}} \mathbf{G}\mathbf{Y} + \beta_{\overline{x}} \mathbf{G}\mathbf{X} + \eta.$$

we can write the outcome in terms of the exogenous variables as

$$\mathbf{Y} = (\mathbf{I} - \beta_{\overline{y}}\mathbf{G})^{-1}\beta_0 \iota_N + (\mathbf{I} - \beta_{\overline{y}}\mathbf{G})^{-1}(\beta_x + \beta_{\overline{x}}\mathbf{G})\mathbf{X} + (\mathbf{I} - \beta_{\overline{y}}\mathbf{G})^{-1}\eta.$$

Following the arguments in Bramoullé, Djebbaria, and Fortin (2009), we can expand this to

$$\mathbb{E}\left[\mathbf{Y}|\mathbf{X},\mathbf{G}\right] = \frac{\beta_0}{1-\beta_{\overline{y}}}\iota + \beta_x\mathbf{X} + \sum_{k=1}^{\infty}\beta_k\mathbf{G}^k\mathbf{X}.$$

Here we estimate an approximation to this conditional expectation based on the first two terms:

$$\mathbb{E}\left[\mathbf{Y}|\mathbf{X},\mathbf{G}\right] \approx \frac{\beta_0}{1-\beta_{\overline{y}}}\iota_N + \beta_x\mathbf{X} + \beta_1\mathbf{G}\mathbf{X} + \beta_2\mathbf{G}\mathbf{G}\mathbf{X}.$$

The linear-in-means model suggest that the term on the friends-in-common average covariates should be

$$\beta_1 = \beta_x\beta_{\overline{y}} + \beta_{\overline{x}}, \qquad \text{and} \quad \beta_2 = (\beta_x\beta_{\overline{y}} + \beta_{\overline{x}})\beta_{\overline{y}}.$$

Two implications emerge. First, if own covariate effect, and the exogenous and endogenous peer effects are positive, the second coefficient in this expansion should be positive. Second, the ratio of the first and second coefficients in this expansion should equal $\beta_{\overline{y}}$.

Table 6 reports estimation results for this model. We find that the effect of the average effect of second-friends is posterior 95% probability interval includes zero. The posterior variance is large, so in fact the data are consistent with wide range of values. The point though is that we cannot conclusively establish that there is any indirect peer effect of friends of friends.

The second attempt to establishing whether there are indeed indirect effects for individuals who are not friends takes a different approach. It adds a moving-average type component to the linear-in-means model so that it allows for the possibility that there is no correlation between outcomes for individuals who are not friends, while allowing for correlations in outcomes between friends. The starting point is the linear in means model

$$\mathbf{Y} = \beta_0\iota_N + \beta_x\mathbf{X} + \beta_{\overline{y}}\mathbf{G}\mathbf{Y} + \beta_{\overline{x}}\mathbf{G}\mathbf{X} + \eta.$$

Now we model the unobserved component η_i as

$$\eta_i = \nu_i + \sum_{j\neq i} D_{ij}\epsilon_{ij},$$

with the ϵ_{ij} independent across i and j, normally distributed with mean zero and variance σ_ϵ^2. The ν_i have a normal distribution with mean zero and variance σ_ν^2. This leads to the following covariance matrix for the η, denoted by Ω:

$$\Omega_{ij} = \begin{cases} \sigma_\nu^2 + M_i \cdot \sigma_\epsilon^2, & \text{if } i = j, \\ \sigma_\epsilon^2 & \text{if } i \neq j, D_{ij} = 1, \\ 0 & \text{if } i \neq j, D_{ij} = 0. \end{cases}$$

In this model if $\sigma_\epsilon^2 > 0$ and $\beta_{\overline{y}} = 0$, there would be no spillovers beyond friends, but there would be positive correlation between outcomes for friends.

Table 7 presents summary statistics for the posterior distribution for this model. The posterior mean for the endogenous peer effects, $\beta_{\overline{y}}$ goes down from 0.15 in the baseline model to 0.09 in the model with the additional flexibility in the error-covariance structure. The posterior standard deviation is 0.06, so there is now considerable uncertainty about the sign of this effect. The posterior mean for the variance component σ_ϵ^2 is 0.10^2, large enough to create a substantial correlation between outcomes for friends.

Both exercises carried out in this section show that the evidence for indirect peer effects, effect not from friends but from their friends, is limited. One explanation may be that the effects we find in the model with peer effects are relatively modest. The direct peer effects may simply be too small to generate substantial indirect effects, or the indirect effects may not be there for outcomes such as grades.

10 Conclusion

In this paper we explore extensions of the linear-in-means model for identifying and estimating peer effects. We study possible evidence for endogeneity of the network, and develop models that allow for endogeneity.

In our application to friendship networks in the Add Health data we find that one's friends's grades are correlated with one's own grades. Whether these are causal peer effects is more difficult to establish. There is limited evidence that the friendships are endogenous to the grades. There is evidence that current friendships are not sufficient for capturing all correlations in outcomes. Correlations in grades with former friends are almost as strong as those with current friends, casting doubt on causal interpretations of the correlations between current friends and current grades. We also explore the evidence for indirect peer effects, effects of friends of friends. We find that the data are inconclusive regarding the presence of such effects.

Appendix: Approximating the Posterior Distribution

Here we provide some details on the evaluation of the posterior distribution. We focus on the model with endogeneity. The model with exogenous network formation is simpler to estimate. In particular, we can separately analyse the network formation model and the model for the primary outcome. We use the results from that model to provide starting values for the model with endogeneity.

The model with endogenous network formation has four components. The first describes the conditional distribution of $\mathbf{Y}$ given $\mathbf{X}$, $\mathbf{G}$, $\mathbf{G}_0$, and the unobserved ξ:

$$\begin{aligned}
&\mathbf{Y}|\mathbf{X}, \mathbf{G}, \mathbf{G}_0, \xi \sim \\
&\qquad \mathcal{N}\left((\mathbf{I} - \beta_{\overline{y}}\mathbf{G})^{-1}\beta_0\iota_N + (\mathbf{I} - \beta_{\overline{y}}\mathbf{G})^{-1}(\beta_x + \beta_{\overline{x}}\mathbf{G})\mathbf{X} + (\mathbf{I} - \beta_{\overline{y}}\mathbf{G})^{-1}\beta_\xi\xi,\right. \\
&\qquad\qquad \left.\sigma^2\left(\mathbf{I} - \beta_{\overline{y}}\mathbf{G}\right)^{-1}\left(\mathbf{I} - \beta_{\overline{y}}\mathbf{G}'\right)^{-1}\right).
\end{aligned}$$

The second part describes the network formation given $\mathbf{X}$ and the unobserved ξ:

$$\begin{aligned}
&D_{ij} = \mathbf{1}_{U_i(j)>0} \cdot \mathbf{1}_{U_j(i)>0}, \\
&U_i(j) = \alpha_0 + \alpha_x|X_i - X_j| + \alpha_d D_{0ij} + \alpha_f F_{0,ij} + \epsilon_{ij}.
\end{aligned} \tag{A.1}$$

The third component gives the conditional distribution of the ξ given $\mathbf{X}$:

$$\begin{aligned}
&\epsilon_{ij} \sim \frac{\exp(-x)}{(1 + \exp(-x))^2}, \\
&\text{pr}(\xi_i = 1) = 1/2, \qquad \xi_i \perp \xi \text{ for } i \neq j.
\end{aligned}$$

This model leads to a likelihood function in terms of $\theta = (\beta, \alpha, \sigma^2)$. The prior distributions for all parameters are independent. The prior distributions for α, β_x are normal with mean zero and variance equal to one. The prior distribution for σ^2 is.

We pick starting values for θ based on the model with exogeneity for all parameters other than β_ξ and α_ξ. We take the starting value for α_ξ from a normal distribution centered at -1 and a variance equal to 0.01. The starting values for β_ξ are drawn from a normal distribution centered at zero with variance 0.01. Then we draw the ξ_i from a binomial distribution with mean 1/2, for $i = 2, \ldots, N$. The first value ξ_1 is set equal to 1. Before changing the values for θ we repeated update the ξ to obtain values more in line with the data. We have found that this leads to faster convergence. We update the ξ_i sequentially, given the starting values for θ. Each time we update a single ξ_i. We cycle through the full set of ξ_i 100 times without changing any of the values for θ.

Next we start cycling through the other parameter values. We divide the updating into three parts. First we update the β given σ^2, α, and ξ. We use a Metropolis step here using for the candidate distribution a normal distribution centered at the current values, with covariance matrix the covariance matrix estimated from the exogenous model times 1/16.

The second step updates σ^2. Here we use the exact posterior distribution given values for the other parameters and given the ξ.

The third step involves updating the α given the ξ. Here we use a Metropolis-Hastings step, with the candidate distribution centered at the current values, and the covariance matrix estimated on the model with exogenous network.

In the fourth step we update the ξ given all the parameter values.

We use five starting values. For each of the five chains we drop the first xxx iterations. We then compute the average value of the elements of θ, and the overall average. We monitor convergence by comparing for each of the elements of θ the ratio of the overall variance and the average of the within-chain variances. Following the suggestion in Gelman and Rubin () we aim for ratios below 1.1. After xxx iterations the convergence criteria for all elements of θ were below xxx, with most below xxx.

References

Blume, L., W. Brock, S. Durlauf, and Y Ioannides, (2011) "Identification of Social Interactions," Chapter 18, in *Social Economics*, Benhabib, Bison and Jackson (eds.), Vol 1B, North-Holland.

Bramoullé, Y., H. Djebbaria, and B. Fortin, (2009), "Identification of peer effects through social networks," *Journal of Econometrics,* Volume 150(1): 41-55.

Calvó-Armengol, A., and M. Jackson, (2004) "The Effects of Social Networks on Employment and Inequality," *American Economic Review*, Vol. 94(3), 426-454.

Easley, D., and J. Kleinberg, (2010) *Networks, Crowds, and Markets*, Cambridge University Press, Cambridge, UK.

Gelfand, A., P. Diggle, M. Fuentes, P. Guttorp, (2010), *Handbook of spatial statistics*, Chapman and Hall.

Jackson, M., (2003) The Stability and Efficiency of Economis and Social Networks, *Advances in Economic Design*, Koray and Sertel (eds.), Springer, Heidelberg.

Jackson, M., (2006) "The Economics of Social Networks," in Volume I of *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*, Richard Blundell, Whitney Newey, and Torsten Persson, (editors) Cambridge University Press.

Jackson, M., (2011) "An Overview of Social Networks and Economic Applications," Chapter 12, in *Social Economics*, Benhabib, Bison and Jackson (eds.), Vol 1B, North-Holland.

Jackson, M., and A. Wolinsky, (1996) A Strategic Model of Social and Economic Networks, *Journal of Economic Theory*, Vol. 71(1), pp 890-915.

Jackson, M., and B. Rogers, (2007) Meeting Strangers and Friends of Friends: How Random are Socially Generated Networks? *American Economic Review*, Vol. 97, No. 3, pp 890-915.

Jackson, M., (2008) *Social and Economic Networks*, Princeton University Press, Princeton, NJ.

Kline, B., (2011), "Identification of Semiparametric Binary Games on Networks with Application to Social Interactions," Department of Economics, Northwestern University

Kolaczyk, E., (2009) *Statistical Analysis of Network Data*, Springer.

Manski, C., (1993), "Identification of Endogenous Social Effects: The Reflection Problem,"*Review of Economic Studies*, 60, 531-542.

Manski, C., (2000), "Economic Analysis of Social Interactions,"*Journal of Economic Perspectives*, 14(3), 115-136.

Schabenberger, O., and C. Gotway, (2004), *Statistical Methods for Spatial Data Analysis*, Chapman and Hall.

Toivonen, R., L. Kovanen, M. Kivelä, J. Onnela, J. Saramäki, and K. Kaski, (2009), "A Comparative Study of Social Network Models: Network Evolution Models and Nodal Attribute Models,"Social Networks, 31: 240-254.

Wooldridge, J., (2002), *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, MA.

Table 1: Summary Statistics Add Health Sample ($N = 534$)

	Average	Standard Deviation
GPA_1 (Y_i)	2.5	(1.0)
GPA_0 (X_i)	2.6	(0.8)
Number of Friends	5.2	(3.0)

Table 2: Distribution of Degree of Separation (Number of Pairs 142,311)

Degree of Separation	Number of Pairs
1	1374
2	5215
3	15573
4	32019
5	39310
6	25852
7	10283
8	2696
9	487
10	70
11	7
12	0
∞	8892

Table 3: Summary Statistics Posterior Distribution: Exogenous Network

	Post Mean	Post Standard Deviation
Outcome Equation		
β_0	-0.13	(0.12)
β_x	0.74	(0.04)
$\beta_{\overline{y}}$	0.16	(0.05)
$\beta_{\overline{x}}$	0.11	(0.07)
Network Model		
α_0	-2.56	(0.04)
α_x	-0.20	(0.03)
α_d	2.52	(0.05)
α_f	1.20	(0.04)

Table 4: Summary Statistics Posterior Distribution: Endogenous Network

	Post Mean	Post Standard Deviation
Outcome Equation		
β_0	-0.10	(0.13)
β_x	0.73	(0.04)
$\beta_{\overline{y}}$	0.15	(0.05)
$\beta_{\overline{x}}$	0.11	(0.06)
β_ξ	-0.01	(0.10)
σ^2	0.37	(0.02)
Network Model		
α_0	-2.26	(0.04)
α_x	-0.21	(0.04)
α_ξ	-1.06	(0.07)
α_d	2.63	(0.03)
α_f	1.22	(0.04)

Table 5: Summary Statistics Posterior Distribution: Exogenous Network With Lagged Network Effects

	Post Mean	Post Standard Deviation
Outcome Equation		
β_0	-0.12	(0.12)
β_x	0.73	(0.04)
$\beta_{\overline{y}}$	0.15	(0.05)
$\beta_{\overline{x}}$	0.09	(0.06)
$\beta_{0,\overline{y}}$	0.13	(0.05)
$\beta_{0,\overline{x}}$	-0.09	(0.06)

Table 6: Summary Statistics Posterior Distribution: Exogenous Network with No Endogenous Peer Effects

	Post Mean	Post Standard Deviation
intercept	-0.25	(0.18)
β_x	0.73	(0.04)
$\beta_{\overline{x}}$	0.21	(0.08)
$\beta_{\text{FIC},\overline{x}}$	0.10	(0.11)

Table 7: Summary Statistics Posterior Distribution: Exogenous Network with No Endogenous Peer Effects

	Post Mean	Post Standard Deviation
intercept	-0.08	(0.12)
β_x	0.74	(0.04)
$\beta_{\overline{y}}$	0.09	(0.06)
$\beta_{\overline{x}}$	0.16	(0.07)
σ^2	0.58^2	
σ_ν^2	0.10^2	