Identifying marginal returns to treatment through social networks*

Firmin Doko Tchatoka[a†] Timothy Hersey[b] Virginie Masson[a]

[a] *School of Economics and Public Policy, The University of Adelaide, Australia*

[b] *Department of Economics, Yale University, USA*

Abstract

We propose an econometric framework that accounts for the effects of both observed and unobserved characteristics of social networks on individuals' decision to undergo treatment in counterfactual models based on potential outcomes. We show that network peer effects provide exogenous variation to identify the marginal treatment effect (MTE) and the marginal policy-relevant treatment effect (MPRTE). Monte Carlo experiments demonstrate that ignoring the influence of social interactions on an individual's decision in counterfactual analysis can lead to misspecification, substantially biasing both the MTE and MPRTE estimates. We apply the proposed methodology to the college attainment model in the US, using Add-Health data that contain high school friendship networks. The results indicate that not controlling for high school friendship network peer effects misidentifies the marginal return to attending college, as well as the marginal effect of policies aimed at improving college attendance.

Key words: Network formation; peer effects; counterfactual analysis; treatment effect measures; propensity scores; Bayesian method; college attendance.

JEL classification: C11; C143; C15; C21; I26.

* We are grateful to the University of North Carolina at Chapel Hill for providing us the Add Health Data. We also thank Propser Dovonon, SeoJeong Lee, Valentyn Panchenko, and Wenying Yao for valuable comments and suggestions. Hersey acknowledges the support from Chih-Sheng Hsieh who provided some guidance and Matlab codes for the earlier draft. This work was supported with super computing resources provided by the Phoenix HPC service at The University of Adelaide. Doko Tchatoka Acknowledges the financial support from the Australian Research Council through the Discovery Grant DP200101498.

†Corresponding author contacts: School of Economics and Public Policy, The University of Adelaide, 10 Pulteney St, Adelaide SA 5005, Australia. Tel:+618 8313 1174, Fax:+618 8223 1460; e-mail: firmin.dokotchatoka@adelaide.edu.au

1 Introduction

The influence of social networks on decision-making behavior is a well-researched topic, with empirical studies covering a wide range of areas such as education (Kramarz and Skans, 2014; Sacerdote, 2001), health (Kremer and Levy, 2008; Christakis and Fowler, 2007), and financial behavior. The literature provides evidence that individuals are influenced by their peers when making decisions, and this influence can have significant implications for their outcomes.

Most studies have focused on a binary decision setting such as college attendance[1] or observed outcome framework such as labour market outcome.[2] However, little is still known on the role of peer effects in counterfactual analysis, specifically in identifying the marginal treatment effect (MTE) and the marginal policy-relevant effect (MRPTE). These effects are important for policymakers and practitioners in evaluating the effectiveness of interventions and policies; e.g., see Heckman and Vytlacil (1999, 2001); Heckman et al. (2006). Further research is needed to understand the mechanisms through which social networks influence decision-making, the conditions under which peer effects are most pronounced, and the extent to which these effects can be leveraged to improve outcomes. Recent advances in machine learning and network analysis techniques offer promising avenues for further research in this area.

This paper focuses on the generalized Roy model (Roy, 1951; Quandt, 1972) and studies the identification of the MTE and MPRTE when peers influence a binary treatment decision. Specifically we incorporate endogenous friendship network formation into the treatment decision, and estimate the latent treatment decision and network formation models simultaneously using Bayesian procedures. This novel approach allows for a better understanding of the role of peer effects in decision-making and outcomes.

The literature on peer effects (e.g., see Moffitt et al., 2001) has identified three types of peer influences: endogenous, exogenous, and correlated effects. The endogenous effect refers to the direct influence of an individual's peers on their outcome, while the exogenous effect captures the influence of the characteristics of an individual's peers on their outcome. The correlated effect embeds homophily and environmental effects. The three peer effects can jointly influence an individual's decision or affect their outcomes (Manski, 1993; Glaeser et al., 2003). However, identifying them simultaneously is challenging due to the "reflection problem,"[3] which assumes that individuals interact in groups where group members have an equal influence on their peers. Manski (2013) discusses the challenges associated with identifying the causal effects of treatment

[1] See Kramarz and Skans (2014); Sacerdote (2001).

[2] See Goldsmith-Pinkham and Imbens (2013b).

[3] i.e, the difficulty to empirically isolate endogenous peer effects from exogenous peer effects as both are often strongly correlated.

in the presence of social interactions, such as peer effects. He argues that standard methods used in the literature for estimating treatment effects, such as instrumental variables or regression discontinuity, may not accurately capture the true effects of treatment when social interactions are present.

Lee (2007); Bramoullé et al. (2009) introduce network fixed effects into the outcome equation to identify in order to control for the correlated effects. However, Goldsmith-Pinkham and Imbens (2013b); Hseih and Lee (2016) show that this approach does not control for the unobserved within-network characteristics that are important factors of both the decision to take on treatment and networks formation. Moreover, controlling for network effects in a binary treatment framework can be challenging due to the nonlinearity of the outcome variable (Blume et al., 2011). The standard probit estimation used in binary choice models normalizes the variance of the errors to unity in the treatment equation, which imposes restrictions on both the network formation and the peer effects. Researchers may instead use the linear probability approach (e.g., see Patacchini and Arduini, 2016; Calabrese and Elkink, 2014). This can be very problematic because it ignores the continuous nature of the underlying latent treatment decision variable.

The paper proposes a Bayesian estimation framework that can be used to jointly estimate the treatment decision and network formation models. The treatment decision model specifies the probability that an individual will receive treatment based on their observed characteristics and the latent treatment decision variable. The network formation model specifies the probability that two individuals will form a connection in the network based on their observed characteristics and the treatment decision variable. The framework allows for the filtering of the latent treatment decision variable based on observed characteristics, which makes it possible to use the Bayesian maximum likelihood method. This approach avoids the limitations imposed by probit estimation and takes into account the continuous nature of the latent treatment decision variable. Overall, the proposed Bayesian estimation framework provides a flexible and powerful approach for jointly modeling the treatment decision and network formation processes in social networks.

The remaining of the paper is organized as follows. Section 2 presents the conceptual theoretical framework, the econometric models for both the treatment decision and the friendship formation. Model assumptions and the estimation method are presented in Section 3. A Monte Carlo experiment is shown in Section 4 and the proposed methodology is applied to the educational attainment model in Section 5. Section 6 concludes.

2 Conceptual framework

We consider the canonical counterfactual model[4] with a binary treatment $S \in \{0,1\}$ and a scalar, real-valued outcome $Y \in \mathscr{Y} \subset \mathbb{R}$. The observed outcome production is modelled as

$$Y \quad = \quad Y_1 S + Y_0(1-S), \tag{2.1}$$

where (Y_0, Y_1) are the potential outcomes under no treatment and under treatment. The treatment decision is latent, i.e., we observe $S = 1$ if $S^* \geq 0$ and $S = 0$ if $S^* < 0$, where S^* represents the net benefit of receiving treatment. Let $X \in \mathscr{X} \subset \mathbb{R}^{d_X}$ denote a vector of exogenous pre-treatment covariates. In the returns to schooling application for instance, S is an indicator for enrollment in college, Y is log wage, and X includes observable individual characteristics which affect wages– such as parental education, location, appearance, age. The potential outcome model is given by

$$Y_j \quad = \quad \mu_j(X) + U_j, \; j = 0, 1; \tag{2.2}$$

$\mu_j(X) := \mathbb{E}[Y_j|X]$ and U_j is the error term affecting Y_j of the causation state $j \in \{0,1\}$.

We assume that the population of interest, $\mathscr{G}$, is comprised of m groups, i.e., $\mathscr{G} = \{\mathscr{G}_1, \ldots, \mathscr{G}_m\}$. For instance, the population of interest can be students and the groups are formed using school enrollments, whereby students from the same school belong to the same group. We suppose that each group $\mathscr{G}_g$, $g = 1, \ldots, m$, has a network (i.e., a group of individuals interacting for their collective or mutual interest) that is observed over time, and that these networks are independent across groups. Let $\boldsymbol{G}_{gt} : N_{gt} \times N_{gt}$ be the network of group g at time $t \in \{1, \ldots, T\}$, where N_{gt} is the number of individuals in network $\boldsymbol{G}_{gt}$ and T is the sample period.

2.1 Treatment model with network formation

2.1.1 Treatment model

Let S^*_{git} denote the latent treatment decision at time t of individual i in network $\boldsymbol{G}_{gt}$. Individual i in network g takes on treatment at time t (i.e., $S_{git} = 1$) if and only if $S^*_{git} \geq 0$. Let $S^*_{gt} = \left(S^*_{g1t}, \ldots, S^*_{gN_{gt}t}\right)'$ be the $N_{gt} \times 1$ vector of latent treatment decision variables. Assuming that the decision to take on treatment for an individual in a given network $\boldsymbol{G}_g, g = 1, \ldots, m$, is influenced by his network's peers, and focusing on the standard linear in parameters setting

[4] See Roy (1951); Quandt (1972); Carneiro et al. (2011).

(see Manski, 1993; Moffitt et al., 2001), we model the treatment decision as:

$$S^*_{gt} = \gamma_1 \boldsymbol{G}_{gt} S^*_{gt} + Z_{gt}\gamma_2 + \boldsymbol{G}_{gt} Z_{gt}\gamma_3 + \alpha_g \iota_{gt} + \gamma_4 \xi_{gt} - V_{gt} \tag{2.3}$$

where $S_{git} = 1$ iff $S^*_{git} \geq 0$, $i = 1, \ldots, N_{gt}$ and $t = 1, \ldots, T$, $Z_{gt} : N_{gt} \times K_2$ contains exogenous covariates (instruments), $\iota_{gt} : N_{gt} \times 1$ is a vector of ones; $\xi_{gt} : N_{gt} \times 1$ contains unobserved within network $\boldsymbol{G}_{gt}$ characteristics affecting the treatment decision, $V_{gt} : N_{gt} \times 1$ is a vector of disturbances. We assume that V_g is continuous with a common strictly increasing cumulative density function (cdf) $F_V(\cdot)$, and that V_g and ξ_g are independent given X_g.[5] Some covwariates in X_g may depend on $U_{jg}, j = 0, 1$, but all instruments in Z_g are strictly exogenous, i.e., Z_g is independent of $U_{jg}, j = 0, 1$, V_g, and ξ_g. In general, Z_{gt} contains all exogenous covariates in X_g.

Model (2.3) captures five types of effects. The first is the endogenous peer effect, measured by γ_1, representing the average treatment benefit that an individual receives from their network peers. The second is the direct effect (γ_2) which measures the direct impact of exogenous instruments in Z_{gt} on the decision to take on treatment. The third is the exogenous peer effect captures by the parameter γ_3. This effect represents the average impact of an individual's peers' characteristics on the outcome being studied. The fourth is the correlated or group fixed effect captures by the parameter α_g. And finally, the fifth is the effect of the unobserved within network characteristics captures by the parameter γ_4. Identifying all five effects simultaneously can be a challenging task. Therefore, it is important to carefully analyze and understand the variables at play, as well as the possible interactions between them, in order to accurately identify and measure each effect. Additionally, depending on the specific situation, some effects may be more important or influential than others, further complicating the identification process.

Standard treatment effect models only control for the direct effect[6] and, to some extent, the correlated effects. Other network features, such as the endogenous peer effect, the exogenous peer effect, and the effect of unobserved within-network characteristics, are often not controlled for. This potentially creates an omitted variable problem. By incorporating the influence of individuals' networks on their decision to take on treatment, specification (2.3) systematically addresses this issue. Importantly, the inclusion of ξ_{gt} not only controls for unobserved group factors but also addresses concerns related to self-selection into groups, as some groups may be formed based on their characteristics (Hseih and Lee, 2016).

Define $\mu_{S,g}(Z_{gt}, \boldsymbol{G}_{gt}) := \gamma_1 \boldsymbol{G}_{gt} S^*_{gt} + Z_{gt}\gamma_2 + \boldsymbol{G}_{gt} Z_{gt}\gamma_3 + \alpha_g \iota_{gt} + \gamma_4 \xi_{gt}$. Then the probability that individual i in network g receives treatment at time t, conditional on Z_{git} and $\boldsymbol{G}_{gt}$, is given

[5] It is important to note that V_g may still depend on the errors $U_{jg}, j = 0, 1$ of the outcome equation (2.2).
[6] See e.g. Heckman and Vytlacil (1999, 2001); Heckman et al. (2006).

from (2.3) by

$$P_{git}(Z_g, \boldsymbol{G}_{gt}) := \mathbb{P}\big[S_{git} = 1 | Z_{git}, \boldsymbol{G}_{gt}\big] = F_V\big[\mu_{S,g}(Z_{git}, \boldsymbol{G}_{gt})\big]. \tag{2.4}$$

$P_{git}(Z_g, \boldsymbol{G}_{gt})$ is the mean scale utility function (see e.g. McFadden, 1974) which plays a crucial role in the identification of MTE and MPRTE.

Let $U_S = F_V(V)$. U_S is uniform on $[0, 1]$ by construction and its realizations correspond to the quantiles of V. An individual i in network g will receive treatment at time t, conditional on Z_{git} and $\boldsymbol{G}_{gt}$, if its scale utility in doing so exceeds its U_S value, i.e.,

$$S_{git} = 1 \text{ iff } P_g(Z_{git}, \boldsymbol{G}_{gt}) \geq U_{S_{git}} = F_V(V_{git}). \tag{2.5}$$

Remark. Note that estimating (2.3) comes with challenges. On one hand, it differs fundamentally from the standard setting (in particular from Eq.(6.1) of Goldsmith-Pinkham and Imbens, 2013b) in the sense that the dependent variable S^*_{gt} is latent. This adds to the econometric challenge of estimating (2.3). Indeed, applying a probit estimation, for example, restricts the covariance matrix of the error $\eta_{gt} = (I_{N_{gt}} - \gamma_1 \boldsymbol{G}_{gt})^{-1} V_{gt}$ in the reduced-form[7] equation (2.6) to an identity matrix;

$$\begin{aligned} S^*_{gt} \;=\; & \alpha_g (I_{N_{gt}} - \gamma_1 \boldsymbol{G}_{gt})^{-1} \iota_{gt} + (I_{N_{gt}} - \gamma_1 \boldsymbol{G}_{gt})^{-1} Z_{gt}\gamma_2 + (I_{N_{gt}} - \gamma_1 \boldsymbol{G}_{gt})^{-1} \boldsymbol{G}_{gt} Z_{gt}\gamma_3 \\ & + \gamma_4 (I_{N_{gt}} - \gamma_1 \boldsymbol{G}_{gt})^{-1} \xi_{gt} - \eta_{gt}. \end{aligned} \tag{2.6}$$

Since η_{gt} depends on γ_1 and $\boldsymbol{G}_{gt}$, such restriction also constrains the formation of network $\boldsymbol{G}_{gt}$ as well as the endogenous peer effect measure γ_1. These identifying restrictions may not be sustained in observed data, and to avoid imposing them, we propose a data dependent method to filter the latent dependent variable S^*_{gt} from a truncated multivariate normal (TMVN) distribution, upon exploiting the Gibbs sampling technique in Geweke (1991).

2.1.2 Network formation

We assume that network $\boldsymbol{G}_g$ in (2.3) is dynamic, and to model friendship formation, we use the following framework of Goldsmith-Pinkham and Imbens (2013b).

Let $\boldsymbol{D}_{gt} = \big[D^{ij}_{gt}\big]_{1 \leq i,j \leq N_{gt}}$ be a symmetric adjacency matrix of network $\boldsymbol{G}_{gt}$: $D^{ii}_{gt} = 0$ for

[7] As long as $|\gamma_1| < 1$, (2.3) implies (2.6) (i.e., (2.6) is a reduced-form equation). To see it, note first that the determinant of the matrix $I_{N_{gt}} - \gamma_1 \boldsymbol{G}_{gt}$ is $\det(I_{N_{gt}} - \gamma_1 \boldsymbol{G}_{gt}) = \prod_{j=1} (1 - \gamma_1 \lambda^j_{gt})$, where $\lambda^j_{gt}, j = 1, \ldots, N_{gt}$, are the eigenvalues of $\boldsymbol{G}_{gt}$ and satisfy $-1 < \lambda^j_{gt} \leq 1$ (see e.g. Case, 1991, Footnote 5). Therefore, $I_{N_{gt}} - \gamma_1 \boldsymbol{G}_{gt}$ is invertible if and only if $|\gamma_1| < 1$; i.e., (2.3) can be expressed as (2.6) if $|\gamma_1| < 1$.

all i; and for all $i \neq j$, $D_{gt}^{ij} = 1$ if individuals i and j are friends and $D_{gt}^{ij} = 0$ otherwise. Let $\boldsymbol{M}_{gt} = \left(M_{gt}^{i}\right)_{1 \leq i \leq N_{gt}}$ denote a N_{gt}-dimensional vector with elements representing the number of individual i's friends[8] at time t (i.e., $M_{gt}^{i} = \sum_{j=1}^{N_{gt}} D_{gt}^{ij}$). At time t and given $\boldsymbol{D}_{gt}$, network $\boldsymbol{G}_{gt}$ is defined as a row-normalized adjacency matrix:

$$\boldsymbol{G}_{gt} \quad = \quad \text{diag}(\boldsymbol{M}_{gt})^{-1}\boldsymbol{D}_{gt} := \left[G_{gt}^{ij}\right]_{1 \leq i,j \leq N_{gt}}; \;\; G_{gt}^{ij} = D_{gt}^{ij}/M_{gt}^{i}. \tag{2.7}$$

Let U_{gt}^{ij} represent the net benefit of individual i to form a friendship link with individual j in network $\boldsymbol{G}_g$ at time t. We assume the decision of individuals i and j 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, so $D_{gt}^{ij} = \mathbb{1}[U_{gt}^{ij} > 0] \cdot \mathbb{1}[U_{gt}^{ji} > 0]$, where $\mathbb{1}[\cdot]$ is the indicator function. Following Goldsmith-Pinkham and Imbens (2013b) we model U_{gt}^{ij} as:

$$U_{gt}^{ij} = \theta_1' c_{gt}^{i} + \theta_2' c_{gt}^{j} + \theta_3' c_{gt}^{ij} + \theta_4 D_{g,t-1}^{ij} + \theta_5 F_{g,t-1}^{ij} + \delta|\xi_{gt}^{i} - \xi_{gt}^{j}| + \varepsilon_{gt}^{ij}, \tag{2.8}$$

where θ_k $(k = 1 \ldots, 5)$ and δ are unknown parameters; c_{gt}^{i} and c_{gt}^{j} represent the observed individual specific variables which may affect friendship formation at time t; and c_{gt}^{ij} represents dyad-specific variables which may be either dummy variables indicating the same characteristics between individuals i and j (e.g. race or sex) or the difference between two continuous individual characteristics (e.g. difference in age, difference in household income); ε_{gt}^{ij} is an error term assumed uncorrelated with the unobserved within network characteristics ξ_{gt}^{i} and ξ_{gt}^{j}, the observed individual specific characteristics c_{gt}^{i}, c_{gt}^{j} and c_{gt}^{ij}, as well as the error V_{gt} of the treatment model (2.3). Both $\boldsymbol{D}_{g,t-1}^{ij}$ and $\boldsymbol{F}_{g,t-1}^{ij}$ characterize the network in the previous period. In particular, $\boldsymbol{D}_{g,t-1}^{ij}$ is a dummy variable indicating whether i and j were friends in the previous period, while $\boldsymbol{F}_{g,t-1}^{ij}$ is a dummy variable representing whether i and j had friends in common in the previous period. Differences in the unobserved characteristics at time t towards every potential friendship (i, j) (i.e., $|\xi_{gt}^{i} - \xi_{gt}^{j}|$) are some of the key factors of friendship formation. In particular, a low value of $|\xi_{gt}^{i} - \xi_{gt}^{j}|$ is likely to result in a friendship pair (i, j) at time t, while a higher value indicates that i and j are unlikely to be friends at time t. As such, the parameter δ measures the intensity of friendship and is expected to be negative. In (2.8), ξ_{gt}^{i} is a scalar for a given i but the model could be generalized to account for multivariate unobserved factors (see e.g., Hseih and Lee, 2016, Eq.(2)).

Remark. Model (2.8) describes a dynamic network formation, so, at least *two data points* on the networks are required to fully identify friendship relations. While this dynamic network

[8]Individuals with no friends are discounted from the model so that $M_{gt}^{i} > 0$.

formation may be simplistic (as highlighted by Jackson, 2013) because it implies that an individual's preference for a relationship with another person does not depend on any of the network structure except for whether they had some friends in common in an earlier period, a more complex specification of (2.8) may not be tractable at least from an econometric estimation viewpoint (Goldsmith-Pinkham and Imbens, 2013a). Despite its simplicity, model (2.8) allows to investigate a range of issues related to endogeneity of networks, measurement error in links, and heterogeneity in peer effects. Many factors contributing to minimize a potential bias in the estimates are accounted for. In particular, the effect of homophily is controlled through the inclusion of the unobserved characteristics ξ_{gt}, while transitivity in link formation is controlled through the inclusion of $\boldsymbol{F}_{g,t-1}$. Also, Graham (2013) argues that a network formation model should also account for degree heterogeneity– some individuals are naturally "good friends" and thus give greater utility to friendship. In (2.8), degree heterogeneity is controlled through the inclusion of the observed characteristics of the networks (i.e., c^i_{gt}, c^j_{gt} for all (i,j)). Finally, although there is no explicit cost for friendship formation in (2.8), it can be introduced by allowing for cost functions that depend on $\boldsymbol{G}_{gt}$ and capture a *capacity constraint* whereby no node has degree exceeding some positive finite constant $\bar{D}$ (see e.g., Leung, 2014).

The identification of notwork formation models often requires the assumptions under which asymptotic theory (Weak Law of Large Numbers (WLLN) and Central Limit Theorem (CLT)) is valid for the network statistics. In particular, let $\mathscr{N}_{gt} = \{(i,j) \in \boldsymbol{G}_{gt}\}$ denote the set of nodes of network $\boldsymbol{G}_g$ at time t and define the node statistic $\frac{1}{N_{gt}} \sum_{i=1}^{N_{gt}} \psi_i(\boldsymbol{G}_{gt})$, where $\psi_i(\boldsymbol{G}_{gt})$ is a function depending on network $\boldsymbol{G}_{gt}$ only through the set of links between all nodes connected to node i. The identification of (2.8) boils down to the condition under which the WLLN and the CLT are valid for $\frac{1}{N_{gt}} \sum_{i=1}^{N_{gt}} \psi_i(\boldsymbol{G}_{gt})$; see Leung (2014). In our network model (2.8), we are particularly interested in the application of the WLLN and the CLT to network moments corresponding to the probability of a friendship linkage, i.e., $\mathbb{P}[D^{ij}_{gt} = 1]$. Chandrasekhar (2016); Leung (2014) show that this generally requires the sequence $\{\psi_i(\boldsymbol{G}_{gt}), i \in \mathscr{N}_{gt}\}$ to be α-mixing, so that individuals' decision to form links in a given network are sufficiently uncorrelated. This α-mixing property can be achieved under the following conditions.

1. **No coordination**. Let $\boldsymbol{W}^{ij}_{gt} = \{c^i_{gt}, c^j_{gt}, c^{ij}_{gt}, D^{ij}_{g,t-1}, D^{ji}_{g,t-1}, F^{ij}_{g,t-1}, F^{ji}_{g,t-1}, \xi^i_{gt}, \xi^j_{gt}\}$. From (2.8), the net benefit of individual i to form a friendship link with individual j in network $\boldsymbol{G}_g$ at time t is $U^{ij}_{gt}(\boldsymbol{W}^{ij}_{gt}; \theta, \delta)$. Given θ, δ and $\boldsymbol{W}_g$, no coordination means that the sets of nodes that are not connected under any equilibrium make friendship linkage decisions independently, i.e., isolated friendship networks have no incentive to coordinate on their friendship decisions.

2. **Homophily**. Irrespective of endogenous network effects on the net benefit of an individual to form a friendship with his peers, it is typically not worthwhile to form a friendship directly to someone who is very far away in homophilic characteristics.

Remark. Let $d(i,j)$ represents differences in characteristics in which individuals display homophily. For example, in our empirical application on high school friendship networks formation model, $d(i,j)$ contains elements such as differences in age ($|AGE_g^i - AGE_g^j|$), Math Grade ($|MathGD_g^i - MathGD_g^j|$), etc., between individuals i and j. As long as individuals are typically unlikely to form a link with those far away in these characteristics, the homophily condition will hold so that we can express $U_{gt}^{ij}(\boldsymbol{W}_{gt}^{ij};\theta,\delta)$ as

$$U_{gt}^{ij}(\boldsymbol{W}_{gt}^{ij};\theta,\delta) = U_{gt}^{ij}(\underbrace{d(i,j)}_{distance}, \underbrace{z(\boldsymbol{G}_{gt},w)}_{endogenous}, \underbrace{f(w)}_{exogenous};\theta,\delta) \tag{2.9}$$

where $z(\cdot)$ represents non-homophilic endogenous network effects (which may depend on the network $\boldsymbol{G}_{gt}$) and $f(\cdot)$ represents non-homophilic exogenous network effects. As such the homophily assumption implies that:

$$\lim_{d(i,j)\to\infty} U_{gt}^{ij}(d(i,j),\bar{z},f(w);\bar{\theta},\bar{\delta}) < 0 \tag{2.10}$$

for some $f(w) \leq \bar{f}$ and the probability of $f(w) > \bar{f}$ is sufficiently low.

3. **Thin Tails**. The distribution of the error term ε_{gt}^{ij} in the network formation model (2.8) has thin tails such that $\mathbb{P}[\varepsilon_{gt}^{ij} > r] \leq Ce^{-\kappa r}$ for some $C > 0$ and $\kappa > 0$. This condition is satisfied in our model as ε_{gt}^{ij} is assume to follow a logistic distribution.

4. **Increasing Domain**. $\sup_g \max_{i\in\mathcal{N}_{gt}}(\#j \in \mathcal{N}_{gt} : |d(i,j) < r_0|) < \infty$, for some $r_0 > 0$. This property says that the largest number of individuals that are at most a distance r_0 away from any other individual is finite. Otherwise, communities are small relative to the overall network.

5. **Diversity**. For any individual i and distance r, there exists a set of nodes S containing i such that any $k \in S$ satisfies $d(i,k) < r$ and $d(k,l) \geq 9\kappa^{-1}\log r$ for $l \notin S$ and κ is given in the **Thin Tails** condition (i.e., condition 3. above). This condition ensures that there is sufficient diversity in the distance characteristics. Diversity ensures that the α-missing coefficient $\alpha(i,j)$ decays at a sufficiently fast rate to guarantee the validity of the CLT. This is a sufficient condition to prove that $\mathbb{P}[D_{gt}^{ij} = 1] \to 0$ as $d(i,j) \to \infty$. In our empirical application on high school friendship networks formation in the US, we

use several homophilic indicators such as parents education, personality variables, School achievement, age, race and gender to ensure that there is sufficient diversity in our network formation model.

3 Estimation

We implement the Bayesian maximum likelihood estimation that requires specifying the likelihood functions of data and parameters.

3.1 Likelihood functions

Let $\boldsymbol{W}_{gt} = \{\boldsymbol{W}_{gt}^{ij} | (i,j) \in \boldsymbol{G}_{gt}\}$ and $\boldsymbol{C}_{gt} = \{c_{gt}^{i}, c_{gt}^{j}, c_{gt}^{ij} | (i,j) \in \boldsymbol{G}_{gt}\}$, where

$$\boldsymbol{W}_{gt}^{ij} = \{c_{gt}^{i}, c_{gt}^{j}, c_{gt}^{ij}, D_{g,t-1}^{ij}, D_{g,t-1}^{ji}, F_{g,t-1}^{ij}, F_{g,t-1}^{ji}, \xi_{gt}^{i}, \xi_{gt}^{j}\}.$$

In order to derive the likelihood functions, we make the following assumptions.

Assumption 3.1.

(*i*) *Given* $\boldsymbol{W}_{gt}$, *each link of network* $\boldsymbol{G}_{gt}$ *is independent of other links.*

(*ii*) *The errors* ε_{gt} *in* (2.8) *are i.i.d. across pairs* (i,j), *and are uncorrelated with both* $\boldsymbol{W}_{gt}$ *and the errors* (U_{0t}, U_{1t}) *in* (2.2). *In addition, the conditional distribution of* ε_{gt}^{ij}, *given* $\boldsymbol{W}_{gt}$, *is a standard logistic distribution for all pairs* (i,j).

(*iii*) $(V_{gt}^{i}, \xi_{gt}^{i})' \mid \boldsymbol{D}_{g,t-1}, \boldsymbol{F}_{g,t-1}, \boldsymbol{C}_{gt} \overset{i.i.d.}{\sim} N\left(0, \begin{bmatrix} \sigma_{V_g}^2 & 0 \\ 0 & 1 \end{bmatrix}\right)$ *for all* g *and* t.

Assumptions 3.1-(i):(iii) are commonly used in the social networks literature; see e.g. (Goldsmith-Pinkham and Imbens, 2013b; Hseih and Lee, 2016). Assumption 3.1-(i) implies that an individual's preference for a relationship with another individual does not depend on any of the network structure except for whether they had some friends in common in an earlier period. The current network formation can thus be inferred link-by-link, which simplifies the computational burden but can be restrictive due to homophily, transitivity of relations, or clustering (Jackson, 2008). To minimize any potential bias this may cause, we address the problems of: (1) homophily by controlling for the unobserved network within characteristics (ξ_{gt}); (2) transitivity of relations by controlling for the characteristics of the network in the previous period ($\boldsymbol{F}_{gt}$); and (3) degree heterogeneity by controlling for the network observed characteristics (c_{gt}^{i}, c_{gt}^{j}).

Under Assumption 3.1, a given network $\boldsymbol{G}_g$ may shrink or enlarge over time, although new friendship relations are assumed not to violate the independence of links. This may also be

restrictive as new friendship relations may be motivated and enabled by dyad-specific characteristics such as race, wealth, or gender considerations (Jackson, 2013), which must then be controlled for in (2.8).

Assumption 3.1-(ii) requires the errors of the network formation model (2.8) be uncorrelated with all other regressors in that model, as well as the errors of the potential outcome equation (2.2) and the latent treatment decision model (2.3). A similar assumption can be found in Goldsmith-Pinkham and Imbens (2013b) and Hseih and Lee (2016). Note that the errors of the potential outcomes model (2.2) may still be correlated with that of the latent treatment decision model (2.3).

Assumption 3.1-(iii) enables the implementation of the ML method. Note that under this assumption, the covariates X_g of the potential outcome equation (2.2) need not be strictly exogenous. The variance of ξ_g^i is normalized to unity because it is not identified neither in the treatment model (2.3) nor in the friendship formation model (2.8); see Hseih and Lee (2016). As such, the parameters of both models can only be identified up to the variance of ξ_g^i.

Let $\theta = (\theta_1', \theta_2', \theta_3', \theta_4, \theta_5)'$ and define

$$\psi_{gt}^{ij}(\theta, \delta) := \theta_1' c_{gt}^i + \theta_2' c_{gt}^j + \theta_3' c_{gt}^{ij} + \theta_4 D_{g,t-1}^{ij} + \theta_5 F_{g,t-1}^{ij} + \delta |\xi_{gt}^i - \xi_{gt}^j|. \tag{3.1}$$

Given $\boldsymbol{W}_{gt}^{ij}$, the probability of friendship formation, $\mathbb{P}[D_{gt}^{ij} = 1 | \boldsymbol{W}_{gt}^{ij}] = \mathbb{P}[U_{gt}^{ij} > 0, U_{gt}^{ji} > 0 | \boldsymbol{W}_{gt}^{ij}$, is captured[9] under Assumption 3.1 by:

$$q_g\big(D_{gt}^{ij} | \boldsymbol{W}_{gt}^{ij}; \theta, \delta\big) := q_g(\theta, \delta) \quad = \quad \boldsymbol{\Lambda}\Big(\psi_{gt}^{ij}(\theta, \delta)\Big)\boldsymbol{\Lambda}\Big(\psi_{gt}^{ji}(\theta, \delta)\Big), \tag{3.2}$$

where $\boldsymbol{\Lambda}(\cdot)$ is the cdf of the standard logistic random variable.

Then, the likelihood function of $\boldsymbol{D}_{gt}$ (or network $\boldsymbol{G}_{gt}$), conditional on $\boldsymbol{W}_{gt}$ is:

$$P^{\xi_g}\big(\boldsymbol{G}_{gt} | \boldsymbol{W}_{gt}; \theta, \delta\big) = \prod_{i \neq j} \Big[q_g(\theta, \delta)\Big]^{D_{gt}^{ij}} \Big[1 - q_g(\theta, \delta)\Big]^{1 - D_{gt}^{ij}}. \tag{3.3}$$

Since $\boldsymbol{W}_{gt}^{ij}$ includes the unobserved within network characteristics ξ_{gt}^i, $P^{\xi_g}\big(\boldsymbol{G}_{gt} | \boldsymbol{W}_{gt}; \theta, \delta\big)$ depends on ξ_{gt}^i. Therefore, the likelihood function of network $\boldsymbol{G}_{gt}$, conditional on $\boldsymbol{W}_{gt} \setminus \{\xi_{gt}\}$, is obtained by integrating (3.3) over ξ_{gt}, i.e.

$$P\big(\boldsymbol{G}_{gt} | \boldsymbol{W}_{gt} \setminus \{\xi_{gt}\}; \theta, \delta\big) = \int_{\xi_{gt}} P^{\xi_g}\big(\boldsymbol{G}_{gt} | \boldsymbol{W}_{gt}; \theta, \delta\big) \phi(\xi_{gt}) d\xi_{gt}, \tag{3.4}$$

[9] See also Goldsmith-Pinkham and Imbens (2013b); Hseih and Lee (2016).

where $\phi(\cdot)$ is the density function (pdf) of a $N(0, I_{N_{gt}})$ random variable.

Similarly, we can use Bayes rule to express the joint likelihood function of the latent treatment decision S^*_{gt} and the network $\boldsymbol{G}_{gt}$ conditional on Z_{gt} and $\boldsymbol{W}_{gt}$ as:

$$\begin{aligned} &P^{\xi_g}\big(S^*_{gt}, \boldsymbol{G}_{gt} | Z_{gt}, \boldsymbol{W}_{gt}; \gamma, \alpha_g, \sigma^2_{V_g}, \theta, \delta\big) \\ &\qquad = P^{\xi_g}\big(S^*_{gt} | Z_{gt}, \boldsymbol{W}_{gt}; \gamma, \alpha_g, \sigma^2_{V_g}\big) \times P^{\xi_g}\big(\boldsymbol{G}_{gt} | \boldsymbol{W}_{gt}; \theta, \delta\big), \end{aligned} \tag{3.5}$$

where $P^{\xi_g}\big(S^*_{gt} | Z_{gt}, \boldsymbol{W}_{gt}; \gamma, \alpha_g, \sigma^2_{V_g}\big)$ is the likelihood function of S^*_{gt} conditional on Z_{gt} and $\boldsymbol{W}_{gt}$. Thus, the joint likelihood function of S^*_{gt} and $\boldsymbol{G}_{gt}$, conditional on Z_{gt} and $\boldsymbol{W}_{gt} \setminus \{\xi_{gt}\}$ is also obtained by integrating (3.5) over ξ_{gt}, i.e.

$$\begin{aligned} &P\big(S^*_{gt}, \boldsymbol{G}_{gt} | Z_{gt}, \boldsymbol{W}_{gt} \setminus \{\xi_{gt}\}; \gamma, \alpha_g, \sigma^2_{V_g}, \theta, \delta\big) \\ &\qquad = \int_{\xi_{gt}} P^{\xi_g}\big(S^*_{gt} | Z_{gt}, \boldsymbol{W}_{gt}; \gamma, \alpha_g, \sigma^2_{V_g}\big) \times P^{\xi_g}\big(\boldsymbol{G}_{gt} | \boldsymbol{W}_{gt}; \theta, \delta\big) \phi(\xi_{gt}) d\xi_{gt}. \end{aligned} \tag{3.6}$$

Note that given $\boldsymbol{W}_{gt}$ and $(\theta', \delta)'$, $P^{\xi_g}\big(\boldsymbol{G}_{gt} | \boldsymbol{W}_{gt}; \theta, \delta\big)$ can be evaluated. However, $P^{\xi_g}\big(S^*_{gt} | Z_{gt}, \boldsymbol{W}_{gt}; \gamma, \alpha_g, \sigma^2_{V_g}\big)$ cannot be evaluated even conditional on $(Z_{gt}, \boldsymbol{W}_{gt})$ and $(\gamma', \alpha_g, \sigma^2_{V_g})'$. This is because S^*_{gt} is not observed. Therefore, we propose a methodology to filter the latent index S^*_{gt} first, and then use these values to evaluate (3.6).

3.2 Filtering the latent treatment variable

We employ the Gibbs sampling technique of Geweke (1991) to simulate S^*_{gt} from a multivariate truncated normal (TMVN) distribution. We then apply the truncated regression method to derive $P^{\xi_g}\big(S^*_{gt} | Z_{gt}, \boldsymbol{W}_{gt}; \gamma, \alpha_g, \sigma^2_{V_g}\big)$. This approach also allows us to filter the unobserved within network characteristics ξ_{gt} within the Bayesian framework, and to condition on these filtered values in the estimation so that no numerical integration over all $\xi^i_{gt}, i = 1, \ldots, N_{gt}$ (which would have been computationally demanding) is required. To expand further, observe first from (2.6) that

$$S^*_{gt} \mid Z_{gt}, \boldsymbol{W}_{gt} \sim TMVN(\mu_{gt}, \Sigma_{\eta_{gt}}) \tag{3.7}$$

under Assumption 3.1-(i)&(iii), where $\Sigma_{\eta_{gt}} := \sigma^2_{V_g}\big[(I_{N_{gt}} - \gamma_1 \boldsymbol{G}_{gt})'(I_{N_{gt}} - \gamma_1 \boldsymbol{G}_{gt})\big]^{-1}$ is the covariance matrix of η_{gt} and $\mu_{gt} = \mathbb{E}\big(S^*_{gt} | Z_{gt}, \boldsymbol{W}_{gt}; \gamma, \alpha_g\big)$ is given by

$$\begin{aligned} \mu_{gt} \;=\; & \alpha_g (I_{N_{gt}} - \gamma_1 \boldsymbol{G}_{gt})^{-1} \iota_{gt} + (I_{N_{gt}} - \gamma_1 \boldsymbol{G}_{gt})^{-1} Z_{gt} \gamma_2 + (I_{N_{gt}} - \gamma_1 \boldsymbol{G}_{gt})^{-1} \boldsymbol{G}_{gt} Z_{gt} \gamma_3 \\ & + \gamma_4 (I_{N_{gt}} - \gamma_1 \boldsymbol{G}_{gt})^{-1} \xi_{gt}. \end{aligned} \tag{3.8}$$

The Geweke's (1991) method is based on the well-known result that sampling S^*_{gt} from a $TMVN(\mu_{gt}, \Sigma_{\eta_{gt}})$ subject to the inequality constraints $a \leq S^*_{gt} \leq b$ is equivalent to sampling τ_{gt} from $N(0, \Sigma_{\eta_{gt}})$ under the linear constraints $\underline{b} \leq \tau_{gt} \leq \bar{b}$, where $\underline{b} = a - \mu_{gt}$, $\bar{b} = b - \mu_{gt}$, and then constructing the sample for S^*_{gt} as $S^*_{gt} = \mu_{gt} + \tau_{gt}$. We apply this algorithm to build the sample for τ_{gt} from the conditional distribution of τ^i_{gt} given $\tau^{(-i)}_{gt}$ for all $i = 1, \ldots, N_{gt}$, where $\tau^{(-i)}_{gt} = \tau_{gt} \setminus \tau^i_{gt}$ denotes the vector formed with the components of τ_{gt} other than τ^i_{gt}. It is easy to show that

$$\mathbb{E}(\tau^i_{gt}|\tau^{(-i)}_{gt}) = \gamma^{(-i)}_{gt}\tau^{(-i)}_{gt} \quad \text{with} \quad \gamma^{(-i)}_{gt} = -(\omega^{ii}_{gt})^{-1}\omega^{(-i)}_{gt}, \tag{3.9}$$

where ω^{ij}_{gt} is the (i,j)th element of $\Sigma^{-1}_{\eta_{gt}}$ and $\omega^{(-i)}_{gt}$ is the ith row of $\Sigma^{-1}_{\eta_{gt}}$ excluding the ith element. Therefore, τ^i_{gt} can be generated as:

$$\tau^i_{gt} = \gamma^{(-i)}_{gt}\tau^{(-i)}_{gt} + h^i_{gt}\nu^i_{gt}, \tag{3.10}$$

where $h^i_{gt} = (\omega^{ii}_{gt})^{-1/2}$ and $\nu^i_{gt} \sim N(0,1)$ for all i and g. As $\underline{b} \leq \tau_{gt} \leq \bar{b}$, it follows from (3.10) that ν^i_{gt} satisfies the constraints:

$$(h^i_{gt})^{-1}(\underline{b}^i - \gamma^i_{gt}\tau^{(-i)}_{gt}) < \quad \nu^i_{gt} < (h^i_{gt})^{-1}(\bar{b}^i - \gamma^{(-i)}_{gt}\tau^{(-i)}_{gt}), \tag{3.11}$$

where $\underline{b}^i = -\infty$, $\bar{b}^i = -\mu^i_{gt}$ if $S^i_{gt} = 0$ (i.e. $S^{*^i}_{gt} \leq 0$) and $\underline{b}^i = -\mu^i_{gt}$, $\bar{b}^i = +\infty$ if $S^i_{gt} = 1$ (i.e. $S^{*^i}_{gt} > 0$). With the restrictions in (3.11), $\nu^i_{gt} \sim N(0,1)$ can be simulated and τ^i_{gt} can be generated following (3.10). Thus the sample for $S^*_{gt} = \mu_{gt} + \tau_{gt}$ can be built using this approach.

Now, we can derive the conditional distribution of S^*_{gt} in (3.6) by applying the above truncation method to (2.6). Specifically, let $\Sigma_{\eta_{gt}} = \left(\Sigma^{ij}_{\eta_{gt}}\right)_{1\leq i,j\leq N_{gt}}$ and $\mu_{gt} = \left(\mu^i_{gt}\right)_{1\leq i\leq N_{gt}}$, where $\Sigma_{\eta_{gt}}$ and μ_{gt} are defined in (3.7)-(3.8). Then, the conditional truncated normal distribution of $S^*_{gt} = \left(S^{*^i}_{gt}\right)_{1\leq i\leq N_{gt}}$ given Z_{gt} and $\boldsymbol{W}_{gt}$ is:

$$S^{*^i}_{gt}|Z_{gt}, \boldsymbol{W}_{gt} \sim \begin{cases} N\left(\mu^i_{gt}, \Sigma^{ii}_{\eta_{gt}}\right) \text{ truncated at the left by 0 if } S^i_{gt} = 1 \\ \\ N\left(\mu^i_{gt}, \Sigma^{ii}_{\eta_{gt}}\right) \text{ truncated at the right by 0 if } S^i_{gt} = 0, \end{cases} \tag{3.12}$$

for all $i = 1, \ldots, N_{gt}$. Letting $\Phi(\cdot)$ denote the cdf of the standard normal random variable, we

can express the conditional density of $S_{gt}^{*^i}|Z_{gt}, \boldsymbol{W}_{gt}$ from (3.12) as:

$$f(s_{gt}^{*^i}|Z_{gt}, \boldsymbol{W}_{gt}; \gamma, \alpha_g, \sigma_{V_g}^2) = \begin{cases} \frac{\phi\left(\tilde{s}_{gt}^{*^i} - \tilde{\mu}_{gt}^i\right)}{\Sigma_{\eta_{gt}}^{ii\,1/2}\Phi(\tilde{\mu}_{gt}^i)} \text{ truncated from above at 0 if } S_{gt}^i = 1 \\ \\ \frac{\phi\left(\tilde{s}_{gt}^{*^i} - \tilde{\mu}_{gt}^i\right)}{\Sigma_{\eta_{gt}}^{ii\,1/2}\left(1-\Phi(\tilde{\mu}_{gt}^i)\right)} \text{ truncated from below at 0 if } S_{gt}^i = 0, \end{cases} \tag{3.13}$$

where $\tilde{\mu}_{gt}^i = \Sigma_{\eta_{gt}}^{ii}\mu_{gt}^i$ and $\tilde{s}_{gt}^{*^i} = \Sigma_{\eta_{gt}}^{ii\,-1} s_{gt}^{*^i}$. The likelihood function $P^{\xi_g}\left(S_{gt}^*|Z_{gt}, \boldsymbol{W}_{gt}; \gamma, \alpha_g, \sigma_{V_g}^2\right)$ is then given by

$$\begin{aligned} P^{\xi_g}\left(S_{gt}^*|Z_{gt}, \boldsymbol{W}_{gt}; \gamma, \alpha_g, \sigma_{V_g}^2\right) &= \prod_{i=1}^{N_{gt}} f(s_{gt}^{*^i}|Z_{gt}, \boldsymbol{W}_{gt}; \gamma, \alpha_g, \sigma_{V_g}^2) \\ &= \prod_{\{i:\, s_{gt}^i=1\}} \frac{\phi\left(\tilde{s}_{gt}^{*^i} - \tilde{\mu}_{gt}^i\right)}{\Sigma_{\eta_{gt}}^{ii\,1/2}\Phi(\tilde{\mu}_{gt}^i)} \prod_{\{i:\, s_{gt}^i=0\}} \frac{\phi\left(\tilde{s}_{gt}^{*^i} - \tilde{\mu}_{gt}^i\right)}{\Sigma_{\eta_{gt}}^{ii\,1/2}\left(1-\Phi(\tilde{\mu}_{gt}^i)\right)}. \end{aligned} \tag{3.14}$$

Given Z_{gt}, $\boldsymbol{W}_{gt}$, γ, α_g, and $\sigma_{V_g}^2$, $P^{\xi_g}\left(S_{gt}^*|Z_{gt}, \boldsymbol{W}_{gt}; \gamma, \alpha_g, \sigma_{V_g}^2\right)$ can be evaluated using (3.14). Then, the joint likelihood function of interest (3.6) can be evaluated given Z_{gt}, $\boldsymbol{W}_{gt}$, γ, α_g, $\sigma_{V_g}^2$, θ and δ.

We implement a joint estimation of the network formation and treatment models using a Bayesian method.[10] The algorithm of this estimation for a two period ($T = 2$) friendship formation model is presented in Section B of the Supplemental Appendix.

3.3 MTE and MPRTE measures

For an individual in network $\boldsymbol{G}_g$, the marginal treatment effect (MTE), conditional on $X_g = x_g$ and $U_{Sg} = u_{Sg}$, is given by

$$MTE_{\boldsymbol{G}_g}\left(x_g, u_{Sg}\right) = \mathbb{E}(Y_{1g} - Y_{0g}|X_g = x_g, U_{Sg} = u_{Sg}), \tag{3.15}$$

where $Y_{1g} - Y_{0g} = \mu_1(X_g) - \mu_0(X_g) + U_{1g} - U_{0g}$ measures the return to treatment.[11] Clearly, $MTE_{\boldsymbol{G}_g}\left(x_g, u_{Sg}\right)$ is influenced by networks' characteristics through u_{Sg} values. As $U_{Sg} \sim \mathscr{U}_{[0,1]}$, $MTE_{\boldsymbol{G}_g}\left(x_g, u_{Sg}\right)$ varies across u_{Sg} values even when evaluated at mean X_g values. Therefore, tracing $MTE_{\boldsymbol{G}_g}\left(x_g, u_{Sg}\right)$ over the values of u_{Sg} given $X_g = x_g$ indicates how much the return to treatment varies with different quantiles of the treatment equation error for individuals willing to take on treatment. clearly, not controlling for the effects of network $\boldsymbol{G}_g$ on the decision to take on treatment can lead to misidentifying u_{Sg}, and thus inconsistently estimating

[10] See also Goldsmith-Pinkham and Imbens, 2013b; LeSage and Pace, 2009; Hseih and Lee, 2016.

[11] Note that $MTE_{\boldsymbol{G}_g}\left(x_g, u_{Sg}\right)$ is the mean return to treatment conditional on $X_g = x_g$ for individuals with mean scale utility value $P(Z_g, \boldsymbol{G}) = u_{Sg}$ who are indifferent between being treated or not.

$MTE_{\boldsymbol{G}_g}(x_g, u_{Sg})$. The practical estimation of the MTE is often done using he method of local instrumental variables (see Heckman and Vytlacil, 1999, 2001), thus we shall adapt this methodology in this study.

The MPRTE is defined by considering a class of policies that change the probability of treatment without affecting directly the potential outcomes and the error term in the treatment decision model. Precisely, let $\tilde{S}$ denote the treatment choice after the policy change and $\tilde{P}$ the corresponding probability of $\tilde{S} = 1$ after the policy change, i.e., $\tilde{S} = \mathbb{1}[\tilde{P} \geq U_S]$. Also, let $\tilde{Y} = \tilde{S}Y_1 + (1 - \tilde{S})Y_0$ be the observed outcome under the alternative policy. The mean effect of the change from a baseline policy to an alternative policy per net individual shift is the policy relevant treatment effect (PRTE) defined (see e.g. Heckman and Vytlacil, 2005), given $\boldsymbol{G}_g$ and $X_g = x_g$, as :

$$PRTE_{\boldsymbol{G}_g}(x_g) = \int_0^1 MTE_{\boldsymbol{G}_g}(x_g, u_{Sg})\omega_{PRTE}(u_{Sg})du_{Sg}, \tag{3.16}$$

where $\omega_{PRTE}(u_{Sg})) \equiv \omega_{PRTE}(u_{Sg}) = \frac{F_P(u_{Sg}) - F_{\tilde{P}}(u_{Sg})}{\mathbb{E}_{F_{\tilde{P}}}(P) - \mathbb{E}_{F_P}(P)}$, $F_P(\cdot)$ and $F_{\tilde{P}}(\cdot)$ denote the distributions of P and $\tilde{P}$ respectively. In (3.16), it is implicitly assumed that $\mathbb{E}_{F_{\tilde{P}}}(P) \neq \mathbb{E}_{F_P}(P)$, i.e., the fraction of individuals switching into treatment is not exactly offset by the fraction of those leaving the treatment, as $PRTE_{\boldsymbol{G}_g}(x_g)$ is ill-defined otherwise. This condition insures that the policy has a monotonic effect on participation. Integrating over the full unit interval $[0, 1]$– i.e., identifying the PRTE in the data– requires that the support of the propensity scores be $[0, 1]$. This is often not possible because PRTE is defined for discrete changes from a baseline policy to a fixed alternative policy. Heckman and Vytlacil (2005) suggest the marginal version of PRTE in (3.16), MPRTE, which measures a marginal change from the baseline policy. Specifically, consider a sequence of policies indexed by a scalar parameter α, where $\alpha = 0$ indicates the baseline policy. Let P_α denote the probability of treatment with the policy α, where P_0 is computed from the data. Define the corresponding PRTE for each α from the baseline policy ($\alpha = 0$) to policy α as per (3.16). Then, MPRTE is the limit of such a sequence of PRTEs as α goes to zero.

3.4 Implementation of the Bayesian method

We implemented a joint estimation of the network formation and treatment models using a Bayesian framework. To simplify the analysis, we used a two-period friendship formation model and labeled the initial and last periods as "period 0" and "period 1," respectively. For ease of notation, we omitted the time index in period 1 for the time-varying variables and parameters.

We normalize the variance of the unobserved network within characteristics as $\sigma^2_{V_g} = 1$. As a result, the unknown parameters of the models are $\theta = (\theta_1', \theta_2', \theta_3', \theta_4, \theta_5)'$, $\gamma = (\gamma_1, \gamma_2', \gamma_3', \gamma_4)'$, α_g and δ. As the estimation method is essentially a full information bayesian maximum likelihood, we specify prior distributions for these parameters following Hseih and Lee (2016); LeSage and Pace (2009). Specifically, we assume that

$$\gamma_1 \sim \mathscr{U}[-1,1],\ \bar{\gamma} = (\gamma_2', \gamma_3', \gamma_4, \alpha_g)' \sim N(\gamma_0, \Gamma_0);\ (\theta', \delta)' \sim N(\phi_0, \Phi_0), \tag{3.17}$$

where ϕ_0 and γ_0 are fixed vectors with appropriate sizes, Φ_0 and Γ_0 are fixed matrices with appropriate dimensions, and $\mathscr{U}[-1,1]$ stands for uniform distribution with support $[-1,1]$. The prior distribution of γ_1 is restricted to $[-1,1]$ to ensure that $I_{N_{gt}} - \gamma_1 \boldsymbol{G}_g$ is nonsingular almost surely. The prior distribution of ξ_g is

$$\xi_g^i \sim N(0,1),\ i = 1, \ldots, N_g. \tag{3.18}$$

Given (3.17)-(3.18) and observed sample on Z_g, $\boldsymbol{C}_g$, $\boldsymbol{D}_g$, the conditional prior distribution of the latent treatment variable S_g^* is given in (3.14). Then, the key posterior distributions needed in the Markov chain Monte Carlo (MCMC) algorithm are constructed sequentially (given covariates Z_g and X_g, and network characteristics $\boldsymbol{C}_g$, $\boldsymbol{D}_{g0}$, and $\boldsymbol{F}_{g0}$) as follows.

1. First, we construct the (conditional) posterior distribution of ξ_g^i as:

$$P(\xi_g^i | S_g^*, \boldsymbol{G}_g, \xi_g \setminus \{\xi_g^i\}, \bar{\gamma}, \theta, \delta, \sigma_v, \alpha_g) \sim \phi(\xi_g^i) \cdot P(S_g^*, \boldsymbol{G}_g | \xi_g, \bar{\gamma}, \theta, \delta, \alpha_g), \tag{3.19}$$

   where $P(S_g^*, \boldsymbol{G}_g | \xi_g, \bar{\gamma}, \theta, \delta, \alpha_g)$ is the likelihood function in (3.6) and $\phi(\cdot) \equiv$ pdf of $N(0,1)$.

2. Similarly, the (conditional) posterior for $(\theta', \delta)'$ can be simplified to

$$P(\theta, \delta | \boldsymbol{G}_1, \ldots, \boldsymbol{G}_m, \xi_1, \ldots, \xi_m) \propto \pi(\theta, \delta) \cdot \prod_{g=1}^{m} P(\boldsymbol{G}_g | \xi_g, \theta, \delta), \tag{3.20}$$

   where $P(\boldsymbol{G}_g | \xi_g, \theta, \delta)$ is the likelihood function in (3.4) and $\pi(\cdot)$ is the marginal prior density of $(\theta', \delta)'$.

3. The (conditional) posterior for γ_1 is constructed as

$$P(\gamma_1 | (S_g^*, \boldsymbol{G}_g, \xi_g, \alpha_g)_{g=1}^m, \bar{\gamma}) \sim \prod_{g=1}^{m} P(S_g^* | \boldsymbol{G}_g, \xi_g, \gamma_1, \bar{\gamma}), \tag{3.21}$$

   where $P(S_g^* | \boldsymbol{G}_g, \xi_g, \gamma_1, \bar{\gamma})$ is obtained by integrating (3.14) over ξ_g.

4. Following Albert and Chib (1993), we construct the posterior of $\bar{\gamma}$ as:

$$P(\bar{\gamma}|(S_g^*, \boldsymbol{G}_g, \xi_g)_{g=1}^m, \gamma_1) \propto N(\bar{\gamma}; \gamma_0, \Gamma_0) \cdot \prod_{g=1}^m P(S_g^*|\boldsymbol{G}_g, \xi_g, \gamma_1, \bar{\gamma}, \alpha_g).$$
$$\therefore P(\bar{\gamma}|(S_g^*, \boldsymbol{G}_g, \xi_g)_{g=1}^m, \gamma_1) \propto N(\bar{\gamma}; \gamma_0^*, \Gamma_0^*), \tag{3.22}$$

where $\gamma_0^* = (\sum_{g=1}^m \tilde{Z}_g' \tilde{Z}_g + \Gamma_0^{-1})^{-1}(\sum_{g=1}^m \tilde{Z}_g' A_g S_g^* + \Gamma_0^{-1}\gamma_0)$, $\tilde{Z}_g = [\iota_g', vec(Z_g)', \xi_g']'$, $\Gamma_0^* = (\sum_{g=1}^m \tilde{Z}_g' \tilde{Z}_g + \Gamma_0^{-1})^{-1}$, $A_g := I_{N_g} - \gamma_1 \boldsymbol{G}_g$, and $N(x; y, \boldsymbol{B})$ is the value at x of a normal distribution with mean y and covariance matrix $\boldsymbol{B}$.

We implement the estimation using the MCMC algorithm $M = 30,000$ iterations where the first $20,000$ iterations were discarded. Specifically, the MCMC algorithm is described as follows.

MCMC algorithm. At the kth iteration:

1. The Metropolis-Hastings (M-H) algorithm is used to draw samples $\xi_g^{i(k)}$ from the posterior distribution $P(\xi_g^i|S_g^{*(k-1)}, \boldsymbol{G}_g, \xi_g^{-i(k-1)}, \bar{\gamma}^{(k-1)}, \theta^{(k-1)}, \delta^{(k-1)}, \alpha_g^{(k-1)})$ given in (3.19), where $\xi_g^{-i(k-1)} = (\xi_g^{1(k-1)}, \ldots, \xi_g^{(i-1)(k-1)}, \xi_g^{(i+1)(k-1)}, \ldots, \xi_g^{N_g^{(k-1)}})$. This occurs for every individual $i = 1, \ldots, N_g$ and network $g = 1, \ldots, m$. Specifically:

   (1) Propose $\tilde{\xi}_g^i \sim N(\xi_g^{i(k-1)}, \kappa_\xi^2)$, where κ_ξ^2 is chosen by the user, and let $\tilde{\xi}_g = (\xi_g^{1(k-1)}, \ldots, \xi_g^{(i-1)(k-1)}, \tilde{\xi}_g^i, \xi_g^{(i+1)(k-1)}, \ldots, \xi_g^{N_g^{(k-1)}})$. The value of κ_ξ^2 is adjusted to achieve an acceptance rate between 20% and 40%.

   (2) With probability equal to $a(\xi_g^{(k-1)}; \tilde{\xi}_g^i) =$

$$\min\left\{\frac{P(S_g^*, \boldsymbol{G}_g|\tilde{\xi}_g, \bar{\gamma}^{(k-1)}, \theta^{(k-1)}, \delta^{(k-1)}, , \alpha_g^{(k-1)})}{P(S_g^*, \boldsymbol{G}_g|\xi_g^{(k-1)}, \bar{\gamma}^{(k-1)}, \theta^{(k-1)}, \delta^{(k-1)}, \alpha_g^{(k-1)})} \cdot \frac{N(\tilde{\xi}_g^i; 0, 1)}{N(\xi_g^{i(k-1)}; 0, 1)}, 1\right\},$$

   set $\xi_g^{i(k)}$ equal to $\tilde{\xi}_g^i$; otherwise, set it to $\xi_g^{i(k-1)}$.

2. M-H procedure is used to sample $(\theta^{(k)'}, \delta^{(k)})'$ from $P(\theta, \delta|\boldsymbol{G}_g, \xi_g^{(k)})$ given in (3.20):

   (1) Propose $(\tilde{\theta}', \tilde{\delta})' \sim N\big((\theta^{(k-1)'}, \delta^{(k-1)'}), \kappa_{\theta,\delta}^2 I\big)$, where $\kappa_{\theta,\delta}^2$ is chosen by the user.

   (2) With probability equal to $a(\theta^{(k-1)}, \delta^{(k-1)}; \tilde{\theta}, \tilde{\delta}) =$

$$\min\left\{\prod_{g=1}^m \frac{P(\boldsymbol{G}_g|\xi_g^{(k)}, \tilde{\theta}, \tilde{\delta})}{P(\boldsymbol{G}_g|\xi_g^{(k)}, \theta^{(k-1)}, \delta^{(k-1)})} \cdot \frac{N(\tilde{\theta}, \tilde{\delta}; \phi_0, \Phi_0)}{N(\theta^{(k-1)}, \delta^{(k-1)}; \phi_0, \Phi_0)}, 1\right\},$$

   set $(\theta^{(k)'}, \delta^{(k)})'$ equal to $(\tilde{\theta}', \tilde{\delta})'$; otherwise, set it to $(\theta^{(k-1)'}, \delta^{(k-1)})'$.

3. γ_1 is sampled from $P(\gamma_1|S_g^{*(k-1)}, \boldsymbol{G}_g, \xi_g^{(k)}, \bar{\gamma}^{(k-1)}, \alpha_g^{(k-1)})$ in (3.21) using M-H:

(1) Propose $\tilde{\gamma}_1 \sim N(\gamma_1^{(k-1)}, \kappa^2_{\tilde{\gamma}_1})$, where $\kappa^2_{\tilde{\gamma}_1}$ is chosen by the user.

(2) With probability equal to $a(\gamma_1^{(k-1)}; \tilde{\gamma}_1) =$

$$\min\left\{\prod_{g=1}^{m}\left(\frac{P(S_g^*|\boldsymbol{G}_g, \xi_g^{(k)}, \bar{\gamma}^{(k-1)}, \tilde{\gamma}_1, \alpha_g^{(k-1)})}{P(S_g^*|\boldsymbol{G}_g, \xi_g^{(k)}, \bar{\gamma}^{(k-1)}, \gamma_1^{(k-1)}, \alpha_g^{(k-1)})}\right) \cdot \frac{\mathbb{1}(\tilde{\gamma}_1 \in [-1,1])}{\mathbb{1}(\gamma_1^{(k-1)} \in [-1,1])}, 1\right\},$$

set $\gamma_1^{(k)}$ to $\tilde{\gamma}_1$; otherwise, set it to $\gamma_1^{(k-1)}$, where $\mathbb{1}(\cdot) \equiv$ indicator function.

4. The Gibbs sampling method is used to draw samples for $\bar{\gamma}^{(k)} = \left(\gamma_2^{(k)\prime}, \gamma_3^{(k)\prime}, \gamma_4^{(k)}, \alpha_g^{(k)}\right)'$ from the posterior distribution $P(\bar{\gamma}|S_g^{*(k-1)}, \boldsymbol{G}_g, \xi_g^{(k)}, \gamma_1^{(k)})$ in (3.22). The sign of γ_4 will not be determined as $|\xi_g^i - \xi_g^j|$ is not affected by a change in sign of ξ_g^i or ξ_g^j. To address this issue, we set γ_4 positive using an acceptance-rejection algorithm.

5. $S_g^{*(k)}$ is sampled from the TMN distribution $P(S_g^*|\boldsymbol{G}_g, \xi_g^{(k)}, \gamma_1^{(k)}, \bar{\gamma}^{(k)}, \alpha_g^{(k)})$ in (4.11) of the main paper.

Computing treatment effect measures after ML estimation

1. **Propensity scores calculations.** The propensity scores for each individual i in network $\boldsymbol{G}_g$, i.e., $P_i(Z_g^i, G_g)$, are calculated using the estimated parameters, along with (4.11) in the main paper.

2. **MTEs and MPRTEs calculations.** They are computed using the pooled sample of all networks, so we omit the reference to the group index "g" to simplify the notations. We distinguish the case where X is independent of (U_0, U_1, V) to the one where it is not as this distinction plays a crucial role in identifying the MTE and PRTE (see Section 2 in the main paper).

   Case 1: ***Strictly exogenous covariates.*** In this case, the MTEs are filtered as follows:

   (a) The Robinson's (1988) partially linear equations method is applied to obtain estimates of β_0 and $\beta_1 - \beta_0$:

      i. The difference between the outcome equation and its expected value is taken to remove the non-linear component in $p = P(Z, \boldsymbol{G})$:

$$Y - \mathbb{E}(Y|p) = [X - \mathbb{E}(X|p)]\beta_0 + p[X - \mathbb{E}(X|p)](\beta_1 - \beta_0). \tag{3.23}$$

      ii. Kernel regressions of the dependent variable and each of the regressors are run on p in order to estimate the expected values in equation (3.23).

iii. The kernel regression residuals of the dependent variable are regressed on the kernel regression residuals of each independent variable to estimate β_0 and $\beta_1 - \beta_0$.

(b) Following equation (2.11) in the main paper, a local polynomial regression of $Y - X\hat{\beta}_0 - pX(\hat{\beta}_1 - \hat{\beta}_0)$ is run on P to estimate the function $K(p)$,[12] and its partial derivative with respect to p is thus computed. Adding this partial derivative to $X(\hat{\beta}_1 - \hat{\beta}_0)$ results in an estimate for the MTE:

$$MTE = X(\hat{\beta}_1 - \hat{\beta}_0) + \frac{dK(p)}{dp}.$$

(c) To compute the MPRTEs, a weighted average of the MTEs is taken across the support of $P(Z, \boldsymbol{G})$. The relevant weight is expressed conditional on the value of X and must be integrated over the distribution of X. Since conditioning on X is computationally demanding when X contain many covariates, we follow the suggestion of Carneiro et al. (2011) to condition on the index $X(\beta_1 - \beta_0)$. We can thus measure the MPRTEs using different weighting functions and corresponding policy change. As stated in Section 2 of the main paper, we focus on three policies indexed by α: (1) a policy α that directly increases the probability of treatment equally for all individuals, i.e., $P_\alpha = P(Z, \boldsymbol{G}) + \alpha$; (2) a policy α that proportionally increases the probability of treatment, i.e., $P_\alpha = (1 + \alpha)P(Z, \boldsymbol{G})$; and (3) a policy α that affects one or more instruments used in the treatment equation, i.e., $Z_\alpha = Z + \alpha$.

Case 2: ***Non exogenous covariates.*** When X is not exogenous, we can only identify the MTEs conditional on X, as discussed extensively in Carneiro et al. (2011). As for **Case 1**, we shall condition on the index $X(\beta_1 - \beta_0)$ rather than X.

(a) The value of $X(\beta_1 - \beta_0)$ is calculated at the $25th$ and $75th$ percentile of the distribution of $X(\beta_1 - \beta_0)$.

(b) Holding $X(\beta_1 - \beta_0)$ constant at this point, the instruments Z and the network characteristics (observed and unobserved) are allowed to vary.

(c) The MTE is calculated over the portions of the support of $P(Z, \boldsymbol{G})$ where both treatment and control are jointly observed as the instruments vary.

(d) The MPRTE is computed as a weighted average of the MTEs for each policy change, placing weights only on those segments of the MTEs that are identified.

[12] Fan and Gijbels (1996) recommend using a local quadratic estimator for fitting a first order derivative. We therefore use a local quadratic estimator with a bandwidth that minimises their proposed residual squared criterion.

It is worth noting that for both **Case 1** & **Case 2**, the MPRTEs are calculated conditional on X, so they are identified even when X is correlated with (U_0, U_1, V), while the MTEs can only be identified for the segments of the support of $P(Z, \boldsymbol{G})$ where treatment and control are jointly observed.

4 Simulation experiment

We study the performance of the proposed methodology through a Monte Carlo experiment. For brevity, we consider a two period ($T = 2$) friendship formation model, where the initial and last periods are labeled 'period 0' and 'period 1' respectively.

The data generation process is such that $(U_0^i, U_1^i, V^i, X_i, Z_{1i}^*, Z_{2i}, \xi^i, c^i)'$ are drawn i.i.d. across i from a multi-normal distribution with $m_0 = (0, 0, 0, 1, 0.8, 0, 0.5, 0.2)'$ and variance Σ, where Σ has ones along the diagonal and cross-elements zeros except:

$$\begin{aligned} \mathbb{E}(U_0^i U_1^i) &= 0.3,\ \mathbb{E}(U_0^i V^i) = 0.3,\ \mathbb{E}(U_1^i V^i) = -0.5, \\ \mathbb{E}(X_i Z_{1i}^*) &= \mathbb{E}(X_i Z_{2i}) = 0.2,\ \mathbb{E}(Z_{1i}^* \xi^i) = 0.1,\ \mathbb{E}(Z_{2i} \xi^i) = -0.2. \\ \mathbb{E}(X_i U_0^i) &= \mathbb{E}(X_i U_1^i) \equiv \rho_x \in \{0, 0.3\}. \end{aligned} \tag{4.1}$$

As such, we consider a two-period network formation model with the two networks $\boldsymbol{G}_0$ and $\boldsymbol{G}_1$ having equal size of 250 individuals. The initial network $\boldsymbol{G}_0$ is generated from the sample, while $\boldsymbol{G}_1$ is constructed using (2.8), where the true parameters of the prior means are:

$$\theta_0^0 = 0.2,\ \theta_1^0 = 0.2,\ \theta_3^0 = (-0.9, 0.5, 0.4)',\ \theta_4^0 = 0.2,\ \theta_5^0 = 0.1,\ \delta^0 = -1.2. \tag{4.2}$$

Individuals with no friends are removed from the sample. The network formation process is calibrated to ensure the number of individuals with no friends is less than 20% of the original sample and the average number of friends is greater than 10.

On average, 7 individuals with no friends are removed from the network, and the remaining individuals have an average of 13.65 friendship connections. The average dynamics of the network are given in Table 1. Of the 863 friendships in period 0, 614 (around 2.1% of possible friendships; see Table 1) remain friends in period 1. Similarly, of the 29,890 untapped friendships in period 0, only 790 (around 2.71% of possible friendships) are tapped in period 1. Due to the large number of friendship possibilities (29100), we see more new friendships in period 1 than existing ones carried from period 0. An example of a generated network is presented in Figure 1. We see that the generated network is relatively dense for the size of the network, due to

the high number of average friendships. We use this high degree in order to establish stronger results for the peer effects.

Figure 1: Generated network using (2.8)

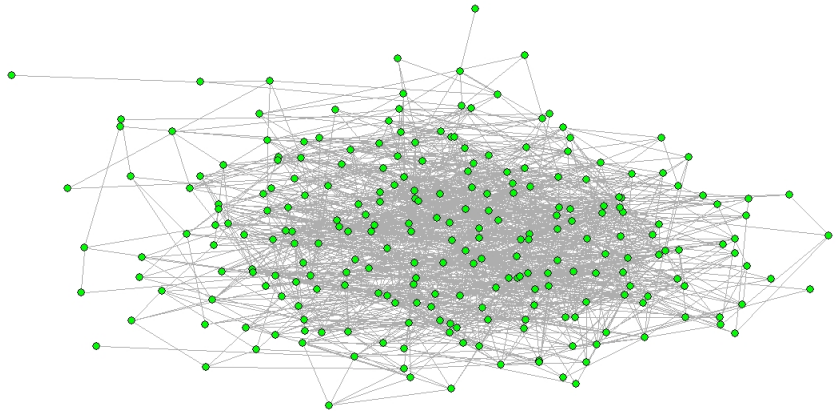


Table 1: Dynamic Friendships

		Period 1	
		Friends	Not Friends
Period 0	Friends	614 (2.1%)	249 (0.85%)
	Not Friends	790 (2.71%)	29100 (94.34%)

4.1 Network formation model

Table 2 displays the estimates of the network formation model (2.8). As the model is estimated using the Bayesian method, the reported estimates are posterior means of the estimated parameter distributions (Columns 2 & 4 of the table). Columns 1 & 3 of the table show the network characteristics and their corresponding coefficients as per model (2.8). The standard errors in parentheses are bootstrapped.

As shown in the table, all estimates have the expected signs. Except θ_4, the estimates of the network formation model parameters are significant at 1% nominal level. The estimates of θ_1 and $\theta_{3k},\ k = 1, 2, 3$ are slightly above the true parameter values, while those of θ_4 and θ_5 are slightly below the true parameter values. Also, similarity in unobserved characteristics estimated coefficient (δ) is negative and highly significant, thus indicating strong homophily effect on friendship choices. While this effect is reasonably large, it is below the true posterior mean. The slight discrepancy between the estimated posterior means and the true values is likely due to the relatively small number of MCMC draws.[13]

[13]Computing limitations have impeded the ability to run the estimation with more iterations or with a greater

Table 2: Posterior Means in Network Formation Model: Model (I)

Variable (coef.)	Posterior mean	Variable (coef.)	Posterior mean
c^i (θ_1)	0.2422***	$\boldsymbol{D}_0^{ij}$ (θ_4)	0.1416*
	(0.0180)		(0.0810)
$\lvert Z_2^i - Z_2^j \rvert$ (θ_{31})	−0.8344***	$\boldsymbol{F}_0^{ij}$ (θ_5)	0.0768***
	(0.0207)		(0.0228)
$\mathbb{1}_{[Z_1^i=Z_1^j=1]}$ (θ_{32})	0.5470***	ξ (δ)	−0.7922***
	(0.0417)		(0.0486)
$\mathbb{1}_{[Z_1^i=Z_1^j=0]}$ (θ_{33})	0.6744***		
	(0.0370)		

Bootstrap standard error estimates are in $(\cdot)$.

* p < 0.1, ** p < 0.05, *** p < 0.01.

4.2 Treatment decision model

In this section, we compare the results of the misspecified model where network effects are omitted from (2.3) (although they are part of the true DGP as described in Section 4) to the full model that contains all of the network components in (2.3). Hereafter, we refer to the full model as Model (I) and the misspecified models as Model (II). Comparing the two model enable us to illustrate the importance of peer effects in identifying treatment effects.

Table 3 presents the estimates (posterior means) of the treatment model for both exogenous ($\rho_x = 0$) and non-exogenous ($\rho_x = 0.3$) covariates. As expected, the results are qualitatively similar regardless of whether X is exogenous ($\rho_x = 0$) or not ($\rho_x = 0.3$). In both cases, the estimated endogenous peer effect γ_1 is highly significant, highlighting the fact that Model (II), which does not control for the endogenous peer effect, is misspecified. The group fixed effects are significant in both Model (I) and Model (II) but are more downward biased in Model (II). None of the exogenous peer effect, nor the effect of the unobserved network within characteristics, appears significant in Model (I) probably due to the high standard error estimates. Note however that with the exception of $\gamma_{3,2}$ estimate (which is quite close to zero; -0.0502 for $\rho_x = 0$ and -0.0499 for $\rho_x = 0.3$), the other estimates have the expected sign and are quite similar for both $\rho_x = 0$ and $\rho_x = 0.3$.

4.3 Marginal treatment effects

We first analyze the results when covariate are strictly exogenous in Section 4.3.1. Section 4.3.2 deals with the case where covariate are not exogenous.

sample size. We run the simulations in parallel on the University of Adelaide's Phoenix High Performance Computing service. The simulations with the two networks of size 250 recorded an average running time of 1 day, 14 hours. We note that while the results are reasonable, larger sample size and number of iterations could improve the accuracy of the estimation. Our algorithm is available and can be adapted accordingly by researchers.

Table 3: Treatment model estimates

	Posterior means: $\rho_x = 0$		Posterior means: $\rho_x = 0.3$	
coefficients	Model (I)	Model (II)	Model (I)	Model (II)
γ_1	0.5895***	-	0.5579 **	-
	(0.1046)		(0.1099)	
$\gamma_{2,1}$	0.6289***	0.5826***	0.6334 ***	0.5811***
	(0.0802)	(0.0640)	(0.0801)	(0.0600)
$\gamma_{2,2}$	0.7155***	0.8603***	0.7265 ***	0.8253***
	(0.1591)	(0.1563)	(0.1566)	(0.1805)
$\gamma_{2,3}$	0.2918*	0.6459***	0.3054*	0.6645***
	(0.1662)	(0.0811)	(0.1566)	(0.0846)
$\gamma_{3,1}$	0.1594	-	0.1717	-
	(0.1607)		(0.1744)	
$\gamma_{3,2}$	−0.0502	-	-0.0499	-
	(0.3433)		(0.3541)	
$\gamma_{3,3}$	0.0925	-	0.0966	-
	(0.2370)		(0.2261)	
γ_4	0.0814	-	0.1155	-
	(0.0827)		(0.0810)	
α_1	−1.1608***	−1.0181***	-1.1579***	−0.9488***
	(0.2050)	(0.2446)	(0.2106)	(0.2395)
α_2	−1.1419***	−0.9824***	-1.1534***	−1.0002***
	(0.1947)	(0.2401)	(0.1943)	(0.2576)

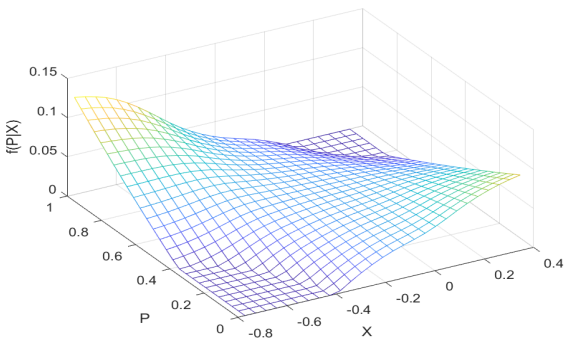
* p < 0.1, ** p < 0.05, *** p < 0.01.

4.3.1 Exogenous covariates

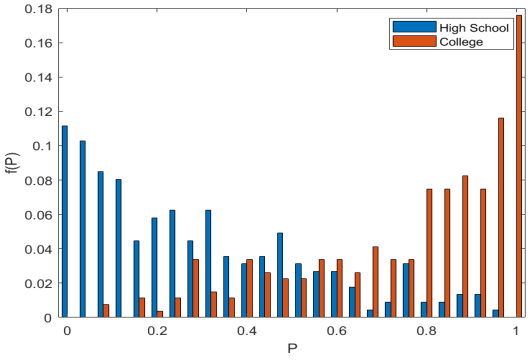
Figure 2 presents the density of the propensity score $P(Z, \boldsymbol{G})$ given X on the left and the support of $P(Z, \boldsymbol{G})$ for $S = 1$ and $S = 0$ on the right when covariate X is exogenous, for both Model (I) and Model (II). The support of $P(Z, \boldsymbol{G})$ where both treatment and control are jointly observed corresponds to the region where the MTE is identified. In Figure 2, this region is interval (0.0443,0.9633) in Model (I), and in the interval (0.0237, 0.8796) in model (II). Therefore, Model (I) identifies the MTE on the interval (0.0443,0.9633), while Model model (II) does only on the interval (0.0237, 0.8796). Clearly, the misspecified model with no network, Model (II), identifies a smaller portion of the MTE, while Model (I), which includes the network formation, induces a broader spread of propensities over which the MTE can be identified. The regions of the MTE that are not identified are those areas of higher variance corresponding to the extreme values of the propensity scores. This illustrates why the traditional treatment effects measures such as the *average treatment effect* (ATE) or *average treatment effect of the treated* (ATET) cannot be identified, as they require that the common support of $P(Z, \boldsymbol{G})$ for both $S = 0$ and $S = 1$ be the full interval [0,1]. This, in turn, makes MTE a more practical and relevant measure of the treatment effect.

The individual MTE estimates can easily be traced across the values of U_S, where U_S is defined in (2.5). Figure 3 shows the plots of these MTEs curves along with the 90% confidence

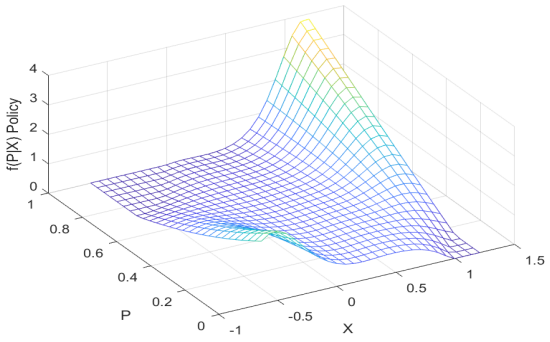
Figure 2: Exogenous covariates– supports of $P(Z, \boldsymbol{G})$: (a) given X (left) & (b) for $S = 0$ and $S = 1$ (right)



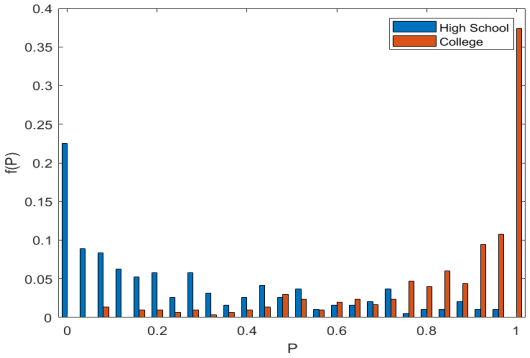
(a) Model (I)



(b) Model (I)

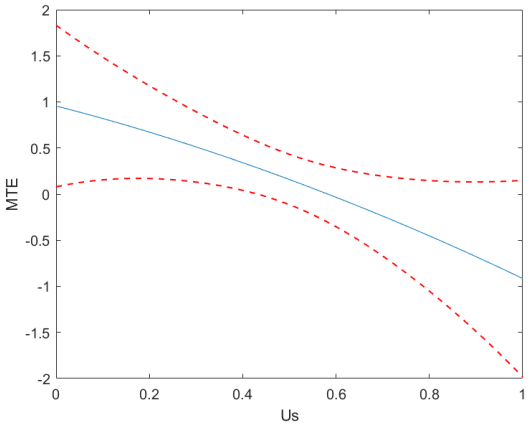


(c) Model (II)

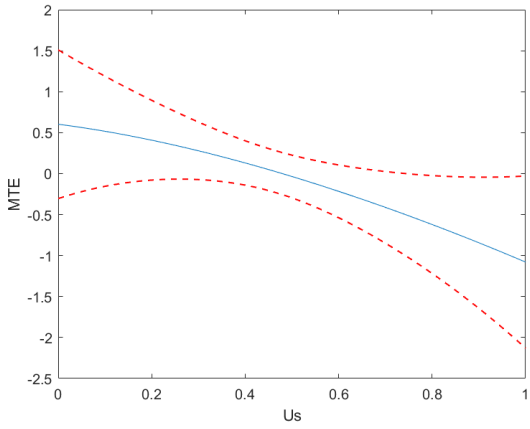


(d) Model (II)

Figure 3: MTE curve with 90% confidence bands



Model (I): Exogenous covariates



Model (II): Exogenous covariates

bands (dashed red lines) for both models, where the values of the MTEs are computed at the mean value of the covariate X. We see that a high value of U_S corresponds to a low probability of taking on treatment, since at those high values, individuals at the margin have corresponding high propensity scores $P(Z, \boldsymbol{G})$, and are thus less likely to take on treatment. Similarly, a

low value of U_S corresponds to a high probability of taking on treatment, as individuals with propensity scores greater than such value are likely to take on treatment. A marginal increase of $P(Z, \boldsymbol{G})$ starting from high values of $P(Z, \boldsymbol{G})$ can induce individuals with high U_S values into treatment, while it has no effect on individuals with low U_S values, since they are likely to already be in treatment at such high values of $P(Z, \boldsymbol{G})$. The MTE at these points is defined as the expected increase in the outcome variable Y when $P(Z, \boldsymbol{G})$ is varied marginally. In this version of the model, we clearly see that a high probability of treatment (i.e., a low value of U_S) is associated with a high return to treatment (around 90% in each model), with the opposite observed for a low probability of treatment (around -90%). It is clear from Figure 3 that the model with network peer effects (i.e., Model (I)) identifies higher returns at low U_S values than the misspecified model with no peer effect (i.e., Model (II)). The marginal treatment effect ranges from 0.9 for low U_S to -0.8 for high U_S for Model (I), compared to 0.6 and -1.1 for Model (II). Most notable is the progressive improvement in the width of the 90% confidence bands for low U_S values moving towards the more robust model. In particular, the 90% confidence interval of the MTEs of Model (II) includes zero at low values of U_S, while that of Model (I) does not. As such, there is evidence that the network model enables a more precise identification of the true returns to treatment where it matters (i.e., at least for low U_S values).[14]

The MTE curves in Figure 3 are clearly downward sloping, thus suggesting that individuals likely select into treatment based on heterogeneous returns in outcomes. We can test the zero slope hypothesis of the MTE curve at each point as well as the joint hypothesis that this slope is constant across all values of U_S; see e.g. Carneiro et al. (2011). To illustrate how these tests are implemented, let us focus on the correctly specified Model (I). The tests results are presented in Table 4. We evaluate the MTE at 26 equally spaced points between 0 and 1 with interval length of 0.04. This enables the computation of the local average treatment effects (LATEs), defined as the mean of the MTEs of each interval, and the bootstrap tests are constructed comparing adjacent LATES (see Table 4). The results indicate that the MTE curve is statistically downward sloping at the 10% nominal level for the middle portion of the curve (around 0.25 to 0.75). There is more dispersion at the ends of the MTEs curve, and we cannot statistically reject that the slope is zero. As a result, the p-value for a constant slope across the whole MTE curve is 0.1333, meaning that we can only reject the constant of the overall slope at nominal size above 13.33%.

[14]Note that individuals with high U_S are unlikely to take on treatment.

Table 4: Test of Equality of LATEs over different intervals

Ranges of U_S for $LATE_j$	(0,0.04)	(0.08,0.12)	(0.16,0.20)	(0.24,0.28)	(0.32,0.36)	(0.40,0.44)
Ranges of U_S for $LATE_{j+1}$	(0.08,0.12)	(0.16,0.20)	(0.24,0.28)	(0.32,0.36)	(0.40,0.44)	(0.48,0.52)
Difference in LATEs	0.0502	0.0723	0.0937	0.1145	0.1346	0.1542
p-value	0.5333	0.3000	0.2333	0.1333	0.0667	0.0667
Ranges of U_S for $LATE_j$	(0.48,0.52)	(0.56,0.60)	(0.64,0.68)	(0.72,0.76)	(0.80,0.84)	(0.88,0.92)
Ranges of U_S for $LATE_{j+1}$	(0.56,0.60)	(0.64,0.68)	(0.72,0.76)	(0.80,0.84)	(0.88,0.92)	(0.96,1)
Difference in LATEs	0.1731	0.1915	0.2093	0.2264	0.2430	0.2589
p-value	0.1333	0.1333	0.1667	0.1667	0.1333	0.1333
joint p-value			0.13333			

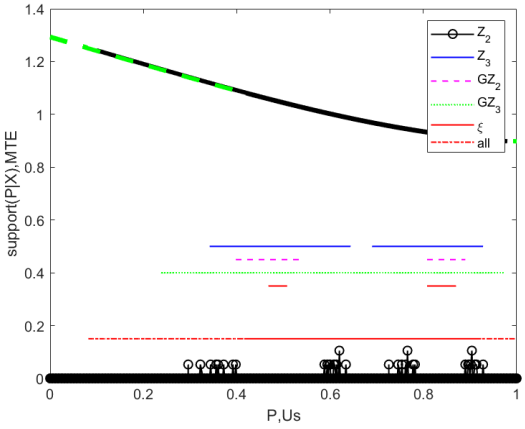
4.3.2 Non exogenous covariates

We now analyze the case where the covariate X is not exogenous. As discussed in Section 3.4, the MTE is only identified conditional on X when the strict exogeneity assumption on covariates does not hold. Given X, one can quantify the contribution of the peer effects to the identification of the MTE.

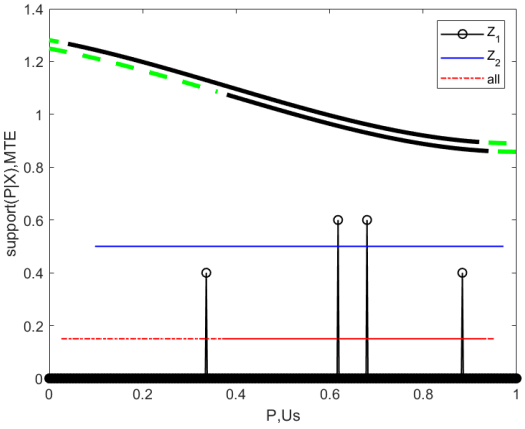
Indeed, when x is fixed at $X = x$, one needs external variations to identify the MTE. In particular, all the variables appearing in the RHS of the treatment model (2.3), except the latent variable S^*, should provide exogenous variations enabling to identify the MTE. That is, variations in the instruments Z and the network peer effects identify the MTE. Figure 4 presents the segments of the support of $P(Z, \boldsymbol{G})$ over which MTE is identified, as well as the contribution of each instrument or peer effects for both models. When Z contains many instruments, we can vary them collectively or individually given value $X = x$ to produce the graphs such as in Figure 4. Each subfigure shows two curves corresponding to the MTE evaluated at $25th$ percentile (bottom) and at the 75th percentile (top) of the distribution of the index $X(\beta_1 - \beta_0)$. Specifically, we start by computing the average X for values of $X(\beta_1 - \beta_0)$ at the the $25th$ and $75th$ percentile of its distribution. The two curves in each subfigure represent the MTE evaluated at these two values of $X(\beta_1 - \beta_0)$, with the black solid segment representing the portion of the MTE that is identified. The dashed green line represents the segments of the MTE we do not identify. In order to draw the line labelled say 'Z_2', we do not only fix X at the two mean values referred above, but we also fix all the other instruments at their corresponding mean values. The red line corresponds to the support of $P(Z, \boldsymbol{G})$ where the MTE is identified when all instruments are varied. Similarly, the other lines correspond to the support of $P(Z, \boldsymbol{G})$ where the MTE is identified when only a single instrument is varied. As Z_1 is a binary variable, the support associated with varying Z_1 is constituted by single points only, represented by black circles.

Looking at the misspecified Model (II), we see that the overall identified segment (red line) of the MTE is short at the extremes values of U_S compared with the preferred Model (I)

Figure 4: Instrument contribution to the identification of MTEs on the support of $P(Z, \boldsymbol{G})$



Model (I): $\rho_x = 0.3$



Model (II): $\rho_x = 0.3$

with dynamic friendship formation. Considering the contribution of each instrument to the identification of the MTEs, we see that the segments of the MTE identified by varying only the peer effects are the ones contributing much to the MTE identification at the extremes values of U_S. This highlights the level of misspecification in the baseline Model (II) due to the exclusion of peer effects. It is worth noting that the contribution of each exogenous peer effect instrument is very important in identifying the MTEs, although the unobserved within network characteristics ξ only contribute marginally to the MTEs identification.

4.4 Marginal policy relevant treatment effects

We now focus on estimating the marginal policy relevant treatment effect (MPRTE). Table 5 reports the estimated MPRTE for both models with three different changes in policy, as described in Section 3.3. We see that for both, $\rho_x = 0$ and $\rho_x = 0.3$, Model (I), which incorporates dynamic network formation, estimates a higher marginal policy effect for individuals induced into treatment compared with the misspecified Model (II). More importantly, while all policy changes are statistically significant at nominal levels ranging from 1% to 5% in Model (I), none of the policy change is statistically significant even at 10% in Model (II).

Table 5: MPRTE estimates

Policy Change	Exogenous covariates		Non-exogenous covariates	
	Model (I)	Model (II)	Model (I)	Model (II)
$Z_\alpha = Z + \alpha$	0.5489**	0.3929	0.6797***	0.1120
	(0.2689)	(0.3430)	(0.2630)	(0.4082)
$P_\alpha = P + \alpha$	0.5303**	0.3801	0.6680**	0.0973
	(0.2637)	(0.3451)	(0.2637)	(0.4030)
$P_\alpha = (1+\alpha)P$	0.4222**	0.2986	0.5561***	0.0554
	(0.2083)	(0.2973)	(0.2083)	(0.3315)

* p < 0.1, ** p < 0.05, *** p < 0.01.

5 Empirical application

We apply the proposed method to the college attainment model in the US using the Add Health data.[15] This data contain high school friendship networks from the United States, and thus allows us to assess the influence of peers on individuals' decision to attend college, and how this decision affects wages after college completion. This paper adds to the existing and growing evidence of peer influence on college attendance decision.[16]

5.1 Add Health data

The dataset was collected through longitudinal surveys across high schools in the US. Surveys were conducted for 90,118 individuals in school years 7-12 in representative high schools during the 1994-1995 school year. A core sample of 20,745 students was selected to take part in a detailed in-home survey across four waves; wave I: 1994-1995, wave II: 1996, wave III: 2001-2002 and wave IV: 2008. A saturated sample, comprising of 16 schools where all students within each school were selected, is part of this core sample. These 16 heterogeneous schools included two large schools (total enrolment exceeding 3100) and 14 smaller schools (enrolment fewer than 300 each). One of the large schools is located in a mid-sized town with a predominantly white enrolment, while the other is located in a metropolitan area and is ethnically diverse. The smaller schools are a mix of public and private schools located in rural and urban areas. In waves I and II, students were asked to name up to 10 of their closest friends (5 males, 5 females). We used these friendship rosters to construct high school networks in waves I and II. Friendships named outside of the selected schools were excluded as the nominated individuals are not in the sample.

Wave II responses were used to construct most variables for the treatment and outcome equations, while observed wages and some contextual variables were collected from wave IV.

[15] The data are restricted and further information is available at "Add Health."
[16] See e.g. Kramarz and Skans (2014).

Individuals who did not take part in all of waves I, II and IV were removed from the sample. We also removed individuals with no friendship nominations (409 individuals) as they experience no measurable peer effect. After removing individuals with incomplete data or no friendship relations, the remaining sample contains 1696 observations across 15 schools. Due to the high-demanding computation capacity required in the estimation of large networks, we limited our analysis to the 13 smaller schools. As the influence of peer effects would strengthen if the larger schools were included, our results can be interpreted as providing lower bounds on peers' effect. We also removed friendships between individuals from different schools as these represent less than 1% of the reported friendships within the overall sample. This further reduced the sample to a total of 631 individuals and 13 networks.[17] While the survey restricts students to naming 10 close friends, only one student in our sample reaches this limit.

The description of the key variables are presented in Table B.1. All variables, except wage and local income (for which logs are taken), are in levels. In addition to the peer exogenous effect variables, other instruments in the treatment equation include number of siblings, innovations in local labour market variables, and the proportion of college education in the local area. Summary statistics of the variables of the 13 constructed networks are given in Table B.2. About 71% of individuals in the sample attended college and the sample is predominantly white individuals (84%) with only 14% of blacks. Mothers of the respondents seem slightly more educated than their fathers: 46% vs. 33% for High School and 39% vs. 35% for College. Around 55% of respondents are married with an average number of children of about 0.9. As with most survey data, potential measurement error is possible. For example, self reported variables such as wage and subject grade are liable to misreporting or unconscious bias. However, most variables of interest are easily verifiable (race, gender, college decision).

Figure 5 illustrates the networks of two of the schools in Wave I (initial period), and the properties of the networks are presented in Table 6. Only two networks are fully connected, but most have a large component dominating the network. On average, a network is composed of about 48 individuals, each with 3.85 friends. This coexists with a link density of about 11%, an average path length of 3.68 and diameter of 9.08. Thus, each individual is on average 3.68 links away from anyone in their network, with a maximum distance of 9.08 per network. Friends tend to cluster together, with few connections across the network. The clustering coefficient indicates that just over a third of possible triplets are closed, i.e. an individual is likely to be connected with his friend's friends. This is very important as it supports the assumption of the network formation model in (2.8) that an individual's preference for a relationship with

[17] Using the University of Adelaide's Phoenix High Performance Computing service, the estimation of these 13 networks takes approximately 23 hours.

another person is mainly based on whether both had some friends in common in an earlier period, thus addressing Jackson's (2013) concern. This clustering pattern is more visible in the second graph of Figure 5, with evidence of three distinct groups. Table 7 gives a transition matrix of friendships from period 0 to period 1. Out of 1387 friendships in period 0, only 661 (47.66%) remain in period 1. So, slightly less than half of friendships in period 0 remains in the next period. In fact, the number of new friendships in period 1 (558) is not far below the number of existing friendships transferred from period 0. A lot of potential friendships however remained untapped, with 16,860 possible links[18] which never formed.

Figure 5: Example of two schools' networks in wave I

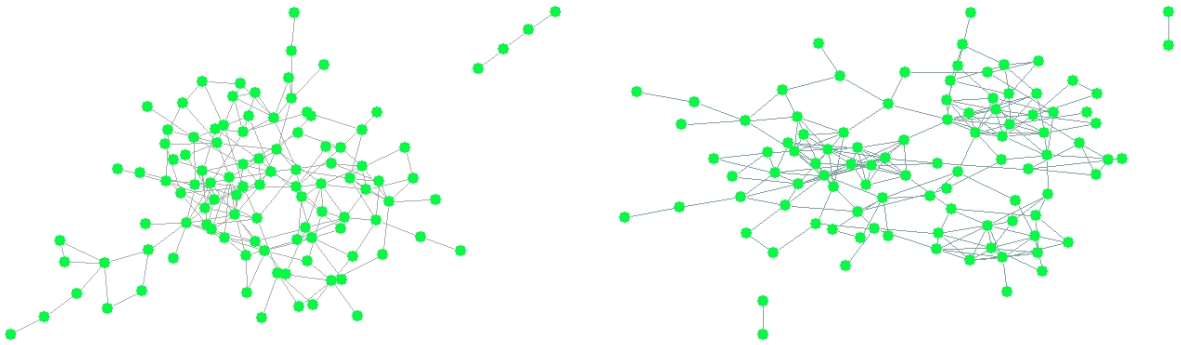


Table 6: Average Network Properties

Property	Mean	S.D.
Number of Nodes	48.54	26.19
Number of Links	93.77	58.08
Link Density	0.11	0.08
Average Degree	3.85	1.02
Clustering Coefficient	0.35	0.14
Number of Components	2.30	1.55
Average Path Length (of largest component)	3.68	0.91
Diameter (of largest component)	9.08	2.63

Number of Networks = 13

Table 7: Transition of friendships of the 13 independent networks

		Period 1	
		Friends	Not Friends
Period 0	Friends	661 (3.52%)	726 (3.86%)
	Not Friends	558 (2.97%)	16860 (89.66%)

We treated friendships as a binary relation and normalized the network so that the average peer effect is equal for every individual regardless of their number of friends. In other words,

[18] As we consider the 13 networks independently of one another, the total number of possible links in the sample is the sum over all 13 networks of the number of all possible links within each network.

each of an individual's friends is given equal weight. These are necessary and reasonable simplifications of the true friendship network, and as a result we may underestimate the true influence of peer effects.

5.2 Main results

We estimate both the model with network formation process (Model (I) of Section 4) and the baseline model with no networks (Model (II) of Section 4). The number of replications in the MCMC and bootstrap algorithms are the same as in Section 4.

5.2.1 Network formation and college attainment models

Table 8 presents the estimates of the parameters in the network formation model (2.8). With the exception of the variable in c^i (shy, independent, number of siblings), all other variables are statistically significant at nominal levels ranging from 1%-10%. Homophily characteristics matter in friendship formation. In particular, individuals select friendships based on similarity in age, race, gender, parent's education, grade in Math and English, and even appearance ($APER$). Having been friends in the previous period ($\boldsymbol{D}_0$), or having friends in common in the previous period ($\boldsymbol{F}_0$) are particularly strong determinants of friendship formation. Dissimilarity in the unobserved within network characteristics has the expected negative effect on friendship choices.

Table 9 presents the estimates of the college attendance decision parameters with both dynamic network formation and no network. Focusing first on the estimates with dynamic network formation, we see that most variables are highly statistically significant. The endogenous effect estimate is small (-0.0790) with a counter-intuitive sign but is highly significant. The estimates of schools' fixed effects that control for correlated effects are all highly significant, with race, gender, Math and English grades, and parents' education also playing a relatively important role in college attendance decision. Most of the exogenous effects are also significant, particularly having high school or college educated parents and having friends with a higher English grade boost the likelihood of college attendance. Although being black and having higher Math grade have a positive effect on treatment, their exogenous effects (i.e. having black friends and friends with high Math grades) appear to negatively affect the decision to attend college. The effect of unobserved within network characteristics, ξ, is small but significant at the 1% nominal level with the expected positive sign. The instruments are also highly significant; *number of siblings*, *local unemployment in 1990* and *proportion of college education* positively influence college attendance decision, while *local income in 1990* has a strong negative effect, as we would

Table 8: Network formation model

variables	coef.	variables	coef.
c^i		c^{ij}	
		Male	0.1690***
Shy	−0.0563		(0.0374)
	(0.0659)	Female	0.0270*
Independent	−0.1452		(0.0147)
	(0.0997)	White	0.4543***
No. of Siblings	−0.0165		(0.1180)
	(0.0202)	Black	0.4668***
c^{ij}			(0.0739)
$\lvert Age^i - Age^j \rvert$	−0.4728***	Other race[19]	0.5088***
	(0.0074)		(0.2379)
$\lvert MathGD^i - MathGD^j \rvert$	−0.0489***	$\boldsymbol{D}_0^{ij}$	2.1126***
	(0.0074)		(0.1187)
$\lvert EngGD^i - EngGD^j \rvert$	−0.0987***	$\boldsymbol{F}_0^{ij}$	0.3986***
	(0.0044)		(0.0365)
$\lvert APER^i - APER^j \rvert$	−0.0651***	$\lvert \xi^i - \xi^j \rvert$	−1.1949*
	(0.0148)		(0.6133)
Parents: No High School	−0.2256*		
	(0.1350)		
Parents: High School	0.0998***		
	(0.0212)		
Parents: College	0.1573***		
	(0.0498)		

* p < 0.1, ** p < 0.05, *** p < 0.01.

expect.

It is worth noting that the parameter estimates of the misspecified model with no network raise concerns of omitted variable bias when compared with the model with dynamic network formation. For example, none of the instruments (number of siblings, local unemployment and income in 1990, proportion of college education) is statistically significant. The correlated effect estimates (group fixed effects)– the only one we can control for in this setting– are very large but imprecise (large standard error estimates), thus rendering them statistically insignificant.

5.2.2 Marginal returns to college education

Figure 6 shows the supports of the propensity scores conditional on covariates for both the model with network formation and the one without. The MTE is identified over some portions of these supports. As expected, low values of the propensity scores correspond to low covariate values[20] and conversely, high covariate values are associated with high propensity scores. Figure 6b shows clearly that the model with network formation correctly assigns high propensity scores to those who attend college. Those who do not attend college have a more diverse spread of propensities compared to those who attend college.

The model with no network has a tendency to assign high propensity scores across all the

[20]Note that X contains many covariates, so in graphs we condition on the index $X(\hat{\beta}_1 - \hat{\beta}_0)$ rather than X.

Table 9: College decision model estimates (posterior means)

With dynamic network formation				**No network**			
vars.	coefs.	vars.	coefs.	vars.	coefs.	vars.	coef.
Endog. effect	−0.0790***	MathGD	−0.0404***	Female	0.3504**	School 6	−3.0399
	(0.0116)		(0.0041)		(0.1426)		(4.7828)
Female	0.3564***	EnglishGD	0.2356***	Black	−0.1834	School 7	−4.1312
	(0.0037)		(0.0067)		(0.3940)		(4.7701)
Black	0.1175***	ξ	0.0585***	Other race	−0.6773**	School 8	−2.6549
	(0.0087)		(0.0154)		(0.3208)		(4.4685)
Other race	−0.4131***	*Fixed effects*		Parents HS	0.5542**	School 9	−2.8556
	(0.0115)	School 1	−0.2590***		(0.2811)		(4.8720)
Parents HS	0.3704***		(0.0079)	Parents College	1.2467***	School 10	−2.7943
	(0.0083)	School 2	−0.0535***		(0.3276)		(4.7991)
Parents College	1.0509***		(0.0082)	Math grade	0.0092	School 11	−2.9624
	(0.0069)	School 3	−0.4934		(0.0550)		(4.7915)
Math grade	0.0178***		(0.0086)	English grade	0.2766***	School 12	−3.3964
	(0.0022)	School 4	−0.4567***		(0.0719)		(4.8447)
English grade	0.2430***		(0.0082)	No. of Siblings	0.0123	School 13	−2.8934
	(0.0023)	School 5	0.2997***		(0.0335)		(4.8954)
No. of Siblings	0.0042***		(0.0112)	Income 1990	0.1643		
	(0.0012)	School 6	0.0016		(0.5584)		
Income 1990	−0.1840***		(0.0082)	Unemployment 1990	10.2607		
	(0.0027)	School 7	−0.5592***		(10.5087)		
Unemployment 1990	0.0299***		(0.0107)	College Prop 1990	0.1236		
	(0.0073)	School 8	−0.0527***		(1.5880)		
College Prop 1990	0.3183***		(0.0151)				
	(0.0161)	School 9	0.1216***	*Fixed Effects*			
Exogenous effects			(0.0084)	School 1	−3.3830		
					(4.8337)		
Female	−0.0130	School 10	0.3517***	School 2	−3.4070		
	(0.0109)		(0.0081)		(4.7833)		
Black	−0.2249***	School 11	0.2838***	School 3	−3.8951		
	(0.0136)		(0.0113)		(4.8190)		
Other race	0.4525***	School 12	0.1629***	School 4	−3.8736		
	(0.0176)		(0.0119)		(4.8138)		
Parents HS	0.1602***	School 13	0.4446	School 5	−2.6730		
	(0.0122)		(0.0095)		(4.8030)		
Parents College	0.5575***						
	(0.0127)						

* p < 0.1, ** p < 0.05, *** p < 0.01.

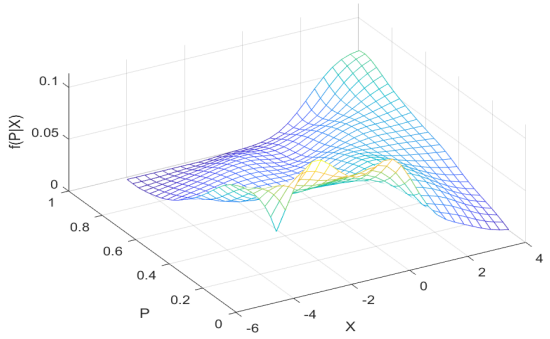
range of covariate values (see Figure 6c), although low propensity scores are assigned to low covariate values. Clearly, without network formation, the model is ineffective at separating those who attend college to those who do not. If covariates were assumed exogenous, the MTEs would be identified[21] over the range (0.1063,0.9695) under Model (I), which corresponds to (0.1435,0.9729) with Model (II). Although both models are comparable at low values of U_S (although neither identifies the MTEs in this region), the model with network formation performs better at high U_S (i.e. low P) values by predicting less negative return.

Figure 7 shows the plots of the annualized[22] estimated marginal returns to college education. Again, we see that the MTE confidence bands in the model with no network are much wider compared to the model with network formation. Focusing on the correct specification, the result is quite striking. We see that individuals with low U_S values –in the neighborhood of $U_S = 0$– (i.e., high probability of college attendance) have a high return (about 70%), while those with high values of U_S–in the neighborhood of $U_S = 1$ (i.e., low probability of college attendance)

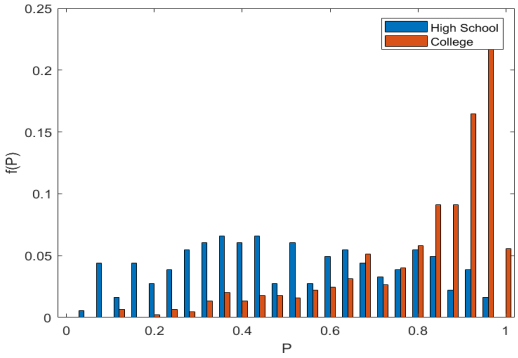
[21] That is, where we observe common support for both $S = 1$ and $S = 0$.

[22] The MTEs are divided by four to obtain annualised estimates as in Carneiro et al. (2011).

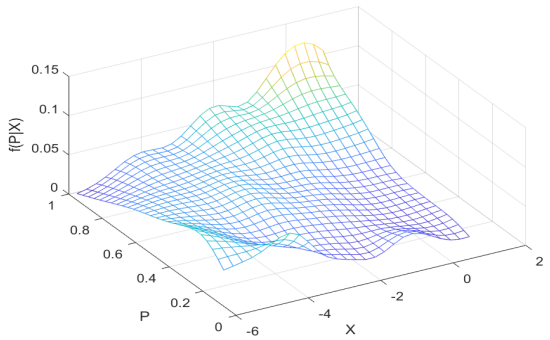
Figure 6: Propensity scores for high school network data



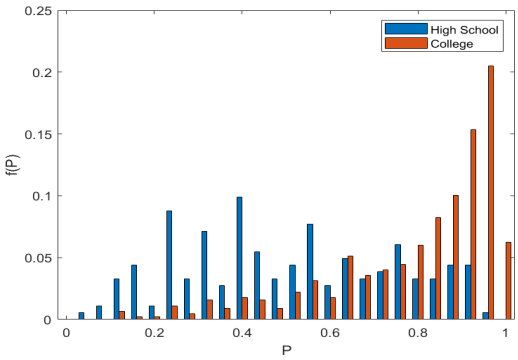
(a) Support of propensity scores given covariates– Model with networks



(b) Support of propensity scores given $S = 1$ and $S = 0$– Model with networks



(c) Support of propensity scores given covariates– Model with no network

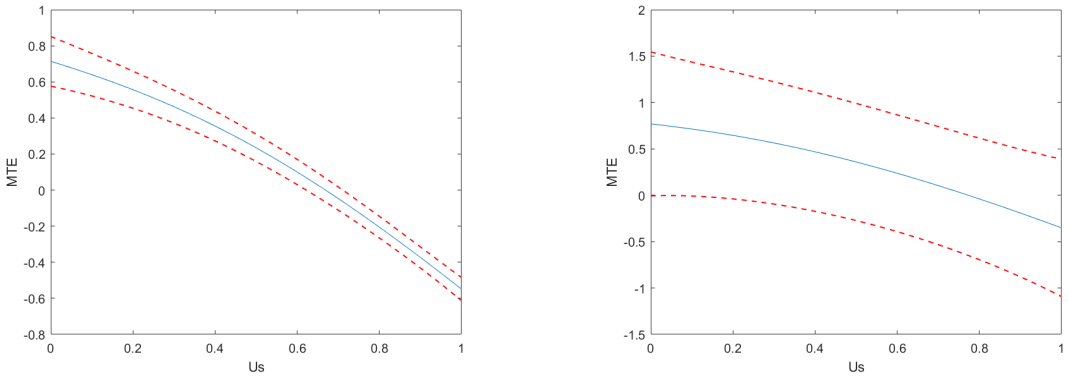


(d) Support of propensity scores given $S = 1$ and $S = 0$– Model with no network

have a negative return (-60%). A 70% return indicates that these individuals would expect to have a wage 70% higher if they attended college than otherwise. We also tested the constant slope hypothesis and the difference in LATEs for a negative slope hypothesis of the MTE curve. The results for the model with network formation are reported in Table D.4 in the Supplemental Appendix and demonstrate an undeniably downward slope at each point of the MTE curve. Clearly, in this model, individuals are self-selecting into college, with those who are likely to attend college receiving much higher returns.

Since it highly unlikely all covariates used are strictly exogenous, the MTEs are not identified over the full support of the propensity scores given in Figure 6. The MTEs can only be identified conditional on covariates, requiring variations in the instruments since covariates are fixed at specific values. To draw the graphs in Figure 8, we start by computing the average covariates (X) for values of $X(\hat{\beta}_1 - \hat{\beta}_0)$ at the the $25th$ and $75th$ percentile of its distribution. The two curves in each subfigure represent the MTEs evaluated at these two values of $X(\hat{\beta}_1 - \hat{\beta}_0)$, with the black solid segment representing the portion of the MTEs that is identified when covariates are not assumed exogenous. The dashed green line represents the sections of the MTEs we

Figure 7: Marginal returns to college education

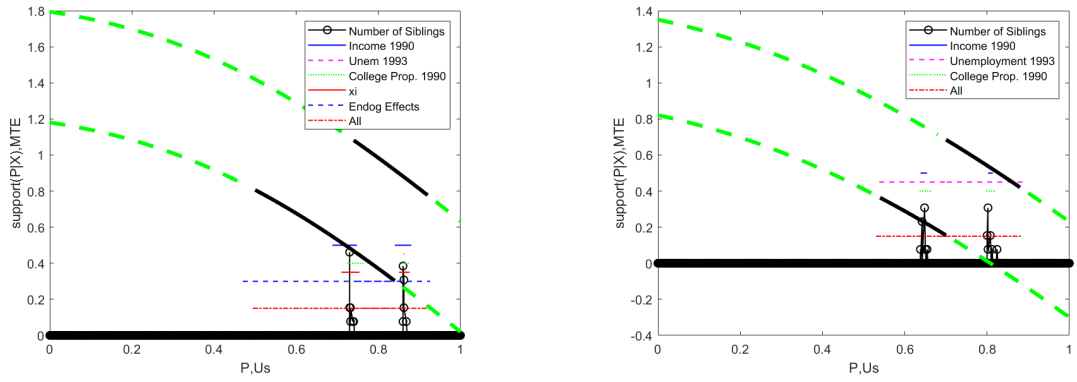


(a) Model with network formation

(b) Model with no network

do not identify. The three labour market variables and ξ are the instruments that are varied to identified the MTEs. It appears from these graphs that the strength of the instruments in identifying the MTEs is low, as the black segment of the MTE curve is short even in the model with network formation. Interestingly, the portion of the MTEs that is identified corresponds to the propensity scores clustering behavior observed around 0.5 to 0.8 in Figure 6b. Note that the MTE curve is still downwards sloping in these identified regions. The MTE curve has drastically different intercepts depending on the value we hold the covariates at. This is to be expected, given the number of covariates in the regression, compared to the number of available instruments. Also, note that the inclusion of the network exogenous peer effects as instruments in Figure 8a has improved the identification of the MTE compared to the model with no network (see Figure 8b). Clearly, using both the unobserved within network characteristics ξ and the

Figure 8: Identified support of MTEs without strict exogeneity assumption of covariates



(a) Model with network formation

(b) Model with no network

exogenous peer effects as instruments offer a greater variation compared with the model with no network. Also noticeable is that the misspecified model with no network only identifies a small

portion of the marginal return to college education at the $25th$ percentile of the distribution of the index $X(\hat{\beta}_1 - \hat{\beta}_0)$, which is not the case for the model with network formation.

5.2.3 Marginal policy relevant returns to college education

Table 10 presents the results of the MPRTEs for the three policies indexed by α: (1) a policy α that directly increases the probability of college attendance equally for all high school students in the sample, i.e., $P_\alpha = P + \alpha$; (2) a policy α that proportionally increases the probability of college attendance, i.e., $P_\alpha = (1+\alpha)P$; and (3) a policy α that affects one or more instruments used in college attendance decision, i.e., $Z_\alpha = Z + \alpha$. As seen, the effect of all policies is positive and highly significant when dynamic network formation is accounted for. However, with no network, the estimated policy impacts, albeit not statistically significant, are all negative.

Focusing on the correct specification with network formation, those induced into college education by a small change in the instruments receive a return of around 182%. This effect is slightly better at 184% for a policy that increases marginally and additively the propensity scores. A policy aiming at increasing the propensity scores marginally but proportionally appears to have the smallest impact (although still large), with around 152% return to college education.

Table 10: Marginal policy relevant returns to college education

Policy Change	**With network formation**	**No network**
$Z_\alpha = Z + \alpha$	1.8196***	−0.1760
	(0.2581)	(1.8736)
$P_\alpha = P + \alpha$	1.8431***	−0.1983
	(0.2667)	(1.8635)
$P_\alpha = (1+\alpha)P$	1.5215***	−0.3061
	(0.2389)	(1.7882)

* p < 0.1, ** p < 0.05, *** p < 0.01.

6 Conclusions

This paper develops an econometric framework that incorporates peer effects into the standard counterfactual model. The identification of standard treatment effect measures such as the marginal return to treatment and the marginal policy relevant treatment effect are explored under social interactions. A Bayesian procedure is developed to estimate the model and quantify the contribution of peer effects to the identification of these treatment effect measures. The proposed methodology is illustrated through simulations and an empirical application on the educational attainment model in the US.

We find significant impacts of peer influence on the decision to take on treatment as well as on the marginal return to treatment and on the policies aiming at encouraging individuals to take on treatment. In particular, the exogenous effects provide strong variations in identifying the treatment effect measures even when covariates in the potential outcome equation are kept constant. As such, failing to account for peer effects in counterfactual analysis leads to model misspecification, as shown in the empirical application. The proposed methodology adequately addresses many of the complications arising from the network estimation. Wider acceptance and use of peer effect models within the econometric field will enable to more widely contemplate and exploit the potential role of networks in enacting and dispersing economic and social policies.

References

Albert, J. H. and S. Chib (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American statistical Association 88*(422), 669–679.

Blume, L. E., W. A. Brock, S. N. Durlauf, and Y. M. Ioannides (2011). Identification of Social Interactions. In J. Benhabib, A. Bisin, and M. O. Jackson (Eds.), *Handbook of Social Economics*, Volume 1, Chapter 18, pp. 853–964. Elsevier.

Bramoullé, Y., H. Djebbari, and B. Fortin (2009). Identification of peer effects through social networks. *Journal of econometrics 150*(1), 41–55.

Calabrese, R. and J. A. Elkink (2014). Estimators of binary spatial autoregressive models: A Monte Carlo study. *Journal of Regional Science 54*(4), 664–687.

Carneiro, P., J. J. Heckman, and E. J. Vytlacil (2011). Estimating marginal returns to education. *The American Economic Review 101*(6), 2754–2781.

Case, A. C. (1991). Spatial patterns in household demand. *Econometrica*, 953–965.

Chandrasekhar, A. (2016). Econometrics of network formation. *The Oxford Handbook of the Economics of Networks*, 303–357.

Christakis, N. A. and J. H. Fowler (2007). The spread of obesity in a large social network over 32 years. *New England journal of medicine 357*(4), 370–379.

Fan, J. and I. Gijbels (1996). *Local polynomial modelling and its applications: monographs on statistics and applied probability 66*, Volume 66. CRC Press.

Geweke, J. (1991). Efficient simulation from the multivariate normal and student-t distributions subject to linear constraints and the evaluation of constraint probabilities. In *Computing science and statistics: Proceedings of the 23rd symposium on the interface*, pp. 571–578. Interface Foundation of North America, Inc., Fairfax Station, VA.

Glaeser, E. L., B. I. Sacerdote, and J. A. Scheinkman (2003). The social multiplier. *Journal of the European Economic Association 1*(2-3), 345–353.

Goldsmith-Pinkham, P. and G. Imbens (2013a). Rejoinder. *Journal of Business & Economic Statistics 31*(3), 279–281.

Goldsmith-Pinkham, P. and G. W. Imbens (2013b). Social networks and the identification of peer effects. *Journal of Business & Economic Statistics 31*(3), 253–264.

Graham, B. S. (2013). Comment. *Journal of Business & Economic Statistics 31*(3), 266–270.

Heckman, J. J., S. Urzua, and E. Vytlacil (2006). Understanding instrumental variables in models with essential heterogeneity. *The Review of Economics and Statistics 88*(3), 389–432.

Heckman, J. J. and E. Vytlacil (2001). Policy-relevant treatment effects. *American Economic Review 91*(2), 107–111.

Heckman, J. J. and E. Vytlacil (2005). Structural equations, treatment effects, and econometric policy evaluation. *Econometrica 73*(3), 669–738.

Heckman, J. J. and E. J. Vytlacil (1999). Local instrumental variables and latent variable models for identifying and bounding treatment effects. *Proceedings of the national Academy of Sciences 96*(8), 4730–4734.

Hseih, C.-S. and L. F. Lee (2016). A Social Interactions Model with Endogenous Friendship Formation and Selectivity. *Journal of Applied Econometrics 31*(2), 301–319.

Jackson, M. O. (2008). *Social and economic networks*. Princeton university press.

Jackson, M. O. (2013). Unraveling Peers and Peer Effects: Comments on Goldsmith-Pinkham and Imbens'"Social Networks and the Identification of Peer Effects". *Journal of Business & Economic Statistics 31*, 270–273.

Kramarz, F. and O. N. Skans (2014). When strong ties are strong: Networks and youth labour market entry. *The Review of Economic Studies 81*(3), 1164–1200.

Kremer, M. and D. Levy (2008). Peer effects and alcohol use among college students. *The Journal of Economic Perspectives 22*(3), 189–189.

Lee, L.-f. (2007). Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics 140*(2), 333–374.

LeSage, J. P. and R. K. Pace (2009). *Introduction to Spatial Econometrics (Statistics, textbooks and monographs)*. CRC Press.

Leung, M. (2014). A random-field approach to inference in large models of network formation. *Available at SSRN*.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies 60*(3), 531–542.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. *Frontiers in Economics*, 105–142.

Moffitt, R. A. et al. (2001). Policy interventions, low-level equilibria, and social interactions. *Social dynamics 4*(45-82), 6–17.

Patacchini, E. and T. Arduini (2016). Residential choices of young americans. *Journal of Housing Economics 34*, 69–81.

Quandt, R. E. (1972). A New Approach to Estimating Switching Regressions. *Journal ofthe American Statistical Association 67*(338), 306–310.

Robinson, P. M. (1988). Root-N-consistent semiparametric regression. *Econometrica: Journal of the Econometric Society*, 931–954.

Roy, A. D. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers New Series 3*(2), 135–146.

Sacerdote, B. (2001). Peer Effects with Random Assignment: Results for Dartmouth Roommates. *Review of Economics and statistics 116*(2), 681–704.

A Appendix

B Additional tables and figures

B.1 Empirical application

Table B.1 presents the description of the variables used in the empirical application. For clarity, we have separated them into distinction groups– outcome variable (Y), covariates (X), and instruments (Z).

Table B.1: Description of variables

Variable name	Description
Outcome variable (Y), wage	Earnings reported at wave IV (individuals who do not report exact earnings are asked to report earnings within a range of values. We take the mid-value of this range.)
Treatment variable (S), college attendance	Dummy variable indicating education level is at least some college
Variables in X and Z	
Age	Age of participant at the time of the first wave (1993)
Female	Dummy variable indicating female, male
Race	Dummy variables indicating white, black or other race
School	Dummy variable indicating the school of the respondent
Mother's Education	Dummy variables indicating the respondent's mother 1. graduated high school and 2. attended college
Father's Education	Dummy variables indicating the respondent's father 1. graduated high school and 2. attended college
GPA	Sum of reported grades in Mathematics and English; A=4, B=3, C=2, D=1, no grade=0
Appearance	Response from the surveyor on a Likert scale to the question "How physically attractive is the respondent?"
Personality variables	Variables on a Likert scale for Shyness and Independence (E.g. Response to the question "How much do you agree with the statement 'You are Shy.'"
Variables in Z, not X	
Number of Siblings	Variable of the number of reported siblings
Local income at wave II	Income per capita at the local tract area level in 1990
Local unemployment at wave II	Unemployment at the local tract area level in 1990
Local proportion with a college degree at wave II	Proportion of residents over 25 who hold a bachelor degree or higher, taken at the local tract area level
Variables in X, not Z	
Married	Dummy variable indicating if the respondent has ever been married
Number of Kids	Variable of the number of reported children
Local income at wave IV	Income per capita at the local tract area level in 2008
Local unemployment at wave IV	Unemployment at the local tract area level in 2008

Table B.2 reports the summary statistics of the key variables in Table B.1. About 71% of individuals in the sample attended college and the sample is predominantly white individuals (84%) with only 14% of blacks. Mothers of the respondents seem slightly more educated than their fathers: 46% vs. 33% for High School and 39% vs. 35% for College. Around 55% of respondents are married with an average number of children of about 0.9.

Table B.2: Data Summary

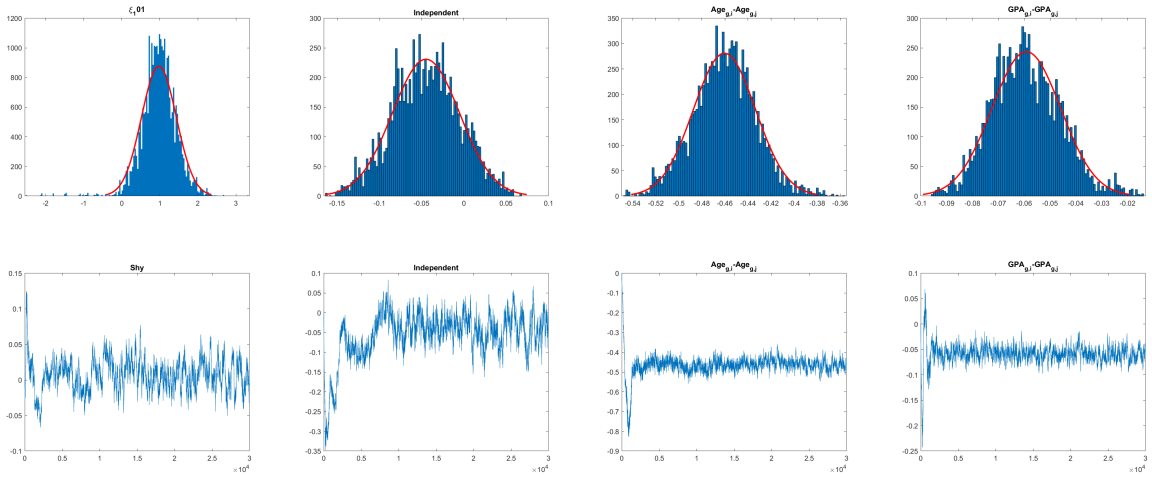
Variable Name	Min	Max	Mean	SD
Wage Y ($)	0	300000	30386	27531
College Attendance (S)	0	1	0.71	0.45
Age (months)	130	222	162.1	18.92
Male	0	1	0.47	0.50
Female	0	1	0.53	0.50
Race:				
White	0	1	0.84	0.36
Black	0	1	0.14	0.35
Other race	0	1	0.05	0.22
Mother's education:				
High School	0	1	0.46	0.50
College	0	1	0.39	0.49
Less than High School	0	1	0.10	0.30
Father's education:				
High School	0	1	0.33	0.48
College	0	1	0.35	0.47
Less than High School	0	1	0.11	0.31
Grade in Maths	0	4	2.54	1.34
Grade in English	0	4	2.84	1.14
Appearance	1	5	3.65	0.78
Shy	1	5	3.31	1.26
Independent	1	5	1.80	0.76
Instruments:				
Number of siblings	0	14	2.59	1.99
Local income 1990	3817	28501	11155	4090
Local unem. 1993	0.02	0.15	0.07	0.02
College ed. 1990	0.03	0.48	0.18	0.11
Variables in X only:				
Local income 2008	8500	94950	23419	10230
Local unem. 2008	0	0.35	0.07	0.04
Married	0	1	0.55	0.50
Number of Children	0	5	0.90	1.08

Table B.3: Test of equality of LATEs over different intervals for high school networks

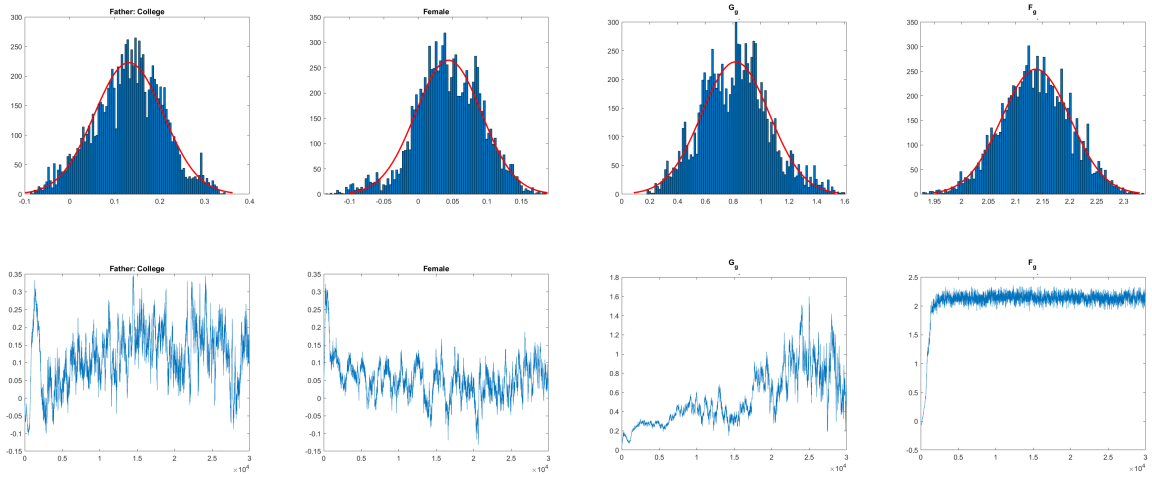
Ranges of U_S for $LATE_j$	(0,0.04)	(0.08,0.12)	(0.16,0.20)	(0.24,0.28)	(0.32,0.36)	(0.40,0.44)
Ranges of U_S for $LATE_{j+1}$	(0.08,0.12)	(0.16,0.20)	(0.24,0.28)	(0.32,0.36)	(0.40,0.44)	(0.48,0.52)
Difference in LATEs	0.2325	0.2497	0.2698	0.2929	0.3194	0.3496
p-value	0	0	0	0	0	0
Ranges of U_S for $LATE_j$	(0.48,0.52)	(0.56,0.60)	(0.64,0.68)	(0.72,0.76)	(0.80,0.84)	(0.88,0.92)
Ranges of U_S for $LATE_{j+1}$	(0.56,0.60)	(0.64,0.68)	(0.72,0.76)	(0.80,0.84)	(0.88,0.92)	(0.96,1)
Difference in LATEs	0.3838	0.4222	0.4650	0.5124	0.5646	0.6216
p-value	0	0	0	0	0	0
joint p-value			0			

Figures B.1-B.2 report the MCMC plots of parameters' posterior distributions.

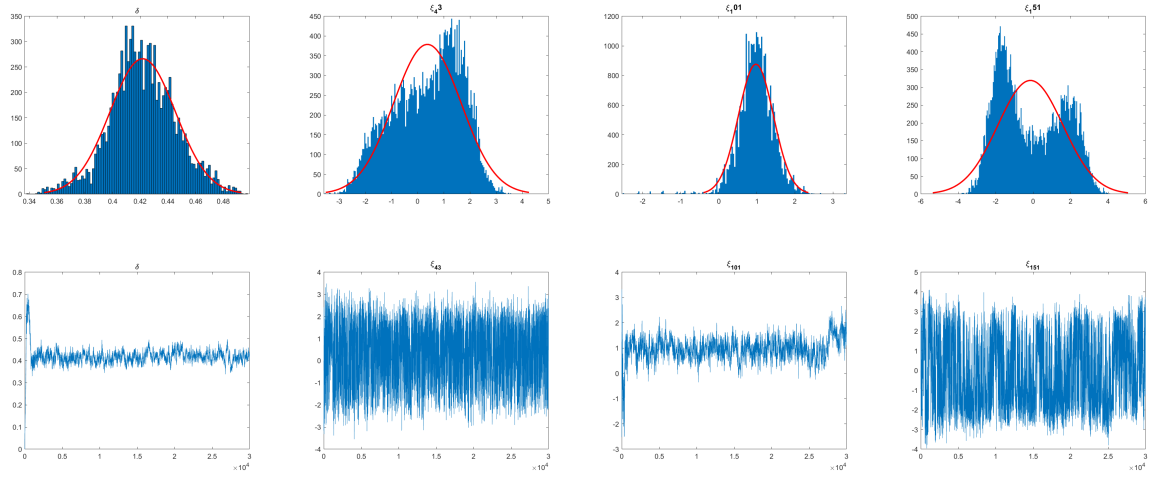
Figure B.1: Network formation– MCMC plots for high school networks



(a) Shy (b) Independent (c) $Age_{g,i} - Age_{g,j}$ (d) $GPA_{g,i} - GPA_{g,j}$

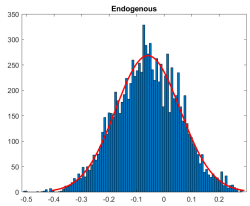


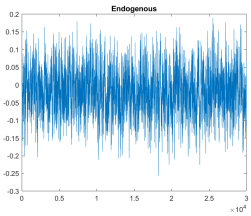
(e) Father College (f) Female (g) $G_{g\_}$ (h) $F_{g\_}$



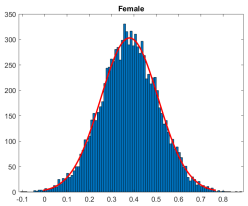
(i) δ (j) ξ_{43} (k) ξ_{101} (l) ξ_{151}

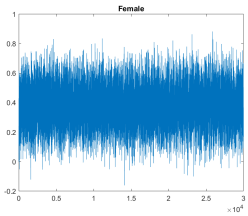
Figure B.2: Treatment decision– MCMC plots for high school networks



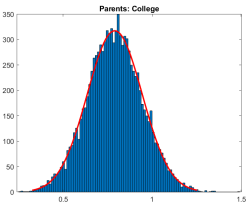


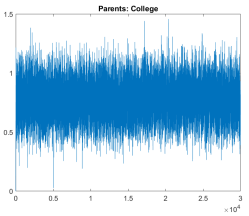
(a) Endogenous



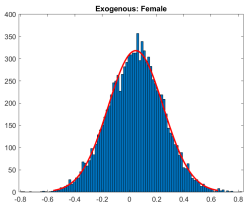


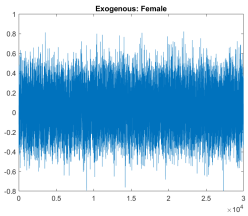
(b) Female



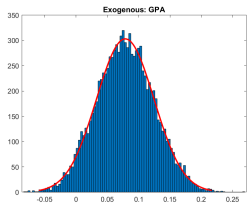


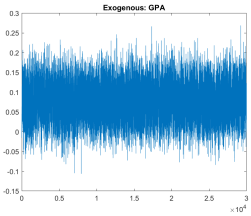
(c) Parents: College



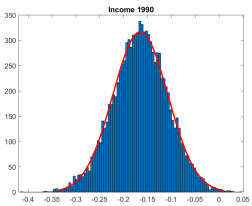


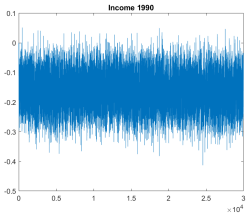
(d) Exogenous: Female



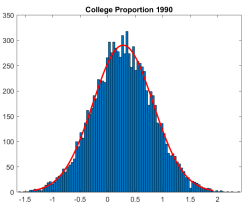


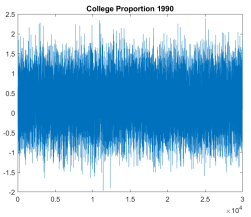
(e) Exogenous: GPA



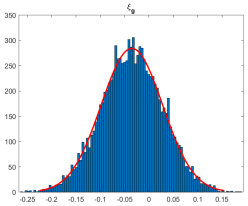


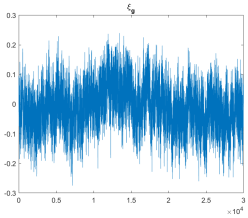
(f) Income 1990



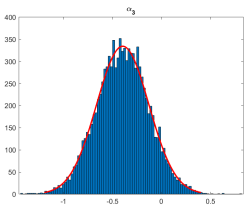


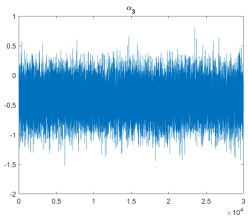
(g) College Proportion 1990



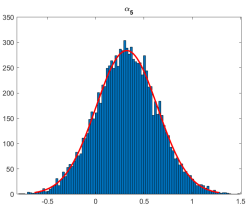


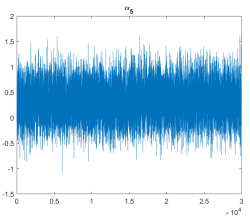
(h) ξ_g





(i) α_3



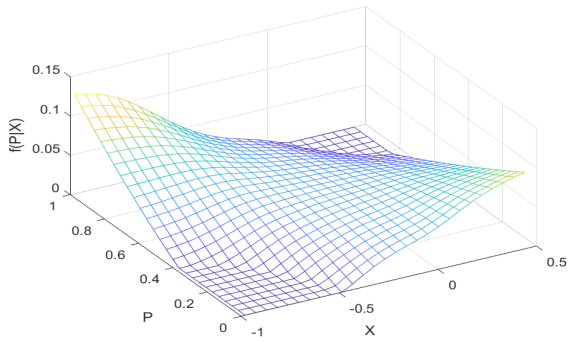


(j) α_5

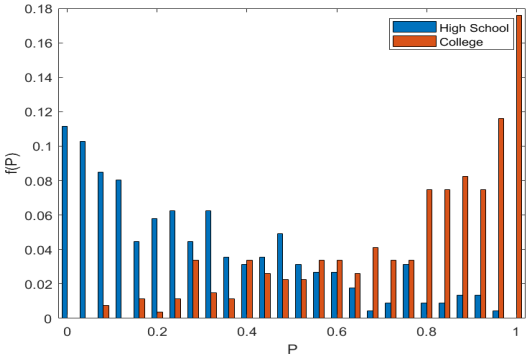
B.2 Monte Carlo experiments

B.2.1 MTE plots

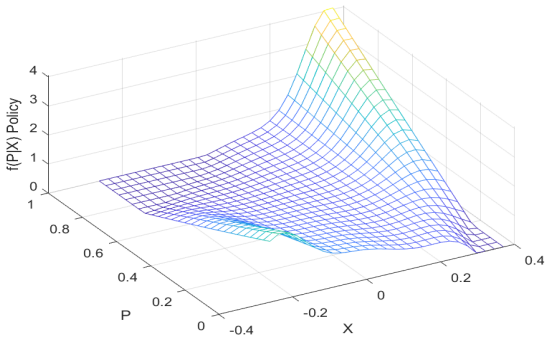
Figure B.3: $\rho_x = 0.3$– Support of $P(Z, \boldsymbol{G})$ given X (left)– Support of $P(Z, \boldsymbol{G})$ for $S = 0$ and $S = 1$ (right)



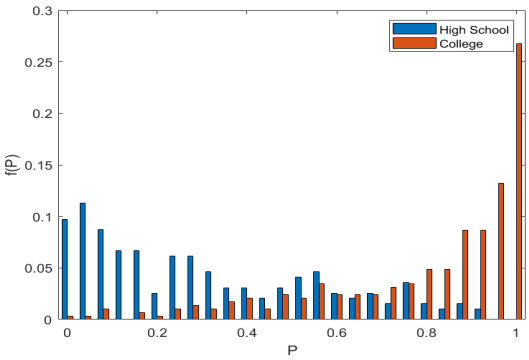
(a) Model (I)



(b) Model (I)

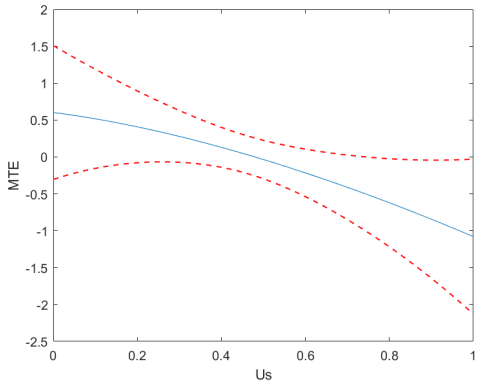


(c) Model (II)

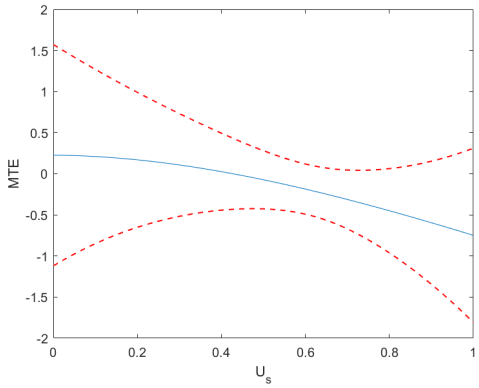


(d) Model (II)

Figure B.4: MTE with $\rho_x = 0.3$



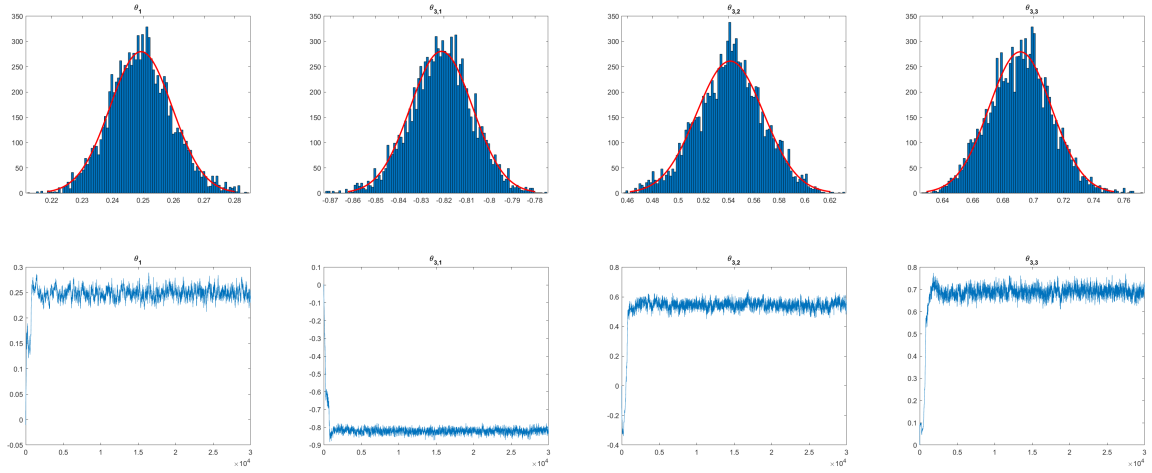
Model (I)



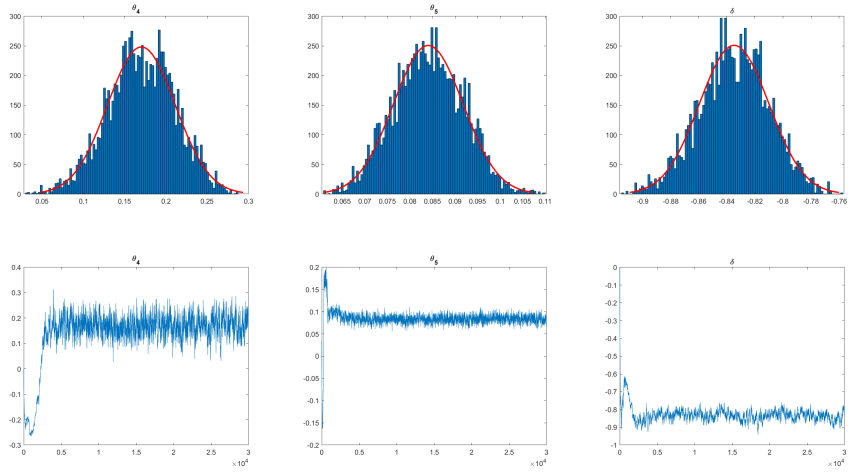
Model (II)

B.3 MCMC plots of parameters' posterior distributions

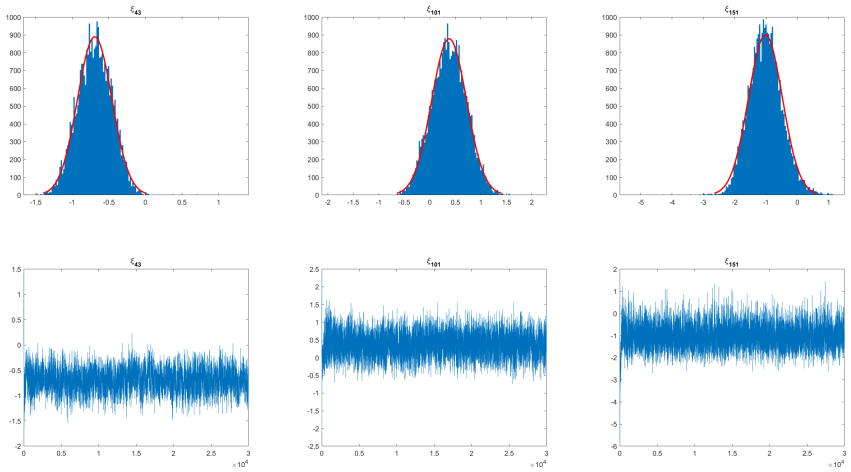
Figure B.5: Network formation model– MCMC plots of posterior distributions



(a) θ_1 (b) $\theta_{3,1}$ (c) $\theta_{3,2}$ (d) $\theta_{3,3}$

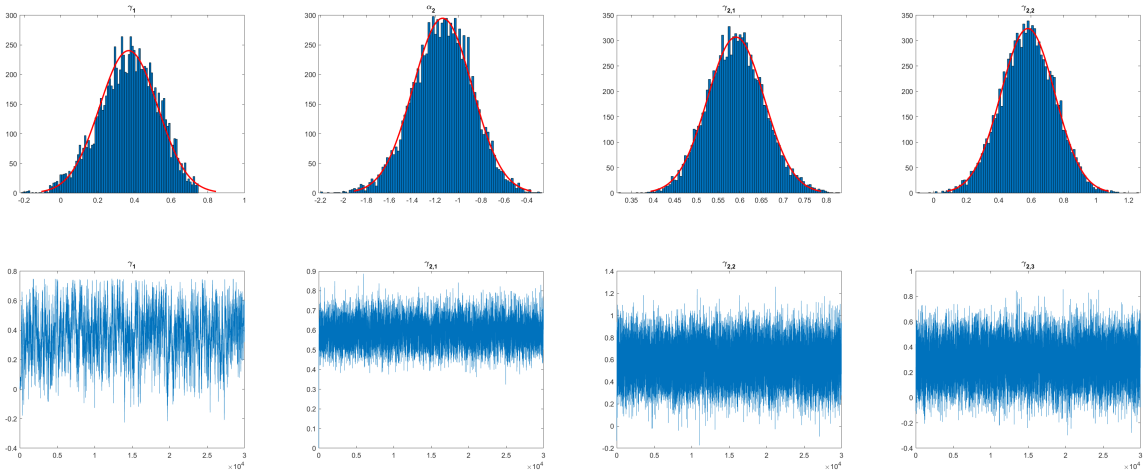


(e) θ_4 (f) θ_5 (g) δ



(h) ξ_{43} (i) ξ_{101} (j) ξ_{151}

Figure B.6: Treatment decision– MCMC plots of posterior distributions

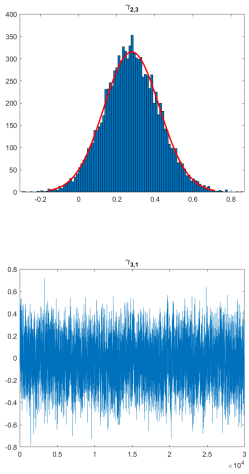


(a) γ_1

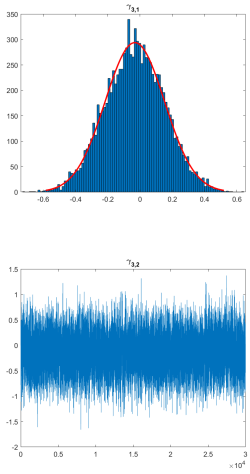
(b) γ_0

(c) $\gamma_{2,1}$

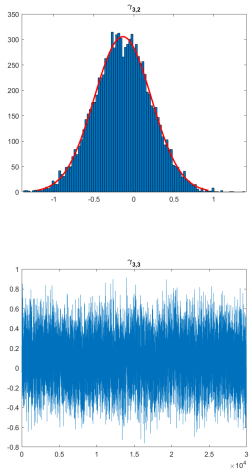
(d) $\gamma_{2,2}$



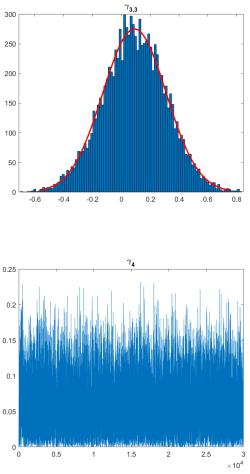
(e) $\gamma_{2,3}$



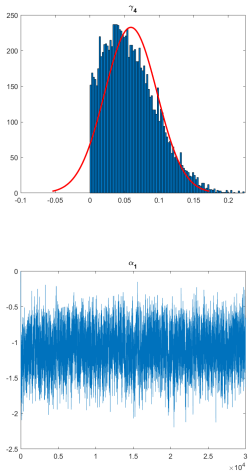
(f) $\gamma_{3,1}$



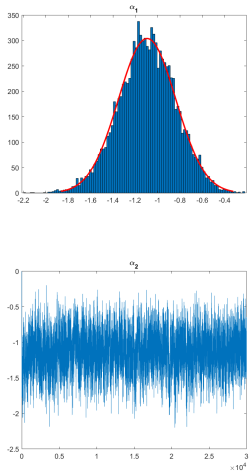
(g) $\gamma_{3,2}$



(h) $\gamma_{3,3}$



(i) α_1



(j) α_2