Supplementary Appendices of “Spatial Modeling Approach for Dynamic Network Formation and Interactions”

Xiaoyi Han    Chih-Sheng Hsieh    Stanley I.M. Ko

Abstract

We provide a few technical details, additional simulation and empirical results in this supplementary appendices. In particular, supplementary Appendix B provides the discussion on the identification of latent variables and their coefficients. Supplementary Appendix C outlines the detailed MCMC estimation algorithm for the proposed model. We explain the model selection criterion for the latent dimension in supplementary Appendix D. Empirical results from other network relationships and robustness checks of model specifications are reported in supplementary Appendix E. We studies a counter-factual policy simulation to examine the multiplier effects from network interactions in Supplementary Appendix F. The goodness-of-fit of our network formation model to the real data is examined in supplementary Appendix G, while supplementary Appendix H depicts the time evolution of unobserved latent variables. Additional empirical tables and figures are relegated to supplementary Appendix I. The MATLAB codes used in this study are available at https://sites.google.com/site/chihshenghsieh/research or available from the author upon request.
B Discussions on the identification of $z_{igt}$’s, $m_{igt}$’s and their coefficients

In this appendix we discuss how we can identify individual latent variables $z_{igt}$’s, $m_{igt}$’s, and their coefficients in the model. We first follow Hsieh and Lee (2016) and Hsieh and van Kippersluis (2018) to justify that the coefficients of $|z_{igp,t} - z_{jgp,t}|$’s and $m_{igp,t}$’s in the network formation model, i.e., $\delta_{p1}$’s, $\xi_{p2}$’s, and $\zeta_{p2}$’s, can be identified under the identification constraints in Section 2.3. We then show that under the same identification constraints, conditional on the identified network formation model, all of the unknown parameters in the SC-SDPD model, as well as $z_{igt}$’s, $m_{igt}$’s, can be identified. Lastly, when $z_{igt}$’s and $m_{igt}$’s do not appear in the network formation model, the “rotational indeterminancy” issue on them can be solved via the Procrustes transformation in Abmann et al. (2012, 2016). We drop the group subscript $g$ in Appendix B.1 for notation simplicity.

B.1 Identification of $\delta$, $\xi$, and $\zeta$ in the network formation model

Consider the general network formation model in Eqs. (2) and (3) of the main text. As in Appendix A it is similar to a standard dichotomous choice model that can be motivated from individual’s utility maximization. Recall that $\nu_{ijt}$ denote $i$’s utility from the link with individual $j$ at time $t$. Individual $i$ would choose to form a link with $j$ ($w_{ijt} = 1$) if $\nu_{ijt}(w_{ijt} = 1) - \nu_{ijt}(w_{ijt} = 0) > 0$, otherwise would choose to be disconnected ($w_{ijt} = 0$). $C_t$ denotes the collection of all observed exogenous variables $c_{it}$’s and $c_{ijt}$’s in the network formation model at time $t$, and $R_t$ is a $n \times n$ matrix, with the $(i,j)^{th}$ element $r_{ijt} = \sum_{l=1}^{n} w_{ilt} \times w_{jlt}$ being the number of common friends shared by $i$ and $j$ at time $t$. The utility difference is

$$\nu_{ijt}(w_{ijt} = 1) - \nu_{ijt}(w_{ijt} = 0) = \eta_{ijt}(W_{t-1}, R_{t-1}, C_t, Y_{t-1}) + \varphi_{ijt}(H_t, \delta, \xi, \zeta),$$

where the deterministic part $\eta_{ijt}(W_{t-1}, R_{t-1}, C_t, Y_{t-1})$ contains $\gamma_0 + \alpha_{it} \gamma_1 + \alpha_{jt} \gamma_2 + \alpha_{ijt} \gamma_3 + \gamma_4 w_{ij,t-1} + \gamma_5 r_{ij,t-1} + \gamma_6 y_{it,t-1} + \gamma_7 y_{jt,t-1} + \gamma_8 |y_{it,t-1} - y_{jt,t-1}|$, namely, all predetermined variables and exogenous variables at time $t$. The error term $\varphi_{ijt}(H_t, \delta, \xi, \zeta) = \tilde{\varphi}_{ijt}(H_t, \delta, \xi, \zeta) + \omega_{ijt}$ includes latent distances $\sum_{p1=1}^{p1} \delta_{p1} |z_{ip1,t} - z_{jip1,t}|$, time-varying degree heterogeneity effects $\sum_{p2=1}^{p2} \xi_{p2} m_{ip2,t} + \sum_{p2=1}^{p2} \xi_{p2} m_{jip2,t}$, and a pure i.i.d disturbance $\omega_{ijt}$ across all $(i,j)$ pair and $t$. According to Ichimura (1993), the dichotomous choice model for network formation can be viewed as a single index equation, namely,

$$E(w_{ijt}|W_{t-1}, R_{t-1}, C_t, Y_{t-1}) = P(w_{ijt} = 1|W_{t-1}, R_{t-1}, C_t, Y_{t-1}) = 1 - F_{\varphi_{ijt}}(-\eta_{ijt}), \quad (B.1)$$
with $F_{\varphi_{ijt}}(.)$ being the distribution function of $\varphi_{ijt}$. Ichimura (1993) shows that, even when $F_{\varphi_{ijt}}(.)$ is unknown, parameters in the linear index $\eta_{ijt}$ are identified up to a scale. With further parametric and normalization assumptions on $\varphi_{ijt}$, parameters in $\eta_{ijt}$ are identified, and hence $\eta_{ijt}$ can be determined. As $\eta_{ijt}$ is identified, the distribution function $F_{\varphi_{ijt}}$ can also be identified (estimated) from the data by a nonparametric kernel regression with $\eta_{ijt}$ as the regressor. Given that $\varphi_{ijt}$ is continuous, if we assume $\eta_{ijt}$ can take values that cover the support of the probability density function $f_{\varphi_{ijt}}(.)$, the moments of $\varphi_{ijt}$ can also be estimated from the data.

Recall that, from Eqs. (4) and (5) of the main text, for each dimension of $z_{it}$’s and $m_{it}$’s, we have $z_{ip_1t} \sim N(0, t\sigma_z^2)$ and $m_{ip_2t} \sim N(0, t\sigma_m^2)$ for $p_1 = 1, 2, \cdots, \tilde{p}_1$, $p_2 = 1, 2, \cdots, \tilde{p}_2$ and $t = 2, 3, \cdots, T$. Also, $z_{ip_1t} \sim N(0, \sigma_z^2)$ and $m_{ip_2t} \sim N(0, \sigma_m^2)$ for $t = 1$. After imposing normalization constraints on $\sigma_z^2$, $\sigma_\eta^2$, $\sigma_\phi^2$ and $\sigma_m^2$, the distributions of $z_{ip_1t}$’s and $m_{ip_2t}$’s simplify to $z_{ip_1t} \sim N(0, t)$ and $m_{ip_2t} \sim N(0, t)$. Also, we assume $z_{ip_1t}$’s and $m_{ip_2t}$’s are independent cross individuals and dimensions. With moments of $\varphi_{ijt}$’s, we can study the identifications of $\delta_{p_1}, \xi_{p_2}$ and $\zeta_{p_2}$ for $p_1 = 1, 2, \cdots, \tilde{p}_1$ and $p_2 = 1, 2, \cdots, \tilde{p}_2$.

First, consider the case where $\tilde{p}_1 = 1$ and $\tilde{p}_2 = 1$. In this case, we have $\varphi_{ijt}(H, \delta, \xi, \zeta) = \delta_1|z_{i1t} - z_{j1t}| + \xi_1m_{i1t} + \zeta_1m_{j1t} + \omega_{ijt}$. The variance of $\varphi_{ijt}$ equals $\delta_1^2\text{var}(|z_{i1t} - z_{j1t}|) + (\xi_1^2 + \zeta_1^2)\text{var}(m_{i1t}) + \sigma_\omega^2$. Note that $\sigma_\omega^2$ is usually normalized to one due to the arbitrary scaling problem in discrete choice models. We can calculate $\text{var}(|z_{i1t} - z_{j1t}|)$’s and $\text{var}(m_{i1t})$ given the distribution of $z_{it}$’s and $m_{it}$’s. So the variance of $\varphi_{ijt}$ gives us one moment equation for $\delta_1^2$, $\xi_1^2$ and $\zeta_1^2$. Furthermore, we can obtain other moment equations by exploring the covariance of $\varphi_{ijt}$’s. Notice that $E(\varphi_{ijt}) = 0$. For any $i \neq \tilde{i}$ and $j \neq \tilde{j}$, we have

\[
\text{Cov}(\varphi_{ijt}, \varphi_{\tilde{i}j\tilde{t}}) = E(\varphi_{ijt}\varphi_{\tilde{i}j\tilde{t}}) = \delta_1^2E(|z_{i1t} - z_{j1t}|z_{i1t} - z_{j1t}) + \delta_1^2E(m_{i1t}^2),
\]

\[
\text{Cov}(\varphi_{ijt}, \varphi_{i\tilde{t}j}) = E(\varphi_{ijt}\varphi_{i\tilde{t}j}) = \delta_1^2E(|z_{i1t} - z_{j1t}|z_{i1t} - z_{j1t}) + \zeta_1^2E(m_{j1t}^2).
\]

So we can have 3 polynomial equations involving $\delta_1^2$, $\xi_1^2$ and $\zeta_1^2$ to identify the magnitudes of $\delta_1$, $\xi_1$ and $\zeta_1$. Regarding their signs, as $\delta_1$ is the coefficient of $|z_{i1t} - z_{j1t}| > 0$, the sign of $\delta_1$ can be determined. In the SC-SDPD model we assume $\kappa_2 > 0$ and it pins down the unobserved signs of $m_{it}$’s. So the signs of $\xi_1$ and $\zeta_1$ are also determined as the coefficients of $m_{it}$’s in the network formation model. Therefore, we can identify both the magnitudes and the signs of $\delta_1$, $\xi_1$ and $\zeta_1$.

Next, we consider the case where $\tilde{p}_1 = 2$ and $\tilde{p}_2 = 1$. We have $\varphi_{ijt}(H, \delta, \xi, \zeta) = \delta_1|z_{i1t} - z_{j1t}| + \delta_2|z_{i2t} - z_{j2t}| + \xi_1m_{i1t} + \zeta_1m_{j1t} + \omega_{ijt}$. Given the distributional assumptions
With the above three moment equations we can only identify \( m_{it}'s \) and \( m_{it}'s \), we have

\[
\begin{align*}
\text{Var}(\varphi_{ijt}) &= (\delta_1^2 + \delta_2^2)\text{Var}(|z_{it} - z_{jt}|) + (\xi_1^2 + \xi_1^2)\text{Var}(m_{it}) + \sigma_\omega^2, \\
\text{Cov}(\varphi_{ijt}, \varphi_{ijt}') &= (\delta_1^2 + \delta_2^2)E(|z_{it} - z_{jt}|^2|z_{jt} - z_{jt}|) + \xi_i^2E(m_{it}^2), \\
\text{Cov}(\varphi_{ijt}, \varphi_{ijt}') &= (\delta_1^2 + \delta_2^2)E(|z_{it} - z_{jt}|^2|z_{jt} - z_{jt}|) + \xi_i^2E(m_{jt}^2).
\end{align*}
\]

Based on the above three equations, we can identify coefficients in the degree heterogeneity component, \( \xi_i^2 \) and \( \xi_i^2 \), and \((\delta_1^2 + \delta_2^2)\) as a summation. To separately identify \( \delta_1^2 \) and \( \delta_2^2 \) we need to look at higher order moments of \( \varphi_{ijt} \). For instance, we may explore \( \text{Cov}(\varphi_{ijt}^2, \varphi_{ijt}') \), which is

\[
\text{Cov}(\varphi_{ijt}^2, \varphi_{ijt}') = E(\varphi_{ijt}^2\varphi_{ijt}') = (\delta_1^3 + \delta_2^3)E((|z_{it} - z_{jt}|^2|z_{jt} - z_{jt}|) + \xi_1^3E(m_{jt}^3).
\]

As we have assumed the distributions of \( z_{it}'s \) and \( m_{it}'s \), \( E(|z_{it} - z_{jt}|^2|z_{jt} - z_{jt}|) \) and \( E(m_{jt}^3) \) can be calculated. So we obtain another polynomial equation on \( \delta_1^2 \) and \( \delta_2^2 \). More polynomial equations can be derived from the fourth or higher order moments of \( \varphi_{ijt} \)'s. These polynomial equations can identify \( \delta_1 \) and \( \delta_2 \) separately. After showing identification of the magnitudes of \( \xi \), \( \zeta \), \( \delta \) and \( \delta \), we follow the same argument in the previous case to argue identification of their signs.

Similarly, consider the case where \( \tilde{p}_1 = 1 \) and \( \tilde{p}_2 = 2 \). Then, \( \varphi_{ijt}(H_i, \tilde{p}_1, \xi, \zeta) = \delta_1|z_{it} - z_{jt}| + \xi_1m_{it} + \xi_1m_{jt} + \xi_2m_{2t} + \xi_2m_{2t} + \omega_{ijt} \). We may first explore the variance and covariance of \( \varphi_{ijt} \)'s, namely,

\[
\begin{align*}
\text{Var}(\varphi_{ijt}) &= \delta_1^2\text{Var}(|z_{it} - z_{jt}|) + (\xi_1^2 + \xi_1^2 + \xi_2^2 + \xi_2^2)\text{Var}(m_{it}) + \sigma_\omega^2, \\
\text{Cov}(\varphi_{ijt}, \varphi_{ijt}') &= \delta_1^2E(|z_{it} - z_{jt}|^2|z_{jt} - z_{jt}|) + (\xi_1^2 + \xi_2^2)E(m_{it}^2), \\
\text{Cov}(\varphi_{ijt}, \varphi_{ijt}') &= \delta_1^2E(|z_{it} - z_{jt}|^2|z_{jt} - z_{jt}|) + (\xi_1^2 + \xi_2^2)E(m_{jt}^2).
\end{align*}
\]

With the above three moment equations we can only identify \( \delta_1 \), and the summations, \((\xi_1^2 + \xi_2^2)\) and \((\xi_1^2 + \xi_2^2)\). To separately identify \( \xi_1 \), \( \xi_2 \), \( \xi_1 \) and \( \xi_2 \), we can further exploit \( \text{Cov}(\varphi_{ijt}^2, \varphi_{ijt}^2) \) and \( \text{Cov}(\varphi_{ijt}^2, \varphi_{ijt}') \), namely,

\[
\begin{align*}
\text{Cov}(\varphi_{ijt}^2, \varphi_{ijt}') &= \delta_1^3E((|z_{it} - z_{jt}|^2|z_{jt} - z_{jt}|) + (\xi_1^3 + \xi_2^3)E(m_{jt}^3), \\
\text{Cov}(\varphi_{ijt}^2, \varphi_{ijt}') &= \delta_1^3E((|z_{it} - z_{jt}|^2|z_{jt} - z_{jt}|) + (\xi_1^3 + \xi_2^3)E(m_{jt}^3).
\end{align*}
\]

and some other higher order moments of \( \varphi_{ijt} \)'s.

Generally, as \( \tilde{p}_1 \) or \( \tilde{p}_2 \) increases, we can derive more polynomial equations from higher-order moments of \( \varphi_{ijt} \)'s. Eventually, the system of these polynomial equations can be used to solve for \( \delta_{p_1}'s \), \( \xi_{p_2}'s \) and \( \zeta_{p_2}'s \). To distinguish between different dimensions of \( z_{it}'s \) and
In this appendix, we discuss the identification of unknown parameters in the SC-SDPD model. In Section 2.3, using the same strategy, we can identify $\delta_i$’s, $\xi_j$’s and $\zeta_{ij}$’s in three or higher dimensions.

### B.2 Identification of unknown parameters in the SC-SDPD model

In this appendix, we discuss the identification of unknown parameters in the SC-SDPD model. We first eliminate the group-time effect $\alpha_{gt}$’s by a difference approach. Let $J_g$ be the corresponding $(n_g - 1) \times n_g$ difference matrix, namely,

$$J_g = \begin{pmatrix} -1 & 1 \\ -1 & 1 \\ \vdots & \vdots \end{pmatrix}.$$  

The variables $Y_{gt}$, $W_{gt}Y_{gt}$, $Y_{g,t-1}$, $W_{g,t-1}Y_{g,t-1}$, $X_{gt}$, $W_{gt}X_{gt}$, $\tau_g$, $H_{gt}$ and $V_{gt}$ are transformed to $J_gY_{gt}$, $J_gW_{gt}Y_{gt}$, $J_gY_{g,t-1}$, $J_gW_{g,t-1}Y_{g,t-1}$, $J_gX_{gt}$, $J_gW_{gt}X_{gt}$, $J_g\tau_g$, $J_gH_{gt}$ and $J_gV_{gt}$. The SC-SDPD model after the first differencing is

$$J_gY_{gt} = \lambda J_gW_{gt}Y_{gt} + \rho J_gY_{g,t-1} + \mu J_gW_{g,t-1}Y_{g,t-1} + J_gX_{gt}\beta_1 + J_gW_{gt}X_{gt}\beta_2 + J_g\tau_g + J_gH_{gt}\kappa + J_gV_{gt},$$  \hspace{1cm} (B.2)

for $t = 1, 2, \cdots, T$. We then take the second differencing on Eq. (B.2) across time to eliminate the individual effect $\tau_g$’s. Specifically,

$$J_g(Y_{gt} - Y_{g,t-1}) = \lambda J_g(W_{gt}Y_{gt} - W_{g,t-1}Y_{g,t-1}) + \rho J_g(Y_{g,t-1} - Y_{g,t-2}) + \mu J_g(W_{g,t-1}Y_{g,t-1} - W_{g,t-2}Y_{g,t-2}) + J_gX_{gt} - X_{g,t-1}\beta_1 + J_gW_{gt}X_{gt} - W_{g,t-1}X_{g,t-1}\beta_2 + J_g(H_{gt} - H_{g,t-1})\kappa + J_g(V_{gt} - V_{g,t-1}), \quad t = 2, 3, \cdots, T.$$  

Denote $Y^*_g = J_g(Y_{gt} - Y_{g,t-1})$, $(W_{gt}Y_{gt})^* = J_g(W_{gt}Y_{gt} - W_{g,t-1}Y_{g,t-1})$, $X^*_g = J_g(X_{gt} - X_{g,t-1})$, $(W_{gt}X_{gt})^* = J_g(W_{gt}X_{gt} - W_{g,t-1}X_{g,t-1})$, $H_{gt}^* = J_g(H_{gt} - H_{g,t-1})$ and $V_{gt}^* = J_g(V_{gt} - V_{g,t-1})$. The SC-SDPD model after the second differencing is

$$Y^*_g = \lambda(W_{gt}Y_{gt})^* + \rho Y_{g,t-1}^* + \mu(W_{g,t-1}Y_{g,t-1})^* + X^*_g\beta_1 + (W_{gt}X_{gt})^*\beta_2 + H_{gt}^*\kappa + V_{gt}^*, \quad (B.3)$$

for $t = 2, 3, \cdots, T$. Taking expectation of Eq. (B.3) conditional on $W_{gt}$ gives

$$E(Y^*_g|W_{gt}) = \lambda E((W_{gt}Y_{gt})^*|W_{gt}) + \rho E(Y_{g,t-1}^*|W_{gt}) + \mu E((W_{g,t-1}Y_{g,t-1})^*|W_{gt}) + E(X^*_g|W_{gt})\beta_1 + E((W_{gt}X_{gt})^*|W_{gt})\beta_2 + E(H_{gt}^*|W_{gt})\kappa,$$  \hspace{1cm} (B.4)
for \( g = 1, 2, \cdots, G \) and \( t = 2, 3, \cdots, T \). As we assume \( \kappa > 0 \), the signs of \( H_{gt} \)’s are fixed. So all terms in Eq. (B.4) can be identified from the data. In particular, we can identify

\[
E(H^*_t|W_t) = \int H^*_tP(H_{gt}, H_{g,t-1}|W_{gt})dH_{g,t-1}dH_{gt} \\
= \int H^*_t \frac{P(H_{gt}|H_{g,t-1})P(W_{gt}|H_{gt}, H_{g,t-1})}{P(W_{gt})}dH_{g,t-1}dH_{gt},
\]

provided that \( P(W_{gt}|H_{gt}, H_{g,t-1}) = P(W_{gt}|H_{gt}) \) and the parameters in \( P(W_{gt}|H_{gt}) \) are identified (estimated) from the network function model, as discussed in Appendix B.1. Let

\[
\mathcal{J}_t = [E((\overline{W}_t Y_t)^*|W_t), E(Y_{t-1}^*|W_t), E((\overline{W}_{t-1} Y_{t-1})^*|W_t), E(X_t^*|W_t), E((\overline{W}_t X_t)^*|W_t), E(H_t^*|W_t)].
\]

Dropping the subscript \( g \) means \( \mathcal{J}_t \) stacks observations across groups for \( t = 2, 3, \cdots, T \). Denote \( \mathcal{J} = (\mathcal{J}_2, \mathcal{J}_3, \cdots, \mathcal{J}_T)' \). The condition that \( \mathcal{J}' \mathcal{J} \) has full rank will identify parameters in Eq. (B.4), including \( \kappa \). Hence, there is no rotational indeterminacy issue on \( H_{gt} \)’s and \( \kappa \).

### B.3 Identification of individual unobservables \( z_{igt} \)’s, \( m_{igt} \)’s and \( \kappa \) through Procrustes transformation

In Appendix B.2, we show that if \( \mathcal{J}' \mathcal{J} \) has full rank, \( z_{igt} \)’s, \( m_{igt} \)’s and \( \kappa \) can be identified from the model. Thus, the rotation matrix \( Q \) in \( H_{gt} Q Q^{-1} \kappa \) can be determined. However, when \( z_{igt} \)’s and \( m_{igt} \)’s do not present in the network formation model, i.e., when \( \delta_{p_1} = 0 \) for \( p_1 = 1, 2, \cdots, \bar{p}_1 \), \( \xi_{p_2} = 0 \) and \( \zeta_{p_2} = 0 \) for \( p_2 = 1, 2, \cdots, \bar{p}_2 \), the full rank condition would be violated because the conditional expectation of \( H^*_t \) simplifies to \( E(H^*_t|W_{gt}) = E(H^*_t) = J_y E(H_{gt} - H_{g,t-1}) = 0 \). In this case, \( h_{igt} \)’s only take place as \( H_{gt} \kappa \)’s in the control function of the SC-SDPD model. Since \( H_{gt} Q Q^{-1} \kappa = H_{gt} \kappa \) for \( g = 1, 2, \cdots, G \), with any \( \bar{p} \times \bar{p} \) nonsingular matrix \( Q \), \( H_{gt} \)’s and \( \kappa \) are not separately identified. (It is possible that only \( Z_{gt} \)’s or \( M_{gt} \)’s (but not both) shows up in the network formation model. Then the full rank condition would also be violated and we only have the identification problem for \( Z_{gt} \)’s and \( \kappa_1 \), or \( M_{gt} \)’s and \( \kappa_2 \). In this case, the Procrustes transformation algorithm in this appendix can still be applied on \( Z_{gt} \)’s or \( M_{gt} \)’s.) Recall that, from distributional assumptions in Eqs. (6) and (7), and the normalization that \( \sigma^2_{z_o}, \sigma^2_{m_o}, \sigma^2_{z} \) and \( \sigma^2_{m} \) all being 1, we have \( h_{igt} = (z_{igt}', m_{igt}')' \sim \mathcal{N}_p(0, t I_{\bar{p}}) \) for \( t = 1, 2, \cdots, T, \). So the distribution of the rotated individual unobservable \( Q'h_{igt} \) is \( \mathcal{N}_p (0, tQ'Q) \). To keep the distribution invariant after rotation, we have \( Q'Q = I_{\bar{p}} \), which restricts \( Q \) to be orthogonal. This provides \( \frac{(t \bar{p} + 1)}{2} \) identification conditions. To determine \( Q \), additional \( \frac{(t \bar{p} - 1)}{2} \) conditions are needed. This
is very similar to the “rotational indeterminacy” problem in common factor models. See, among others, Bai and Ng (2013), Bai and Wang (2015) and Aßmann et al. (2016), for more discussions.

In this subsection, we utilize the orthogonal Procrustean transformation in Aßmann et al. (2012, 2016) to impose extra identification conditions on the posterior draws of \( h_{igt} \). Let \( \tilde{n} = \sum_{g=1}^{G} n_g \), \( H_t = (H_{1t}, H_{2t}, \ldots, H_{Gt})' \) and \( \mathbf{H} = (H_1', H_2', \ldots, H_T')' \) be the \((\tilde{n}T) \times \tilde{p}\) matrix of individual unobservables across all groups and periods. Denote \( \mathbf{H}^{(s)} \) as the posterior draw of \( \mathbf{H} \) at iteration \( s \). As suggested by Aßmann et al. (2012, 2016), the posterior sampler gives orthogonally mixing samples of \( \mathbf{H}^{(s)}'s \). Without further restrictions, \( \mathbf{H}^{(s)} \) is subjected to an \( \tilde{p} \times \tilde{p} \) unknown orthogonal transformation of \( \mathbf{H}^{(s)}Q^{(s)} \).

So as long as we can pin down all \( Q^{(s)}'s \), identification is reached up to orientation. With solution \( Q^{(s)} \), original posterior draws of \( \mathbf{H}^{(s)} \) and \( \kappa^{(s)} \) can be transformed as “identified draws” \( \mathbf{H}^{(s)}Q^{(s)} \) and \( Q^{(s)}\kappa^{(s)} \). Below is the detailed algorithm.

**Step 0:** Set the initial values of \( \mathbf{H}^{(s)}'s \) to be their original MCMC draws at iteration \( s \). Let the initial value of \( \tilde{\mathbf{H}} \) be the last draw of the posterior sample of \( \mathbf{H} \).

**Step 1:** Conditional on \( \tilde{\mathbf{H}} \) and \( \mathbf{H}^{(s)} \), solve the following minimization problem for \( Q^{(s)} \):

\[
\min \text{tr}[(\mathbf{H}^{(s)}Q^{(s)} - \tilde{\mathbf{H}})'(\mathbf{H}^{(s)}Q^{(s)} - \tilde{\mathbf{H}})], \text{ s.t. } Q^{(s)}Q^{(s)'} = I_{\tilde{p}}. \tag{B.6}
\]

The detailed derivation of the solution to this orthogonal Procrustes transformation problem can be found in Schönenmann (1966) and Borg and Groenen (2005), among others. It can be implemented by the following sub-steps:

1. **Compute** \( L_2^{(s)} = H^{(s)'}\tilde{\mathbf{H}} \).
2. **Conduct** the eigenvalue decomposition \( L_2^{(s)'}L_2^{(s)} = J_sD_sJ'_s \) and \( L_2^{(s)}L_2^{(s)'} = K_sD_sK'_s \).
3. **Denote** \( Re_s \) as a \( \tilde{p} \times \tilde{p} \) diagonal matrix, with +1 or −1 on the diagonal. Find the
unique $Re_s$, such that the $\tilde{p}$ main diagonals of $J_sL_2(s)'(K_sRe_s)$ are non-negative.

1.4 Derive the orthogonal transformation matrix $Q(s) = (K_sRe_s)J_s$. Transform $H(s)$ as $
abla(s) = H(s)Q(s)$.

**Step 2:** Conditional on $Q(s)$ and $H(s)$, derive $\tilde{H} = \frac{1}{s} \sum_{s=1}^{S} H(s) = \frac{1}{s} \sum_{s=1}^{S} H(s)Q(s)$. Set $H(s) = \tilde{H}(s)$. Go back to Step 1.

These two steps goes iteratively until a fixed point $\tilde{H}$ is reached. The corresponding $H(s)$ and $\kappa(s)$ can be identified up to orientation at iteration $s$. Let $H_u(s)$ and $\kappa_u(s)$ be the original MCMC draws before transformation. If we run the above two steps for $B$ times, let $\tilde{Q}(s)$ be the derived orthogonal transformation matrix at time $j$, for iteration $s$, then $\tilde{Q}(s) = \prod_{j=1}^{B} Q_j(s)$ is the orthogonal transformation matrix at iteration $s$. Thus, $H_B(s) = H_u(s)\tilde{Q}(s)$ and $\kappa_B(s) = \tilde{Q}(s)'\kappa_u(s)$, where $H_B(s)$ and $\kappa_B(s)$ are identified draws at iteration $s$, after $B$ times of the algorithm.

Note that Hoff et al. (2002) and Sewell and Chen (2015) also apply the Procrustes transformation to post-screen the posterior draws of latent positions $z_t$'s because the Euclidean distance in their network formation model is invariant to rotation, reflection, and translation of $z_t$'s. But the transformation algorithm they use is not an iterative one. They just fix $\tilde{H}$ at some meaningful initial values, such as the MLE, and do one time Procrustes transformation.

### C The MCMC algorithm

In this appendix we provide details of the MCMC algorithm for the SC-SDPD model with the general dynamic network formation model. The algorithm consists of sampling steps for parameters in the network formation equation, namely, $\Gamma = (\gamma', \Phi')'$ with $\Phi = (\delta', \xi', \zeta')'$; parameters in the SC-SDPD model, namely, $\theta = (\Psi', \beta', \kappa, \sigma')'$ with $\Psi = (\lambda, \rho, \mu)'$, $\alpha_{gt}$'s and $\tau_{gt}$'s, and for latent variables $H_{gt} = (Z_{gt}, M_{gt})'$s. Recall from Eq. (14) in the main text that the posterior distribution takes the following form:

$$P(\theta, \Gamma, \{\tau_{gt}\}, \{\alpha_{gt}\}, \{H_{gt}\}|\{Y_{gt}\}, \{W_{gt}\}) \propto \pi(\theta) \times \pi(\Gamma) \times \pi(\{\tau_{gt}\}) \times \pi(\{\alpha_{gt}\}) \times \pi(\{H_{gt}\})$$

$$\times P(\{Y_{gt}\}, \{W_{gt}\}|\{H_{gt}\}, \{\tau_{gt}\}, \{\alpha_{gt}\}, \theta, \Gamma).$$

The likelihood function of $Y_{gt}$ and $W_{gt}$ is

$$P(Y_{gt}, W_{gt}|H_{gt}, \tau_{gt}, \alpha_{gt}, \theta, \Gamma) = P(Y_{gt}|W_{gt}, H_{gt}, \tau_{gt}, \alpha_{gt}, \theta) \times P(W_{gt}|H_{gt}, \Gamma)$$

$$= (\sigma_v^2)^{-n_y} \times |S_{gt}(\lambda)| \times \exp(-\frac{1}{2\sigma_v^2}V_{gt}V_{gt}) \times \prod_{i=1}^{n_y} \prod_{j=1}^{n_y} \left( \exp(w_{ijgt}\psi_{ijgt}) \right) \frac{1}{1 + \exp(\psi_{ijgt})},$$

where
for \( g = 1, 2, \cdots, G \) and \( t = 1, 2, \cdots, T \).

Below we list the set of conditional posterior distributions required in the MCMC sampler. To simplify notation, exogenous variables \( X_{gt} \)'s, \( c_{iqt} \)'s and \( c_{ijgt} \)'s, lagged dependent variable \( Y_{g,t−1} \)'s and initial values \( W_{g0} \) and \( Y_{g0} \) are suppressed from the conditional set. For each step, the full conditional is conditioned on the rest of parameters and latent variables with the most updated values at the current iteration.

**Step 1:** Sample \( h_{ig1} = (z_{ig1}, m_{ig1})' \) from \( P(h_{ig1}|Y_{g1}, W_{g1}, h_{−i,g1}, h_{i,g2}, \tau_g, \theta, \Gamma) \) for \( i = 1, 2, \cdots, n_g \) and \( g = 1, 2, \cdots, G \).

By Bayes' theorem,

\[
P(h_{ig1}|Y_{g1}, W_{g1}, h_{−i,g1}, h_{i,g2}, \tau_g, \theta, \Gamma) \propto \pi(h_{ig1}) \times \pi(h_{i,g2}|h_{ig1}) \times P(Y_{g1}, W_{g1}|H_{g1}, \tau_g, \theta, \Gamma).
\]

At the \( q^{th} \) iteration, we apply a M-H step to sample \( h_{ig1} \)'s for \( i = 1, 2, \cdots, n_g \) and \( g = 1, 2, \cdots, G \).

1.1: Propose \( \tilde{h}_{ig1} = (\tilde{z}_{ig1}, \tilde{m}_{ig1})' \sim N_p(h_{ig1}^{(q−1)}, c_h I_p) \), where \( c_h \) is chosen by the user.

1.2: Let \( \tilde{H}_{g1|t} = (h_{1g1}, \cdots, h_{i−1,g1}, \tilde{h}_{ig1}, h_{i+1,g1}, \cdots, h_{n_g,g1}) \) and

\[
H_{g1|t}^{(q−1)} = (h_{1g1}, \cdots, h_{i−1,g1}, h_{ig1}^{(q−1)}, h_{i+1,g1}, \cdots, h_{n_g,g1}).
\]

Also let

\[
\tilde{\psi}_{ijg1} = \gamma_0 + \gamma_1 c_{i1} \gamma_1 + \gamma_2 c_{ij1} \gamma_2 + \gamma_3 c_{ijg1} \gamma_3 + \gamma_4 w_{ijg0} + \gamma_5 r_{ijg0} + \gamma_6 y_{i,g0} + \gamma_7 y_{j,g0} + \gamma_8 y_{i,g0} - y_{j,g0} + \sum_{p_1=1}^{\tilde{p}_1} \delta_{p_1} |\tilde{z}_{ip1,g1} - \tilde{z}_{jp1,g1}| + \sum_{p_2=1}^{\tilde{p}_2} \xi_{p2} \tilde{m}_{ip2g1} + \sum_{p_2=1}^{\tilde{p}_2} \xi_{p2} \tilde{m}_{jp2g1},
\]

\[
\tilde{\psi}_{ijg1} = \gamma_0 + \gamma_1 c_{j1} \gamma_1 + \gamma_2 c_{ijg1} \gamma_2 + \gamma_3 c_{ijg1} \gamma_3 + \gamma_4 w_{ijg0} + \gamma_5 r_{ijg0} + \gamma_6 y_{j,g0} + \gamma_7 y_{i,g0} + \gamma_8 y_{i,g0} - y_{j,g0} + \sum_{p_1=1}^{\tilde{p}_1} \delta_{p_1} |\tilde{z}_{jp1,g1} - \tilde{z}_{ip1,g1}| + \sum_{p_2=1}^{\tilde{p}_2} \xi_{p2} \tilde{m}_{jp2g1} + \sum_{p_2=1}^{\tilde{p}_2} \xi_{p2} \tilde{m}_{ip2g1}.
\]

\( \psi_{ijg1}^{(q−1)} \) (\( \tilde{\psi}_{ijg1}^{(q−1)} \)) can be defined similarly as \( \tilde{\psi}_{ijg1} \) (\( \tilde{\psi}_{ijg1} \)), with \( \tilde{z}_{ip1,g1} \)'s and \( \tilde{m}_{ip2g1} \)'s replaced, respectively, by \( \tilde{z}_{ip1,g1}^{(q−1)} \) and \( \tilde{m}_{ip2g1}^{(q−1)} \) for \( p_1 = 1, 2, \cdots, \tilde{p}_1 \) and \( p_2 = 1, 2, \cdots, \tilde{p}_2 \). With
Let $\tilde{h}_{i,g,t}^{(q-1)}$, $\bar{h}_{i,g,t}$

By Bayes' theorem, the probability

Step 2: Sample $h_{i,g,t} = (z'_{i,g}, m'_{i,g}, y'_{i,g})$ from $P(h_{i,g,t}|Y_{g,t}, W_{g,t}, h_{i,g,t+1}, h_{i-1,g,t}, h_{i,g,t-1}, \tau_{g}, \alpha_{g,t}, \theta, \Gamma)$ for $i = 1, 2, \cdots, n_{g}$, $g = 1, 2, \cdots, G$ and $t = 2, \cdots, T - 1$.

At the $q^{th}$ iteration, we apply a M-H step to sample $h_{i,g,t}$'s.

2.1: Propose $\tilde{h}_{i,g,t} = (\tilde{z}_{i,g,t}, \tilde{m}_{i,g,t}, \tilde{y}_{i,g,t}) \sim N_{p}(\tilde{h}_{i,g,t}^{(q-1)}, d_{h}I_{p})$, where $d_{h}$ is chosen by the user.

2.2: Let $\tilde{H}_{g,t} = (h_{1,g,t}, \cdots, h_{i-1,g,t}, h_{i,g,t}, h_{i+1,g,t}, \cdots, h_{n_{g},g,t})$ and

update $h_{i,g,t}^{(q)} = \bar{h}_{i,g,t}$, else set $h_{i,g,t}^{(q)} = h_{i,g,t}^{(q-1)}$.

and $\tilde{y}_{i,g,t}$ (and $\tilde{y}_{i,g,t}$) be defined similarly to $\tilde{y}_{i,g,t}(\tilde{y}_{i,g,t})$, with $\tilde{z}_{i,g,t}$'s and $\tilde{m}_{i,g,t}$'s replaced by, respectively, $z_{i,g,t}$ and $m_{i,g,t}$ for $p_{1} = 1, 2, \cdots, \tilde{p}_{1}$ and $p_{2} = 1, 2, \cdots, \tilde{p}_{2}$. With
probability

\[ \Pr(h_{igt}^{(q-1)}, \tilde{h}_{igt}) = \min \left\{ \frac{P(Y_{gt}|W_{gt}, H_{igt}; \tau_g, \alpha_{gt}, \theta)}{P(Y_{gt}|W_{gt}, H_{igt}^{(q-1)}; \tau_g, \alpha_{gt}, \theta)} \times \prod_{j:j \neq i} \frac{\exp(w_{igt}\tilde{\psi}_{igt})(1 + \exp(\psi_{igt}^{(q-1)}))}{\exp(w_{igt}\tilde{\psi}_{igt})(1 + \exp(\psi_{igt}^{(q-1)}))} \cdot \frac{\exp(w_{igt}\tilde{\psi}_{igt})(1 + \exp(\psi_{igt}^{(q-1)}))}{\exp(w_{igt}\tilde{\psi}_{igt})(1 + \exp(\psi_{igt}^{(q-1)}))} \right\} \]

update \( h_{igt}^{(q)} = \tilde{h}_{igt} \), else set \( h_{igt}^{(q)} = h_{igt}^{(q-1)} \).

**Step 3:** Sample \( h_{igt} = (z_{igt}', m_{igt}') \) from \( P(h_{igt}|Y_{gt}, W_{gt}, h_{igt-1}, \tau_g, \alpha_{gt}, \theta, \Gamma) \) for \( i = 1, 2, \ldots, n_g \) and \( g = 1, 2, \ldots, G \).

By Bayes’ theorem,

\[ P(h_{igt}|Y_{gt}, W_{gt}, h_{igt-1}, \tau_g, \alpha_{gt}, \theta, \Gamma) \propto \pi(h_{igt}|h_{igt-1}) \times P(Y_{gt}|W_{gt}, h_{igt}, \tau_g, \alpha_{gt}, \theta, \Gamma). \]

At the \( q^{th} \) iteration, we apply a M-H step to sample \( h_{igt} \).

3.1: Propose \( \tilde{H}_{igt} = (\tilde{z}_{igt}', \tilde{m}_{igt}') \) \( \sim \mathcal{N}_p(h_{igt}^{(q-1)}, e_h I_p) \), where \( e_h \) is chosen by the user.

3.2: Let \( \tilde{H}_{igt} = (\tilde{h}_{igt}; \tilde{h}_{igt-1}, \tilde{h}_{igt}, \tilde{h}_{igt+1}, \tilde{h}_{igt+1, gT}, \tilde{h}_{igt}) \) and \( H_{igt}^{(q-1)} = (h_{igt}; h_{igt-1}, h_{igt}, h_{igt+1, gT}, \ldots, h_{igt}) \). Also let

\[ \tilde{\psi}_{igt} = \gamma_0 + \tilde{c}_{igt}^T \gamma_1 + \tilde{c}_{igt}^T \gamma_2 + \tilde{c}_{igt}^T \gamma_3 + \gamma_4 w_{igt} + \gamma_5 z_{igt} - \gamma_7 z_{igt} - \gamma_8 y_{igt} \]

and \( \psi_{igt}^{(q-1)}(\tilde{\psi}_{igt}) \) be defined similarly to \( \tilde{\psi}_{igt} \), with \( \tilde{z}_{igt} \)’s and \( \tilde{m}_{igt} \)’s replaced by, respectively, \( z_{igt}^{(q-1)} \) and \( m_{igt}^{(q-1)} \) for \( p_1 = 1, 2, \ldots, \tilde{p}_1 \) and \( p_2 = 1, 2, \ldots, \tilde{p}_2 \). With
By Bayes' theorem, probability

\[
\Pr(h_{igT}^{(q-1)}, \tilde{h}_{igT}) = \min \left\{ \frac{P(Y_{gt}|W_{gt}, H_{gT}; \tau_g, \alpha_{gT}, \theta)}{P(Y_{gt}|W_{gt}, H_{gT}^{(q-1)}; \tau_g, \alpha_{gT}, \theta)} \times \prod_{j,j \neq i} \frac{\exp(w_{ijgT}\tilde{\psi}_{ijgT})(1 + \exp(\psi_{ijgT}^{(q-1)}))}{\exp(w_{ijgT}\psi_{ijgT}^{(q-1)})(1 + \exp(\psi_{ijgT}))}, \frac{\exp(w_{ijgT}\tilde{\psi}_{ijgT})(1 + \exp(\psi_{ijgT}^{(q-1)}))}{\exp(w_{ijgT}\psi_{ijgT}^{(q-1)})(1 + \exp(\psi_{ijgT}))} \right\} \frac{N_p(h_{igT}; h_{igT-1}, I_{\bar{p}})}{N_p(h_{igT}; h_{igT-1}, I_{\bar{p}}), 1}
\]

update \(h_{igT}^{(q)} = \tilde{h}_{igT}\); else set \(h_{igT}^{(q)} = h_{igT}^{(q-1)}\).

**Step 4:** Sample \(\Gamma = (\gamma', \Phi')\) from \(P(\Gamma|\{W_{gt}\}, \{H_{gt}\})\).

By Bayes' theorem,

\[
P(\Gamma|\{W_{gt}\}, \{H_{gt}\}) \propto \pi(\Gamma) \times P(\{W_{gt}\}|\{H_{gt}\}, \Gamma),
\]

where \(\pi(\Gamma) = \pi(\gamma) \times \pi(\Phi)\) is the prior of \(\Gamma\). We apply a M-H step to sample \(\Gamma\).

**Step 4.1:** Propose \(\Gamma = N_{6+2\bar{l}_1+\bar{l}_2+\bar{p}_1+2\bar{p}_2}(\Gamma^{(q-1)}, c_T I_{6+2\bar{l}_1+\bar{l}_2+\bar{p}_1+2\bar{p}_2})\), where \(c_T\) is chosen by the user. Check whether \(\gamma\) satisfies the identification constraints \(|\delta_1| \geq |\delta_2| \geq \cdots \geq |\delta_{\bar{p}}|\) and \(|\xi_1| \geq |\xi_2| \geq \cdots \geq |\xi_{\bar{p}}|\). If not, redraw \(\Gamma\) until it meets those constraints.

**Step 4.2:** With probability

\[
Pr(\Gamma^{(q-1)}, \Gamma) = \min \left\{ \frac{G \prod_{g=1}^G \prod_{t=1}^T \left( \frac{P(W_{gt}|H_{gt}, \Gamma)}{P(W_{gt}|H_{gt}, \Gamma^{(q-1)})} \right)}{\pi(\Gamma)} \right\}
\]

set \(\Gamma^{(q)} = \Gamma\), else set \(\Gamma^{(q)} = \Gamma^{(q-1)}\).

**Step 5:** Sample \(\kappa\) from \(P(\kappa|\{Y_{gt}\}, \{W_{gt}\}, \{H_{gt}\}, \{\tau_g\}, \{\alpha_{gt}\}, \Psi, \beta, \sigma^2_v)\).

By Bayes' theorem,

\[
P(\kappa|\{Y_{gt}\}, \{W_{gt}\}, \{H_{gt}\}, \{\tau_g\}, \{\alpha_{gt}\}, \Psi, \beta, \sigma^2_v)
\]

\[
\propto \pi(\kappa) \times P(\{Y_{gt}\}|\{W_{gt}\}, \{H_{gt}\}, \{\tau_g\}, \{\alpha_{gt}\}, \Psi, \beta, \sigma^2_v)
\]

\[
\propto \exp\left(-\frac{1}{2}(\kappa - \kappa_0)'^K_0^{-1}(\kappa - \kappa_0)\right) \times \prod_{g=1}^G \prod_{t=1}^T \exp\left(-\frac{1}{2\sigma^2_v}V_{gt}'V_{gt}\right)
\]

\[
\sim T \mathcal{N}_p(T\kappa, \Sigma_\kappa),
\]

11
where $\Sigma_\kappa = \left( K_0^{-1} + \sigma_\nu^{-2} \sum_{g=1}^G \sum_{t=1}^T H'_{gt} H_{gt} \right)^{-1}$ and

$$T_\kappa = \Sigma_\kappa \left( K_0^{-1} \kappa_0 + \sigma_\nu^{-2} \sum_{g=1}^G \sum_{t=1}^T H'_{gt} (S_{gt}(\lambda) Y_{gt} - \rho Y_{g,t-1} - \mu W_{g,t-1} Y_{g,t-1} - X_{gt} \beta - \tau_g - I_g \alpha_{gt}) \right).$$

Note that as we need to ensure $\kappa > 0$, the posterior distribution of $\kappa$ turns out to be a multivariate truncated normal.

**Step 6:** Sample $\Psi$ from $P(\Psi | \{Y_{gt}\}, \{W_{gt}\}, \{H_{gt}\}, \{\tau_g\}, \{\alpha_{gt}\}, \beta, \kappa, \sigma^2_\nu)$.

By Bayes’ theorem,

$$P(\Psi | \{Y_{gt}\}, \{W_{gt}\}, \{H_{gt}\}, \{\tau_g\}, \{\alpha_{gt}\}, \beta, \kappa, \sigma^2_\nu) \propto \pi(\Psi) \times P(\{Y_{gt}\} | \{W_{gt}\}, \{H_{gt}\}, \{\tau_g\}, \{\alpha_{gt}\}, \Psi, \beta, \kappa, \sigma^2_\nu),$$

where $P(\{Y_{gt}\} | \{W_{gt}\}, \{H_{gt}\}, \{\tau_g\}, \{\alpha_{gt}\}, \Psi, \beta, \kappa, \sigma^2_\nu) \propto \prod_{g=1}^G \prod_{t=1}^T \exp \left( - \frac{V_{gt}' V_{gt}}{2 \sigma^2_\nu} \right)$.  

6.1: Propose $\tilde{\Psi} \sim \mathcal{N}_3(\Psi^{(q-1)}, c_\Psi I_3)$, where $c_\Psi$ is chosen by the user. Check whether $\tilde{\Psi}$ satisfies the stability condition implied by its prior. If not, redraw $\tilde{\Psi}$ until it meets those conditions.

6.2: With acceptance probability

$$\text{Pr}(\Psi^{(q-1)}, \tilde{\Psi}) = \min \left( 1, \frac{P(\{Y_{gt}\} | \{W_{gt}\}, \{H_{gt}\}, \{\tau_g\}, \{\alpha_{gt}\}, \tilde{\Psi}, \beta, \kappa, \sigma^2_\nu) \times \pi(\tilde{\Psi})}{P(\{Y_{gt}\} | \{W_{gt}\}, \{H_{gt}\}, \{\tau_g\}, \{\alpha_{gt}\}, \Psi, \beta, \kappa, \sigma^2_\nu) \times \pi(\Psi^{(q-1)})} \right)$$

update $\Psi^{(q)} = \tilde{\Psi}$, else set $\Psi^{(q)} = \Psi^{(q-1)}$.

**Step 7:** Sample $\beta$ from $P(\beta | \{Y_{gt}\}, \{W_{gt}\}, \{H_{gt}\}, \{\tau_g\}, \{\alpha_{gt}\}, \Psi, \kappa, \sigma^2_\nu)$.

By Bayes theorem,

$$P(\beta | \{Y_{gt}\}, \{W_{gt}\}, \{H_{gt}\}, \{\tau_g\}, \{\alpha_{gt}\}, \Psi, \kappa, \sigma^2_\nu) \propto \pi(\beta) \times f(\{Y_{gt}\} | \{W_{gt}\}, \{H_{gt}\}, \{\tau_g\}, \{\alpha_{gt}\}, \Psi, \beta, \kappa, \sigma^2_\nu) \times \exp \left( - \frac{1}{2} \beta' B_0^{-1} \beta \right) \times \exp \left( - \frac{1}{2} \sigma^2_\nu \sum_{g=1}^G \sum_{t=1}^T V_{gt}' V_{gt} \right) \times \mathcal{N}_{2k}(T_\beta, \Sigma_\beta),$$

where $\Sigma_\beta = \left( B_0^{-1} + \sigma_\nu^{-2} \sum_{g=1}^G \sum_{t=1}^T X_{gt}' X_{gt} \right)^{-1}$ and

$$T_\beta = \Sigma_\beta \left( B_0^{-1} \beta_0 + \sigma_\nu^{-2} \sum_{g=1}^G \sum_{t=1}^T X_{gt}' (S_{gt}(\lambda) Y_{gt} - \rho Y_{g,t-1} - \mu W_{g,t-1} Y_{g,t-1} - H_{gt} \kappa - \tau_g - I_g \alpha_{gt}) \right).$$
Step 8: Sample $\sigma^2_v$ from $P(\sigma^2_v|\{Y_{gt}\}, \{W_{gt}\}, \{H_{gt}\}, \{\tau_g\}, \{\alpha_{gt}\}, \Psi, \beta, \kappa)$.

By Bayes’ theorem,

$$P(\sigma^2_v|\{Y_{gt}\}, \{W_{gt}\}, \{H_{gt}\}, \{\tau_g\}, \{\alpha_{gt}\}, \Psi, \beta, \kappa) \propto \pi(\sigma^2_v) \times P(\{Y_{gt}\}|\{W_{gt}\}, \{H_{gt}\}, \{\tau_g\}, \{\alpha_{gt}\}, \Psi, \beta, \kappa, \sigma^2_v)$$

$$\propto (\sigma^2_v)^{-\frac{1}{2}} \times \exp\left(-\frac{b}{2\sigma^2_v}\right) \times (\sigma^2_v)^{-\frac{\tau(\sum_{g=1}^{G} n_g)}{2}} \times \exp\left(-\frac{1}{2\sigma^2_v} \sum_{g=1}^{G} \sum_{t=1}^{T} V'_{gt} V_{gt}\right)$$

$$\sim IG\left(\frac{a_{new}}{2}, \frac{b_{new}}{2}\right),$$

with $a_{new} = a + T \cdot \sum_{g=1}^{G} n_g$ and $b_{new} = b + \sum_{g=1}^{G} \sum_{t=1}^{T} V'_{gt} V_{gt}$.

Step 9: Sample $\tau_g$ from $P(\tau_g|\{Y_{gt}\}, \{W_{gt}\}, \{H_{gt}\}, \{\alpha_{gt}\}, \Psi, \beta, \kappa, \sigma^2_v)$ for $g = 1, 2, \cdots, G$.

By Bayes’ theorem,

$$P(\tau_g|\{Y_{gt}\}, \{W_{gt}\}, \{H_{gt}\}, \{\alpha_{gt}\}, \Psi, \beta, \kappa, \sigma^2_v) \propto \pi(\tau_g) \times \prod_{t=1}^{T} P(Y_{gt}|W_{gt}, H_{gt}, \tau_g, \alpha_{gt}, \Psi, \beta, \kappa, \sigma^2_v)$$

$$\propto \exp\left(-\frac{1}{2}(\tau_g - \tau_0)' E_0^{-1}(\tau_g - \tau_0)\right) \times \exp\left(-\frac{1}{2\sigma^2_v} \sum_{t=1}^{T} V'_{gt} V_{gt}\right)$$

$$\sim N_{n_g}(T_{\tau_g}, \Sigma_{\tau_g}),$$

with $\Sigma_{\tau_g} = \left(E_0^{-1} + T \Sigma_{\tau_g}^{-1} \right)^{-1}$ and

$$T_{\tau_g} = \Sigma_{\tau_g} \left(E_0^{-1} \tau_0 + \sigma^2_v \sum_{t=1}^{T} (S_{gt}(\lambda) Y_{gt} - \mu Y_{gt} - \mu W_{gt-1} Y_{gt} - X_{gt} - H_{gt} - l_g \alpha_{gt})\right).$$

Step 10: Sample $\alpha_{gt}$ from $P(\alpha_{gt}|Y_{gt}, W_{gt}, H_{gt}, \tau_g, \Psi, \beta, \kappa, \sigma^2_v)$ for $g = 1, 2, \cdots, G$ and $t = 2, \cdots, T$.

By Bayes’ theorem,

$$P(\alpha_{gt}|Y_{gt}, W_{gt}, H_{gt}, \tau_g, \Psi, \beta, \kappa, \sigma^2_v) \propto \pi(\alpha_{gt}) \times P(Y_{gt}|W_{gt}, H_{gt}, \tau_g, \alpha_{gt}, \Psi, \beta, \kappa, \sigma^2_v)$$

$$\propto \exp\left(-\frac{1}{2F_0}(\alpha_{gt} - \alpha_0)^2\right) \times \exp\left(-\frac{1}{2\sigma^2_v} V'_{gt} V_{gt}\right)$$

$$\sim N(T_{\alpha_{gt}}, \Sigma_{\alpha_{gt}}),$$

where $\Sigma_{\alpha_{gt}} = \left(\frac{1}{F_0} + \frac{\nu_{gt}}{\sigma^2_v}\right)^{-1}$ and
\[ T_{a,t} = \Sigma_{a,t} \left( \frac{a_0}{F_0} + \sigma_v^{-2}(l'_{a}(S_{a,t}(\lambda))Y_{a,t} - \rho Y_{a,t-1} - \mu W_{a,t-1}Y_{a,t-1} - X_{a,t}\beta - H_{a,t}\kappa - \tau_{a}) \right). \]

**D Derivation of the AICM**

The conventional AIC (Akaike, 1973) is defined as

\[ \text{AIC} = 2\ell_{\text{max}} - 2d, \]  

(D.7)

where \( \ell_{\text{max}} \) is the maximum log-likelihood and \( d \) is the dimension of the parameters in the model. However, \( \ell_{\text{max}} \) is not directly observable in Bayesian estimation approach because \( \ell_{\text{max}} \) may not be reached during the MCMC sampling procedure. Raftery et al. (2007) propose the posterior simulation-based analogue of AIC, namely the AICM. Their key insight is that given the MCMC draws from the posterior and suppose that the log-likelihoods \( \{\ell_s: s = 1, \cdots, S\} \) corresponding to the MCMC samplers are approximately independent, \( \ell_{\text{max}} - \ell_s \) would asymptotically follow

\[ \ell_{\text{max}} - \ell_s \sim \text{Gamma}(d/2, 1), \]  

(D.8)

where \( \ell_{\text{max}} \) is the maximum achievable log-likelihood, and \( d \) is the *effective* number of the parameters. The asymptotic distribution in Eq. (D.8) follows when the amount of data underlying the likelihoods increases to infinity (Bickel and Ghosh, 1990; Dawid, 1991). With the Gamma distribution in Eq. (D.8), we have \( E[\ell_{\text{max}} - \ell_s] = d/2 \) and \( \text{Var}(\ell_s) = d/2 \).

Therefore, we can obtain the moment estimators \( \hat{d} = 2s^2_{\ell} \) and \( \hat{\ell}_{\text{max}} = \bar{\ell} + s^2_{\ell} \), where \( \bar{\ell} \) and \( s^2_{\ell} \) are the sample mean and variance of the \( \ell_s \)'s, respectively. Note that \( \hat{d} \) is the estimate of the *effective* number of parameters which is in general different from the dimension of parameters in the MCMC draw. This is because the total number of parameters includes hyperparameters and high dimensional latent variables (e.g. the case in our paper) which may not be effective in model selection. Therefore, the simulation-based (Monte Carlo) version of AIC is given as

\[ \text{AICM} = 2\hat{\ell}_{\text{max}} - 2\hat{d} = 2(\bar{\ell} - s^2_{\ell}), \]  

(D.9)

and its standard error can be calculated as

\[ \text{S.E.}(\text{AICM}) = \sqrt{4\hat{d}/(2S) + 4\hat{d}(11\hat{d}/4 + 12)/S}, \]  

(D.10)

by using the facts that \( \text{Var}(\bar{\ell}) \approx d/(2S) \), \( \text{Var}(s^2_{\ell}) \approx d(11d/4 + 12)/S \), and the approximate independence between \( \bar{\ell} \) and \( s^2_{\ell} \). Note that the estimator \( \hat{d} \) is also derived by Gelman.
et al. (2003, Section 6.7) to replace the $p_D$ in their alternative definition of DIC. Thus, AICM is equivalent to Gelman et al. (2003)’s definition of Deviance Information Criterion (DIC).

E  Network Interaction Effects from Other Relationships and Robustness Checks of Model Specifications.

Note that the peer effects reported in Table 3 of the main text are estimated based on general friendships. Given that our data provides information on antipathetic relationship, study mate, and cram school mate, we also report the peer effects based on these other relationships. Table I.3 summarizes the estimation results under the assumption of exogenous networks (for comparison with Table 3 in the main text). We can see neither contemporary nor temporal peer effect exist in the foe (antipathetic relationship) networks. This finding shows that peer effects on students’ academic outcomes do not operate through antipathetic relationships and serves as a placebo test to support that the significant peer effects among friends found in Table 3 exist. When we focus on friends who study together, the peer effect obtained among study mates is slightly higher than the one in Table 3. By contrast, the estimated peer effect on academic performance decreases by 40% when focusing on friends who go to cram schools together. Our explanation for this finding is that students who go to cram schools are often the ones who fall behind in school learning and need additional tutoring. There is no much gain on own academic learning from interacting with poorly performing friends.

In the empirical results based on networks of other relationships, we can see that the peer effect estimated from study mates seems stronger than that from general friendships (comparing Table I.3 Column 3 with Table 3 Column (IV) in the main text). This specific friendship might capture more relevant peer groups. Thus, we further conduct the full model estimation with endogenous friendship formation to study peer effects from the study mates network. The full model estimates of the contemporary peer effect $\lambda$, the persistent effect $\rho$, and the temporal peer effect $\mu$ for the study mates network are reported in Table I.4. It shows a correction of the endogeneity bias after controlling for network formation. Similar to the results from friendship networks, the contemporary peer effect and the persistency effect are both significant in all specifications of latent dimensions ((I) to (III)). The contemporary peer effect ($\lambda$) in the study mates network is still stronger than that of friendship networks after correcting the endogeneity of network formation (comparing Table I.4 with Table 4 in the main text). Furthermore, the AICM
selects the two-dimensional latent variables specification to be the best fit model for the study mates network data.

We further conduct two robustness checks for our model specification. First, we consider the non-row-normalized $W_{gt}$ specification in our SC-SDPD model. For row-normalized $W_{gt}$, the coefficient $\lambda$ can be interpreted as the local average effect, i.e., the influence of the average behavior of one’s peers; while for non-row-normalized $W_{gt}$, $\lambda$ can be interpreted as the local aggregate effect, i.e., the influence of the aggregate behavior of one’s peers (Liu et al., 2014). The left panel of Table I.5 shows the estimates of contemporary peer effect $\lambda$, time persistence effect $\rho$, and temporal peer effect $\mu$ under the non-row-normalized specification. Compared to the row-normalized version (see Table 4 in the main text), $\lambda$ and $\rho$ are both significant and positive in either specifications. Time persistency effects are all similar in magnitude which shows different specifications of network matrix do not affect $\rho$. The estimated contemporary peer effect $\lambda$ is 0.022 under non-row-normalized specification. Meanwhile, increasing latent variables from one to three dimensions ((I) to (III)) does not change the estimates of $\lambda$. Overall, the AICM still selects the row-normalized two-dimensional (II) specification as the best fit to the data. Hence, in our data set, individual’s behavior is more likely to be affected by the average behavior (social norm) of their friends.

The second robustness check is about extending the control function in the SC-SDPD model with additional contextual effects from latent variables, i.e., $W_{gt}Z_{gt}$ and $W_{gt}M_{gt}$. We report the estimation results of $\lambda$, $\rho$ and $\mu$ with such specifications in the right panel of Table I.5. Compared to the original specification (see Table 4 in the main text), we conclude that there are no significant differences between the two specifications in estimating $\lambda$, $\rho$ and $\mu$. In particular, according to the AICM, the original latent specification with two dimension is still selected as the best fit model to the data.

F Multiplier Effects from Policy Intervention

The advantage of our structural model is that it enables us to explore the detailed channels in which networks and economic activities affect each other. This model likewise enables us to simulate and evaluate the policy impact on network formation and interaction of economic outcomes with higher accuracy. We follow the empirical framework of this study and analyze a policy scenario where government (or school) agencies provide financial assistance to students’ families who experience financial difficulties. From the empirical results, we determine that the variable “family in financial difficulty” has significant negative effects on both students’ friendship formation and academic performance. Consequently, we expect that the financial releasing policy (program) will assist students
to improve their social networking and school academic performance.

When studying the policy impact of releasing families’ financial difficulties, we focus on the multiplier effect generated through network interactions. In particular, the multiplier effect on academic outcome implied by our dynamic model is different from that of the cross-sectional social interactions model. It is jointly determined by the contemporary and temporal peer effects, subjected to endogenous network rewiring.

Note that the contemporary peer effect $\lambda$ in the structural econometric model (8) in the main text is a function of both the social multiplier coefficient $\lambda_1$ and the social conformity coefficient $\lambda_2$ (see Appendix A Eq. (A.3)). Even though $\lambda$ can be identified, $\lambda_1$ and $\lambda_2$ may not be separately identified (See also the discussion in Boucher and Fortin (2016)). However, when discussing the multiplier effects from policy intervention for our model, we are not aiming at determining the exact source of network interaction nor identifying the social multiplier. Instead, our focus is the multiplier effect that arises from the network interaction effect $\lambda$ as a whole. As long as $\lambda$ can be identified and estimated, the source of network interaction would be irrelevant.

We use the networks and academic outcomes observed at the last time period ($T$) as bases to analyze the marginal effect (ME) of the policy on the one-period out-of-sample expected network outdegree, $D_{g,T+1}$ (the network outdegree of individual $i$ is the number of nominated friends of $i$), and academic outcome, $Y_{g,T+1}$. We compute as follows:

$$
\text{ME}^{D}_{g,T+1} = E(D_{g,T+1}|W_{g,T}, Y_{g,T}, X_{g,T+1}^{treated}, \Gamma) - E(D_{g,T+1}|W_{g,T}, Y_{g,T}, X_{g,T+1}^{untreated}, \Gamma), \quad (F.11)
$$

$$
\text{ME}^{Y}_{g,T+1} = E(Y_{g,T+1}|W_{g,T}, Y_{g,T}, X_{g,T+1}^{treated}, \theta) - E(Y_{g,T+1}|W_{g,T}, Y_{g,T}, X_{g,T+1}^{untreated}, \theta), \quad (F.12)
$$

where $\text{ME}^{D}_{g,T+1}$ is the vector of marginal effects on the network outdegrees and $\text{ME}^{Y}_{g,T+1}$ is the vector of marginal effects on the academic outcome for school $g$; $X_{g,T+1}^{treated}$ refers to the case that we turn all non-zero financial difficulty dummies to zero at $T + 1$ and retain all other exogenous variables unchanged from $T$, and $X_{g,T+1}^{untreated}$ refers to the case where we retain the financial difficulty dummy as well as all other exogenous variables the same values as at $T$. We compute the marginal effect by the difference formula in Eqs. (F.11) and (F.12) because the policy instrument is discrete. For continuous instruments, one can follow Lee and Yu (2012) to compute the marginal effect from the SDPD model by the space-time multiplier. We take the parameter estimates in Column (II) of Table 4 in the main text and compute ME separately for each school. To focus on the policy effect, we assume that the group-time effects, individual latent variables and latents’ coefficients are unchanged from $T$ to $T + 1$. So the normalization constraints on the variances of those latent variables would not affect the interpretation of the marginal effects. For the
marginal effect on academic outcome, we consider two scenarios – one fixes the network $W_{g,T+1}$ as $W_{g,T}$ and the other enables the network $W_{g,T+1}$ to be rewired. Our network formation model determines the endogenous network rewiring for $W_{g,T+1}$. In the network rewiring scenario, we report the ME as the average value of Eq. (F.12), which is calculated from 1,000 simulated networks.

We pool the ME values across all schools and plot their distributions for network outdegree and academic outcome in Figures I.2 and I.3, respectively. In both figures, Panel (a) focuses on the group of the 64 treated students, i.e., whose families experience financial difficulties; Panel (b) focuses on the rest untreated student; and Panel (c) combines the students of the two groups. For the network outdegree, Figure I.2 shows that the policy improves the students’ social networking. Overall, each student will own an average of 0.66 additional outdegree links because of the policy. The policy effect on treated students is stronger. In particular, each treated student will own an average of 1.7 additional outdegree links and each untreated student will own 0.49 additional outdegree links.

For academic outcome, the multiplier effect of the policy intervention can be clearly identified in Panels (a) and (b) in Figure I.3. If no multiplier effects were present for the treated students, then the ME values in Panel (a) should concentrate at 2.504, which is the corresponding coefficient of the variable Lessmoney in Table 4. However, we observe a few ME values above 2.504. Similarly, if no multiplier effects were present for the untreated students, then all the ME values in Panel (b) should be zero. Instead, we observe many positive ME values for those who do not experience financial difficulty. The comparison between the scenarios of the fixed and rewired networks reveals that network rewiring would change the distributions of ME for the treated and untreated students. In Panel (a) for the treated students, a minor change is shown on the mean values of ME (3.04 and 3.06, respectively) between the fixed and rewired network cases. However, the standard deviation decreases from 0.75 in the fixed network case to 0.50 in the rewired network case. A similar pattern can be observed in Panel (b) for the untreated students. Although means are similar (0.44 and 0.46, respectively), a huge decline is observed from the standard deviation from the fixed network case (0.58) to the rewired network case (0.36). Thus, the network rewiring results in a mean-preserving contraction in the distribution of ME for academic outcome.

We also conduct a policy impact analysis under the potentially misspecified SAR model in Column (II) of Table I.2. Given that the estimation results of the SAR model show stronger own and contextual effects of financial difficulty, and a stronger contemporary endogenous peer effect on academic outcome, the policy impact analysis based on the SAR model in Figure I.4 is not surprising in showing stronger marginal and mul-
tiplier effects. Given that the SAR model is potentially misspecified and the result on endogenous peer effect is possibly upward biased because of the omission of the persistency and temporal peer effects, our study raises a caution for using the SAR model as a policy evaluation tool when the policy environment is dynamic. Under such case, the SAR model may exaggerate the policy impact and mislead the policy implementation.

G Network Goodness-of-fit Examination

To investigate whether the network formation model that we propose fits the observed network data well, we follow Hunter et al. (2008) to conduct a model goodness-of-fit examination. For illustration, we select the observed networks of periods 2 to 4 from school 7 as the examination benchmarks. The results from other schools are generally the same and are available upon request. We pass over the network of period 1 because it is assumed to be exogenously given. For each period, we simulate 100 artificial networks from our network formation model with the parameter estimates reported in Column (II) of Table 4 in the main text. Model fitness is examined by evaluating the similarity between the simulated and observed networks in the distribution of four network statistics, namely, degree, edge-wise shared partner, minimum geodesic distance, and average nearest neighbor connectivity. Figures I.5, I.6 and I.7 show the examination results of periods 2, 3 and 4. We present the distribution of statistics for the observed network by solid curves, distributions for simulated networks by box plots and the corresponding 5th and 95th percentiles by dotted lines. From the figures, the simulated and observed networks display similar distributions over these four statistics. These results suggest that our estimated model is able to capture the unobserved network generating process.

H Time Evolution of Unobserved Latent Variables

When researchers study network formation by the latent space models, they have an intention of visualizing the position of each node in the latent space. In a dynamic setting, one can further show the time evolving trajectory of the node positions (Sewell and Chen, 2015). For the current empirical study, we choose one school (school 7) to conduct this visualization exercise on students’ latent network positions. The visualization results of other schools are available upon request.

The discussion of identification in Appendix B indicates that the latent variables $z_{igt}$’s and $m_{igt}$’s are identified under the constraints in Section 2.3 when $\kappa$ is identified from the full rank condition of the SC-SDPD model. When the latent variables are identified, we can visualize their positions by their posterior means from the MCMC draws. However,
from the estimation results in Column (II) of Table 4, we do not obtain all the coefficients of the latent variables in the network formation model to be significant. In particular, the estimate of $\zeta_1$ is insignificant. One may worry that a non-significant coefficient of the latent variables would cause an identification problem. Therefore, we apply the iterative Procrustes transformation (Aßmann et al., 2012, 2016) on the posterior draws of $z_{igt}$’s and $m_{igt}$’s to secure their identifications from the potential rotation problem. To be explicit, we take the posterior MCMC draws from the estimation results and conduct transformation with the procedures outlined in Appendix B.3. We calculate the posterior mean of the transformed draws of the latent variables as the point estimate of each student’s latent position at the second, third, and fourth time periods. Figure I.8 shows the visualization results based on the transformed $z_{igt}$’s and $m_{igt}$’s.

Panels (a) to (c) in the first row of Figure I.8 show the positions of $z_{igt}$’s for homophily in each time period and Panel (d) summarizes the time-evolving trajectory. In each panel, the horizontal and vertical axes separately stand for the first and the second latent dimension. The size of each node represents the number of nominated friends (i.e., outdegree). The plots of $z_{igt}$’s show that the students’ positions are getting closer over time in the latent space of homophily and students with more nominated friends (i.e., larger nodes) are clustered in the center of the space. Panels (e) to (h) in the second row show the positions of $m_{igt}$ for degree heterogeneity and the time evolving trajectory. Contrary to the case of $z_{igt}$, a diverging pattern of students’ positions exists in the space of degree heterogeneity. In particular, students with higher outdegrees move to the right hand side of the space and students with lower outdegrees move to the top left corner of the space. The diverging latent positions for high-outdegree and low-outdegree students are consistent with the positive estimate of $\xi_1$ and the negative estimate of $\xi_2$ from the perspective of link senders in network formation (see Table 4). We similarly show the correspondence between the latent position of $m_{igt}$ and the number of friendship nominations that each student received (i.e., indegree) from the perspective of link receivers in Panels (i) to (l) of the third row. The plots show that the nodes that move to the top left corner of the space are in fact high-indegree students. This result is consistent with the negative estimate of $\zeta_1$ and the positive estimate of $\zeta_2$ in Table 4.

To analyze whether the non-significant $\zeta_1$ in the network formation model causes the identification problem, we plot the positions of the latent variables based on the original MCMC draws in Figure I.9. The results show that the visualized latent positions in Figure I.8 and in Figure I.9 look extremely similar. Thus, we can confirm that the identification of the latent variables in our model is achieved and their positions are robust to the transformation.
I Additional Tables and Figures
| Para. | DGP mean | s.d. | Full-D2 mean | s.d. | Full-D1 mean | s.d. | Unobs. homo. mean | s.d. | Unobs. Deg. mean | s.d. | SDPD w/ factor mean | s.d. | SDPD mean | s.d. | SAR mean | s.d. |
|-------|----------|------|--------------|------|--------------|------|-----------------|------|-----------------|------|-----------------|------|-----------|------|---------|------|
| λ     | 0.3000   | 0.3092 | 0.0145    | 0.3711 | 0.0172    | 0.3114 | 0.0147    | 0.2039 | 0.0201    | 0.3574 | 0.0097    | 0.3518 | 0.0138    | 0.7246 | 0.0167 |
| ρ     | 0.2000   | 0.1919 | 0.0144    | 0.3160 | 0.0172    | 0.3088 | 0.0124    | 0.2118 | 0.0193    | 0.2115 | 0.0106    | 0.5641 | 0.0115    | -  | - |
| µ     | -0.1000  | -0.0910 | 0.0156    | -0.1329 | 0.0134    | -0.1350 | 0.0153    | -0.1484 | 0.0211    | -0.0903 | 0.0097    | -0.0831 | 0.0089    | -  | - |
| β₁    | 1.0000   | 1.0071 | 0.0124    | 1.0260 | 0.0104    | 1.0279 | 0.0144    | 1.0111 | 0.0136    | 1.0065 | 0.0108    | 1.0345 | 0.0111    | 0.9437 | 0.0167 |
| β₂    | 1.0000   | 0.9933 | 0.0275    | 0.9364 | 0.0282    | 1.0132 | 0.0400    | 1.1194 | 0.0413    | 0.9426 | 0.0221    | 0.9770 | 0.0345    | 0.4449 | 0.0379 |
| κ₁₁   | 1.2000   | 1.1563 | 0.0834    | 0.4949 | 0.2951    | 0.7580 | 0.1850    | -  | -         | -  | -         | -  | -         | -  | - |
| κ₁₂   | 0.4000   | 0.3328 | 0.2021    | -  | -         | -  | 0.6878 | 0.2143    | -  | -         | -  | -         | -  | - |
| κ₂₁   | 1.0000   | 0.9337 | 0.0610    | 1.1508 | 0.0482    | -  | -         | -  | 0.6422    | 0.0599 | -  | -         | -  | - |
| κ₂₂   | 0.5000   | 0.5427 | 0.1021    | -  | -         | -  | 0.6631 | 0.0812    | -  | -         | -  | -         | -  | - |
| γ₀    | -1.0000  | -0.9245 | 0.1531    | -1.2497 | 0.1996    | 0.6013 | 0.1265    | -3.5375 | 0.3238    | -  | -         | -  | -         | -  | - |
| γ₁    | 0.5000   | 0.5125 | 0.1004    | 0.4100 | 0.1545    | 0.3875 | 0.0624    | 0.3434 | 0.0928    | -  | -         | -  | -         | -  | - |
| γ₂    | 0.5000   | 0.5114 | 0.0222    | 0.4077 | 0.0226    | 0.3838 | 0.0185    | 0.3480 | 0.0220    | -  | -         | -  | -         | -  | - |
| γ₃    | 0.5000   | 0.5121 | 0.0205    | 0.4052 | 0.0193    | 0.3854 | 0.0157    | 0.3462 | 0.0157    | -  | -         | -  | -         | -  | - |
| γ₄    | 0.5000   | 0.5145 | 0.0956    | 0.8149 | 0.1765    | 1.4170 | 0.0771    | 1.2018 | 0.1157    | -  | -         | -  | -         | -  | - |
| γ₅    | 0.3000   | 0.3188 | 0.0236    | 0.2291 | 0.0143    | 0.3709 | 0.0120    | 0.2385 | 0.0104    | -  | -         | -  | -         | -  | - |
| γ₆    | 0.3000   | 0.3081 | 0.0193    | 0.1831 | 0.0120    | 0.3141 | 0.0127    | 0.1941 | 0.0130    | -  | -         | -  | -         | -  | - |
| γ₇    | -1.0000  | -1.0340 | 0.0368    | -0.4387 | 0.0140    | -0.6509 | 0.0240    | -0.6810 | 0.0242    | -  | -         | -  | -         | -  | - |
| δ₁    | -2.0000  | -2.0256 | 0.1128    | -2.1018 | 0.1519    | -1.7934 | 0.0953    | -  | -         | -  | -         | -  | -         | -  | - |
| δ₂    | -1.0000  | -1.1253 | 0.1990    | -1.6174 | 0.0940    | -  | -         | -  | -         | -  | -         | -  | -         | -  | - |
| ξ₁    | 1.0000   | 1.0235 | 0.0686    | 0.9518 | 0.0584    | -  | -         | 0.8447 | 0.0531    | -  | -         | -  | -         | -  | - |
| ζ₁    | 0.5000   | 0.5065 | 0.0941    | 0.8761 | 0.0609    | -  | -         | 0.5054 | 0.0797    | -  | -         | -  | -         | -  | - |
| ξ₂    | 0.7000   | 0.6931 | 0.1142    | -  | -         | -  | -         | 0.8322 | 0.0954    | -  | -         | -  | -         | -  | - |
| ξ₂    | 1.0000   | 1.0056 | 0.0703    | -  | -         | -  | -         | 0.9599 | 0.0839    | -  | -         | -  | -         | -  | - |
| σ₂ v  | 1.0000   | 0.9820 | 0.0840    | 2.1328 | 0.2734    | 5.0684 | 0.1104    | 5.1769 | 0.1308    | 2.0350 | 0.0445    | 5.0524 | 0.0964    | 9.0081 | 0.3972 |
| AICM   | 4680.4  | 453.00 | 5518.7    | 493.10 | 5666.9    | 592.70 | 5035.10    | 535.80 | -         | -  | -         | -  | -         | -  | - |

Note: This Monte Carlo study contains 100 repetitions. The mean and standard deviation of the point estimates across repetitions are reported. Column “Full-D2” refers to the true model that generates the artificial data, which has latent variables in two dimensions. Column “Full-D1” refers to the model that has only one-dimensional latent variable. Columns “Unobs. homo.” and “Unobs. Deg.” refer to the two encompassed network formation models with only unobserved homophily or unobserved degree heterogeneity. Column “SDPD w/ factor” refers to the SDPD model with the common factor structure where the number of common factors is set to 4. Column “SDPD” refers to the SDPD model which neglects endogenous network formation. Column “SAR” refers to the SAR model with individual and time effects.
Table I.2: Estimation Results of Peer Effects on Academic Performance Under SAR Model with Dynamic Friendship Network Formation

|                | (I)              | (II)             | (III)            |
|----------------|------------------|------------------|------------------|
|                | Λ (0.045)        | Λ (0.030)        | Λ (0.034)        |
| Male           | Own Contex.      | Own Contex.      | Own Contex.      |
| Male           | 1.392 −4.277    | 0.638 1.462      | 1.388 −0.355     |
| Sibling        | (4.043) (4.819) | (2.618) (3.228)  | (2.508) (3.056)  |
| Felh           | −1.613 −3.341** | −2.459*** −0.654 | −2.770*** −6.712*** |
| Felh           | (1.035) (1.639) | (0.603) (0.969)  | (0.600) (1.054)  |
| Mehh           | −1.062 −2.910   | −1.227 −1.109    | −0.263 −0.309    |
| Mehh           | (2.164) (4.118) | (1.307) (2.625)  | (1.232) (2.449)  |
| Meh 0          | 4.456** 3.129   | 2.203* −4.574    | 1.969 6.919***   |
| Meh 0          | (2.147) (4.399) | (1.245) (2.796)  | (1.138) (2.687)  |
| Mer 0          | 5.038** −9.036**| 4.876*** −11.374 | 8.166*** −2.702  |
| Mer 0          | (2.036) (4.059) | (1.193) (2.613)  | (1.047) (2.410)  |
| Munemp         | 0.803 −6.876    | 1.901 1.789      | 2.330* 0.961     |
| Munemp         | (2.419) (4.767) | (1.604) (3.196)  | (1.263) (2.832)  |
| Funemp         | −0.875 1.955    | −4.542** 2.346   | −6.516*** 7.511** |
| Funemp         | (3.008) (5.898) | (1.829) (3.961)  | (1.694) (3.530)  |
| Fretired       | 4.085 9.471     | −0.624 7.657     | 1.156 0.980      |
| Fretired       | (4.269) (6.912) | (2.612) (4.977)  | (2.570) (4.482)  |
| Mumem 0        | 0.877 −13.175*  | 9.069*** −7.337  | 10.638*** −10.511*** |
| Mumem 0        | (4.493) (7.089) | (2.839) (4.932)  | (2.934) (5.054)  |
| Mretired       | 6.316 −2.413    | 6.777 1.571      | 4.653 −4.820     |
| Mretired       | (6.775) (8.909) | (4.516) (7.638)  | (4.605) (7.525)  |
| Housewife      | 1.660 −2.482    | −1.961* 0.014    | 1.715* 0.363     |
| Housewife      | (1.778) (3.344) | (1.104) (2.028)  | (0.922) (1.882)  |
| Height         | 0.417*** −0.408** | 0.349*** −0.487*** | 0.374*** −0.623** |
| Height         | (0.113) (0.202) | (0.069) (0.122)  | (0.062) (0.122)  |
| Weight         | −0.220** 0.142   | −0.084 0.149     | −0.240*** 0.403*** |
| Weight         | (0.090) (0.199) | (0.057) (0.117)  | (0.051) (0.122)  |
| Divorce        | −9.364** 14.443**| −9.000*** 4.384  | −12.118*** 7.102* |
| Divorce        | (3.428) (6.236) | (2.093) (4.283)  | (2.003) (4.169)  |
| Age            | −0.463 4.222**  | 1.925*** 4.058***| 1.216*** 6.948*** |
| Age            | (0.985) (1.756) | (0.615) (1.064)  | (0.556) (1.006)  |
| Lessmoney      | −7.326*** −6.348| −3.723*** −13.394***| −4.780*** −14.195*** |
| Lessmoney      | (2.224) (4.587) | (1.273) (2.904)  | (1.120) (2.580)  |
| Parentfight    | 8.129* −6.535   | 1.540 −2.372     | 0.998 −4.217     |
| Parentfight    | (4.230) (7.481) | (2.695) (5.528)  | (2.412) (5.530)  |
|                   | 0.308 0.425***   | 4.177***         |
|                   | (0.266) (0.287)  | (0.447)          |
| κ_11           | 2.714***         | 0.501            |
| κ_11           | (0.452) (0.341)  |                |
| κ_12           | 1.064***         | 0.336            |
| κ_21           | 9.524***         | 14.204***        |
| κ_21           | (0.805) (0.295)  | (0.294)          |
| κ_22           | 2.532 4.723***   | 0.236            |
| κ_22           | (0.360) (0.236)  |                |
| κ_23           | 14.214***        | 0.287            |
| σ^2            | 514.643 135.118  | 101.143          |
|                 | (25.994) (11.990)| (19.377)        |

Network Formation

Continued on Next Page
| Constant       | -0.830*** | 2.017*** | 3.568*** |
|----------------|-----------|----------|----------|
| (0.125)        | (0.168)   | (0.170)  |
| Same gender    | 1.473***  | 1.820*** | 2.165*** |
| (0.081)        | (0.080)   | (0.092)  |
| Lessmoney (i,t)| -0.364*** | -0.208***| -0.386***|
| (0.067)        | (0.068)   | (0.057)  |
| Lessmoney (j,t)| -0.167*** | -0.188** | -0.128*  |
| (0.064)        | (0.088)   | (0.078)  |
| Parentfight (i,t) | -0.033     | -0.215   | -0.176   |
| (0.134)        | (0.136)   | (0.112)  |
| Partenfight (j,t) | -0.116    | -0.607***| -0.523***|
| (0.115)        | (0.140)   | (0.115)  |
| Friend (t-1)   | 2.931***  | 3.327*** | 3.777*** |
| (0.052)        | (0.066)   | (0.076)  |
| Common friend (t-1) | 0.220***   | 0.242*** | 0.296*** |
| (0.016)        | (0.014)   | (0.019)  |
| Common enemy (t-1) | 0.368**   | 0.883*** | 0.985*** |
| (0.167)        | (0.219)   | (0.195)  |
| Acad. ranking (i,t-1) | -0.011***   | -0.030***| -0.007***|
| (0.001)        | (0.001)   | (0.001)  |
| Acad. ranking (j,t-1) | -0.001    | 0.004*** | -0.004** |
| (0.001)        | (0.001)   | (0.002)  |
| Diff. of ranking | -0.006***  | -0.006***| -0.008***|
| (0.001)        | (0.001)   | (0.001)  |
| $\delta_1$    | -1.992*** | -2.486***| -2.638***|
| (0.060)        | (0.059)   | (0.056)  |
| $\delta_2$    | -         | -2.374***| -2.442***|
| -              | (0.062)   | (0.057)  |
| $\delta_3$    | -         | -        | -2.299***|
| -              | (0.055)   |          |
| $\xi_1$       | 0.608***  | 0.692*** | 0.701*** |
| (0.040)        | (0.027)   | (0.028)  |
| $\zeta_1$     | 0.209***  | -0.143   | -0.476***|
| (0.022)        | (0.024)   | (0.027)  |
| $\xi_2$       | -         | -0.272***| 0.148*** |
| -              | (0.025)   | (0.015)  |
| $\zeta_2$     | -         | 0.339*** | 0.315*** |
| -              | (0.019)   | (0.022)  |
| $\xi_3$       | -         | -        | -0.129***|
| -              |         | (0.017)  |
| $\zeta_3$     | -         | -        | 0.082*** |
| -              |         |          |

| AICM           | 26491     | 18201    | 16006    |
|----------------|-----------|----------|----------|
| se(AICM)       | 79.718    | 47.920   | 55.879   |

Note: The dynamic network formation with latent variables of one dimension in Column (I), the two dimensions in Column (II), and the three dimensions in Column (III). The asterisks *** (** , *) indicate that its 99% (95%, 90%) highest posterior density range does not cover zero. We perform the MCMC sampling for 100,000 iterations, discard the initial 20,000 iterations for burn-in, and compute the posterior means and standard deviations from the remaining draws as point estimates. The posterior standard deviations are reported in parentheses.
Table I.3: Estimation Results of Peer Effects on Academic Performance from Other Relationship Networks assumed Exogenous

|           | Foe       | Study mate | Cram school mate |
|-----------|-----------|------------|------------------|
| $\lambda$ | -0.060    | 0.115***   | 0.065*           |
|           | (0.074)   | (0.032)    | (0.038)          |
| $\rho$    | 0.810***  | 0.798***   | 0.809***         |
|           | (0.018)   | (0.018)    | (0.019)          |
| $\mu$     | -0.017    | -0.006     | -0.020           |
|           | (0.024)   | 0.017      | (0.020)          |

|           | Own      | Contex.    | Own      | Contex.    | Own      | Contex. |
|-----------|----------|------------|----------|------------|----------|---------|
| Male      | -0.553   | -0.052     | 1.350    | -2.238     | -0.657   | 0.019   |
|           | (1.518)  | (2.725)    | (1.772)  | (2.005)    | (1.549)  | (2.152) |
| Sibling   | -0.104   | 1.407      | -0.114   | -0.861     | -0.189   | -2.916  |
|           | (0.641)  | (1.884)    | (0.634)  | (1.074)    | (0.638)  | (1.382) |
| Fehl      | -0.439   | 1.377      | -0.337   | 0.483      | -0.249   | -0.050  |
|           | (1.241)  | (2.641)    | (1.239)  | (2.031)    | (1.237)  | (2.263) |
| Fehh      | 0.423    | 1.522      | 0.465    | 1.100      | 0.449    | 0.136   |
|           | (1.223)  | (2.403)    | (1.221)  | (1.874)    | (1.225)  | (1.997) |
| Mehl      | 0.373    | 1.094      | 0.951    | -1.782     | 0.556    | -0.284  |
|           | (1.192)  | (2.486)    | (1.194)  | (1.941)    | (1.183)  | (2.256) |
| Mehh      | 1.074    | 1.935      | 1.345    | -0.313     | 1.217    | 0.326   |
|           | (1.372)  | (2.516)    | (1.364)  | (2.070)    | (1.369)  | (2.194) |
| Funemp    | -0.257   | 0.680      | -0.083   | 1.003      | -0.198   | 0.163   |
|           | (1.683)  | (2.885)    | (1.678)  | (2.526)    | (1.677)  | (2.656) |
| Retired   | 1.412    | 0.686      | 0.936    | 0.982      | 0.854    | -0.114  |
|           | (2.174)  | (3.015)    | (2.150)  | (2.720)    | (2.150)  | (2.733) |
| Mumemp    | 0.450    | 0.237      | 0.820    | -0.488     | 0.251    | -1.262  |
|           | (2.222)  | (3.168)    | (2.220)  | (2.797)    | (2.205)  | (2.884) |
| Mretired  | 0.203    | 1.069      | 0.224    | -0.223     | 0.308    | 0.052   |
|           | (2.755)  | (3.097)    | (2.781)  | (3.013)    | (2.775)  | (2.985) |
| Housewife | 1.140    | -3.330     | 0.910    | -1.308     | 0.901    | -1.701  |
|           | (1.061)  | (2.440)    | (1.062)  | (1.722)    | (1.056)  | (1.953) |
| Height    | 0.120*   | -0.174     | 0.160*** | -0.084     | 0.146**  | -0.195  |
|           | (0.068)  | (0.175)    | (0.068)  | (0.112)    | (0.069)  | (0.119) |
| Weight    | -0.068   | -0.010     | -0.072   | 0.001      | -0.073   | 0.087   |
|           | (0.055)  | (0.185)    | (0.055)  | (0.110)    | (0.055)  | (0.127) |
| Divorce   | -3.063*  | 0.166      | -3.299*  | 0.084      | -3.299*  | -0.614  |
|           | (1.838)  | (2.952)    | (1.832)  | (2.527)    | (1.855)  | (2.480) |
| Age       | -0.240   | 1.733      | -0.805   | 0.804      | -0.506   | 1.774   |
|           | (0.577)  | (1.360)    | (0.585)  | (0.943)    | (0.579)  | (1.001) |
| Lessmoney | -2.989*  | 1.357      | -2.442** | -0.403     | -2.209*  | -0.542  |
|           | (1.302)  | (2.662)    | (1.282)  | (2.135)    | (1.293)  | (2.354) |
| Parentfight | 2.833 | 0.755    | 2.912    | 0.543      | 2.914    | 0.380   |
|           | (2.140)  | (2.944)    | (2.139)  | (2.847)    | (2.145)  | (2.934) |

$\sigma^2$ | 213.594 | 210.324 | 212.523 |
|           | (9.800) | (9.605) | (9.705) |

Note: We perform the MCMC sampling for 100,000 iterations with the initial 20,000 iterations discarded for burn-in, and compute the posterior means and standard deviations from the remaining draws as point estimates. The posterior standard deviations are reported in parentheses. The asterisks "***(*, *)" indicate that its 99% (95%, 90%) highest posterior density range does not cover zero.
Table I.4: Estimation Results of SC-SDPD Model for Peer Effects on Academic Performance from Study Mates Network

|       | (I)            | (II)            | (III)           |
|-------|----------------|-----------------|-----------------|
| \(\lambda\) | 0.107*** (0.032) | 0.106*** (0.032) | 0.107*** (0.031) |
| \(\rho\)   | 0.799*** (0.018) | 0.789*** (0.018) | 0.785*** (0.018) |
| \(\mu\)    | 0.003 (0.018)   | -0.008 (0.018)  | -0.007 (0.018)  |
| AICM       | 14,923          | 13,099          | 19,304          |
| se(AICM)   | 40.454          | 52.648          | 173.867         |

Note: Dynamic network formation model with latent variables of one dimension is in Column (I); two dimensions in Column (II); and three dimensions in Column (III). Each model controls own and contextual effects of exogenous regressors. We perform the MCMC sampling for 100,000 iterations with the initial 20,000 iterations discarded for burn-in, and compute the posterior means and standard deviations from the remaining draws as point estimates. The posterior standard deviations are reported in parentheses. The asterisks *** (**, *) indicate that its 99% (95%, 90%) highest posterior density range does not cover zero.

Table I.5: Robustness Check of SC-SDPD Model Estimation Results with Non-row-normalized \(W_{gt}\) or Additional Contextual Latent Variables

|       | Non-row-normalized \(W_{gt}\) | Additional contextual latent effects \(\overline{W}_{gt}Z_{gt}\) and \(\overline{W}_{gt}M_{gt}\) |
|-------|-------------------------------|-----------------------------------|
| \(\lambda\) | (I) 0.023*** (0.007) | (I) 0.109*** (0.035) |
|        | (II) 0.022*** (0.007) | (II) 0.102*** (0.033) |
|        | (III) 0.022*** (0.008) | (III) 0.101*** (0.037) |
| \(\rho\)   | (I) 0.796*** (0.018) | (I) 0.787*** (0.018) |
|        | (II) 0.790*** (0.018) | (II) 0.791*** (0.017) |
|        | (III) 0.797*** (0.017) | (III) 0.767*** (0.018) |
| \(\mu\)    | (I) -0.003 (0.003) | (I) 0.001 (0.030) |
|        | (II) -0.004 (0.003) | (II) 0.001 (0.031) |
|        | (III) 0.002 (0.003) | (III) 0.000 (0.032) |
| AICM       | 19,912           | 20,217            |
| se(AICM)   | 26,197           | 26,956            |
|           | 55,418           | 59,000            |
|           | 134,856          | 137,431           |

Note: Dynamic network formation model with latent variables of one dimension is in Column (I); two dimensions in Column (II), and three dimensions in Column (III). All regressions include own and contextual effects of exogenous \(X\). We perform the MCMC sampling for 100,000 iterations with the initial 20,000 iterations discarded for burn-in, and compute the posterior means and standard deviations from the remaining draws as point estimates. The posterior standard deviations are reported in parentheses. The asterisks *** (**, *) indicate that its 99% (95%, 90%) highest posterior density range does not cover zero.
Figure I.1: Time variations of network outdegrees in schools 1 to 12. Schools 8 and 9 only appear in the first and second time periods. Schools 10, 11, and 12 only appear in the third and fourth time periods.
Figure I.2: Marginal effects of the policy of removing financial hardship on network outdegree
Figure I.3: Marginal effects of the policy of removing financial hardship on academic outcome

(a) treated group: family in financial hardship

(b) untreated group: family not in financial hardship

(c) all students
Figure I.4: Marginal effects of the policy of removing financial hardship on academic outcome (based on the SAR model)
Figure I.5: Network goodness of fit plot for School 7 at the 2\textsuperscript{nd} time period
Figure I.6: Network goodness of fit plot for School 7 at the 3rd time period
Figure I.7: Network goodness of fit plot for School 7 at the 4th time period
Figure I.8: Time evolution of latent variables (based on transformed MCMC draws): panels (a)-(d) draw $z_{igt}$ and the node size represents outdegree. Panels (e)-(h) draw $m_{igt}$ and the node size represents outdegree. Panels (i)-(l) draw $m_{igt}$ and the node size represents indegree.
Figure I.9: Time evolution of latent variables (based on raw MCMC draws). Panels (a)-(d) draw $z_{igt}$ and the node size represents outdegree. Panels (e)-(h) draw $m_{igt}$ and the node size represents outdegree. Panels (i)-(l) draw $m_{igt}$ and the node size represents indegree.
References

Akaike, H. (1973), “Information theory and an extension of the maximum likelihood principle.” in Second international symposium on information theory, eds. Petrov, B. N. and Csaki, F., Budapest: Academiai Kiado, pp. 267–281.

Aßmann, C., Boysen-Hogrefe, J., and Pape, M. (2012), “The Directional Identification Problem in Bayesian Factor Analysis: An Ex-post Approach,” Working Paper No.1799, Kiel Institute for the World Economy, Hindenburgufer 66, 24105 Kiel, Germany.

— (2016), “Bayesian analysis of static and dynamic factor models: An ex-post approach towards the rotation problem,” Journal of Econometrics, 192, 190–206.

Bai, J. S. and Ng, S. (2013), “Principal components estimation and identification of static factors,” Journal of Econometrics, 176(1), 18–29.

Bai, J. S. and Wang, P. (2015), “Identification and Bayesian Estimation of Dynamic Factor Models,” Journal of Business & Economic Statistics.

Bickel, P. J. and Ghosh, J. K. (1990), “Decomposition of the likelihood ratio statistic and the Bartlett correction - A Bayesian argument.” Annals of Statistics, 18, 1070–1090.

Borg, I. and Groenen, P. J. F. (2005), Modern Multidimensional Scaling: Theory and Applications, New York: Springer.

Boucher, V. and Fortin, B. (2016), “Some challenges in the empirics of the effects of networks,” Oxford Handbook on the Economics of Networks, Oxford: Oxford University Press.

Dawid, A. P. (1991), “Fisherian inference in likelihood and prequential frames of reference,” Journal of Royal Statistical Society B (with discussion), 53, 79–109.

Gelman, A., Carlin, J., Stern, H., and Rubin, D. (2003), Bayesian Data Analysis, Second Edition, Chapman & Hall/CRC Texts in Statistical Science, Taylor & Francis.

Hoff, P., Raftery, A. E., and Handcock, M. S. (2002), “Latent Space Approaches to Social Network Analysis,” Journal of the American Statistical Association, 97.

Hsieh, C.-S. and Lee, L.-F. (2016), “A social Interaction model with endogenous friendship and selectivity,” Journal of Applied Econometrics, 31, 301–319.

Hsieh, C.-S. and van Kippersluis, H. (2018), “Smoking Initiation: Peers and Personality,” Quantitative Economics, 9, 825–863.
Hunter, D. R., Goodreau, S. M., and Handcock, M. S. (2008), “Goodness of fit of social network models,” *Journal of the American Statistical Association*, 103, 248–258.

Ichimura, H. (1993), “Semiparametric least squares (SLS) and weighted SLS estimation of single-index models,” *Journal of Econometrics*, 58, 71–120.

Lee, L.-f. and Yu, J. (2012), “QML estimation of spatial dynamic panel data models with time varying spatial weights matrices,” *Spatial Economic Analysis*, 7, 31–74.

Liu, X., Patacchini, E., and Zenou, Y. (2014), “Endogenous peer effects: local aggregate or local average?” *Journal of Economic Behavior & Organization*, 103, 39–59.

Raftery, A. E., Newton, M. A., Satagopan, J. M., and Krivitsky, P. N. (2007), “Estimating the integrated likelihood via posterior simulation using the harmonic mean identity,” *Bayesian Statistics*, 8, 1–45.

Schönemann, P., H. (1966), “A generalized solution to the orthogonal Procrustes Problem,” *Psychometrika*, 31(1), 1–10.

Sewell, D. K. and Chen, Y. (2015), “Latent space models for dynamic networks,” *Journal of the American Statistical Association*, 110, 1646–1657.