Econ 7217 Economic Analysis of Social Networks
Peer (Spillover) Effects from Social Networks

Chih-Sheng Hsieh

Department of Economics
National Taiwan University

March 11, 2024

Studies on social interactions

There is a huge literature studying peer (spillover) effects in different contexts:

- primary school reading literacy progress (Ammermueller and Pischke, 2009); high school academic performance (Calvó-Armengol et al., 2009; Lin, 2010); college achievement (Carrell et al., 2009; Zimmerman, 2003; Sacerdote, 2001); academic cheating (Carrell et al., 2008);
- smoking (Powell et al., 2005; Nakajima, 2007; Fletcher, 2010a; Hsieh and Van Kippersluis, 2018) and substance use (Lundborg, 2006; Kawaguchi, 2004; Clark and Lohéac, 2007); crime and juvenile behavior (Glaeser et al., 1996; Gaviria and Raphael, 2001; Bayer et al., 2009);

Studies on social interactions

- school enrollment decision (Bobonis and Finan, 2009; Lalive and Cattaneo, 2009);
- occupation choice (Arcidiacono and Nicholson, 2005);
- productivity in workplace (Mas and Moretti, 2009; Guryan et al., 2009; Nair et al., 2010);
- scientific research (Waldinger, 2011);
- program participation (Dahl et al., 2014; Duflo and Saez, 2003; Bertrand et al., 2000);
- monetary shocks on real economy (Ozdagli and Weber, 2017);
- spatial dependence on FDI (Blonigen et al., 2007)

Linear-in-means model

- Finding correlation on individuals' economic activities is not difficult, but to argue that it results from peer effects is difficult. One must deal with several identification challenges, such as the reflection problem, self-selection, unobserved confounding effects, etc.
- The conventional model employed to study social interactions is the linear-in-means model (Manski, 1993).
- Group mean outcome, $\mathbb{E}(Y_g)$, and group mean characteristics, $\mathbb{E}(X_g)$, for group g are included to capture the endogenous effect (λ_0) and contextual effects (β_{20}), respectively:

$$y_{ig} = \lambda_0 \mathbb{E}\left(Y_g\right) + \beta_{10} x_{ig} + \beta_{20} \mathbb{E}\left(X_g\right) + \varepsilon_{ig}. \tag{1}$$

- Assuming $\mathbb{E}(\varepsilon_g) = 0$, we can show that the two regressors $\mathbb{E}\left(Y_g\right)$ and $\mathbb{E}\left(X_g\right)$ are linearly dependent:

$$\mathbb{E}\left(Y_g\right) = \frac{\beta_{10} + \beta_{20}}{1 - \lambda_0} \mathbb{E}\left(X_g\right). \tag{2}$$

Linear-in-means model

- The reduced form of the model is
$$y_{ig} = \beta_{10} x_{ig} + \left(\frac{\lambda_0 \beta_{10} + \beta_{20}}{1 - \lambda_0} \right) \mathbb{E}\left(X_g\right) + \varepsilon_{ig}. \tag{3}$$
- The "reflection problem": only β_{10} and $\frac{\lambda_0 \beta_{10} + \beta_{20}}{1 - \lambda_0}$ are identified from the reduced form. One cannot separately identify the endogenous effect coefficient λ_0 and the contextual effects coefficient β_{20}.
- To circumvent the "reflection problem", empirical studies usually focus on either the endogenous effect or the contextual effect, assuming the other type of social effects is not present.
- The alternative way to solve the "reflection problem" is the **exclusion** of individual's own outcome/characteristics from the group mean as well as **the variation of group size**, which breaks down the linear relationship between the endogenous and contextual effects in Equation (2).
- One can also prevent the reflection problem by using nonlinear models, or panel data, etc.

Reflection Problem

- The model with both endogenous and contextual effects:

$$y_{ig} = \lambda_0 \overline{Y}_{-ig} + \beta_{10} x_{ig} + \beta_{20} \overline{X}_{-ig} + \varepsilon_{ig}, \quad g = 1, \cdots, G, \tag{4}$$

where $\overline{Y}_{-ig} = \left(\frac{1}{m_g - 1} \sum_{j \neq i}^{m_g} y_{jg} \right)$ and $\overline{X}_{-ig} = \left(\frac{1}{m_g - 1} \sum_{j \neq i}^{m_g} x_{jg} \right)$ and m_g is the size of group g.

- Solving for $\overline{Y}_{-ig}$ one can get a different version of Equation (2):

$$\overline{Y}_{-ig} = \frac{\lambda_0}{(m_g - 1) - \lambda_0 (m_g - 2)} y_{ig} + \frac{\beta_{10} (m_g - 1) + \beta_{20} (m_g - 2)}{(m_g - 1) - \lambda_0 (m_g - 2)} \overline{X}_{-ig} + \frac{\beta_{20}}{(m_g - 1) - \lambda_0 (m_g - 2)} x_{ig} + \frac{(m_g - 1) \overline{\varepsilon}_{-ig}}{(m_g - 1) - \lambda_0 (m_g - 2)},$$

which is certainly not linear in $\overline{X}_{-ig}$.

Reflection Problem

The reduced form of the model becomes:

$$y_{ig} = \frac{\lambda_0\beta_{20} + \left[(m_g - 1) - \lambda_0\,(m_g - 2)\right]\beta_{10}}{(1-\lambda_0)\,(m_g - 1 + \lambda_0)} x_{ig} + \frac{(m_g - 1)\,(\lambda_0\beta_{10} + \beta_{20})}{(1-\lambda_0)\,(m_g - 1 + \lambda_0)} \overline{X}_{-ig} + \frac{(m_g - 1)}{(m_g - 1 + \lambda_0)} \varepsilon_{ig} + \frac{\lambda_0 m_g}{(1-\lambda_0)\,(m_g - 1 + \lambda_0)} \overline{\varepsilon}_g . \tag{5}$$

- As long as the group size, m_g, is different across groups, valuable information from the differences in the social interaction patterns among the groups resolves the "reflection problem".

Correlated Effects

- There is another endogeneity issue due to self-selection (sorting) into groups and the unobserved environmental factors (correlated effects) that confound with the endogenous and contextual peer effects.
- For example, rich families will move in good school districts that their children can go to these good schools.
- Standard solution in the literature is to include group fixed effects.
- A model with endogenous and contextual effects, as well as group fixed effects, is specified as follows:

$$y_{ig} = \lambda_0 \overline{Y}_{-ig} + \beta_{10} x_{ig} + \beta_{20} \overline{X}_{-ig} + \alpha_g + \epsilon_{ig}, \quad g = 1, \cdots, G, \qquad (6)$$

where α_g denotes the fixed effect for group g.

Correlated Effects

- Based on Equation (6), Lee (2007) derives the between equation

$$\overline{Y}_g = \frac{(\beta_{10} + \beta_{20})}{(1 - \lambda_0)}\overline{X}_g + \frac{1}{(1 - \lambda_0)}\alpha_g + \frac{1}{(1 - \lambda_0)}\bar{\epsilon}_g, \quad g = 1, \cdots, G, \quad (7)$$

and the within equation

$$y_{ig} - \overline{Y}_g = \frac{(m_g - 1)\left(\beta_{10} - \frac{\beta_{20}}{m_g - 1}\right)}{(m_g - 1 + \lambda_0)}(x_{ig} - \overline{X}_g) + \frac{(m_g - 1)}{(m_g - 1 + \lambda_0)}(\epsilon_{ig} - \bar{\epsilon}_g), \quad (8)$$

where $\overline{Y}_g$, $\overline{X}_g$, and $\bar{\epsilon}_g$ are computed from all m_g individuals in the g^{th} group.

Correlated Effects

- Because Equation (7) contains the unknown group fixed effect α_g, there is no way to identify λ_0, β_{10}, and β_{20} separately from α_g. The possible identification only relies on Equation (8).
- If all group sizes are the same, i.e., $m_g = m$ for all g, identification fails because there are three structural form parameters but there is only one reduced-form equation. Thus, group size variation – at least three different group sizes – is again a necessary condition for identification.
- Nonetheless, it is important to note that the group size variation may only provide a tenuous source of identification. When the group size is large, the ratio $\frac{m_g-1}{m_g-1+\lambda_0}$ is close to one and λ_0 cannot be easily estimated from Equation (8).

SAR Model – Identification

- Information on the social network of each individual in a group may be captured by a spatial weights matrix in a spatial autoregressive (SAR) model (Anselin, 1988; Kelejian and Prucha, 1998; Lee, 2004), which introduces additional non-linearity for identification of various social interaction effects beyond the variation of group sizes.
- The network interaction model with the endogenous and contextual peer effects, and the group fixed effects is specified as:

$$y_{ig} = \lambda_0 \sum_{j=1}^{m_g} \overline{w}_{ij,g} y_{jg} + \beta_{10} x_{ig} + \beta_{20} \sum_{j=1}^{m_g} \overline{w}_{ij,g} x_{jg} + \alpha_g + \epsilon_{ig}, \quad g = 1, \cdots, G, \tag{9}$$

where $\overline{w}_{ij,g} = \frac{w_{ij,g}}{\sum_j w_{ij,g}}$, $w_{ij,g} = 1$ if individual j is i's friend;

SAR Model – Identification

- To see where the identification comes from, consider a simplified model without group fixed effects and in the matrix form:

$$Y_g = \lambda_0 \overline{W}_g Y_g + X_g \beta_{10} + \overline{W}_g X_g \beta_{20} + \epsilon_g, \quad g = 1, \ldots, G. \tag{10}$$

- This model is identified if and only if some instruments can be found for the endogenous vector $\overline{W}_g Y_g$.
- Bramoullé et al. (2009) and Calvó-Armengol et al. (2009) found useful IVs under the condition that the matrices, $I_{m_g}, \overline{W}_g, \overline{W}_g^2$ are linearly independent, where I_{m_g} denotes the identity matrix of dimension m_g.
- This will be true as long as the network links are intransitive, i.e., some individuals are not friends with her friends' friends. Then the characteristics of peers of peers, $\overline{W}_g^2 X_g$, may serve as instruments for peers' outcomes, $\overline{W}_g Y_g$.

SAR Model – estimation with 2SLS

- From Equation (10), the reduced form is

$$
\begin{aligned}
Y_g &= \left(I_{m_g} - \lambda_0 \overline{W}_g\right)^{-1} \left(\beta_{10} X_g + \beta_{20} \overline{W}_g X_g + \epsilon_g\right) \\
&= \left(I_{m_g} - \lambda_0 \overline{W}_g\right)^{-1} \left(H_g \beta_0 + \epsilon_g\right), \qquad (11)
\end{aligned}
$$

where $H_g = (X_g, \overline{W}_g X_g)$ and $\beta_0 = (\beta_{10}, \beta_{20})'$.

$$
\mathbb{E}(\overline{W}_g Y_g | \overline{W}_g, X_g) = \overline{W}_g \left(I_{m_g} - \lambda_0 \overline{W}_g\right)^{-1} (H_g \beta_0). \qquad (12)
$$

- If $|\lambda_0| < 1$, the matrix $\left(I_{m_g} - \lambda_0 \overline{W}_g\right)$ is nonsingular and one can write

$$
\left(I_{m_g} - \lambda_0 \overline{W}_g\right)^{-1} = I_{m_g} + \lambda_0 \overline{W}_g + \lambda_0^2 \overline{W}_g^2 + \cdots \qquad (13)
$$

$$
\begin{aligned}
\mathbb{E}(\overline{W}_g Y_g) &= \overline{W}_g H_g \beta_0 + \lambda_0 \overline{W}_g^2 H_g \beta_0 + \lambda_0^2 \overline{W}_g^3 H_g \beta_0 + \cdots \\
&= (\overline{W}_g H_g, \overline{W}_g^2 H_g, \overline{W}_g^3 H_g, \cdots)(\beta_0', \lambda_0 \beta_0', \lambda_0^2 \beta_0', \cdots)' \qquad (14)
\end{aligned}
$$

and thus $(\overline{W}_g H_g, \overline{W}_g^2 H_g, \overline{W}_g^3 H_g, \cdots)$ can be used as IVs for $\overline{W}_g Y_g$.

SAR Model – estimation with 2SLS

- Define $U_g = (H_g, \overline{W}_g H_g, \overline{W}_g^2 H_g, \cdots, \overline{W}_g^p H_g)$ as the IV matrix, then

$$\mathbb{E}(U_g' \epsilon_g) = 0. \tag{15}$$

- Let $\epsilon_g(\lambda_0, \delta_0) = Y_g - \lambda_0 \overline{W}_g Y_g - H_g \beta_0$, the empirical counterpart of Equation (15) is

$$U_g' \epsilon_g(\lambda_0, \beta_0) = 0. \tag{16}$$

- One can further incorporate quadratic moment conditions in the form, $\epsilon_g(\lambda_0, \beta_0)' A \epsilon_g(\lambda_0, \beta_0)$ with A denotes a constant square matrix in GMM to improve estimation efficiency (Liu and Lee, 2010) so that GMM can be asymptotically as efficient as MLE.

SAR Model – estimation with 2SLS

- Define $Y = (Y_1', \cdots, Y_G')'$, $H = (H_1', \cdots, H_G')'$, $\epsilon = (\epsilon_1', \cdots, \epsilon_G')'$, $W = D(\overline{W}_1, \cdots, \overline{W}_G)$, where D denotes the block diagonal matrix, and $Z = (WY, H)$ is the matrix of regressors.
- Use the instrument matrix $U = (H, WH, W^2H)$, the 2SLS estimator for $\delta = (\lambda, \beta')'$ is

$$\hat{\delta}_{2sls} = \left(Z'PZ\right)^{-1} Z'PY, \tag{17}$$

where $P = U(U'U)^{-}U'$ and A^{-} is the generalized inverse of a square matrix A.

$$\mathrm{var}(\hat{\delta}_{2sls}) = \hat{\sigma}^2 \left(Z'PZ\right)^{-1}. \tag{18}$$

SAR Model – estimation with 2SLS

- Consider a general form of the model with group fixed effect as in Equation (9):

$$Y_g = \lambda_0 \overline{W}_g Y_g + X_g \beta_{10} + \overline{W}_g X_g \beta_{20} + \ell_{m_g} \alpha_g + \epsilon_g, g = 1, \ldots, G. \quad (19)$$

- As shown in Lee et al. (2010) and Lin (2010), the group fixed effect can be eliminated by the de-group-mean transformation, i.e., by multiplying Equation (19) with the matrix

$$J_{m_g} = I_{m_g} - \frac{1}{m_g} \ell_{m_g} \ell'_{m_g}. \quad (20)$$

$$\hat{Y}_g = \lambda_0 \overline{W}_g \hat{Y}_g + \hat{X}_g \beta_{10} + \overline{W}_g \hat{X}_g \beta_{20} + \hat{\epsilon}_g, \quad (21)$$

where $\hat{Y}_g = J_{m_g} Y_g, \hat{X}_g = J_{m_g} X_g,$ and $\hat{\epsilon}_g = J_{m_g} \epsilon_g$.

- This model is identified if and only if $\mathbb{E}\left(\overline{W}_g \hat{Y}_g | X_g\right)$ is not perfectly collinear with the regressors $\left(\hat{X}_g, \overline{W}_g \hat{X}_g\right)$ which is equivalent to the matrices, $I_{m_g}, \overline{W}_g, \overline{W}_g^2, \overline{W}_g^3$ are linearly independent.

SAR Model – estimation with MLE

- Assume $\epsilon_g | W_g, X_g \sim N(0, \sigma^2 I_{m_g})$, then from Equation(11) we have

$$Y_g \sim N(\mu_{Y_g}, \Sigma_{Y_g}), \tag{22}$$

where $\mu_{Y_g} = (I_{m_g} - \lambda \overline{W}_g)^{-1} H_g \delta$ and
$\Sigma_{Y_g} = \sigma^2 (I_{m_g} - \lambda W_g)^{-1} (I_{m_g} - \lambda W_g')^{-1}$.

- Let $\theta = (\lambda, \delta', \sigma^2)'$, the log-likelihood is given by

$$\mathcal{L}(\theta) = -\frac{\sum_g m_g}{2} \ln(2\pi) - \frac{\sum_g m_g}{2} \ln \sigma^2 + \sum_g \ln |(I_{m_g} - \lambda \overline{W}_g)| - \frac{1}{2\sigma^2} \sum_g \epsilon_g(\lambda, \delta)' \epsilon_g(\lambda, \delta). \tag{23}$$

- MLE is inconsistent when the error terms are heteroskedastic.
- Lin and Lee (2010) propose GMM estimation for SAR models with unknown heteroskedasticity.

Social Interactions and Social Preference in Social Networks (Hsieh and Lin, 2021)

- The experimental literature on social preferences has recently started to uncover the connection between peer effects and social preferences.
- As social interactions occur on a regular basis and within small groups, altruism is expected to play an important role since people may intrinsically care about the well-being of their social contacts and take into account their preferences when making decisions.
- In this paper, we provide the first empirical analysis on social interactions and social preferences in social networks, building a bridge between the two strands of rapidly growing yet unrelated literature. We investigate the model specification issues that emerge from extending the standard assumption of self-interest to a more evolutionary foundation in which individuals may be altruistic.

Social Interactions and Social Preference in Social Networks (Hsieh and Lin, 2021)

Consider that individuals belong to groups, $g = 1, \cdots, G$.

- $y_{i,g}$: activity outcome of individual i in group g.
- $x_{i,g}$: a $k \times 1$ individual characteristics.
- W_g: an adjacency matrix that summarizes the friendship links between m_g individuals in group g. $w_{ij,g} = 1$ if individual i chooses individual j as a friend and $w_{ij,g} = 0$ otherwise. **We assume W_g is not symmetric.**

$$y_{i,g} = \lambda \sum_{j=1}^{m_g} w_{ij,g} y_{j,g} + x_{i,g}\beta_1 + \sum_{j=1}^{m_g} w_{ij,g} x_{j,g}\beta_2 + \tau_g + \epsilon_{i,g}, \quad \text{for } i = 1, \cdots, m_g. \tag{24}$$

We call Equation (24) or its vector form "conventional social interactions model",

$$Y_g = \lambda W_g Y_g + X_g\beta_1 + W_g X_g \beta_2 + \ell_g \tau_g + \epsilon_g. \tag{25}$$

Social Interactions and Social Preference in Social Networks

- The quadratic specification of individual payoff function from Ballester et al. (2006) and Calvó-Armengol et al. (2009):

$$v_{i,g}(y_{i,g}, Y_{-i,g}, W_g) = \mu_{i,g} y_{i,g} - \frac{1}{2} y_{i,g}^2 + \lambda y_{i,g} \sum_{j=1}^{m_g} w_{ij,g} y_{j,g} + \eta \sum_{j=1}^{m_g} w_{ij,g} y_{j,g}. \tag{26}$$

- $\mu_{i,g}$ in the first term reflects individual heterogeneity.
- $y_{i,g} \sum_{j=1}^{m_g} w_{ij,g} y_{j,g}$ in the third term captures the complementary effect from peers' activities, which provides the source of peer influences.
- $\sum_{j=1}^{m_g} w_{ij,g} y_{j,g}$ in the fourth term captures the direct externality effect from peers' activity outcomes.

Social Interactions and Social Preference in Social Networks

- If individuals directly choose $y_{i,g}$ to maximize Eq. (26), then the first order condition gives

$$\mu_{i,g} - y_{i,g} + \lambda \sum_{j=1}^{m_g} w_{ij,g} y_{j,g} = 0,$$

which implies the optimal choice of $y_{i,g}$ satisfies

$$y_{i,g} = \mu_{i,g} + \lambda \sum_{j=1}^{m_g} w_{ij,g} y_{j,g}.$$

- The unique Nash equilibrium vector of outcomes is given by $Y_g = (I_{m_g} - \lambda W_g)^{-1} \mu_g$ if $|\lambda|$ satisfies a bound constraint.
- By specifying $\mu_{i,g} = \left(x_{i,g}\beta_1 + \sum_{j=1}^{m_g} w_{ij,g} x_{j,g}\beta_2 + \tau_g + \epsilon_{i,g}\right)$, the outcome Y_g can be modelled by the SAR in Equation (25).

Social Interactions and Social Preference in Social Networks

- Now assume individuals are altruistic to others, which means individuals care about others' payoffs.
- One simplest form of the utility function $U_{i,g}$ which illustrates the above altruistic preference is a linear function of individual's own and others' payoffs

$$U_{i,g} = v_{i,g}(y_{i,g}, Y_{-i,g}, W_g) + \alpha_1 \sum_{j=1, j \neq i}^{m_g} v_{j,g}(y_{j,g}, Y_{-j,g}, W_g). \tag{27}$$

- The coefficient α_1 reflects how much individuals care about others' payoffs, i.e., the altruism level.
- We assume α_1 to be bounded by [-1,1]. When α_1 is positive, individuals are altruistic; when α_1 equals to zero, individuals are selfish; and when α_1 is negative, individuals are spiteful.

Social Interactions and Social Preference in Social Networks

- Based on Equation (27), the first order condition $\frac{\partial U_{i,g}}{\partial y_{i,g}} = 0$ determines individual's choice of $y_{i,g}$.
- What we get is

$$\mu_{i,g} - y_{i,g} + \lambda \sum_{j=1}^{m_g} w_{ij,g} y_{j,g} + \alpha_1 (\lambda y_{j,g} w_{ji,g} + \eta w_{ji,g}) = 0.$$

- The vector of Nash equilibrium strategy played by individuals is

$$Y_g = \lambda W_g Y_g + \lambda^{\mathrm{I}} W_g^{\mathrm{T}} Y_g + \eta^{\mathrm{I}} W_g^{\mathrm{T}} \ell_g + X_g \beta_1 + W_g X_g \beta_2 + \ell_g \tau_g + \epsilon_g, \tag{28}$$

where $\lambda^{\mathrm{I}} = \alpha_1 \lambda$; $\eta^{\mathrm{I}} = \alpha_1 \eta$ and W^{T} denotes the transpose of W_g. We refer Equation (28) as the "altruistic social interactions model."

Social Interactions and Social Preference in Social Networks

- This extension justifies the peer effects from both outward and inward friendship links and why the in-degree centrality ($W_g^{\mathrm{T}}\ell_g$) (or popularity) would matter for individual outcome.
- It also allows us to infer the degree of altruism toward others' payoffs and the direct externality effect from peers' activity outcome.
- The estimate of λ in the conventional social interactions model would be
  1. (upward) biased when omitting the altruism term, $W_g^{\mathrm{T}} Y_g$
  2. (upward or downward) biased when omitting the externality term, $W_g^{\mathrm{T}}\ell_g$

Social Interactions and Social Preference in Social Networks

- We can further enrich the utility specification of Equation (27) as follows:

$$U_{i,g} = v_{i,g}(y_{i,g}, Y_{-i,g}, W_g) + \sum_{j=1, j\neq i}^{m_g} (\alpha_1 + \alpha_2 w_{ij,g}) v_{j,g}(y_{j,g}, Y_{-j,g}, W_g), \tag{29}$$

to reflects the directed altruism in Leider et al. (2009) – the altruism level should be stronger among friends than among random strangers (heterogeneous altruism among pairs of individuals).

Social Interactions and Social Preference in Social Networks

- Based on Equation (29), the vector of Nash equilibrium strategy can be modeled by the "directed altruistic social interactions model."

$$\begin{aligned} Y_g = {} & \lambda W_g Y_g + \lambda^{\mathrm{I}} W_g^{\mathrm{T}} Y_g + \lambda^{\mathrm{R}} W_g^{\mathrm{R}} Y_g + \eta^{\mathrm{I}} W_g^{\mathrm{T}} \ell_g + \eta^{\mathrm{R}} W_g^{\mathrm{R}} \ell_g \\ & + X_g \beta_1 + W_g X_g \beta_2 + \ell_g \tau_g + \epsilon_g, \end{aligned} \tag{30}$$

  where W_g^{R} stands for the network of reciprocal links, i.e., $w_{ij,g}^{\mathrm{R}} = 1$ if $w_{ij,g} = w_{ji,g} = 1$. $\lambda^{\mathrm{I}} = \alpha_1 \lambda$, $\eta^{\mathrm{I}} = \alpha_1 \eta$, $\lambda^{\mathrm{R}} = \alpha_2 \lambda$, $\eta^{\mathrm{R}} = \alpha_2 \eta$.
- It provides the microfoundations to explain the results of heterogeneous peer effects between chosen, unchosen, and mutual friends in Lin and Weinberg (2014).

Social Interactions and Social Preference in Social Networks

- Our empirical study employs the Add Health data, which is a longitudinal study on a nationally representative sample in U.S. covering adolescents in grade 7 through 12 (average age from 12 to 17) from 132 schools.
- The Add Health data contains detailed information about respondents' demographic backgrounds, academic performance, and health related behaviors. Most uniquely, the Add Health asked each respondent to nominate their male and female friends so that researchers can use the information to construct students' friendship networks.

Social Interactions and Social Preference in Social Networks

- We define groups at the school level and consider a sample of 24 schools with each size less than 300. There is a total of 2,926 students in this sample.
- We study four different activity outcomes
    1. academic achievement (measured by GPA)
    2. smoking (measured by number of smoking days per month)
    3. extracurricular activities (measured by the number of school clubs attended)
    4. misconducts (measured by frequency of doing dangerous activity, lying, and school skipping)
- We don't find significant estimates of altruism for club participation and misconduct.

Social Interactions and Social Preference in Social Networks

Variables	Min.	Max.	Mean	S.D.
GPA	1	4	2.9020	0.7379
Smoking	0	30	4.4142	9.8641
Club	0	6	2.5772	1.8540
Misconduct	0	6	1.2230	1.0267
Male	0	1	0.4836	0.4998
Age	11	19	15.5533	1.2383
White	0	1	0.6141	0.4869
Black	0	1	0.2409	0.4277
Asian	0	1	0.0208	0.1429
Hispanic	0	1	0.0687	0.2530
Other race	0	1	0.0554	0.2287
Both parents	0	1	0.7310	0.4435
Less HS	0	1	0.1063	0.3083
HS	0	1	0.3486	0.4766
More HS	0	1	0.4070	0.4914
Edu missing	0	1	0.0660	0.2483
Welfare	0	1	0.0109	0.1040
Job missing	0	1	0.0738	0.2615
Professional	0	1	0.2642	0.4410
Other job	0	1	0.3452	0.4755
Homemaker	0	1	0.2338	0.4233
group size	15	245	171.9986	66.1802
Number of groups (schools)			24	
Observations			2,926	

Social Interactions and Social Preference in Social Networks

Table: Estimation Result: Conventional and Altruistic Social Interactions Models

	Conventional Model		Altruistic Model		Altruistic Model w/ direct externality	
	GPA	Smoking	GPA	Smoking	GPA	Smoking
λ	0.0700***	0.0924***	0.0610***	0.0759***	0.0450***	0.0592***
	(0.0069)	(0.0047)	(0.0055)	(0.0135)	(0.0058)	(0.0099)
λ^{I}			0.0074***	0.0184	0.0302***	0.0367***
			(0.0015)	(0.0127)	(0.0062)	(0.0097)
η^{I}					-0.0701***	-0.1915***
					(0.0194)	(0.0707)

Social Interactions and Social Preference in Social Networks

Table: Estimation Result: Directed Altruistic Social Interactions Models

	Directed Altruistic Model		Directed Altruistic Model w/ externality	
	GPA	Smoking	GPA	Smoking
λ	0.0648***	0.0765***	0.0458***	0.0564***
	(0.0071)	(0.0116)	(0.0075)	(0.0110)
λ^{I}	0.0087***	0.0025	0.0273***	0.0306***
	(0.0021)	(0.0134)	(0.0093)	(0.0119)
λ^{R}	-0.0044	0.0260*	0.0012	0.0148
	(0.0046)	(0.0144)	(0.0118)	(0.0138)
η^{I}			-0.0563**	-0.3007***
			(0.0287)	(0.0945)
η^{R}			-0.0239	0.3792
			(0.0388)	(0.2085)

Smoking Initiation – Peers and Personality (Hsieh and Van Kippersluis, 2018)

- Smoking continues to be the leading preventable cause of death, killing nearly 6 million people each year (Mokdad et al., 2004; Danaei et al., 2009; OECD, 2013).
- The economics literature has traditionally focused on price, taxation, and addiction as determinants of smoking (Chaloupka and Warner, 2000; DeCicca et al., 2002).
- In recent years, considerably more attention is paid to social interactions in smoking and other unhealthy behaviors (Cawley and Ruhm, 2011; DeCicca et al., 2008).
- Social interactions and peer effects are not just often-cited determinants of smoking initiation, but – when present – additionally capable of generating social multiplier effects of policy interventions (Cutler and Glaeser, 2010; Fletcher, 2010b; Cawley and Ruhm, 2011).

Smoking Initiation – Peers and Personality

Most widely-accepted taxonomy of personality (also known as non-cognitive skills) is the so-called "Big Five" (acronym OCEAN, Digman, 1990; Matthews et al., 2003):

1. Openness to experience ("the degree to which a person needs intellectual stimulation, change, and variety")
2. Conscientiousness ("the degree to which a person is willing to comply with conventional rules, norms, and standards")
3. Extraversion ("the degree to which a person needs attention and social interaction")
4. Agreeableness ("the degree to which a person needs pleasant and harmonious relations with others")
5. Emotional stability (or Neuroticism, "the degree to which an individual experiences the world as threatening and beyond his/her control")

Smoking Initiation – Peers and Personality

- When your behavior is correlated with the behaviors of your peers, this could occur for three different reasons:
  - Endogenous effect: individual behavior is truly affected by the peers' behaviors.
  - Contextual effect: individual behavior is affected by the (observed and unobserved) exogenous characteristics of the peers.
  - Correlated effect (a.k.a. selection, common environment): behaviors are similar due to (i) self-selection into the group, or (ii) share similar environments (e.g. rich neighborhood, good teacher quality).
- The reflection problem (Manski, 1993): the endogenous and contextual effects are not separately identified in a linear model when peers of individuals are perfectly identical, e.g., all school students are your peers.

Smoking Initiation – Peers and Personality

Incorporate personality effects into the SAR model:

$$
\begin{aligned}
y_{i,g} = {} & \lambda_{11} I(s_{ir,g} < \overline{S}_{r,g}) \sum_{j \neq i} w_{ij,g} I(s_{jr,g} < \overline{S}_{r,g}) y_{j,g} \\
& + \lambda_{12} I(s_{ir,g} < \overline{S}_{r,g}) \sum_{j \neq i} w_{ij,g} I(s_{jr,g} \geq \overline{S}_{r,g}) y_{j,g} \\
& + \lambda_{21} I(s_{ir,g} \geq \overline{S}_{r,g}) \sum_{j \neq i} w_{ij,g} I(s_{jr,g} < \overline{S}_{r,g}) y_{j,g} \\
& + \lambda_{22} I(s_{ir,g} \geq \overline{S}_{r,g}) \sum_{j \neq i} w_{ij,g} I(s_{jr,g} \geq \overline{S}_{r,g}) y_{j,g} \\
& + x_{i,g}\beta_1 + \sum_{j \neq i} w_{ij,g} x_{i,g}\beta_2 + s_{i,g}\beta_3 + \sum_{j \neq i} w_{ij,g} s_{i,g}\beta_4 + \alpha_g + \epsilon_{i,g},
\end{aligned}
$$

where $I(A)$ denotes an indicator function which equals one if A is satisfied and equals zero, otherwise. $s_{ir,g}$ denote the r^{th} dimension of personality. $\overline{S}_{r,g}$ denotes the average of the r^{th} personality measure in group g.

Smoking Initiation – Peers and Personality

The vector form of our SAR model specification:

$$
\begin{aligned}
Y_g = {} & \lambda_{11} W_{11,g} Y_g + \cdots + \lambda_{22} W_{22,g} Y_g \\
& + X_g \beta_1 + W_g X_g \beta_2 + S_g \beta_3 + W_g S_g \beta_4 + \ell_g \alpha_g + \epsilon_g,
\end{aligned}
$$

W_g is now divided into 2×2 blocks, where each block, W_g^{pq}, $p, q = 1, 2$, represents the subnetwork between individuals in the personality subgroups 1 and 2. $W_{pq,g}$ is a $m_g \times m_g$ matrix with the corresponding $(p, q)^{\mathrm{th}}$ block equal to W_g^{pq} and 0 elsewhere.

Smoking Initiation – Peers and Personality

- Add Health is a nationally representative survey of adolescents enrolled in grades 7-12 from 132 schools. Four waves of surveys were conducted from 1994 to 2008.
- In Wave I, 90,000 Students were interviewed and each respondent was asked to nominate up to five male and five female friends. We use such information to construct students' friendship networks.
- To study the relationship between personality and smoking, we use the Wave I in-home survey data, which contains more detailed questions on family background than the in-school survey data, and includes information on individual's personality characteristics.
- There are 20,745 students who participated in the Wave I in-home survey. We focus on small- and mid-size schools that have less than three hundred students.

Smoking Initiation – Peers and Personality

Construct three out of the big five personality measures: emotional stability, conscientiousness, and extraversion from the factor analysis.

Item identifier	Description	Factor loadings			Regression Coef.		
		ES	C	E	ES	C	E
H1PF30	You have a lot of good qualities	0.6207			0.3164		
H1PF32	You have a lot to be proud of	0.6602			0.3930		
H1PF33	You like yourself just the way you are	0.4145			0.0150		
H1PF34	You feel like you are doing everything just about right	0.3502			-0.0371		
H1PF35	You feel socially accepted	0.4191			0.0408		
H1PF36	You feel wanted and loved	0.4903			0.1137		
H1PF18	You get as many facts about the problem as possible when you have problems to be solved		0.6300			0.2778	
H1PF19	You think of as many different ways to approach a problem as possible when you are attempting to find a solution		0.6700			0.3195	
H1PF20	You generally use a systematic method for judging and comparing alternatives When making decisions		0.6245			0.2741	
H1PF21	You usually try to analyze what went right and what went wrong after carrying out a solution to a problem		0.5680			0.2295	
S62B*	I feel close to people at school			0.7014			0.3543
S62E*	I feel like I am a part of this school			0.7175			0.3786
S62O*	I feel socially accepted			0.6245			0.2696

Note: These 13 items are selected by the Lexical approach and the exploratory factor analysis according to Young et al.(2011). We conduct the principle component analysis on these items and identify one main factor for each personality measure, which explains more than 90% of variation. We report the factor loadings for each item after rotation and the regression coefficients for predicting factor scores. ES: emotional stability; C:conscientiousness; E: extroversion. $*$ denotes that data source is Wave I in-school survey.

Smoking Initiation – Peers and Personality

	Whole Sample				Emotional Stability				Conscientiousness			
					high		low		high		low	
	mean	s.d.	min	max	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
smoke dummy	0.224	0.417	0.000	1.000	0.184	0.388	0.261	0.439	0.205	0.404	0.247	0.431
smoke frequency	0.904	2.162	0.000	7.000	0.721	1.955	1.070	2.321	0.832	2.095	0.989	2.236
drunk	0.311	0.463	0.000	1.000	0.272	0.445	0.346	0.476	0.296	0.457	0.327	0.469
emotional stability	0.015	0.715	-4.033	1.021	0.636	0.285	-0.547	0.485	0.146	0.670	-0.140	0.735
conscientiousness	-0.032	0.831	-3.790	1.565	0.188	0.877	-0.232	0.733	0.536	0.499	-0.701	0.619
extraversion	0.009	0.840	-2.385	1.260	0.193	0.806	-0.157	0.835	0.082	0.838	-0.077	0.833
male	0.466	0.499	0.000	1.000	0.517	0.500	0.420	0.494	0.475	0.499	0.456	0.498
white	0.541	0.498	0.000	1.000	0.538	0.499	0.544	0.498	0.530	0.499	0.554	0.497
black	0.228	0.419	0.000	1.000	0.258	0.438	0.200	0.400	0.239	0.427	0.215	0.411
asian	0.110	0.313	0.000	1.000	0.100	0.301	0.119	0.324	0.107	0.309	0.115	0.319
hisp	0.064	0.244	0.000	1.000	0.048	0.214	0.078	0.267	0.068	0.253	0.058	0.233
other race	0.057	0.232	0.000	1.000	0.055	0.229	0.059	0.235	0.056	0.230	0.059	0.235
school taught	0.934	0.248	0.000	1.000	0.942	0.233	0.927	0.261	0.940	0.238	0.928	0.259
smoke parent	0.643	0.479	0.000	1.000	0.632	0.482	0.652	0.476	0.634	0.482	0.654	0.476
prof	0.275	0.447	0.000	1.000	0.299	0.458	0.253	0.435	0.278	0.448	0.272	0.445
home	0.134	0.341	0.000	1.000	0.124	0.330	0.143	0.350	0.138	0.345	0.130	0.337
nonprof	0.427	0.495	0.000	1.000	0.416	0.493	0.437	0.496	0.428	0.495	0.426	0.495
low parent control	0.741	0.217	0.000	1.000	0.738	0.217	0.744	0.217	0.741	0.219	0.742	0.215
maternal care	4.550	0.526	1.000	5.000	4.627	0.485	4.481	0.552	4.596	0.500	4.497	0.550
sample size	9728				4619		5109		5258		4470	

Smoking Initiation – Peers and Personality

	Smoke Dummy				Smoke Frequency			
	Homogeneous		ES	CO	Homogeneous		ES	CO
Endogenous effect	0.0922***	0.0842***			0.0947***	0.0911***		
	(0.0074)	(0.0077)			(0.0040)	(0.0057)		
low-to-low			0.1119***	0.0914***			0.1243***	0.0965***
			(0.0127)	(0.0156)			(0.0129)	(0.0164)
high-to-low			0.0614***	0.0866***			0.1257***	0.1025***
			(0.0188)	(0.0175)			(0.0224)	(0.0205)
low-to-high			0.0802***	0.0807***			0.0475***	0.0948***
			(0.0175)	(0.0176)			(0.0194)	(0.0209)
high-to-high			0.0700***	0.0802***			0.0610***	0.0907***
			(0.0171)	(0.0137)			(0.0188)	(0.0137)
Own effect								
emotional stability		−0.0268***	−0.0256***	−0.0276***		−0.0946***	−0.0580	−0.0945***
		(0.0064)	(0.0066)	(0.0063)		(0.0327)	(0.0341)	(0.0327)
conscientiousness		−0.0170***	−0.0170***	−0.0161***		−0.0477*	−0.0473*	−0.0451*
		(0.0051)	(0.0051)	(0.0053)		(0.0267)	(0.0267)	(0.0275)
extraversion		−0.0476***	−0.0477***	−0.0475***		−0.2775***	−0.2792***	−0.2776***
Contextual effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Group fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
σ_ϵ^2	0.1577	0.1551	0.1550	0.1551	4.2562	4.1819	4.1750	4.1817

Note: We report the posterior mean of each parameter and the standard deviation in the parenthesis. The asterisks ***(**,*) indicates that its 99% (95%, 90%) highest posterior density range does not cover zero. The MCMC sampling is running for 50,000 iterations with the first 5,000 iterations dropped for burn-in. ES: emotional stability; CO: conscientiousness. In the heterogeneous peer effect case, A-to-B denotes the peer effect that B receives from A.

Smoking Initiation – Peers and Personality

	Smoke Dummy			Smoke Frequency		
	Homogeneous	ES	CO	Homogeneous	ES	CO
Endogenous Effect						
	0.0374^{***}			0.0715^{***}		
	(0.0098)			(0.0091)		
low-to-low		0.0668^{***}	0.0442^{***}		0.0904^{***}	0.0672^{***}
		(0.0147)	(0.0174)		(0.0129)	(0.0164)
high-to-low		0.0247	0.0382^{**}		0.0854^{***}	0.0636^{***}
		(0.0191)	(0.0181)		(0.0210)	(0.0182)
low-to-high		0.0316	0.0357^{**}		0.0139	0.0682^{***}
		(0.0186)	(0.0180)		(0.0183)	(0.0196)
high-to-high		0.0360^{**}	0.0322^{**}		0.0304^{*}	0.0588^{***}
		(0.0183)	(0.0152)		(0.0186)	(0.0137)
Own and Contextual effect	Yes	Yes	Yes	Yes	Yes	Yes
Group fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
σ_u^2	0.1341	0.1344	0.1343	3.5772	3.6319	3.6609
Link formation						
constant	1.0340^{***}	1.0395^{***}	1.0333^{***}	1.0131^{***}	1.0280^{***}	1.0370^{***}
	(0.0789)	(0.0799)	(0.0788)	(0.0813)	(0.0780)	(0.0825)
grade	2.7500^{***}	2.7533^{***}	2.7553^{***}	2.7405^{***}	2.7457^{***}	2.7509^{***}
	(0.0380)	(0.0392)	(0.0390)	(0.0381)	(0.0385)	(0.0384)
sex	0.3511^{***}	0.3498^{***}	0.3506^{***}	0.3516^{***}	0.3508^{***}	0.3533^{***}
	(0.0342)	(0.0348)	(0.0335)	(0.0339)	(0.0347)	(0.0325)
race	1.1213^{***}	1.1216^{***}	1.1190^{***}	1.1222^{***}	1.1175^{***}	1.1225^{***}
	(0.0415)	(0.0413)	(0.0403)	(0.0394)	(0.0406)	(0.0414)
emotional stability	-0.0563^{*}	-0.0582^{*}	-0.0583^{*}	-0.0519^{*}	-0.0591^{*}	-0.0603^{*}
	(0.0291)	(0.0291)	(0.0301)	(0.0290)	(0.0294)	(0.0294)
conscientiousness	-0.0410	-0.0400	-0.0397	-0.0382	-0.0380	-0.0390
	(0.0255)	(0.0255)	(0.0256)	(0.0257)	(0.0254)	(0.0249)
extraversion	-0.2636^{***}	-0.2654^{***}	-0.2626^{***}	-0.2606^{***}	-0.2585^{***}	-0.2592^{***}
	(0.0277)	(0.0269)	(0.0274)	(0.0260)	(0.0275)	(0.0280)

Smoking Initiation – Peers and Personality

One may worry that our dichotomization of personality in strong/weak is arbitrary. We have additionally estimated a model with linear interaction terms between the peer effect and personality.

$$y_{i,g} = \lambda_1 \sum_{j \neq i} w_{ij,g} y_{j,g} s_{i,g} + \lambda_2 \sum_{j \neq i} w_{ij,g} y_{j,g} s_{j,g} + \lambda_3 \sum_{j \neq i} w_{ij,g} y_{j,g} s_{i,g} s_{j,g} + \cdots$$

	Smoke Dummy		Smoke Frequency	
	ES	CO	ES	CO
Endogenous Effect				
λ_1	-0.0168^{*}	-0.0099	-0.0524^{***}	-0.0170
	(0.0101)	(0.0094)	(0.0114)	(0.0101)
λ_2	-0.0261^{**}	-0.0069	-0.0204	-0.0074
	(0.0133)	(0.0123)	(0.0135)	(0.0116)
λ_3	0.0255^{**}	0.0147	0.0169	0.0281^{***}
	(0.0129)	(0.0105)	(0.0141)	(0.0104)
Own & Contextual effects	Yes	Yes	Yes	Yes
Endogenous link formation	Yes	Yes	Yes	Yes
Group fixed effects	Yes	Yes	Yes	Yes
σ_u^2	0.1298	0.1335	3.7602	3.5071

Smoking Initiation – Peers and Personality

- The results indicate that individuals who are emotionally unstable face larger peer effects compared to individuals who are emotionally stable.
- For conscientiousness we do not find a similar pattern.
- Although it seems extremely difficult to manipulate the composition of peer groups on basis of personality, the results do suggest that interventions aimed at groups of emotionally unstable individuals have the largest scope in reducing the uptake of smoking and other unhealthy behaviors in adolescence.

Gender and Racial Peer Effects with Endogenous Network Formation (Hsieh and Lin, 2017)

- Many existing studies focus only on the average magnitude of peer influences, without paying attention to the possible heterogeneous nature of peer effects along the dimensions of race and gender.
- To shed light on the debate on single-sex versus coeducational schooling, we need to know the mechanism of peer effects along the gender dimension.
- Another example is the debate on school segregation versus desegregation, which calls for the analysis on the mechanism of peer interactions within and across racial groups.

Gender and Racial Peer Effects with Endogenous Network Formation

Table: Gender Peer Effects on GPA

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Endogenous effects:					
λ_{ff}	0.140***	0.136***	0.136***	0.109***	0.102***
	(0.024)	(0.023)	(0.024)	(0.024)	(0.022)
λ_{fm}	0.025	0.025	0.019	0.013	0.002
	(0.020)	(0.020)	(0.020)	(0.020)	(0.019)
λ_{mm}	0.111***	0.109***	0.101***	0.077***	0.061**
	(0.024)	(0.025)	(0.024)	(0.025)	(0.024)
λ_{mf}	0.081***	0.078***	0.079***	0.067***	0.060***
	(0.021)	(0.020)	(0.021)	(0.020)	(0.020)

Gender and Racial Peer Effects with Endogenous Network Formation

Table: Gender Peer Effects on smoking

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Endogenous effects:					
λ_{ff}	0.479***	0.476***	0.476***	0.477***	0.487***
	(0.031)	(0.031)	(0.032)	(0.031)	(0.031)
λ_{fm}	0.198***	0.198***	0.198***	0.198***	0.203***
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
λ_{mm}	0.320***	0.322***	0.320***	0.318***	0.325***
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
λ_{mf}	0.233***	0.232***	0.232***	0.231***	0.238***
	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)

Gender and Racial Peer Effects with Endogenous Network Formation

Table: Racial Peer Effects on GPA

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Endogenous effects:					
λ_{ww}	0.253^{***}	0.255^{***}	0.243^{***}	0.194^{***}	0.175^{***}
	(0.028)	(0.029)	(0.026)	(0.028)	(0.029)
λ_{wb}	−0.025	−0.027	−0.029	−0.025	−0.033
	(0.026)	(0.027)	(0.027)	(0.027)	(0.027)
λ_{wo}	-0.036^{**}	-0.036^{**}	-0.034^{**}	-0.033^{**}	-0.029^{**}
	(0.017)	(0.018)	(0.018)	(0.017)	(0.017)
λ_{bb}	0.112^{***}	0.106^{***}	0.105^{**}	0.079^{*}	0.058
	(0.040)	(0.040)	(0.042)	(0.042)	(0.039)
λ_{bw}	0.124^{***}	0.118^{***}	0.120^{***}	0.098^{***}	0.087^{**}
	(0.037)	(0.037)	(0.037)	(0.036)	(0.037)
λ_{bo}	0.036	0.039	0.032	0.013	0.016
	(0.033)	(0.034)	(0.033)	(0.033)	(0.033)
λ_{oo}	0.022	0.021	0.019	0.010	0.006
	(0.032)	(0.030)	(0.029)	(0.029)	(0.030)
λ_{ow}	0.189^{***}	0.189^{***}	0.180^{***}	0.139^{***}	0.141^{***}
	(0.034)	(0.036)	(0.034)	(0.035)	(0.037)
λ_{ob}	−0.014	−0.014	−0.024	−0.029	−0.027
	(0.040)	(0.040)	(0.040)	(0.041)	(0.039)

Gender and Racial Peer Effects with Endogenous Network Formation (Hsieh and Lin, 2017)

Table: Racial Peer Effects on smoking

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Endogenous effects:					
λ_{ww}	0.507***	0.508***	0.513***	0.511***	0.520***
	(0.029)	(0.029)	(0.028)	(0.029)	(0.029)
λ_{wb}	−0.019	−0.017	−0.024	−0.028	−0.012
	(0.056)	(0.056)	(0.057)	(0.056)	(0.057)
λ_{wo}	0.126***	0.124***	0.128***	0.127***	0.134***
	(0.030)	(0.030)	(0.030)	(0.030)	(0.031)
λ_{bb}	0.078	0.069	0.083	0.097	0.055
	(0.102)	(0.102)	(0.100)	(0.100)	(0.100)
λ_{bw}	0.372***	0.364***	0.385***	0.382***	0.378***
	(0.087)	(0.085)	(0.085)	(0.085)	(0.086)
λ_{bo}	0.203	0.196	0.206	0.176	0.160
	(0.169)	(0.171)	(0.172)	(0.172)	(0.170)
λ_{oo}	0.330***	0.328***	0.332***	0.321***	0.337***
	(0.073)	(0.072)	(0.073)	(0.072)	(0.072)
λ_{ow}	0.365***	0.360***	0.380***	0.370***	0.373***
	(0.070)	(0.069)	(0.069)	(0.072)	(0.070)
λ_{ob}	−0.185*	−0.177	−0.173	−0.164	−0.175
	(0.110)	(0.109)	(0.108)	(0.110)	(0.111)

Reference

Ammermueller, Andreas and Jörn-Steffen Pischke (2009) "Peer effects in European primary schools: Evidence from the progress in international reading literacy study," *Journal of Labor Economics*, Vol. 27, pp. 315–348.

Anselin, Luc (1988) *Spatial Econometrics: Methods and Models*, Vol. 4: Springer Science & Business Media.

Arcidiacono, Peter and Sean Nicholson (2005) "Peer effects in medical school," *Journal of public Economics*, Vol. 89, pp. 327–350.

Ballester, Coralio, Antoni Calvó-Armengol, and Yves Zenou (2006) "Who's who in networks. Wanted: The key player," *Econometrica*, Vol. 74, pp. 1403–1417.

Bayer, Patrick, Randi Hjalmarsson, and David Pozen (2009) "Building criminal capital behind bars: Peer effects in juvenile corrections," *The Quarterly Journal of Economics*, Vol. 124, pp. 105–147.

Bertrand, Marianne, Erzo FP Luttmer, and Sendhil Mullainathan (2000)

"Network effects and welfare cultures," *The Quarterly Journal of Economics*, Vol. 115, pp. 1019–1055.

Blonigen, Bruce A, Ronald B Davies, Glen R Waddell, and Helen T Naughton (2007) "FDI in space: Spatial autoregressive relationships in foreign direct investment," *European Economic Review*, Vol. 51, pp. 1303–1325.

Bobonis, Gustavo J and Frederico Finan (2009) "Neighborhood peer effects in secondary school enrollment decisions," *The Review of Economics and Statistics*, Vol. 91, pp. 695–716.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin (2009) "Identification of peer effects through social networks," *Journal of Econometrics*, Vol. 150, pp. 41–55.

Calvó-Armengol, Antoni, Eleonora Patacchini, and Yves Zenou (2009) "Peer effects and social networks in education," *The Review of Economic Studies*, Vol. 76, pp. 1239–1267.

Carrell, Scott E, Richard L Fullerton, and James E West (2009) "Does your cohort matter? Measuring peer effects in college achievement," *Journal of Labor Economics*, Vol. 27, pp. 439–464.

Carrell, Scott E, Frederick V Malmstrom, and James E West (2008) "Peer effects in academic cheating," *Journal of Human Resources*, Vol. 43, pp. 173–207.

Cawley, John and Christopher J. Ruhm (2011) "The economics of risky health behaviors," in Thomas G. Mcguire Mark V. Pauly and Pedro P. Barros eds. *Handbook of Health Economics*, Vol. 2 of Handbook of Health Economics: Elsevier, pp. 95 – 199.

Chaloupka, Frank J and Kenneth E Warner (2000) "The economics of smoking," *Handbook of Health Economics*, Vol. 1, pp. 1539–1627.

Clark, Andrew E and Youenn Lohéac (2007) ""It wasn't me, it was them!" Social influence in risky behavior by adolescents," *Journal of Health Economics*, Vol. 26, pp. 763–784.

Cutler, David M and Edward L Glaeser (2010) "Social interactions and smoking," in *Research Findings in the Economics of Aging*: University of Chicago Press, pp. 123–141.

Dahl, Gordon B, Katrine V Løken, and Magne Mogstad (2014) "Peer effects in program participation," *The American Economic Review*, Vol. 104, pp. 2049–2074.

Danaei, Goodarz, Eric L Ding, Dariush Mozaffarian, Ben Taylor, Jürgen Rehm, Christopher JL Murray, and Majid Ezzati (2009) "The preventable causes of death in the United States: comparative risk assessment of dietary, lifestyle, and metabolic risk factors," *PLoS medicine*, Vol. 6, p. e1000058.

DeCicca, Philip, Donald Kenkel, and Alan Mathios (2002) "Putting out the fires: will higher taxes reduce the onset of youth smoking?" *Journal of political Economy*, Vol. 110, pp. 144–169.

DeCicca, Philip, Donald Kenkel, Alan Mathios, Yoon-Jeong Shin, and Jae-Young Lim (2008) "Youth smoking, cigarette prices, and anti-smoking sentiment," *Health Economics*, Vol. 17, pp. 733–749.

Digman, John M. (1990) "Personality structure: Emergence of the five-factor model," *Annual Review of Psychology*, Vol. 41, pp. 417–440.

Duflo, Esther and Emmanuel Saez (2003) "The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment," *The Quarterly Journal of Economics*, Vol. 118, pp. 815–842.

Fletcher, Jason M (2010a) "Social interactions and smoking: Evidence using multiple student cohorts, instrumental variables, and school fixed effects," *Health Economics*, Vol. 19, pp. 466–484.

——— (2010b) "Social interactions and smoking: Evidence using multiple student cohorts, instrumental variables, and school fixed effects," *Health Economics*, Vol. 19, pp. 466–484.

Gaviria, Alejandro and Steven Raphael (2001) "School-based peer effects and juvenile behavior," *The Review of Economics and Statistics*, Vol. 83, pp. 257–268.

Glaeser, Edward L, Bruce Sacerdote, and Jose A Scheinkman (1996) "Crime and social interactions," *The Quarterly Journal of Economics*, Vol. 111, pp. 507–548.

Guryan, Jonathan, Kory Kroft, and Matthew J Notowidigdo (2009) "Peer effects in the workplace: Evidence from random groupings in professional golf tournaments," *American Economic Journal: Applied Economics*, Vol. 1, pp. 34–68.

Hsieh, Chih-Sheng and Xu Lin (2017) "Gender and racial peer effects with endogenous network formation," *Regional Science and Urban Economics*, Vol. 67, pp. 135–147.

——— (2021) "Social interactions and social preferences in social networks," *Journal of Applied Econometrics*, Vol. 36, pp. 165–189.

Hsieh, Chih-Sheng and Hans Van Kippersluis (2018) "Smoking initiation: Peers and personality," *Quantitative Economics*, Vol. 9, pp. 825–863.

Kawaguchi, Daiji (2004) "Peer effects on substance use among American teenagers," *Journal of Population Economics*, Vol. 17, pp. 351–367.

Kelejian, Harry H and Ingmar R Prucha (1998) "A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances," *Journal of Real Estate Finance and Economics*, Vol. 17, pp. 99–121.

Lalive, Rafael and M Alejandra Cattaneo (2009) "Social interactions and schooling decisions," *The Review of Economics and Statistics*, Vol. 91, pp. 457–477.

Lee, Lung-Fei (2004) "Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models," *Econometrica*, Vol. 72, pp. 1899–1925.

Lee, Lung-fei (2007) "Identification and estimation of econometric models with group interactions, contextual factors and fixed effects," *Journal of Econometrics*, Vol. 140, pp. 333–374.

Lee, Lung-fei, Xiaodong Liu, and Xu Lin (2010) "Specification and estimation of social interaction models with network structures," *The Econometrics Journal*, Vol. 13, pp. 145–176.

Leider, Stephen, Markus M Möbius, Tanya Rosenblat, and Quoc-Anh Do (2009) "Directed altruism and enforced reciprocity in social networks," *The Quarterly Journal of Economics*, pp. 1815–1851.

Lin, Xu (2010) "Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables," *Journal of Labor Economics*, Vol. 28, pp. 825–860.

Lin, Xu and Lung-fei Lee (2010) "GMM estimation of spatial autoregressive models with unknown heteroskedasticity," *Journal of Econometrics*, Vol. 157, pp. 34–52.

Lin, Xu and Bruce A Weinberg (2014) "Unrequited friendship? How reciprocity mediates adolescent peer effects," *Regional Science and Urban Economics*, Vol. 48, pp. 144–153.

Liu, Xiaodong and Lung-fei Lee (2010) "GMM estimation of social interaction models with centrality," *Journal of Econometrics*, Vol. 159, pp. 99–115.

Lundborg, Petter (2006) "Having the wrong friends? Peer effects in adolescent substance use," *Journal of Health Economics*, Vol. 25, pp. 214–233.

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

Mas, Alexandre and Enrico Moretti (2009) "Peers at work," *The American Economic Review*, Vol. 99, pp. 112–145.

Matthews, Gerald, Ian J. Deary, and Martha C. Whiteman (2003) *Personality Traits*: Cambridge University Press.

Mokdad, A.H., J.S. Marks, D.F. Stroup, and J.L. Gerberding (2004) "Actual causes of death in the United States, 2000," *JAMA*, Vol. 291, pp. 1238–1245.

Nair, Harikesh S, Puneet Manchanda, and Tulikaa Bhatia (2010) "Asymmetric social interactions in physician prescription behavior: The role of opinion leaders," *Journal of Marketing Research*, Vol. 47, pp. 883–895.

Nakajima, Ryo (2007) "Measuring peer effects on youth smoking behaviour," *The Review of Economic Studies*, Vol. 74, pp. 897–935.

OECD (2013) "Health at a glance 2013: OECD indicators," *Organization of Economic Cooperation and Development*.

Ozdagli, Ali and Michael Weber (2017) "Monetary policy through production networks: Evidence from the stock market."

Powell, Lisa M, John A Tauras, and Hana Ross (2005) "The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behavior," *Journal of Health Economics*, Vol. 24, pp. 950–968.

Sacerdote, Bruce (2001) "Peer effects with random assignment: Results for Dartmouth roommates," *The Quarterly journal of economics*, Vol. 116, pp. 681–704.

Waldinger, Fabian (2011) "Peer effects in science: Evidence from the dismissal of scientists in Nazi Germany," *The Review of Economic Studies*, Vol. 79, pp. 838–861.

Zimmerman, David J (2003) "Peer effects in academic outcomes: Evidence from a natural experiment," *The Review of Economics and Statistics*, Vol. 85, pp. 9–23.