SOCIAL THRESHOLD REGRESSION

Antri Konstantinidi (University of Cyprus)

Andros Kourtellos (University of Cyprus)

Yiguo Sun (University of Guelph)

20th Conference on Research on Economic Theory and Econometrics
University of Cyprus

July 2022
The study of social influences on individual behavior has attracted a lot of interest in economics.

- Conditional on group membership, individual choices are driven by social factors.
- This relationship is determined by complementarities.
- Mechanisms: peer effects, social norms, social learning, social identity, etc.
• The typical econometric models of social interactions rely on linear models.

• Linear models assume that both individual effects and social effects are invariant over different groups of observations.

• They rule out interesting phenomena such as parameter heterogeneity, poverty traps, and persistent inequality.

• A growing body of empirical work suggests the presence of nonlinearity and heterogeneity of peer effects on test scores and health behavior (e.g., Carell et al. (2013), Sacerdote (2014), Booij, et al. (2016), Hsieh and Lin (2017), Hsieh and van Kippersluis (2018)).

• Estimators of linear models can be biased and inconsistent.
In this paper, we propose a threshold model of social interactions.

Why threshold-type nonlinearities?

- Certain theoretical models of social interactions have emphasized the existence of non-linearities and multiple equilibria.
  - Neighborhood effects; Benabou (1996) and Durlauf (1996).
  - Peer effects; Tincani (2018)
  - Spatial interactions; Kim et al. (2020);
- Distinct equilibria can be well-approximated by distinct linear approximations.
To fix some ideas consider a model of peer effects in academic achievement teenagers.

- Consider a network of $N$ individuals and assume one observes all ties within this network generating a $n \times n$ spatial (network) fixed/predetermined weight matrix $W_n$, which has zero diagonal elements. Let $w_{ij,n}$ be the $(i,j)$th element.

- Assume that ties are binary-valued and directed.

- For each individual we observe a $k \times 1$ vector of attributes and the outcome variable or level of an action $y_{i,n}$.

- Define the (weighted) average outcome of individual’s $i$ peers and the average of her peers’ attribute vectors:

$$
\tilde{y}_{n(i)} = \sum_{j \neq i} w_{ij}y_j \quad \text{and} \quad \tilde{x}_{n(i)} = \sum_{j \neq i} w_{ij}x_j
$$
Equilibrium behavior generates the linear social interactions model

\[ y_{i,n} = \kappa + \alpha \bar{y}_{n(i)} + x_{i,n}^' \beta + \bar{x}_{n(i)}^' \gamma + e_{i,n} \]

- The spatial autoregressive parameter of the spatially lagged dependent variable \( \bar{y}_{n(i)} \) is known as the endogenous effect, \( \alpha \).
- The coefficient of \( \bar{x}_{n(i)} \) is the contextual effect, \( \gamma \).
- This model defines a \( n \times 1 \) system of simultaneous equations.
- The LS estimator will be inconsistent due to the simulteneity bias. The error \( e_{i,n} \) is correlated with \( \bar{y}_{n(i)} \) since \( y_{i,n} \) is a component of each of the players' best response functions. This issue is known as the reflection problem (Manski (1993)).
- It is generically identified (Blume et al. (2015)).
For example, consider the question of peer effects in student academic achievement

- Peer effects in student academic achievement is an important mechanism (e.g., Sacerdote (2011) and Epple and Romano (2011)).

- These peer effect studies are typically based on the linear-in-means model.


- The evidence is mixed and the exact mechanisms are not well understood.
• Here we focus on the heterogeneous peer effects?

• One possible source of heterogeneity is personality.

• Are individuals with higher levels of stock of personality more susceptible to peer pressures?

• Could teens who are more emotionally stable be more susceptible to social influences?
Suppose the observations are organized into regimes according to emotional stability $z$.

Define the regime specific weights

$$ w_{ij,n}^- (\lambda) = w_{ij,n} I(z_{j,n} \leq \lambda) \quad \text{and} \quad w_{ij,n}^+ (\lambda) = w_{ij,n} I(z_{j,n} > \lambda) $$

$$ I(z_{i,n} \leq \gamma) = \begin{cases} 
1, & \text{when } z_{i,n} \leq \lambda \\
0, & \text{when } z_{i,n} > \lambda 
\end{cases} $$

This means emotional stability classifies individuals into two categories: if individual $i$ belongs to $z_{i,n} \leq \lambda$ is emotionally unstable and emotionally stable if $i$ belongs to $z_{i,n} > \lambda$. 
THRESHOLD PEER EFFECTS

We propose a threshold model with two types of endogenous effects

\[
y_{i,n} = \begin{cases} 
\alpha_1^- \sum_{j \neq i} w_{ij,n,2}^-(\lambda) y_{j,n} + \alpha_2^- \sum_{j \neq i} w_{ij,n,2}^+(\lambda) y_{j,n} + e_{i,n}, & z_{i,n} \leq \lambda \\
\alpha_1^+ \sum_{j \neq i} w_{ij,n,2}^-(\lambda) y_{j,n} + \alpha_2^+ \sum_{j \neq i} w_{ij,n,2}^+(\lambda) y_{j,n} + e_{i,n}, & z_{i,n} > \lambda 
\end{cases}
\]

- Heterogeneity of the peer effects depends nonlinearly on the degree of a personality skill such as emotional stability.
THRESHOLD PEER EFFECTS

We propose a threshold model with two types of endogenous effects

\[
y_{i,n} = \begin{cases} 
\alpha_1^- \sum_{j \neq i} w_{ij,n,2}^-(\lambda) y_{j,n} + \alpha_2^- \sum_{j \neq i} w_{ij,n,2}^+(\lambda) y_{j,n} + e_{i,n}, & z_{i,n} \leq \lambda \\
\alpha_1^+ \sum_{j \neq i} w_{ij,n,2}^-(\lambda) y_{j,n} + \alpha_2^+ \sum_{j \neq i} w_{ij,n,2}^+(\lambda) y_{j,n} + e_{i,n}, & z_{i,n} > \lambda 
\end{cases}
\]

- When the receiving individual is emotionally stable, we would expect the presence of positive peer effects if the peers are also emotionally stable (within-regime peer effects).
**Threshold Peer Effects**

We propose a threshold model with two types of endogenous effects

\[
y_{i,n} = \begin{cases} 
  \alpha_1^- \sum_{j \neq i} w_{ij,n,2}(\lambda) y_{j,n} + \alpha_2^- \sum_{j \neq i} w_{ij,n,2}(\lambda) y_{j,n} + e_{i,n}, & z_{i,n} \leq \lambda \\
  \alpha_1^+ \sum_{j \neq i} w_{ij,n,2}(\lambda) y_{j,n} + \alpha_2^+ \sum_{j \neq i} w_{ij,n,2}(\lambda) y_{j,n} + e_{i,n}, & z_{i,n} > \lambda 
\end{cases}
\]

- However, when her peers are emotionally unstable, it is plausible that the peer pressure is reduced or even becomes negative as she tries to distance herself (between-regime peer effects).
SPECIAL CASES

When \( \alpha_2^- = \alpha_1^+ = 0 \), we obtain a pure within-regime threshold social interaction model.

\[
y_{i,n} = \begin{cases} 
\alpha_1^- \sum_{j \neq i} w_{ij,n,2}(\lambda) y_{j,n} + e_{i,n}, & z_{i,n} \leq \lambda \\
\alpha_2^+ \sum_{j \neq i} w_{ij,n,2}(\lambda) y_{j,n} + e_{i,n}, & z_{i,n} > \lambda 
\end{cases}
\]

When \( \alpha_1^- = \alpha_2^- \) and \( \alpha_1^+ = \alpha_2^+ \) we obtain a threshold social interaction model where the spatial weight matrix is formed globally.

\[
y_{i,n} = \begin{cases} 
\alpha_1^- \sum_{j \neq i} w_{ij,n} y_{j,n} + e_{i,n}, & z_{i,n} \leq \lambda \\
\alpha_2^+ \sum_{j \neq i} w_{ij,n} y_{j,n} + e_{i,n}, & z_{i,n} > \lambda 
\end{cases}
\]

When \( \alpha \equiv \alpha_2^+ \), \( \delta_{\alpha_1} \equiv \alpha_2^- - \alpha_2^+ = 0 \), \( \delta_{\alpha_2} \equiv \alpha_1^+ - \alpha_2^+ = 0 \), \( \delta_{\alpha_3} \equiv (\alpha_1^- - \alpha_2^-) - (\alpha_1^+ - \alpha_2^+) = 0 \) we obtain the linear SAR model.

\[
y_{i,n} = \alpha \sum_{j \neq i} w_{ij,n} y_{j,n} + e_{i,n}
\]
Our objective is to

• Develop a two-step GMM estimation for the threshold and regression parameters

• Derive asymptotic theory and propose bootstrap inference.

• Assess the performance our methods using a Monte Carlo simulation.

• Provide an empirical application.
Econometric theory on threshold regression


Plethora of applications in macroeconomics, labor economics, finance, political economy.
Our work is related to several studies in spatial econometrics

- Cliff and Ord (1973)
INTRODUCTION

OUTLINE

1. Model
2. Estimation Method
3. Limiting Results
4. Bootstrap
5. Monte Carlo Evidence
6. Empirical Application
7. Conclusion and work-in-progress
Define the following regime specific variables

\[ x_{i,n}(\lambda) = x_{i,n}I(z_{i,n} \leq \lambda) \]

\[ w_{ij,n,1}(\lambda) = w_{ij,n}I(z_{i,n} \leq \lambda) \]

\[ w_{ij,n,2}(\lambda) = w_{ij,n}I(z_{j,n} \leq \lambda) \]

\[ w_{ij,n,3}(\lambda) = w_{ij,n}I(z_{i,n} \leq \lambda)I(z_{j,n} \leq \lambda) \]
Then we can express our model

\[
y_{i,n} = \alpha \sum_{j \neq i} w_{ij,n} y_{j,n} + \beta' x_{i,n} + \gamma' \sum_{j \neq i} w_{ij,n} x_{j,n} \\
+ \delta_{\alpha_1} \sum_{j \neq i} w_{ij,n,1} (\lambda) y_{j,n} + \delta_{\alpha_2} \sum_{j \neq i} w_{ij,n,2} (\lambda) y_{j,n} + \delta_{\alpha_3} \sum_{j \neq i} w_{ij,n,3} (\lambda) y_{j,n} \\
+ \delta'_{\beta} x_{i,n}(\lambda) \\
+ \delta'_{\gamma_1} \sum_{j \neq i} w_{ij,n,1} (\lambda) x_{j,n} + \delta'_{\gamma_2} \sum_{j \neq i} w_{ij,n,2} (\lambda) x_{j,n} + \delta'_{\gamma_3} \sum_{j \neq i} w_{ij,n,3} (\lambda) x_{j,n} \\
+ e_{i,n},
\]

(1)
Or, in a matrix notation

\[ Y_n = \alpha_0 W_n Y_n + X_n(W_n)\theta_{\beta_0} + Y_n(W_n, \lambda_0)\delta_{\alpha_0} + X_n(W_n, \lambda_0)\delta_{\theta_{\beta_0}} + e_n, \]

where

\[ X_n(W_n) = [X_n, W_nX_n], \]
\[ Y_n(W_n, \lambda_0) = [W_n^{-1}(\lambda_0)Y_n, W_n^{-2}(\lambda_0)Y_n, W_n^{-3}(\lambda_0)Y_n], \]
\[ X_n(W_n, \lambda_0) = [X_n(\lambda_0), W_n^{-1}(\lambda_0)X_n, W_n^{-2}(\lambda_0)X_n, W_n^{-3}(\lambda_0)X_n], \]

\[ \theta_{\beta_0} = (\beta_0', \gamma_0')' : (2k) \times 1 \]
\[ \delta_{\alpha_00} = (\delta_{\alpha_{10}}, \delta_{\alpha_{20}}, \delta_{\alpha_{30}})' : 3 \times 1 \]
\[ \delta_{\theta_{\beta_0}} = (\delta_{\beta}', \delta_{\gamma_{10}}', \delta_{\gamma_{20}}', \delta_{\gamma_{30}}')' : (4k) \times 1. \]
Let
\[ S_n(\theta_y, \lambda) = I_n - \alpha W_n - \delta_{\alpha_1} W_{n,1}^- (\lambda) - \delta_{\alpha_2} W_{n,2}^- (\lambda) - \delta_{\alpha_3} W_{n,3}^- (\lambda), \]
where \( \theta_y = (\alpha, \delta'_\alpha)' \).

Then, the reduced form is given by
\[ Y_n = S_n^{-1} X_n(W_n) \theta_{\beta_0} + S_n^{-1} X_n(W_n, \lambda_0) \delta_{\theta_{\beta_0}} + S_n^{-1} e_n, \quad (3) \]
where \( \theta_y = (\alpha, \delta'_\alpha)' \) and \( S_n = S_n(\theta_{y_0}, \lambda_0) \).

- \( E [(A_n Y_n)' e_n] = \text{tr} [(A_n S_n^{-1})' E (e_n e_n')] \neq 0 \) for \( A_n = W_n, W_{n,\lambda_0}, W_n(\lambda_0), \) and \( W_n^- (\lambda_0) \) where \( \text{tr}(\cdot) \) is the trace operator.

- Expansion of \( S_n(\theta_y, \lambda)^{-1} \) implies that not only powers of the weight matrices but also their interactions are possible instruments.
Assumption 1.

\[(1.1) \quad \rho \left( \alpha_0 W_n + \delta_{\alpha_1,0} W_{n,1}(\lambda_0) + \delta_{\alpha_2,0} W_{n,2}(\lambda_0) + \delta_{\alpha_3,0} W_{n,3}(\lambda_0) \right) < 1, \]
where \( \rho(A) \) is the spectral radius of matrix \( A \), and the spatial weight matrix \( W_n \) and \( S_n^{-1} \) have finite and none zero row- and column-sum norm.

\[(1.2) \quad \text{The parameter vector } \theta_0 = (\alpha_0, \delta_{\alpha_0}, \theta_{\beta_0}, \delta_{\theta_\beta_0}, \lambda_0)' \text{ is an interior point of a compact set } \Theta = \Theta^* \times \Lambda \text{ in the Euclidean Space } R^{k_\theta}, \text{ where } k_\theta = 6k + 5 \text{ and } \Lambda = [\lambda, \bar{\lambda}] \text{. Also, for } j = 1, 2, 3, \]
\[\delta_{\beta_0} = c_\beta n^{-a_0}, \delta_{\alpha_{j,0}} = c_{\alpha_{j}} n^{-b_{j}}, \text{ and } \delta_{\gamma_{j,0}} = c_{\gamma_{j}} n^{-a_{j+1}}, \text{ where at least one of } c_\beta, c_{\alpha_{j}}, c_{\gamma_{j}} \text{ is non-zero, } 0 \leq a_0 < \frac{1}{2}, 0 \leq a_{j} < \frac{1}{2}, \text{ and } 0 \leq b_{j} < \frac{1}{2}. \]
Assumption 2

(2.1) \( \{e_{i,n}\} \) is a sequence of independent errors with zero mean and variance \( \sigma^2_i \), independent of \( \{z_{i,n}\} \) and \( E(e_{i,n}|x_{i,n}) = 0 \) for all \( i \), where \( \max_{1 \leq i \leq n} E|e_{i,n}|^{4+\eta} < M < \infty \) for some \( \eta > 0 \).

(2.2) The threshold variable \( z_{i,n} \) is i.i.d. with a continuous and bounded density, \( f(\cdot) \), such that \( f(\lambda_0) > 0 \).

(2.3) \( \{x_{i,n}\} \) is an independent sequence, uniformly bounded, and independent of \( \{e_{i,n}\} \).

(2.4) We consider a linear transformation of the moment equations \( \mathcal{A}_ng_n(\theta) \), where \( g_n(\theta) \) is a set of moments described below in equation (4) and \( \mathcal{A}_n \) is a matrix with a full row rank greater than or equal to \( k_\theta \) and converges to a constant full row rank matrix \( \mathcal{A}_0 \). Also, the elements of \( \mathcal{A}_n \) are uniformly bounded in absolute values.
INSTRUMENTS

Let $Q_n$ be an $n \times k_Q$ matrix of initial instrumental variables (IV) with $k_Q \geq k_\theta$. For example, we can use $X_n$, $W_n X_n$, $W_n^2 X_n$, $W_n^3 X_n$, . . . after removing linearly dependent components.

In certain cases, these instruments have an interpretation.

For example, $W_n^2 X_n$ denotes the average of your friends’ friends’ average attributes (Bramoulle et al. (2009)).

The moment conditions corresponding to the orthogonality conditions of $Q_n$ and $e_n$ are

$$E(Q_n' e_n) = 0.$$
Following Kelejian and Prucha (1998, 2010) we also employ quadratic moments based on a class of constant $n \times n$ matrices denoted by $\mathcal{P}_n = \{ P_n : \text{diag} (P_n) = 0 \}$.

The corresponding quadratic moments are defined as follows:

$$E \left[ (P_{jn} e_n)' e_n \right] = 0,$$

where matrices $P_{jn}$ are selected from $\mathcal{P}_n$, for $j = 1, \ldots, m$. 
Assumption 3.

The matrices in $P_n$ are uniformly bounded in both row and column sums in absolute values. The elements of $Q_n$ are uniformly bounded. Both $X_n$ and $Q_n$ have full rank.
The set of moment equations $E(g_n(\theta)) = 0$ is given by a $(k_Q + m) \times 1$ vector where the moment functions $g_n(\theta)$ are given by

$$
g_n(\theta) = 
\begin{bmatrix}
e_n(\theta)' P_{1n} e_n(\theta) \\
\vdots \\
e_n(\theta)' P_{mn} e_n(\theta) \\
Q_n' e_n(\theta),
\end{bmatrix}
$$

with $e_n(\theta) = S_n(\theta_y, \lambda) Y_n - X_n(W_n) \theta_\beta - X_n(W_n, \lambda) \delta_\theta$. 

**MOMENT EQUATIONS**
IDENTIFICATION

Following Hansen (1992), the parameter $\theta_0$ is identified as long as

$$A_0 \lim_{n \to \infty} n^{-1} E(g_n(\theta)) = 0$$

has a unique root at $\theta_0 \in \Theta$, where $A_0$ is a constant full rank matrix.

First, we compute

$$E(g_n(\theta)) =
\begin{bmatrix}
E(d_n(\theta)'P_{1n}d_n(\theta)) + \text{tr} \{ \Gamma_n E(A_n(\theta_y, \lambda)'P_{1n}^s) \} + \text{tr} \{ \Gamma_n E(A_n(\theta_y, \lambda)'P_{1n}A_n(\theta_y, \lambda) \} \\
\vdots \\
E(d_n(\theta)'P_{mn}d_n(\theta)) + \text{tr} \{ \Gamma_n E(A_n(\theta_y, \lambda)'P_{mn}^s) \} + \text{tr} \{ \Gamma_n E(A_n(\theta_y, \lambda)'P_{mn}A_n(\theta_y, \lambda) \} \\
E(Q_n'P_{1n}d_n(\theta))
\end{bmatrix},$$

where $P_{jn} = P_{jn} + P_{jn}'$ for $j = 1, 2, \ldots, m$ and $\Gamma_n = \text{diag}(\sigma_1^2, \ldots, \sigma_n^2)$. 
Applying Taylor expansion gives

\[ E(g_n(\theta)) = E(g_n(\theta_0)) + \frac{\partial E(g_n(\bar{\theta}))}{\partial \theta'} (\theta - \theta_0), \]

where \( \bar{\theta} \) lies between \( \theta \) and \( \theta_0 \). Evidently, there will be a unique \( \theta_0 \) satisfying

\[ \mathcal{A}_0 \lim_{n \to \infty} n^{-1} E(g_n(\theta)) = 0 \text{ if } \frac{\partial E(g_n(\theta))}{\partial \theta'} \text{ has a full rank } k_\theta \text{ over } \theta \in \Theta. \]

However, the above rank condition can fail in non-trivial cases. Nevertheless, we show that identification is still possible using the quadratic moments.

To explore the identification for this special case, it is therefore necessary to first discuss the identification of the pure threshold spatial autoregressive model which contains no exogenous regressors, \( x_{i,n} \), next.
Proposition 1

If \( \theta_{x,0} = 0 \), model (3) becomes a pure threshold spatial autoregressive model. The necessary conditions for the identification of \((\theta'_y, \lambda_0)'\) include

(I) \( z_{i,n} \) is independent but not identically distributed with different cdf \( F_i(z) \) and pdf \( f_i(z) \) across \( i \);

(II) \( \text{tr} (\Gamma_nP^s_{jn}G_n), \text{tr} (\Gamma_nP^s_{jn}D_F(\lambda_0)G_n), \text{tr} (\Gamma_nP^s_{jn}W_nD_F(\lambda_0)S_n^{-1}), \) and \( \text{tr} [\Gamma_nP^s_{jn}D_F(\lambda_0)W_nD_F(\lambda_0)S_n^{-1}] \) are linearly independent;

(III) \( \text{tr} (\Gamma_nP^s_{jn}D_f(\lambda_0)G_n), \text{tr} (\Gamma_nP^s_{jn}W_nD_f(\lambda_0)S_n^{-1}), \) and \( \text{tr} [\Gamma_nP^s_{jn} (D_F(\lambda_0)W_nD_f(\lambda_0) + D_f(\lambda_0)W_nD_F(\lambda_0)) S_n^{-1}] \) are linearly independent, where we denote an \( n \times n \) diagonal matrix \( D_m(\lambda) = \text{diag}(m_1(\lambda), \ldots, m_n(\lambda)) \).
Assumption 4

(4.1) \( \lim_{n \to \infty} n^{-1} G(Q_n, \theta) \) has a full rank over \( \theta \in \Theta \) and \( F(\lambda_0) < 1 \), where \( F(\lambda) \) is the cdf of \( z_{i,n} \).

(4.2) (i) \( \lim_{n \to \infty} n^{-1} E(Q'_{n}X^*_n) \) has a full rank 6k and is linearly independent of \( \lim_{n \to \infty} n^{-1} G_\lambda(\delta_{\alpha_0}, \delta_{\beta_0}, \lambda_0)' \) where we denote

\[
D_n = \begin{bmatrix}
\text{tr} \left[ \Gamma_n E \left(P^s_{1n} G_n \right) \right] & \ldots & \text{tr} \left[ \Gamma_n E \left(P^s_{mn} G_n \right) \right] \\
\text{tr} \left\{ \Gamma_n E \left[P^s_{1n} G_{n,2}^-(\lambda_0) \right] \right\} & \ldots & \text{tr} \left\{ \Gamma_n E \left[P^s_{mn} G_{n,2}^-(\lambda_0) \right] \right\} \\
\text{tr} \left\{ \Gamma_n E \left[P^s_{1n} G_{n,1}^-(\lambda_0) \right] \right\} & \ldots & \text{tr} \left\{ \Gamma_n E \left[P^s_{mn} G_{n,1}^-(\lambda_0) \right] \right\} \\
\text{tr} \left\{ \Gamma_n E \left[P^s_{1n} G_{n,3}^-(\lambda_0) \right] \right\} & \ldots & \text{tr} \left( \Gamma_n E \left\{ \Gamma_n E \left[P^s_{mn} G_{n,3}^-(\lambda_0) \right] \right\} \right)
\end{bmatrix}
\]

(ii) conditions (i)-(iii) in Proposition 1 hold.

• We consider both linear and quadratic moments.
First step GMM estimator

Consider the criterion

$$J_n(\theta) = g_n(\theta)'A_n'An g_n(\theta)$$

Given an initial weight matrix $A_n'An$, the GMM estimator of $\theta$ is given by

$$\hat{\theta} = \arg \min_{\theta \in \Theta} J_n(\theta)$$

The estimation is based on concentration.
Conditional on $\lambda$, we can obtain the GMM estimator of $\theta_*$ as

$$\hat{\theta}_*(\lambda) = \arg \min_{\theta_*} J_n(\theta^*, \lambda),$$

where $\theta^* = (\alpha, \delta'_\alpha, \theta'_\beta, \delta'_\theta\beta)'$ and $J_n(\theta^*, \lambda) = g_n(\theta^*, \lambda)'A_n'A_ng_n(\theta^*, \lambda)$.

Then since the threshold parameter takes values in a bounded set of values

$\Lambda = [\lambda, \bar{\lambda}]$

$$\hat{\lambda} = \arg \min_{\lambda \in \Lambda} J_n(\hat{\theta}^*(\lambda), \lambda) \quad \text{and} \quad \hat{\theta}^* = \arg \min_{\theta^*} J_n(\theta^*, \hat{\lambda})$$
Second-Step GMM estimator

Given the first-step GMM estimator we improve the efficiency of our initial estimator by considering regime specific moment functions.

\[
\hat{g}_n(\theta) = \begin{bmatrix}
e_n(\theta)' \hat{P}_{1n} e_n(\theta) \\
\vdots \\
e_n(\theta)' \hat{P}_{4n} e_n(\theta) \\
\hat{Q}'_n e_n(\theta)
\end{bmatrix}
\]
\[
\hat{P}_{1n} = W_nS_n^{-1}(\hat{\theta}_y, \hat{\lambda}) - \text{diag} \left( W_nS_n^{-1}(\hat{\theta}_y, \hat{\lambda}) \right)
\]
\[
\hat{P}_{2n} = W_{n,1}(\hat{\lambda})S_n^{-1}(\hat{\theta}_y, \hat{\lambda}) - \text{diag} \left( W_{n,1}(\hat{\lambda})S_n^{-1}(\hat{\theta}_y, \hat{\lambda}) \right)
\]
\[
\hat{P}_{3n} = W_{n,2}(\hat{\lambda})S_n^{-1}(\hat{\theta}_y, \hat{\lambda}) - \text{diag} \left( W_{n,2}(\hat{\lambda})S_n^{-1}(\hat{\theta}_y, \hat{\lambda}) \right)
\]
\[
\hat{P}_{4n} = W_{n,3}(\hat{\lambda})S_n^{-1}(\hat{\theta}_y, \hat{\lambda}) - \text{diag} \left( W_{n,3}(\hat{\lambda})S_n^{-1}(\hat{\theta}_y, \hat{\lambda}) \right)
\]

and

\[
\hat{Q}_n = \left[ W_nS_n^{-1}(\hat{\theta}_y, \hat{\lambda})[X_n(W_n)\hat{\theta}_\beta, X_n(W_n, \hat{\lambda})\hat{\delta}_{\theta\beta}],
\right.
\]
\[
W_{n,1}(\hat{\lambda})S_n^{-1}(\hat{\theta}_y, \hat{\lambda})[X_n(W_n)\hat{\theta}_\beta, X_n(W_n, \hat{\lambda})\hat{\delta}_{\theta\beta}],
\]
\[
W_{n,2}(\hat{\lambda})S_n^{-1}(\hat{\theta}_y, \hat{\lambda})[X_n(W_n)\hat{\theta}_\beta, X_n(W_n, \hat{\lambda})\hat{\delta}_{\theta\beta}],
\]
\[
W_{n,3}(\hat{\lambda})S_n^{-1}(\hat{\theta}_y, \hat{\lambda})[X_n(W_n)\hat{\theta}_\beta, X_n(W_n, \hat{\lambda})\hat{\delta}_{\theta\beta}],
\]
\[
X_n(W_n), X_n(W_n, \hat{\lambda})]\]
In addition, given the initial consistent estimator calculated in the first step, we can construct a matrix \( \hat{\Omega}_n \) to estimate

\[
\Omega_n = E \left[ g_n(\theta_0)g_n(\theta_0)' \right],
\]

where \( g_n(\theta_0) \) equals \( \hat{g}_n(\theta) \) with \( \hat{\theta} \) replaced with \( \theta_0 \) and the mathematical expression of \( \Omega_n \) as we show next section with \( m = 4 \), we define

\[
\hat{\Omega}_n = \begin{bmatrix}
\text{tr} \left( \hat{\Gamma}_n \hat{P}_{1n} \left( \hat{\Gamma}_n \hat{P}_{1n} \right)^s \right) & \ldots & \text{tr} \left( \hat{\Gamma}_n \hat{P}_{1n} \left( \hat{\Gamma}_n \hat{P}_{4n} \right)^s \right) & 0_{kQ}' \\
\vdots & \ddots & \ddots & \vdots \\
\text{tr} \left( \hat{\Gamma}_n \hat{P}_{4n} \left( \hat{\Gamma}_n \hat{P}_{1n} \right)^s \right) & \ldots & \text{tr} \left( \hat{\Gamma}_n \hat{P}_{4n} \left( \hat{\Gamma}_n \hat{P}_{4n} \right)^s \right) & 0_{kQ}' \\
0_{kQ} & \ldots & 0_{kQ} & \hat{Q}_n \hat{\Gamma}_n \hat{Q}_n
\end{bmatrix}
\]

(5)

where \( \hat{\Gamma}_n = \text{diag} \left\{ \hat{e}_{1,n}^2, \ldots, \hat{e}_{n,n}^2 \right\} \) is an \( n \times n \) diagonal matrix and \( \hat{e}_i \) is the \( i \)th element of the estimated residual \( \hat{e}_n = S_n(\hat{\theta}_y, \hat{\lambda}) Y_n - X_n(W_n)\hat{\beta} - X_n(W_n, \hat{\lambda})\hat{\delta}_{\theta \beta} \).
Then, the second-step GMM estimator of $\theta$ is given by

$$\tilde{\theta} = \arg \min_{\theta \in \Theta} \hat{g}_n(\theta)'\hat{\Omega}_n^{-1}\hat{g}_n(\theta)$$

but is evaluated through the concentration of the objective function in a similar way as in the first step.
**Limiting Results**

Proposition 2.

Under Assumptions 1-4, the identification condition holds which implies that $A_0 \lim_{n \to \infty} n^{-1} E(g_n(\theta)) = 0$ has a unique root at $\theta_0 \in \Theta$, and the GMM estimator is consistent.
Theorem 1.

Under Assumptions 1-4, we have

$$\sqrt{n}H_n^{-1} \begin{pmatrix} \hat{\theta}^* - \theta_0^* \\ \hat{\lambda} - \lambda_0 \end{pmatrix} = \begin{bmatrix} \sqrt{n} (\hat{\theta}^* - \theta_0^*) \\ n^{1-a} \left( \hat{\lambda} - \lambda_0 \right) \end{bmatrix} \overset{d}{\rightarrow} N(0, \Sigma),$$

(6)

where $H_n = \text{diag}(\iota_{6k+4}, na)$, $a = \min(a_0, a_1, a_2, a_3, b_1, b_2, b_3)$,
\[ \Sigma = \lim_{n \to \infty} \left[ \frac{1}{n} H_n \Lambda_n' A_n' A_n \frac{1}{n} \Lambda_n H_n \right]^{-1} \left[ \frac{1}{n} H_n \Lambda_n' A_n' A_n \frac{1}{n} \Omega_n A_n' A_n \frac{1}{n} \Lambda_n H_n \right] \left[ \frac{1}{n} H_n \Lambda_n' A_n' A_n \frac{1}{n} \Lambda_n H_n \right]^{-1}, \]

\[ \Omega_n = E \left[ g_n(\theta_0) g_n(\theta_0)' \right] \]

\[
\begin{bmatrix}
\text{tr} \left( \Gamma_n P_{1n} (\Gamma_n P_{1n})^s \right) & \ldots & \text{tr} \left( \Gamma_n P_{1n} (\Gamma_n P_{mn})^s \right) & 0'_{k_Q} \\
\vdots & \ddots & \vdots & \ddots \\
\text{tr} \left( \Gamma_n P_{mn} (\Gamma_n P_{1n})^s \right) & \ldots & \text{tr} \left( \Gamma_n P_{mn} (\Gamma_n P_{mn})^s \right) & 0'_{k_Q} \\
0'_{k_Q} & \ldots & 0'_{k_Q} & E(Q_n' \Gamma_n Q_n)
\end{bmatrix},
\]

(7)

and \( \Lambda_n = -\partial E(g_n(\theta_0))/\partial \theta' \) and \( \lim_{n \to \infty} n^{-1} \Lambda_n H_n \) has full column rank.
\[
\Lambda_n = \\
\begin{bmatrix}
\text{tr}(\Gamma_n P_{1n}^s G_n) & \text{tr}\left\{\Gamma_n E\left[P_{1n}^s G_{n,1}^-(\lambda_0)\right]\right\} & \text{tr}\left\{\Gamma_n E\left[P_{1n}^s G_{n,2}^-(\lambda_0)\right]\right\} & \text{tr}\left\{\Gamma_n E\left[P_{1n}^s G_{n,3}^-(\lambda_0)\right]\right\} & 0_{6k} & \sum_{j=1}^{3} \delta_{\alpha_j} \text{tr}\left\{\Gamma_n P_{1n}^s \frac{\partial E(G_{n1}^-(\lambda_0))}{\partial \lambda}\right\} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\text{tr}(\Gamma_n P_{mn}^s G_n) & \text{tr}\left\{\Gamma_n E\left[P_{mn}^s G_{n,1}^-(\lambda_0)\right]\right\} & \text{tr}\left\{\Gamma_n E\left[P_{mn}^s G_{n,2}^-(\lambda_0)\right]\right\} & \text{tr}\left\{\Gamma_n E\left[P_{mn}^s G_{n,3}^-(\lambda_0)\right]\right\} & 0_{6k} & \sum_{j=1}^{3} \delta_{\alpha_j} \text{tr}\left\{\Gamma_n P_{mn}^s \frac{\partial E(G_{n1}^-(\lambda_0))}{\partial \lambda}\right\} \\
E(Q_n^s G_n X_n^s \theta_{x,0}) & E(Q_n^s G_n X_n^s \theta_{x,0}^2) & E(Q_n^s G_n X_n^s \theta_{x,0}) & E(Q_n^s G_n X_n^s \theta_{x,0}) & E(Q_n^s G_n X_n^s \theta_{x,0}) & E(Q_n^s G_n \theta_{x,0}) & G_n(\delta_{\alpha_0}, \delta_{\theta_{0}}, \lambda_0)
\end{bmatrix}
\]

(8)

- As \(n^{-1} G_\lambda(\delta_{\alpha_0}, \delta_{\theta_{0}}, \lambda_0) = O(n^{-a})\) and the first \(m\) elements in the last column of \(n^{-1} \Lambda_n\) are of order \(O(n^{-b_{min}})\), where \(b_{min} = \min(b_1, b_2, b_3)\), multiplying \(n^{-1} \Lambda_n\) by \(H_n\) ensures that all the elements in the last column of \(n^{-1} \Lambda_n H_n\) are of order \(O_p(1)\).

- The convergence rate of \(\hat{\lambda}\) will be \(n^{-\frac{1}{2} + b_{min}}\), if there are no exogenous threshold effects.

- The convergence rate becomes \(n^{-\frac{1}{2} + a_{min}}\) if there are no endogenous threshold effects, where \(a_{min} = \min(a_0, a_1, a_2, a_3)\).
Remarks

- Our model nests both the limiting results obtained by Lee (2007) for the linear SDM and Seo and Shin (2016) for the threshold regression model.

- This theorem says that our estimator follows the normal distribution asymptotically regardless of whether \( a = 0 \) or \( 0 < a < \frac{1}{2} \).

- For the estimation of the threshold parameter, \( \lambda_0 \), the GMM estimator converges at a slower speed than what one would expect from a least-square based estimator, e.g., Chan (1993) and Hansen (2000). The estimator for the threshold parameter \( \lambda_0 \) converges at the superconsistent rate \( n^{1-a} \) for the least square method, which converges at a faster rate than the GMM convergence rate \( n^{1/2-a} \). The GMM estimator of \( \lambda_0 \) enjoys the convenience of constructing the classic t-statistic in making inference of \( \lambda = \lambda_0 \).
• While the GMM estimator in Theorem 1 allows for an arbitrary unknown heteroskedasticity, it is not efficient.

• The best selection of $P_{jn}$ and $Q_n$ is not available because the best selection involves matrix $\Gamma_n$, which is generally unknown. Thus an optimal GMM is not feasible.

• In this sense, the second-step estimator is quasi-optimal and using a consistent estimator of $\Omega_n^{-1}$ we obtain a feasible quasi-optimal GMM estimator. The following proposition is used to support the use of $\hat{\Omega}_n$ in step 2.
Proposition 3

Under Assumptions 1-4,

\[ n^{-1}(\hat{\Omega}_n - \Omega_n) = o_P(1), \]

where \( n^{-1}\hat{\Omega}_n \) is the estimator of \( n^{-1}\Omega_n \) with \( \theta_0 \) replaced by a consistent estimator \( \hat{\theta} \), \( \Gamma_n \) by \( \hat{\Gamma}_n \), \( P_{jn} \) by \( \hat{P}_{jn} \) for \( j = 1, \ldots, 4 \), and \( n^{-1}E(Q'_n\Gamma_nQ_n) \) by \( n^{-1}\hat{Q}'_n\hat{\Gamma}_n\hat{Q}_n \).
Assumption 5

The matrix, \( \lim_{n \to \infty} n^{-1} \Omega_n \), exists and is nonsingular.
Theorem 2

Under Assumptions 1-5, we obtain the quasi-optimal GMM estimator from minimizing $\hat{g}_{n(\theta)'}\hat{\Omega}_n^{-1}\hat{g}_n(\theta)$, which has the limiting distribution

$$\begin{bmatrix} \sqrt{n}(\tilde{\theta}^* - \theta^*_0) \\
\frac{1}{n^2-a}(\tilde{\lambda} - \lambda_0) \end{bmatrix} \overset{d}{\rightarrow} N(0, \Sigma^*), \quad (9)$$

with $\Sigma^* = (\lim_{n \to \infty} n^{-1} H_n \Lambda_n' \Omega_n^{-1} \Lambda_n H_n)^{-1}$. 

\[ \text{Theorem 2} \]

\[ \text{Under Assumptions 1-5, we obtain the quasi-optimal GMM estimator from minimizing } \hat{g}_{n(\theta)'}\hat{\Omega}_n^{-1}\hat{g}_n(\theta), \text{ which has the limiting distribution} \]

\[ \begin{bmatrix} \sqrt{n}(\tilde{\theta}^* - \theta^*_0) \\
\frac{1}{n^2-a}(\tilde{\lambda} - \lambda_0) \end{bmatrix} \overset{d}{\rightarrow} N(0, \Sigma^*), \quad (9) \]

\[ \text{with } \Sigma^* = (\lim_{n \to \infty} n^{-1} H_n \Lambda_n' \Omega_n^{-1} \Lambda_n H_n)^{-1}. \]
Remarks

• To make statistical inference on \( \hat{\theta} \), we need to find consistent estimator for \( \Sigma^* \).

• However, \( \Lambda_n \) contains terms like \( f(\lambda_0) \) and \( E \left( P_{jn}^s G_n | \lambda_0 \right) \),
  \( E \left( P_{jn}^s G_{n,-}^0 | \lambda_0 \right) \), and \( E \left( P_{jn}^s G_n (\lambda_0) | \lambda_0 \right) \), and these unknown terms have to be estimated nonparametrically by kernel method, which makes the estimate of \( \Lambda_n \) depend on a bandwidth parameter value.

• The overall accuracy of an estimate of \( \Lambda_n \) may not be robust to the choice of the bandwidth parameter.

• Therefore, bootstrap-based standard error for \( \hat{\theta} \), although time consuming, is recommended.
(1) Using the second-step estimator $\tilde{\theta}$ defined in equation (38), we obtain the demeaned residuals $\tilde{e}_n = (I_n - n^{-1} \omega_n \omega_n^\prime) \tilde{e}_n(\tilde{\theta})$, where $\tilde{e}_n(\tilde{\theta}) = S_n(\tilde{\theta}_y, \tilde{\lambda}) Y_n - X_n (W_n) \tilde{\theta}_\beta - X_n (W_n, \tilde{\lambda}) \delta_{\theta_\beta}$ and $\omega_n$ is the $n \times 1$ vector of ones.

(II) The bootstrap sample of residuals $e_n^\# = (e_{1,n}^\#, \ldots, e_{n,n}^\#)^\prime$ is generated from a two-point distribution $P (e_{i,n}^\# = c_1 \hat{e}_{i,n}) = r$ and $P (e_{i,n}^\# = c_2 \hat{e}_{i,n}) = 1 - r$, where $\hat{e}_{i,n}$ is the $i$th demeaned residual, $c_1 = (1 + \sqrt{5})/2$, $c_2 = (1 - \sqrt{5}/2)$, and $r = (\sqrt{5} - 1)/(2\sqrt{5})$. Then, the bootstrap residuals are used to generate the bootstrap dependent variable $Y_n^\# = S_n(\tilde{\theta}_y, \tilde{\lambda})^{-1} (X_n (W_n) \tilde{\theta}_\beta + X_n (W_n, \tilde{\lambda}) \delta_{\theta_\beta} + e_n^\#).

(III) Using the bootstrap sample $\{y_{i,n}^\#, x_{i,n}, z_{i,n}\}, i=1, \ldots, n$ and the weight matrix $W_n$ and applying the same estimation as for $\tilde{\theta}$, we obtain the second-step GMM estimator $\tilde{\theta}^\#$. 

Bootstrap Confidence Intervals

Bootstrap Procedure
Theorem 3

Under Assumptions 1-5, we obtain

$$\sqrt{n}H_n^{-1} \left( \tilde{\theta}^\# - \hat{\theta} \right) \xrightarrow{d} N(0, \Sigma^*) ,$$

where $\Sigma^*$ is defined in Theorem 2.


- It implies that $\sqrt{n}H_n^{-1} \left( \tilde{\theta}^\# - \hat{\theta} \right)$ converges weakly to the limiting distribution of $\sqrt{n}H_n^{-1} \left( \tilde{\theta} - \theta_0 \right)$ in probability, so that the percentiles of the bootstrap distribution can be used to construct confidence intervals with correct asymptotic coverage probabilities.
Next we employ a bootstrap sup-LR type $supD$ test for testing the null hypothesis

$$H_0: \delta_0 = 0 \text{ vs. } H_1: \delta_0 \neq 0,$$

where $\delta_0 = (\delta_{\alpha_0}', \delta_{\theta_0}')'$.

- The threshold parameter $\lambda$ is not identified under the null of the linear SAR model (the Davies problem).

- Following Hansen (1996) we employ a bootstrap sup-LR type test for the following statistic.
**Bootstrap Procedure**

(A) Our sup-D test statistic is defined as

\[
\sup_{\lambda} D_n (\lambda) = \sup_{\lambda} \left( J_n \left( \tilde{\theta}^* (\lambda), \lambda \right) - J_n \left( \hat{\theta}^* (\lambda), \lambda \right) \right)
\]

with

\[
J_n (\theta^*, \lambda) = \hat{g}_n (\theta^*, \lambda)' \hat{\Omega}_n^{-1} (\lambda) \hat{g}_n (\theta^*, \lambda)
\]

and

\[
\tilde{\theta}^* (\lambda) = \arg\min_{\theta^* \in \Omega^*} J_n (\theta^*, \lambda) \quad \text{and} \quad \hat{\theta}^* (\lambda) = \arg\min_{\theta^* \in \Omega^*} J_n (\theta^*, \lambda),
\]

s.t. \( \delta_{\alpha} = 0, \delta_{\beta} = 0 \)

where the weight matrix \( \hat{\Omega}_n (\lambda) \) for both restricted and unrestricted problems are based on the unrestricted GMM weight matrix to ensure that \( J_n \left( \tilde{\theta}^* (\lambda), \lambda \right) - J_n \left( \hat{\theta}^* (\lambda), \lambda \right) \geq 0 \) for any given \( \lambda \).

(B) Randomly drawing \( u^*_{i,n}, i = 1, \ldots, n \), from the standard normal distribution, we construct the bootstrap data \( y_{i,n}^* = \tilde{e}_{i,n}(\hat{\theta}) u^*_{i,n}, \) where \( \tilde{e}_{i,n}(\hat{\theta}) \) is the second-stage residual.

(C) Then using the bootstrap sample \( \{ y_{i,n}^*, x_{i,n}, z_{i,n} \}, i = 1, \ldots, n \) and weight matrix \( W_n \), we obtain the unrestricted and restricted bootstrap GMM estimates and construct the bootstrap sup-D statistic

\[
\sup_{\lambda} D_n^* (\lambda) = \sup_{\lambda} \left( J_n^* \left( \tilde{\theta}_n^* (\lambda), \lambda \right) - J_n^* \left( \hat{\theta}_n^* (\lambda), \lambda \right) \right),
\]

where \( J_n^* \left( \tilde{\theta}_n^* (\lambda), \lambda \right) \) and \( J_n^* \left( \hat{\theta}_n^* (\lambda), \lambda \right) \) are the bootstrap objective function for restricted and unrestricted estimation problems, respectively.

(D) Repeat steps (b)-(c) B times and index the bootstrap samples by \( b = 1, \ldots, B \). Denoting \( \sup_{\lambda} D_{n,b}^* (\lambda) \) as the bootstrap sup-D test statistic for the bootstrap sample \( b \), we obtain the bootstrap p-value by

\[
p^* = \frac{1}{n} \sum_{b=1}^{B} I \{ \sup_{\lambda} D_{n,b}^* (\lambda) < \sup_{\lambda} D_n (\lambda) \}.
\]
The following theorem supports the usage of the above bootstrap test.

**Theorem 4** Under Assumptions 1-5 and under $H_0$, we have

$$\sup_{\lambda} D_n(\lambda) \overset{d}{\to} \sup_{\lambda} D(\lambda)$$

and

$$\sup_{\lambda} D_n^#(\lambda) \overset{d}{\to} \sup_{\lambda} D(\lambda),$$

where $D$ is a chi-square process with the degrees of freedom equal to $4k + 3$. 
We consider the following DGP using three spatial weight matrices that correspond to three sample sizes.

- The Toledo spatial matrix $WO$ ($98 \times 98$) based on the 5 nearest neighbors of 98 census tracts in Toledo, Ohio.

- For larger sample sizes of $n = 196$ and 392 we use block diagonal matrices with the Toledo spatial matrix as their diagonal blocks.

Finally, we use 1000 Monte Carlo replications.
We consider the following data generating process

\[ y_{i,n} = \alpha \sum_{j \neq i} w_{ij,n} y_{j,n} + \beta_1 x_{i,n} + \delta_{\beta_1} x_{i,n}(\lambda) + \delta_{\alpha_1} \sum_{j \neq i} w_{ij,n,1}(\lambda) y_{j,n} + \delta_{\alpha_2} \sum_{j \neq i} w_{ij,n,2}(\lambda) y_{j,n} + \delta_{\alpha_3} \sum_{j \neq i} w_{ij,n,3}(\lambda) y_{j,n} + e_{i,n} \]  

(12)

where the regressor \( x_{i,n} \), threshold variable \( z_{i,n} \), and regression error \( e_{i,n} \) are independent i.i.d. \( N(0, 1) \) random variables. The threshold parameter \( \lambda \) is set to zero.

e explore various Monte Carlo simulation experiments with 1,000 iterations. We fix \( \alpha = 0.4 \) and \( \beta_1 = 3 \) and vary the threshold effects
\[ \delta_{\alpha} = (\delta_{\alpha_1}, \delta_{\alpha_2}, \delta_{\alpha_3}) = \{(0, 0, 0), (0.4, 0, 0), (0, 0.4, 0), (0, 0, 0.4)\} \] and
\[ \delta_{\beta_1} = 0, 1, 2. \]
## RMSE

### Panel A: $\delta_{\alpha} = (0, 0, 0)$

<table>
<thead>
<tr>
<th></th>
<th>One-step GMM</th>
<th>Two-step GMM</th>
<th>One-step GMM</th>
<th>Two-step GMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{\beta_1}$</td>
<td>0.00 1.00 2.00</td>
<td>0.00 1.00 2.00</td>
<td>0.00 1.00 2.00</td>
<td>0.00 1.00 2.00</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.42 0.40 0.38</td>
<td>0.49 0.41 0.34</td>
<td>0.42 0.38 0.35</td>
<td>0.45 0.36 0.28</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>1.08 1.02 1.14</td>
<td>0.89 0.61 0.69</td>
<td>1.00 0.94 0.85</td>
<td>0.66 0.69 0.53</td>
</tr>
<tr>
<td>$\delta_{\alpha_1}$</td>
<td>1.62 1.36 1.42</td>
<td>1.23 0.80 0.90</td>
<td>1.38 1.26 1.12</td>
<td>0.94 0.86 0.69</td>
</tr>
<tr>
<td>$\delta_{\alpha_2}$</td>
<td>1.69 1.55 1.64</td>
<td>1.12 0.90 0.86</td>
<td>1.48 1.45 1.31</td>
<td>0.88 0.81 0.67</td>
</tr>
<tr>
<td>$\delta_{\alpha_3}$</td>
<td>2.68 2.18 2.13</td>
<td>1.82 1.39 1.22</td>
<td>2.26 2.04 1.82</td>
<td>1.44 1.16 0.97</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.73 0.70 0.79</td>
<td>0.35 0.46 0.66</td>
<td>0.82 0.75 0.82</td>
<td>0.33 0.39 0.47</td>
</tr>
<tr>
<td>$\delta_{\beta_1}$</td>
<td>1.35 1.32 1.50</td>
<td>0.55 0.64 0.80</td>
<td>1.52 1.47 1.52</td>
<td>0.53 0.53 0.66</td>
</tr>
</tbody>
</table>

$n = 98$

### Panel B: $\delta_{\alpha} = (0.4, 0, 0)$

<table>
<thead>
<tr>
<th></th>
<th>One-step GMM</th>
<th>Two-step GMM</th>
<th>One-step GMM</th>
<th>Two-step GMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{\beta_1}$</td>
<td>0.00 1.00 2.00</td>
<td>0.00 1.00 2.00</td>
<td>0.00 1.00 2.00</td>
<td>0.00 1.00 2.00</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.43 0.41 0.37</td>
<td>0.52 0.43 0.32</td>
<td>0.39 0.38 0.33</td>
<td>0.44 0.29 0.22</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>1.01 1.01 0.92</td>
<td>0.93 0.58 0.38</td>
<td>0.88 0.92 0.86</td>
<td>0.79 0.48 0.32</td>
</tr>
<tr>
<td>$\delta_{\alpha_1}$</td>
<td>1.46 1.27 1.15</td>
<td>1.21 0.79 0.46</td>
<td>1.28 1.21 1.13</td>
<td>1.05 0.60 0.42</td>
</tr>
<tr>
<td>$\delta_{\alpha_2}$</td>
<td>1.52 1.52 1.47</td>
<td>1.12 0.78 0.56</td>
<td>1.34 1.38 1.30</td>
<td>0.93 0.65 0.41</td>
</tr>
<tr>
<td>$\delta_{\alpha_3}$</td>
<td>2.34 2.06 1.90</td>
<td>1.70 1.13 0.77</td>
<td>1.99 1.84 1.75</td>
<td>1.44 0.87 0.57</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.76 0.75 0.78</td>
<td>0.26 0.43 0.53</td>
<td>0.73 0.77 0.81</td>
<td>0.26 0.29 0.34</td>
</tr>
<tr>
<td>$\delta_{\beta_1}$</td>
<td>1.36 1.36 1.45</td>
<td>0.38 0.58 0.70</td>
<td>1.37 1.43 1.47</td>
<td>0.37 0.41 0.43</td>
</tr>
</tbody>
</table>

$n = 196$

### Panel C: $\delta_{\alpha} = (0.4, 0, 0)$

<table>
<thead>
<tr>
<th></th>
<th>One-step GMM</th>
<th>Two-step GMM</th>
<th>One-step GMM</th>
<th>Two-step GMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{\beta_1}$</td>
<td>0.00 1.00 2.00</td>
<td>0.00 1.00 2.00</td>
<td>0.00 1.00 2.00</td>
<td>0.00 1.00 2.00</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.41 0.40 0.36</td>
<td>0.56 0.43 0.32</td>
<td>0.40 0.35 0.32</td>
<td>0.40 0.21 0.15</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>1.04 0.95 0.90</td>
<td>0.90 0.43 0.37</td>
<td>0.90 0.98 0.91</td>
<td>0.49 0.41 0.22</td>
</tr>
<tr>
<td>$\delta_{\alpha_1}$</td>
<td>1.39 1.32 1.10</td>
<td>1.12 0.58 0.45</td>
<td>1.28 1.28 1.15</td>
<td>0.67 0.55 0.28</td>
</tr>
<tr>
<td>$\delta_{\alpha_2}$</td>
<td>1.62 1.51 1.45</td>
<td>1.08 0.59 0.50</td>
<td>1.48 1.48 1.33</td>
<td>0.62 0.53 0.32</td>
</tr>
<tr>
<td>$\delta_{\alpha_3}$</td>
<td>2.36 2.17 1.86</td>
<td>1.54 0.89 0.73</td>
<td>2.21 2.01 1.69</td>
<td>1.03 0.75 0.48</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.72 0.73 0.73</td>
<td>0.21 0.49 0.63</td>
<td>0.83 0.84 0.95</td>
<td>0.13 0.19 0.23</td>
</tr>
<tr>
<td>$\delta_{\beta_1}$</td>
<td>1.30 1.30 1.31</td>
<td>0.29 0.59 0.74</td>
<td>1.53 1.58 1.78</td>
<td>0.23 0.27 0.33</td>
</tr>
</tbody>
</table>
### Panel C: $\delta_{\alpha} = (0, 0.4, 0)$

<table>
<thead>
<tr>
<th>Model Type</th>
<th>One-step GMM</th>
<th>Two-step GMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{\beta_1}$</td>
<td>0.00 1.00 2.00</td>
<td>0.00 1.00 2.00</td>
</tr>
</tbody>
</table>

$n = 98$

<table>
<thead>
<tr>
<th>Model Type</th>
<th>One-step GMM</th>
<th>Two-step GMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>0.40 0.38 0.36</td>
<td>0.41 0.34 0.29</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>1.09 1.06 1.32</td>
<td>1.07 0.64 0.65</td>
</tr>
<tr>
<td>$\delta_{\alpha_1}$</td>
<td>1.48 1.37 1.58</td>
<td>1.34 0.78 0.74</td>
</tr>
<tr>
<td>$\delta_{\alpha_2}$</td>
<td>1.59 1.54 1.79</td>
<td>1.25 0.80 0.83</td>
</tr>
<tr>
<td>$\delta_{\alpha_3}$</td>
<td>2.29 2.15 2.25</td>
<td>1.81 1.21 1.04</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.69 0.74 0.88</td>
<td>0.52 0.37 0.47</td>
</tr>
<tr>
<td>$\delta_{\beta_1}$</td>
<td>1.40 1.47 1.53</td>
<td>0.66 0.59 0.73</td>
</tr>
</tbody>
</table>

$n = 196$

<table>
<thead>
<tr>
<th>Model Type</th>
<th>One-step GMM</th>
<th>Two-step GMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>0.40 0.38 0.34</td>
<td>0.44 0.36 0.27</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>1.01 1.02 1.01</td>
<td>0.88 0.85 0.38</td>
</tr>
<tr>
<td>$\delta_{\alpha_1}$</td>
<td>1.35 1.27 1.21</td>
<td>1.07 0.86 0.44</td>
</tr>
<tr>
<td>$\delta_{\alpha_2}$</td>
<td>1.50 1.47 1.51</td>
<td>1.04 0.94 0.48</td>
</tr>
<tr>
<td>$\delta_{\alpha_3}$</td>
<td>2.16 1.88 1.88</td>
<td>1.45 1.06 0.63</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.75 0.75 0.84</td>
<td>0.27 0.39 0.50</td>
</tr>
<tr>
<td>$\delta_{\beta_1}$</td>
<td>1.40 1.43 1.54</td>
<td>0.40 0.53 0.65</td>
</tr>
</tbody>
</table>

$n = 392$

<table>
<thead>
<tr>
<th>Model Type</th>
<th>One-step GMM</th>
<th>Two-step GMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>0.41 0.37 0.33</td>
<td>0.46 0.35 0.28</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.99 0.99 1.03</td>
<td>0.61 0.40 0.30</td>
</tr>
<tr>
<td>$\delta_{\alpha_1}$</td>
<td>1.25 1.23 1.20</td>
<td>0.79 0.45 0.34</td>
</tr>
<tr>
<td>$\delta_{\alpha_2}$</td>
<td>1.51 1.48 1.50</td>
<td>0.76 0.50 0.43</td>
</tr>
<tr>
<td>$\delta_{\alpha_3}$</td>
<td>2.07 1.94 1.76</td>
<td>1.13 0.69 0.57</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.80 0.86 0.84</td>
<td>0.21 0.30 0.44</td>
</tr>
<tr>
<td>$\delta_{\beta_1}$</td>
<td>1.41 1.50 1.56</td>
<td>0.29 0.43 0.62</td>
</tr>
</tbody>
</table>
Relative RMSE of two-step GMM estimator with the quasi-optimal 2SLS

<table>
<thead>
<tr>
<th></th>
<th>$\delta_\alpha = (0, 0, 0)$</th>
<th>$\delta_\alpha = (0.4, 0, 0)$</th>
<th>$\delta_\alpha = (0, 0.4, 0)$</th>
<th>$\delta_\alpha = (0, 0, 0.4)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{\beta_1}$</td>
<td>0.00 1.00 2.00</td>
<td>0.00 1.00 2.00</td>
<td>0.00 1.00 2.00</td>
<td>0.00 1.00 2.00</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1.11 1.78 2.86</td>
<td>1.01 1.43 1.91</td>
<td>0.97 1.37 1.98</td>
<td>1.02 1.40 2.34</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.41 0.22 0.06</td>
<td>0.05 0.33 0.08</td>
<td>0.49 0.00 0.27</td>
<td>0.14 0.09 0.12</td>
</tr>
<tr>
<td>$\delta_{\alpha_1}$</td>
<td>0.27 0.30 0.23</td>
<td>0.33 0.28 0.49</td>
<td>0.49 0.00 0.28</td>
<td>0.15 0.07 0.10</td>
</tr>
<tr>
<td>$\delta_{\alpha_2}$</td>
<td>0.43 0.30 0.07</td>
<td>0.45 0.28 0.15</td>
<td>0.34 0.00 0.33</td>
<td>0.18 0.13 0.14</td>
</tr>
<tr>
<td>$\delta_{\alpha_3}$</td>
<td>0.26 0.46 0.30</td>
<td>0.06 0.29 0.15</td>
<td>0.51 0.01 0.34</td>
<td>0.23 0.12 0.14</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.53 1.12 0.16</td>
<td>0.09 0.71 1.32</td>
<td>0.74 0.01 0.56</td>
<td>0.26 0.17 1.55</td>
</tr>
<tr>
<td>$\delta_{\beta_1}$</td>
<td>0.52 1.28 0.20</td>
<td>0.02 0.82 0.67</td>
<td>0.83 0.01 0.82</td>
<td>0.40 0.24 1.52</td>
</tr>
</tbody>
</table>

$n = 98$

<table>
<thead>
<tr>
<th></th>
<th>$\delta_\alpha = (0, 0, 0)$</th>
<th>$\delta_\alpha = (0.4, 0, 0)$</th>
<th>$\delta_\alpha = (0, 0.4, 0)$</th>
<th>$\delta_\alpha = (0, 0, 0.4)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>1.21 2.59 5.20</td>
<td>1.03 1.67 2.55</td>
<td>1.07 2.33 3.61</td>
<td>1.04 2.26 3.35</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.39 0.44 0.33</td>
<td>0.35 0.40 0.29</td>
<td>0.48 0.42 0.16</td>
<td>0.10 0.10 0.35</td>
</tr>
<tr>
<td>$\delta_{\alpha_1}$</td>
<td>0.42 0.49 0.30</td>
<td>0.40 0.38 0.27</td>
<td>0.39 0.30 0.29</td>
<td>0.12 0.16 0.30</td>
</tr>
<tr>
<td>$\delta_{\alpha_2}$</td>
<td>0.34 0.40 0.41</td>
<td>0.16 0.08 0.06</td>
<td>0.40 0.39 0.15</td>
<td>0.11 0.12 0.28</td>
</tr>
<tr>
<td>$\delta_{\alpha_3}$</td>
<td>0.44 0.44 0.41</td>
<td>0.13 0.09 0.11</td>
<td>0.32 0.30 0.19</td>
<td>0.14 0.20 0.28</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.75 2.04 2.50</td>
<td>0.69 1.29 0.96</td>
<td>0.46 0.84 1.18</td>
<td>0.25 0.81 1.56</td>
</tr>
<tr>
<td>$\delta_{\beta_1}$</td>
<td>0.90 1.67 2.42</td>
<td>0.11 0.35 0.09</td>
<td>0.55 0.94 0.45</td>
<td>0.33 1.01 0.37</td>
</tr>
</tbody>
</table>

$n = 196$

<table>
<thead>
<tr>
<th></th>
<th>$\delta_\alpha = (0, 0, 0)$</th>
<th>$\delta_\alpha = (0.4, 0, 0)$</th>
<th>$\delta_\alpha = (0, 0.4, 0)$</th>
<th>$\delta_\alpha = (0, 0, 0.4)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>1.34 5.04 10.95</td>
<td>0.94 1.88 3.10</td>
<td>1.12 3.32 5.63</td>
<td>1.08 3.27 4.77</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.02 0.16 0.07</td>
<td>0.23 0.06 0.10</td>
<td>0.16 0.25 0.31</td>
<td>0.24 0.16 0.22</td>
</tr>
<tr>
<td>$\delta_{\alpha_1}$</td>
<td>0.03 0.10 0.08</td>
<td>0.23 0.05 0.12</td>
<td>0.19 0.23 0.25</td>
<td>0.14 0.19 0.21</td>
</tr>
<tr>
<td>$\delta_{\alpha_2}$</td>
<td>0.03 0.17 0.10</td>
<td>0.25 0.09 0.14</td>
<td>0.17 0.22 0.23</td>
<td>0.21 0.11 0.22</td>
</tr>
<tr>
<td>$\delta_{\alpha_3}$</td>
<td>0.04 0.12 0.12</td>
<td>0.27 0.11 0.17</td>
<td>0.19 0.21 0.22</td>
<td>0.16 0.13 0.20</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.03 1.00 3.06</td>
<td>0.43 0.13 1.14</td>
<td>0.31 1.24 1.90</td>
<td>0.54 1.21 1.76</td>
</tr>
<tr>
<td>$\delta_{\beta_1}$</td>
<td>0.04 1.14 2.84</td>
<td>0.52 0.12 1.04</td>
<td>0.43 1.43 1.66</td>
<td>0.16 0.41 1.30</td>
</tr>
</tbody>
</table>
## Two-step GMM for $\delta_\alpha = (0.4, 0, 0)$

$n = 98$

<table>
<thead>
<tr>
<th></th>
<th>Bias</th>
<th>Std Dev</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_\beta_1$</td>
<td>0.00</td>
<td>1.00</td>
<td>2.00</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>-0.04</td>
<td>-0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>-0.02</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>$\delta_\alpha_1$</td>
<td>0.05</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>$\delta_\alpha_2$</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.01</td>
</tr>
<tr>
<td>$\delta_\alpha_3$</td>
<td>0.09</td>
<td>0.06</td>
<td>-0.01</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.01</td>
<td>-0.07</td>
<td>-0.10</td>
</tr>
<tr>
<td>$\delta_\beta_0$</td>
<td>-0.02</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>$\delta_\beta_1$</td>
<td>-0.01</td>
<td>0.13</td>
<td>0.19</td>
</tr>
</tbody>
</table>
## Two-step GMM for $\delta_\alpha = (0.4, 0.0, 0.0)$

$n = 196$

<table>
<thead>
<tr>
<th></th>
<th>Bias</th>
<th>Std Dev</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_\beta_1$</td>
<td>0.00</td>
<td>1.00</td>
<td>2.00</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>-0.02</td>
<td>-0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>$\delta_\alpha_1$</td>
<td>0.04</td>
<td>0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td>$\delta_\alpha_2$</td>
<td>-0.04</td>
<td>0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td>$\delta_\alpha_3$</td>
<td>0.10</td>
<td>-0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.00</td>
<td>-0.06</td>
<td>-0.07</td>
</tr>
<tr>
<td>$\delta_\beta_0$</td>
<td>-0.01</td>
<td>-0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>$\delta_\beta_1$</td>
<td>-0.01</td>
<td>0.10</td>
<td>0.09</td>
</tr>
</tbody>
</table>
**TWO-STEP GMM FOR $\delta_\alpha = (0.4, 0, 0)$**

$n = 392$

<table>
<thead>
<tr>
<th></th>
<th>Bias</th>
<th>Std Dev</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_\beta_1$</td>
<td>0.00</td>
<td>1.00</td>
<td>2.00</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>-0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>-0.03</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>$\delta_\alpha_1$</td>
<td>0.07</td>
<td>-0.01</td>
<td>-0.02</td>
</tr>
<tr>
<td>$\delta_\alpha_2$</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
</tr>
<tr>
<td>$\delta_\alpha_3$</td>
<td>0.01</td>
<td>-0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.00</td>
<td>-0.04</td>
<td>-0.05</td>
</tr>
<tr>
<td>$\delta_\beta_0$</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$\delta_\beta_1$</td>
<td>-0.01</td>
<td>0.06</td>
<td>0.06</td>
</tr>
</tbody>
</table>

$\delta_\alpha = (0.4, 0, 0)$
## Bootstrap Threshold Test: Size and Power

<table>
<thead>
<tr>
<th>$\delta_{\beta_1}$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A:</strong> $(\delta_{\alpha_1}, \delta_{\alpha_2}, \delta_{\alpha_3}) = (0, 0, 0)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n = 98$</td>
<td>0.053</td>
<td>0.110</td>
<td>0.306</td>
<td>0.560</td>
</tr>
<tr>
<td>$n = 196$</td>
<td>0.066</td>
<td>0.119</td>
<td>0.291</td>
<td>0.514</td>
</tr>
<tr>
<td>$n = 392$</td>
<td>0.057</td>
<td>0.126</td>
<td>0.287</td>
<td>0.484</td>
</tr>
<tr>
<td><strong>Panel B:</strong> $(\delta_{\alpha_1}, \delta_{\alpha_2}, \delta_{\alpha_3}) = (0.4, 0, 0)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n = 98$</td>
<td>0.104</td>
<td>0.463</td>
<td>0.663</td>
<td>0.819</td>
</tr>
<tr>
<td>$n = 196$</td>
<td>0.191</td>
<td>0.709</td>
<td>0.899</td>
<td>0.942</td>
</tr>
<tr>
<td>$n = 392$</td>
<td>0.302</td>
<td>0.922</td>
<td>0.987</td>
<td>0.993</td>
</tr>
<tr>
<td><strong>Panel C:</strong> $(\delta_{\alpha_1}, \delta_{\alpha_2}, \delta_{\alpha_3}) = (0, 0.4, 0)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n = 98$</td>
<td>0.112</td>
<td>0.213</td>
<td>0.458</td>
<td>0.642</td>
</tr>
<tr>
<td>$n = 196$</td>
<td>0.096</td>
<td>0.208</td>
<td>0.447</td>
<td>0.627</td>
</tr>
<tr>
<td>$n = 392$</td>
<td>0.083</td>
<td>0.229</td>
<td>0.434</td>
<td>0.580</td>
</tr>
<tr>
<td><strong>Panel D:</strong> $(\delta_{\alpha_1}, \delta_{\alpha_2}, \delta_{\alpha_3}) = (0, 0, 0.4)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n = 98$</td>
<td>0.080</td>
<td>0.279</td>
<td>0.516</td>
<td>0.705</td>
</tr>
<tr>
<td>$n = 196$</td>
<td>0.115</td>
<td>0.408</td>
<td>0.626</td>
<td>0.749</td>
</tr>
<tr>
<td>$n = 392$</td>
<td>0.200</td>
<td>0.638</td>
<td>0.830</td>
<td>0.897</td>
</tr>
</tbody>
</table>
Empirical Illustration

- There is a huge body of work that studies how the behavior of peers and their characteristics affect student academic achievement; Sacerdote (2011) and Epple and Romano (2011).

- Yet, the evidence for peer effects in education is generally mixed due to several challenges.

- These peer effect studies are typically based on the linear-in-means model.

• We study the heterogeneity of peer effects in students’ grade point average (GPA) by extending the linear SDM model of Lin (2010) which allows individuals to interact through a network of friends using Add Health data.

• We model the heterogeneity of peer effects by allowing for threshold peer effects based on three alternative variables: consciousness, emotional stability, and extraversion.
We estimate

\[ y_{i,g} = \mu_g + \alpha \sum_{j \neq i} w_{ij,n,g} y_{j,g} + \beta' x_{i,g} + \gamma' \sum_{j \neq i} w_{ij,n,g} x_{j,g} \]

\[ + \delta_{\alpha_1} \sum_{j \neq i} w_{ij,n,g,1} (\lambda) y_{j,g} + \delta_{\alpha_2} \sum_{j \neq i} w_{ij,n,g,2} (\lambda) y_{j,g} + \delta_{\alpha_3} \sum_{j \neq i} w_{ij,n,g,3} (\lambda) y_{j,g} \]

\[ + e_{i,g}, \]

where \( i = 1, \ldots, N_g \) denotes a student in school, \( g = 1, \ldots, G \). The total number of schools is denoted by \( G \) and the total sample size is denoted by \( N = \sum_{g=1}^{G} N_g \).
DATA

• Wave I in-school and in-home Add Health survey, a longitudinal study on a nationally representative sample covering adolescents in grades 7-12 during the 1994-95 school year.

• Add Health survey provides information on students’ friendship links which is used to construct the friendship network of each student.

• Exclude individuals that do not name a friend and are never nominated by others and focus on suburban schools.

• These restrictions yield a sample size of 924 observations and 62 school-grade groups.
• The endogenous variable is the GPA calculated from the respondent’s grades in several subjects, including English or language arts, history or social science, mathematics, and science.

• Our threshold variables include the personality variables of conscientiousness, emotional stability, and extraversion.

• Each personality variable represents the main factor of each personality dimension using factor analysis based on 13 items from the survey according to the Lexical approach.

• Following Lin (2010), we also control for individual characteristics that capture family type (that is, whether the student lives with both parents), mother’s education and occupation, and demographic characteristics (age, race, and gender).
## SUMMARY STATISTICS

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Own Characteristics</th>
<th>Friends' Mean Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPA</td>
<td>Average grade in mathematics, science, English or language arts, and history or social science</td>
<td>3.079 0.720</td>
<td>5.841 4.087</td>
</tr>
<tr>
<td>Age</td>
<td>Age</td>
<td>0.032 0.674</td>
<td>0.118 0.978</td>
</tr>
<tr>
<td>Emotional stability (or neuroticism)</td>
<td>The degree to which an individual experiences the world as threatening and beyond his/her control</td>
<td>-0.058 0.825</td>
<td>-0.056 1.176</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>The degree to which a person is willing to comply with conventional rules, norms, and standards.</td>
<td>0.202 0.791</td>
<td>0.538 1.356</td>
</tr>
<tr>
<td>Extraversion</td>
<td>The degree to which a person needs attention and social interaction</td>
<td>14.580 1.529</td>
<td>27.117 18.178</td>
</tr>
<tr>
<td>Years in school</td>
<td>Number of years in current school</td>
<td>2.767 1.644</td>
<td>5.840 6.516</td>
</tr>
<tr>
<td>Male</td>
<td>1 if Male; 0 otherwise</td>
<td>0.400 0.490</td>
<td>0.731 0.895</td>
</tr>
<tr>
<td>Black</td>
<td>1 if Black; 0 otherwise</td>
<td>0.184 0.388</td>
<td>0.400 1.012</td>
</tr>
<tr>
<td>Asian</td>
<td>1 if Asian; 0 otherwise</td>
<td>0.061 0.239</td>
<td>0.090 0.369</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1 if Hispanic; 0 otherwise</td>
<td>0.065 0.247</td>
<td>0.090 0.297</td>
</tr>
<tr>
<td>Other race</td>
<td>1 if American Indian or other race; 0 otherwise</td>
<td>0.060 0.237</td>
<td>0.089 0.292</td>
</tr>
<tr>
<td>Family type</td>
<td>1 if living with both parents; 0 otherwise</td>
<td>0.808 0.394</td>
<td>1.510 1.161</td>
</tr>
<tr>
<td>Mom edu. less than HS</td>
<td>1 if mom's education less than high school; 0 otherwise</td>
<td>0.065 0.247</td>
<td>0.088 0.291</td>
</tr>
<tr>
<td>Mom edu. more than HS</td>
<td>1 if mom's education more than high school; 0 otherwise</td>
<td>0.521 0.500</td>
<td>1.053 1.110</td>
</tr>
<tr>
<td>Mom edu. missing</td>
<td>1 if the information about mom's education is missing; 0 otherwise.</td>
<td>0.078 0.268</td>
<td>0.122 0.341</td>
</tr>
<tr>
<td>Mom job missing</td>
<td>1 if the information about mom's job is missing; 0 otherwise.</td>
<td>0.079 0.270</td>
<td>0.139 0.367</td>
</tr>
<tr>
<td>Professional</td>
<td>1 if mom is a doctor, lawyer, scientist, teacher, executive, director, and the like; 0 otherwise.</td>
<td>0.340 0.474</td>
<td>0.703 0.923</td>
</tr>
<tr>
<td>Other jobs</td>
<td>1 if mom’s occupation is not among the Professional or Stay home; 0 otherwise.</td>
<td>0.366 0.482</td>
<td>0.696 0.852</td>
</tr>
<tr>
<td>Sports club member</td>
<td>1 if sports club member; 0 otherwise.</td>
<td>0.682 0.466</td>
<td>1.324 1.209</td>
</tr>
</tbody>
</table>
### Peer Effects in Student Academic Achievement

<table>
<thead>
<tr>
<th>Threshold Variable</th>
<th>Conscientiousness</th>
<th>Emotional stability</th>
<th>Extraversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold test (Boot. p-value)</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>Threshold parameter</td>
<td>0.278</td>
<td>-0.198</td>
<td>0.392</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.299</td>
<td>-0.325</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.265</td>
<td>-0.253</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.275</td>
<td>-0.211</td>
</tr>
</tbody>
</table>

### Endogenous Effects

<table>
<thead>
<tr>
<th>Global ($\alpha$)</th>
<th>0.068</th>
<th>0.030</th>
<th>-0.035</th>
<th>0.133</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-to-low ($\alpha_{11}$)</td>
<td>0.218</td>
<td>0.054</td>
<td>0.719</td>
<td></td>
</tr>
<tr>
<td>High-to-low ($\alpha_{21}$)</td>
<td>0.024</td>
<td>-0.177</td>
<td>0.441</td>
<td></td>
</tr>
<tr>
<td>Low-to-high ($\alpha_{12}$)</td>
<td>0.053</td>
<td>-0.467</td>
<td>0.194</td>
<td></td>
</tr>
<tr>
<td>High-to-high ($\alpha_{22}$)</td>
<td>0.386</td>
<td>0.248</td>
<td>0.554</td>
<td></td>
</tr>
</tbody>
</table>
## Linear SDM

<table>
<thead>
<tr>
<th>COEF (1)</th>
<th>SE (2)</th>
<th>95% Boot. CI (3)</th>
<th></th>
<th>COEF (5)</th>
<th>95% Boot. CI (6)</th>
<th></th>
<th>COEF (8)</th>
<th>95% Boot. CI (9)</th>
<th></th>
<th>COEF (11)</th>
<th>95% Boot. CI (12)</th>
<th></th>
<th>COEF (13)</th>
<th>95% Boot. CI (14)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Threshold SDM

<table>
<thead>
<tr>
<th>Threshold Variable</th>
<th>Conscientiousness</th>
<th>Emotional stability</th>
<th>Extraversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotional stability</td>
<td>0.137</td>
<td>0.037</td>
<td>0.064</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>-0.001</td>
<td>0.027</td>
<td>-0.047</td>
</tr>
<tr>
<td>Extraversion</td>
<td>0.025</td>
<td>0.031</td>
<td>-0.041</td>
</tr>
<tr>
<td>Age</td>
<td>-0.180</td>
<td>0.046</td>
<td>-0.285</td>
</tr>
<tr>
<td>Years in school</td>
<td>-0.030</td>
<td>0.023</td>
<td>-0.075</td>
</tr>
<tr>
<td>Male</td>
<td>-0.221</td>
<td>0.048</td>
<td>-0.317</td>
</tr>
<tr>
<td>Black</td>
<td>-0.207</td>
<td>0.110</td>
<td>-0.426</td>
</tr>
<tr>
<td>Asian</td>
<td>0.534</td>
<td>0.106</td>
<td>0.322</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.020</td>
<td>0.101</td>
<td>-0.227</td>
</tr>
<tr>
<td>Other race</td>
<td>-0.139</td>
<td>0.097</td>
<td>-0.315</td>
</tr>
<tr>
<td>Family type</td>
<td>0.054</td>
<td>0.061</td>
<td>-0.074</td>
</tr>
<tr>
<td>Mom ed less than HS</td>
<td>-0.035</td>
<td>0.104</td>
<td>-0.246</td>
</tr>
<tr>
<td>Mom ed more than HS</td>
<td>0.141</td>
<td>0.051</td>
<td>0.028</td>
</tr>
<tr>
<td>Mom ed missing</td>
<td>-0.097</td>
<td>0.091</td>
<td>-0.285</td>
</tr>
<tr>
<td>Mom job missing</td>
<td>-0.047</td>
<td>0.090</td>
<td>-0.209</td>
</tr>
<tr>
<td>Professional</td>
<td>0.110</td>
<td>0.064</td>
<td>-0.010</td>
</tr>
<tr>
<td>Other jobs</td>
<td>0.042</td>
<td>0.061</td>
<td>-0.058</td>
</tr>
<tr>
<td>Sports club member</td>
<td>0.202</td>
<td>0.052</td>
<td>0.110</td>
</tr>
</tbody>
</table>

Sample size: 924
Sample size for low regime: 692
Sample size for high regime: 232
CONCLUSION

• We propose a new class of social interaction models that generalize the linear Spatial Durbin Model to allow for threshold effects that capture the heterogeneity in social interaction effects.

• We employ a GMM estimation method and develop an asymptotic distribution theory for the GMM estimators of the regression parameters as well as for the threshold parameter.

• We provide Monte Carlo simulations that show the finite sample performance of our estimators.

• methodology to study students’ grade point average by uncovering the presence of threshold effects on peer influences based on personality skills using Add Health data.
• In terms of future work, we believe that our proposed paper is immediately applicable to a wide range of questions with policy significance in neighborhood effects, intergenerational mobility, and child development.

• For example, lagged feedbacks can capture neighborhood effects during childhood while contemporaneous feedbacks capture the idea that disadvantaged localities act as barriers.

• Extending the existing framework is to account for the endogeneity of the sociomatrix by relaxing the assumption of predetermined and fixed network structure.