

Identification of Heterogeneous Peer Effects*

Eyo I. Herstad[†] Aristide Houndetoungan[‡] Myungkou Shin[§]

July 24, 2026

Abstract

We develop a model of peer effects where each peer has a separate effect depending on their rank in the distribution of peers' outcomes. Our model admits a unique equilibrium, and model parameters can be identified using peers' exogenous characteristics. To obtain a more parsimonious model of peer effects, we introduce a tractable specification based on quantile-dependent peer effect coefficients, and develop a specification test. Applying the model to several student outcomes in the Add Health data, we uncover heterogeneous and often non-monotonic spillovers that cannot be captured by existing models. Our results have direct implications for counterfactual analysis, suggesting that a student's influence in a network depends not only on network structure, but also on that student's position in the outcome distribution of their peers.

*We are grateful to Yann Bramoullé, Vincent Boucher, Stephan Bonhomme, and Mathieu Lambert for their helpful comments and suggestions. We also thank participants at the Econometrics Seminars at Washington University in St. Louis, New York University, University of Exeter, and Concordia University, as well as participants at the Africa Meeting of the Econometric Society (AFES, 2026) and the World Congress of Econometrics, 2026. This research uses data from the National Longitudinal Study of Adolescent to Adult Health (Add Health), a program that is directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by Grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other US federal agencies and foundations. Special acknowledgment is given to Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain Add Health data files is available on the Add Health website. No direct support was received from Grant P01-HD31921 for this research.

[†]Erasmus University Rotterdam and Tindbergen Institute. email: herstad@ese.eur.nl

[‡]Département d'économie, Université Laval. email: ahoundetoungan@ecn.ulaval.ca

[§]School of Social Sciences, University of Surrey. email: m.shin@surrey.ac.uk

1 Introduction

Economic agents are often connected. These connections raise important questions about how linked units influence one another through spillovers and conformity, known as peer effects. For example, we may expect a student to perform better academically if they have a friend with high grades, or to get in trouble at school if the majority of their friends do. These types of spillovers are inherently heterogeneous. The effect of each peer’s outcome on the student’s outcome depends on the outcomes of all the peers of the student. However, the existing models of peer effects either assume homogeneous effects from peers through the linear-in-means (LiM) model (Manski, 1993) or impose strong parametric restrictions on the underlying peer effects (Boucher et al., 2024). These models, when misspecified, may lead to inaccurate counterfactual analyses. This greatly limits the scope for peer effect estimation in empirical settings, where it’s well known that peer effects are heterogeneous (Sacerdote, 2011; Masten, 2018; Bramoullé et al., 2020).

Our paper fills this gap between theory and empirical reality by introducing a flexible framework for the identification and estimation of heterogeneous peer effects. Our framework, in its most general form, allows for each peer’s outcome to have a differential effect on the individual, determined by their rank in the peer group of the said individual, in terms of the outcome. The framework, therefore, allows for a type of endogenous heterogeneity where the effect of one individual on another depends on the full distribution of outcomes among each individual’s peers. The framework covers several special cases of endogenous peer effects that are of empirical interest, such as quantile peer effects and extreme order statistics.

In its most general form, our model can be motivated as a regression of an individual’s outcome on the distribution of their peers’ outcomes. This is because ordered peer outcomes have a one-to-one relationship with the (empirical) distribution of peer outcomes. As such, our model allows for a wide range of peer effect patterns that are functions of the empirical distribution of peer outcomes.

The generality of our model captures many economically relevant features of peer effects. Firstly, our model allows for substitutability and complementarity among peers. Perfectly substitutable peers, i.e., linear-in-means, and perfectly complementary peers, i.e., linear-in-mins, are nested as special cases of our model.¹ Secondly, our model allows for returns to scale in terms of the number of peers, as the total peer effect can be increasing in the number of peers.

¹By a linear-in-mins model we mean a model where the outcome is a linear function of the minimum outcome of their peers. This model has previously been discussed in Tao and Lee (2014).

Importantly, our model allows for a non-monotonic marginal effect of peer outcomes. This means that we do not impose that peers with higher outcomes necessarily have a greater impact on an individual’s outcome than peers with lower outcomes, or vice versa. This is a particularly distinct feature of our model, which is not shared by any existing models for heterogeneous peer effects. For example, a recent and influential paper, [Boucher et al. \(2024\)](#), models peer effects as linear in either convex or concave functions of peer outcomes. This model recommends using the Constant Elasticity of Substitution (CES) aggregator to measure the effect of peer outcomes. Since the CES function is either convex or concave, the marginal effect of peer outcomes in this model is necessarily monotonic.² This is not true in our model, as can be seen from a special case of it where an individual is only affected by the peer whose outcome is the median among said individual’s peers. In this case, the marginal effect of a peer’s outcome is zero for individuals who are not the median, meaning the extreme outcomes have no impact on the peer effect, while the median has a non-zero effect. These types of non-monotonic peer effects have been heavily discussed in theories on social interactions, such as the theory of social comparison where peer influence is driven by the outcome of a representative peer ([Festinger, 1954](#)). In this framework, the outcome of this representative peer serves as a focal point for the group, and high outcome peers may cause negative spillovers due to discouragement effects ([Rogers and Feller, 2016](#)).

This flexibility in our model does not just allow us to detect more varied patterns of peer effects, but has direct implications for policy. Often, researchers want to estimate spillovers to identify key players in the network ([Ballester et al., 2006](#)), optimal treatment assignment ([Galeotti et al., 2020](#)), or to design optimal network allocations to improve outcomes in a network ([Carrell et al., 2013](#)). In all of these cases, having a misspecified model for spillovers can lead to sub-optimal or even damaging policies. For example, if a researcher uses a linear-in-means model when the true model is linear-in-median, they may mistakenly target treatment to the wrong individuals, when in reality these individuals have little influence on the median and therefore the true spillover.

In the general version of our model, the dimension of the model parameter grows quadratically in the maximum number of peers, making interpretation and estimation challenging. We therefore show how to construct lower-dimensional restricted versions of the model, with a focus on interpolated quantiles. The quantile model allows us to capture peer effect heterogeneity by allowing peers to exert distinct influences depending on their location in the peer outcome distribution. As the researcher needs to select the number of quantiles to include, we develop a specification test that allows practitioners to decide between two potentially misspecified models that do not nest each other. We then provide guidance for researchers

²This model also requires all outcomes to be strictly positive, which our model does not require.

in how to apply this test to select model specifications in practice. Our simulations indicate that this model selection method performs well and has good power properties in realistic settings. In our empirical application, we apply the quantile version of our model and use the test to select the number of quantiles to use.

The primary contribution of this paper is to introduce and show the identification of a general model of heterogeneous peer effects. This model allows for any type of spillovers that can be represented through the empirical peer outcome distribution. Our identification result makes use of the conditional distribution of the outcomes given the network connections and the individual characteristics, with the network assumed to be exogenous. The steps of the identification argument are as follows. First, we show that the general model admits a unique equilibrium when the total incoming spillover from all peers is bounded, in absolute value, by one for each individual. As in most nonlinear peer effects models, our identification conditions rely on a high-level rank restriction (Brock and Durlauf, 2007; Yang and Lee, 2017; Boucher et al., 2024). We show that this restriction is satisfied when the network includes individuals who have friends-of-friends who are not their direct friends. This condition is the same as that used to obtain identification in the LiM model (Bramoullé et al., 2009). Under such a network structure, we show that the distribution of the exogenous characteristics of these friends-of-friends can be used as instruments. We then construct moment conditions to identify all the parameters of the model. This step relies on an instrument relevance condition closely related to the equivalent condition of the LiM model. We also provide a specification test of the model that allows practitioners to choose a model specification in a data-driven way.

Our second contribution is to use our model to show that non-monotonic peer effects are highly prevalent in real world network data. We do this by estimating a four-quantile version of our model on 11 different outcomes in the National Longitudinal Study of Adolescent to Adult Health (Add Health) dataset.³ We then apply our estimator and find a wide range of different spillovers. We find non-monotonic marginal effects of peers for most of our outcomes, with the middle of the peer outcome distribution being the most important, consistent with psychological models of peer effects (Festinger, 1954; Rogers and Feller, 2016). For example, we find that for smoking behavior, it is the middle of the distribution that matters, while the minimum and maximum have no effect. Similarly, for academic achievement (GPA), what matters is the 66th quantile, while the maximum has a small negative effect. Existing methods such as LiM and the CES approach of Boucher et al. (2024) would not be able to capture these spillover patterns as they show a pattern of non-monotonic marginal peer effects.

³We choose four quantiles as our specification test selects four quantiles for almost all outcomes.

Beyond the parameter estimates, we show that the choice of peer effect model has strong effects on estimates of policy objects of interest, like student influence. Using the empirical results, we identify key players as the individuals whose removal of ties in the network yields the largest changes in the outcome distribution. Identifying such nodes is important, for example, for a social planner who is implementing targeted policies (Ballester et al., 2006; Lee et al., 2021). The results indicate that node rankings based on influence differ between the proposed model and the LiM and CES specifications, especially for outcomes that are characterized by non-monotonic peer effects, which these specifications fail to capture. Analyzing such outcomes using the LiM or CES models can distort the identification of key players, thereby reducing the effectiveness of targeted policy interventions.

The study of peer effects has been a large literature in economics (Bramoullé et al., 2020; De Paula, 2020), with seminal contributions by Manski (1993); Sacerdote (2001); Bramoullé et al. (2009). A strand of this literature has focused on estimating peer effects as non-linear functions of peer outcomes, such as extreme values (Tao and Lee, 2014; Tatsi, 2015), the proportion of peers above a cutoff (Kang, 2007; Guryan et al., 2009) or through quantile regressions (Kang, 2007).⁴ However, these contributions remain sporadic and disconnected, lacking both a unifying framework and a formal theoretical treatment of the identification and estimation challenges involved. This paper provides a rigorous treatment of peer effect heterogeneity that the prior work has postulated informally, unifying the intuitions underlying the empirical practice in a tractable framework.

The rest of this paper is structured as follows. Section 2 discusses the model and its implications, and Section 3 presents our formal identification results, specification tests, and comparisons to other models. Section 4 presents and discusses our simulation results, and Section 5 applies our methodology to the Add Health dataset.

2 Model

We consider a many-network setup with m networks in total, indexed by $g = 1, \dots, m$. In the g -th network, there are n_g units, indexed by $i = 1, \dots, n_g$. We assume independence across networks, but not across units within a network. The network structure of the g -th network is represented by an adjacency matrix $\mathbf{A}_g = [a_{i,j,g}]$. We define $a_{i,j,g} = 1$ if the j -th unit in the g -th network is a friend (or peer) of the i -th unit in the g -th network, and

⁴Another strand of this literature has also estimated heterogeneous peer effects based on latent traits or observable characteristics, such as race or gender (Boucher et al., 2022; Comola et al., 2025). These models, while conceptually related, do not directly map into our framework.

$a_{i,j,g} = 0$ otherwise.⁵ The network may be directed; i.e., $a_{i,j,g} = 1$ does not necessarily imply $a_{j,i,g} = 1$. We also assume there are no self-links, meaning $a_{i,i,g} = 0$ for every i and g . The network adjacency matrices $\{\mathbf{A}_g\}_{g=1}^m$ are observed. In addition to the network adjacency matrices, we observe unit-level outcome $y_{i,g} \in \mathbb{R}$ and unit-level control covariates $\mathbf{x}_{i,g} \in \mathbb{R}^l$. In sum, the network dataset given in the econometric framework of this paper is

$$\{\mathbf{A}_g, \{y_{i,g}, \mathbf{x}_{i,g}\}_{i=1}^{n_g}\}_{g=1}^m.$$

The theoretical results of Section 2 are discussed with regard to a single network. Therefore, for notational brevity, we will suppress the network subscript g for the rest of Section 2.

In our model, the outcome y_i of unit i depends on the outcomes of all their peers, i.e., all j such that $a_{i,j} = 1$. We characterize this dependence in a highly flexible way, with every peer having a separate effect on unit i depending on the ranking of their outcome among unit i 's friends. Specifically, let $d_i = \sum_{j=1}^n a_{i,j}$ denote the number of peers of unit i , i.e., d_i is the row-wise sum of $\mathbf{A}$. Let also $\bar{d} = \max_i d_i$ be the maximum number of peers in the data.

Then, for each individual i and $k = 1, \dots, d_i$, let $\tilde{y}_{i,k}$ denote the outcome of the k -th lowest-performing peer of unit i ; that is, $\tilde{y}_{i,k}$ is the k -th lowest value in the set $\{y_j : a_{i,j} = 1\}$.⁶ Our model for y_i can then be written as:

$$y_i = \sum_{k=1}^{d_i} \beta_{k,d_i} \tilde{y}_{i,k} + \mathbf{x}_i^\top \gamma + \bar{\mathbf{x}}_i^\top \delta + \varepsilon_i, \quad (1)$$

for $i = 1, \dots, n$ where $\bar{\mathbf{x}}_i$ is the average of peers' control covariates⁷ and ε_i is an error term. The peer effect coefficient β_{k,d_i} captures the effect of the peer with the k -th lowest outcome among the d_i peers.⁸ For example, if unit i has three peers, $\beta_{2,3}$ corresponds to the effect of the peer at the median of the peer outcome distribution, whereas $\beta_{1,3}$ and $\beta_{3,3}$ correspond to the effects of the peers with the lowest and highest outcomes, respectively.

The flexibility of our peer effect model in (1) enables us to address key questions in the peer effect literature concerning the complementarity and substitutability of peers' contributions to y_i (see, for example, Kang, 2007). When only the lowest-ranked peer matters (i.e.,

⁵In weighted networks, $a_{i,j,g}$ can take nonnegative values (not necessarily binary) to reflect the intensity of the outgoing link from i to j . The results derived in this paper can also be extended to such networks.

⁶In case of ties, any peer may be selected as the lowest without loss of generality.

⁷We include this average to control for *contextual effects* (Calvó-Armengol et al., 2009), as it is the most common specification in the literature. However, the framework can easily be extended to have a similar structure for peers' covariates as we have for peer outcomes. Furthermore, our identification results do not impose restrictions on δ , allowing for $\delta = 0$.

⁸These effects depend on d_i because, as the number of peers increases, the influence of each individual peer may decrease.

$\beta_{1,d} \neq 0$ and $\beta_{k,d} = 0$ for all $k \geq 2$), the peer effect reflects perfect complementarity, in the sense that an individual's outcome depends entirely on the weakest performer in the group. Conversely, when all coefficients are equal across ranks (i.e., $\beta_{k,d} = \beta_{k',d}$ for all k, k'), peers are perfect substitutes, meaning that every peer outcome contributes equally.

The model of this paper nests many widely used specifications in the literature. For example, the LiM model is a special case of our model, by letting $\beta_{k,d} = \frac{\beta}{d}$ for all k , where β is the single peer effect parameter in the LiM specification, capturing the effect of the peers' average outcome. The linear-in-sums (LiS) model is a special case of our model as well, by letting $\beta_{k,d} = \beta$ for all k , where β is the single peer effect parameter in the LiS specification. In addition, our model nests models where the peer effect is decided by an extreme value, such as the minimal and maximal peer outcome, by letting $\beta_{k,d} = 0$ for every $k > 1$ or $k < d$ (Tao and Lee, 2014; Tatsi, 2015).

While our model does not nest the specification of Boucher et al. (2024), in which the peer effect is linear in a CES aggregation of peer outcomes, it allows us to capture similar patterns as the CES model, as well as patterns that it cannot capture. The CES specification is given by:

$$y_i = \beta \left(\sum_{j=1}^n \frac{a_{i,j}}{d_i} y_j^\rho \right)^{\frac{1}{\rho}} + \mathbf{x}_i^\top \gamma + \bar{\mathbf{x}}_i^\top \delta + \varepsilon_i,$$

where ρ is a substitution parameter. This specification nests the LiM model and allows for non-linear effects. However, there are several patterns of peer effects that it cannot capture, but that our model can. Since the CES function is either concave or convex, the marginal effect of peer outcomes, meaning the first derivative of the CES function with respect to a specific peers outcome, is always either monotonically increasing or decreasing. Specifically, when $\rho > 1$, peers with higher outcomes exert greater influence on individual i , whereas when $\rho < 1$, individual i is influenced more by peers with lower outcomes. This means that it must always be the case that a more extreme peer outcome has a larger impact than a less extreme peer outcome. As such, the CES specification cannot capture non-monotonic peer effects, which arise when individuals are most influenced by peers in the middle of the outcome distribution.

In contrast, our model allows for both monotonic and non-monotonic marginal peer effects, as can be seen through a simple example, the linear-in-median model.⁹ In this model, a peer's outcome has a marginal peer effect equal to zero when it is so high or low that the peer is not the median, but is non-zero when it is the median outcome among i 's peers.

⁹To map the linear-in-median model to our framework, we set $\beta_{k,d} = \beta^{\text{med}}$ whenever d is odd and $k = \{\frac{d+1}{2}\}$, and $\beta_{k,d} = \frac{\beta^{\text{med}}}{2}$ when d is even and $k \in \{\frac{d}{2}, \frac{d}{2} + 1\}$, and set $\beta_{k,d} = 0$ otherwise.

Importantly, the CES function, or any aggregator discussed in [Boucher et al. \(2024\)](#), cannot allow for a pattern of marginal effects where neither the minimum nor maximum matters while the total peer effect is non-zero. These types of non-monotonic marginal effects of peer outcomes are natural in many settings with spillovers as they align with theories of social comparison, which suggest that aligning with moderately performing peers, rather than with outliers, can be socially or psychologically optimal ([Festinger, 1954](#); [Rogers and Feller, 2016](#); [Diel and Hofmann, 2019](#)).

Imposing monotonicity, as in the CES and LiM specifications, when peer effects are non-monotonic can have important policy implications. Since peer effects generate social multipliers, they are often leveraged to increase the effectiveness of interventions. For example, optimally assigning treatment within a network can substantially increase policy impact ([Galeotti et al., 2020](#)). Likewise, identifying key players can facilitate the efficient diffusion of information or behaviors throughout the network ([Ballester et al., 2006](#)). The optimal targeting strategy, however, depends critically on the underlying structure of peer effects. In the example above, where peers at the tails of the outcome distribution exert less influence, the most effective individuals to target are likely to be those whose outcomes lie closer to the center of the distribution. In contrast, the CES and LiM specifications instead predict that peers at the tails of the outcome distribution are the most influential and therefore the optimal targets for intervention.

Similarly, researchers have used peer effects to motivate interventions in the network directly. A cautionary tale of such interventions is [Carrell et al. \(2013\)](#), who estimated a linear-in-means using data on air force cadets to create optimal peer groups. However, they then found negative effects for individuals assigned to these optimal groups. The authors attribute this to endogenous network formation. Our model suggests it may also be due to the underlying peer effect model being misspecified as this would make the treatment groups non-optimal, even if the network formation behavior was not affected by treatment.

In our empirical application, we find non-monotonic peer effect patterns in over half of the student outcomes we study. We then conduct a counterfactual analysis to identify key players within the network and find that the CES and LiM specifications systematically misidentify them. Consequently, our heterogeneous peer effect specification is not simply a more granular representation of aggregate peer effects. It instead recovers features of social interactions that are ruled out by existing models, and therefore leads to qualitatively different policy prescriptions. Accounting for this heterogeneity is essential. Not just for accurately characterizing peer influence, but more importantly, to support the design of effective network-based interventions.

Furthermore, our model allows the outcome to take any value, while the CES specification

requires strictly positive outcomes. Additionally, there is no ‘return-to-scale’ element in the CES aggregator.

2.1 Microfoundations

In this section, we present microfoundations for our peer effect model introduced above using a complete-information game. In this game, each individual i selects a strategy y_i (e.g., academic effort) and derives utility from the distribution of their peers’ strategies. Preferences are represented by a standard linear-quadratic utility function, which is commonly used in the context of the linear-in-means (LiM) model (see [Ballester et al., 2006](#); [Calvó-Armengol et al., 2009](#)), in which we allow for each peer to exert a specific influence. The utility function is given by:

$$U_i(y_i, \mathbf{y}_{-i}) = \underbrace{\alpha_i \cdot y_i - \frac{1}{2}y_i^2}_{\text{private benefit}} + \underbrace{\sum_{k=1}^{d_i} \beta_{k,d_i} \tilde{y}_{i,k} \cdot y_i}_{\text{social benefit}}, \quad (2)$$

where α_i is individual i ’s type observable to all players, and $\mathbf{y}_{-i} = (y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_n)^\top$. The type α_i is generally specified as a linear function of observable characteristics $\mathbf{x}_i$ and an error term (e.g., see [Blume et al., 2015](#)).

The utility function is additively separable into a private benefit, which depends only on individual i ’s own effort y_i , and a social benefit, which is a function of both y_i and the effort vector of other individuals, $\mathbf{y}_{-i}$. The private utility is linear in the individual’s type α_i , and the private cost (or disutility) of exerting effort y_i , $\frac{1}{2}y_i^2$, is quadratic in y_i . The social benefit captures spillover effects on an individual’s effort through the peer effect $\sum_{k=1}^{d_i} \beta_{k,d_i} \tilde{y}_{i,k}$, a weighted sum of ranked peer efforts with the peer effect coefficients $\{\beta_{k,d_i}\}_{k=1}^{d_i}$ as weights. The parameters $\{\beta_{k,d_i}\}_{k=1}^{d_i}$ govern the degree of substitutability among peer efforts in the peer effect construction as discussed above, and also the direction of spillovers onto individual i ’s own effort. Specifically, when individual j ’s effort is ranked k -th among i ’s peers, the sign of β_{k,d_i} determines the local spillover: $\beta_{k,d_i} > 0$ indicates that individual j ’s effort raises individual i ’s incentive to exert effort, while a negative weight indicates that it discourages effort. Crucially, since the weight applied to individual j ’s effort depends on its rank, the spillover from any given peer is not fixed but varies with the efforts of all other peers.

The utility function shown in Equation (2) assumes that social benefits take the form of *spillover*. We can further extend this by adding *conformity* to the social benefit as well. While the extension yields a similar reduced-form equation, the implication for counterfactual analysis is not the same (see [Boucher and Fortin, 2016](#)). In Appendix B.1, we discuss the

extension where both spillover and conformity coexist. As in [Boucher et al. \(2024\)](#), we require the presence of isolated individuals (individuals without friends) to estimate the shares of spillover and conformity effects.

Individual i chooses their effort y_i by maximizing the utility function $U_i(y_i, \mathbf{y}_{-i})$. By solving the first-order condition of this maximization problem, we obtain the best response function:

$$BR_i(\mathbf{y}_{-i}) = \sum_{k=1}^{d_i} \beta_{k,d_i} \tilde{y}_{i,k} + \alpha_i. \quad (3)$$

By defining $\alpha_i = \mathbf{x}_i^\top \gamma + \bar{\mathbf{x}}_i \delta + \varepsilon_i$, this best response function corresponds to our peer effect specification in Equation (1). A strategy profile $\mathbf{y} = (y_1, \dots, y_n)^\top$ is a Nash equilibrium if $y_i = BR_i(\mathbf{y}_{-i})$ for all i . To establish the existence and the uniqueness of the Nash equilibrium, we impose the following condition.

Assumption 1. *The peer effect parameters satisfy the condition $\max_{d \leq \bar{d}} \sum_{k=1}^d |\beta_{k,d}| < 1$.*

Assumption 1 ensures that a one-unit increase in peer outcomes does not lead to an increase in individual i 's effort greater than one, regardless of their number of friends d_i . This requirement is standard in peer effect models to establish the existence and uniqueness of the Nash equilibrium. In the LiM model, an equivalent condition imposes that the absolute value of the (single) peer effect parameter is strictly less than one (see [Bramoullé et al., 2009](#)).

Proposition 1. *Assume that the preferences of each individual are characterized by the utility function (2). Under Assumption 1, there exists a unique Nash equilibrium $\mathbf{y}^* = (y_1^*, \dots, y_n^*)^\top$ such that $y_i^* = BR_i(\mathbf{y}_{-i}^*)$.*

Proof. See Appendix A.1. □

The key part of the proof of Proposition 1 is to show that the mapping

$$BR(\mathbf{y}) = (BR_1(\mathbf{y}_{-1}), \dots, BR_n(\mathbf{y}_{-n}))^\top$$

is a contraction. That is, for any two strategy profiles $\mathbf{y}, \mathbf{y}' \in \mathbb{R}^n$, $|BR(\mathbf{y}) - BR(\mathbf{y}')| \leq \bar{\beta} \|\mathbf{y} - \mathbf{y}'\|_\infty$, for some constant $\bar{\beta} < 1$.¹⁰ Under this condition, BR has a unique fixed point, which corresponds to the Nash equilibrium.

¹⁰For any $\mathbf{u} = (u_1, u_2, \dots, u_n) \in \mathbb{R}^n$, the infinity norm of $\mathbf{u}$ is given by $\|\mathbf{u}\|_\infty = \max_{1 \leq i \leq n} |u_i|$.

2.2 Empirical Specification

The number of peer effect parameters in the fully saturated model is $\frac{\bar{d}(\bar{d}+1)}{2}$, which is challenging for estimation in most real-world applications where the maximum number of peers, $\bar{d}$, is large. This is the case in our empirical application where $\bar{d} = 10$, which means we would have 55 endogenous variables if we want to estimate the fully saturated model. As this would make both interpretation and estimation challenging, we impose restrictions on the parameter space to reduce the number of parameters. While there are many ways to do this,¹¹ we will focus on equally spaced, linearly interpolated quantiles as they are highly interpretable, easy to implement, and used extensively in economics.

The quantile version of our model can be written as:

$$y_i = \sum_{k=1}^M \beta^k \tilde{y}_i^k + \mathbf{x}_i^\top \gamma + \bar{\mathbf{x}}_i^\top \delta + \varepsilon_i, \quad (4)$$

where $\tilde{y}_i^k$ is the $(k-1)/(M-1)$ quantile among the peer outcomes for individual i for $k = 1, \dots, M$.¹² For example, when there are three peers, i.e. $d_i = 3$, and three quantile variables, i.e. $M = 3$, we have $\tilde{y}_1^1 = \tilde{y}_{i,1}$, $\tilde{y}_1^2 = \tilde{y}_{i,2}$ and $\tilde{y}_1^3 = \tilde{y}_{i,3}$, corresponding to the minimum, median and maximum peer outcomes. When $d_i \neq M$, we linearly interpolate.¹³ Appendix E discusses the specifics of the linear interpolation in constructing $\{\tilde{y}_i^k\}_{k=1}^M$ in a generalized manner.

There are several merits to the quantile model. Firstly, the quantile-based model is nested in the fully saturated version of our peer effect model (1), since $\{\tilde{y}_i^k\}_{k=1}^M$ are defined to be linear combinations of $\{\tilde{y}_{i,k}\}_{k=1}^{d_i}$. Therefore, the microfoundation from Section 2.1 and the identification results that will be discussed in Section 3.1 naturally extend to the quantile-based model as well. Secondly, quantiles capture important features of the ordered peer outcomes $\{\tilde{y}_{i,k}\}_{k=1}^{d_i}$, the median captures centrality, while the minimum and maximum capture tail behavior, etc. Thirdly, with equally spaced quantile levels, the peer outcomes that are not included in the restricted model lie between the chosen quantiles, making the information loss minimal. In our empirical application, we show that the quantile specification can reveal many interesting patterns of peer effects.

¹¹For example, there are seven different ways of calculating quantiles in the literature (Hyndman and Fan, 1996), all of which can be implemented using our results. The model also allows for non-quantile restrictions, such as imposing different linear slopes for different sets of peers. Since our model also nests the linear-in-means model, one could use such specifications to construct tests of the linear-in-means assumption.

¹²When $M = 1$, $\tilde{y}_i^1 = \tilde{y}_{i,1}$.

¹³For example, to compute the 0.3 quantile among peer outcomes for individual i such that $d_i = 5$, first we find that $\tilde{y}_{i,2}$ is the 0.25 quantile and $\tilde{y}_{i,3}$ is the 0.5 quantile. Then, 0.3 quantile can be computed as $(0.8 \cdot \tilde{y}_{i,2} + 0.2 \cdot \tilde{y}_{i,3})$, since 0.3 lies 20% of the way from 0.25 to 0.5.

The choice of whether to use a quantile model or a different specification is inherently context-dependent. Even within the quantile specification, it is still necessary to pick the number and the locations of the included quantiles. We therefore develop a specification test to help practitioners make these decisions in Section 3.2.

3 Identification and Inference

In this section, we discuss the identification of the structural parameter $(\{\{\beta_{k,d}\}_{k=1}^d\}_{d=1}^{\bar{d}}, \gamma, \delta)$ in terms of network-level moments, under some normalization due to uneven network sizes n_g . As customary in the literature, we assume independence across networks and assume that the network adjacency matrix and the network-level collection of individual-level control covariates are exogenous. To provide empirical guidance on the tractable specification in Section 2.2, we develop an encompassing test.

3.1 Identification

To simplify notation, we use $\mathbb{E}_g$ to denote the conditional expectation of a random variable given exogenous variables from the same network throughout the paper. For an arbitrary random variable w_g in network g , we write

$$\mathbb{E}_g[w_g] := \mathbb{E}[w_g | \mathbf{X}_g, \mathbf{A}_g]$$

where $\mathbf{X}_g = (\mathbf{x}_{1,g} \ \cdots \ \mathbf{x}_{n_g,g})^\top$. Note that $\mathbb{E}_g[w_g] = 0$ also implies

$$\mathbb{E}[w_g | \{\mathbf{x}_{i,g}, \bar{\mathbf{x}}_{i,g}, a_{i,j,g}\}_{1 \leq i,j \leq n_g}] = 0$$

as $\bar{\mathbf{x}}_{i,g}$ is a deterministic function of $\mathbf{x}_{i,g}$ and $\mathbf{A}_g$. Also, we vectorize the peer effect coefficients and the ranked peer outcomes

$$\begin{aligned} \beta &= (\beta_{1,1} \ \beta_{2,1} \ \beta_{2,2} \ \cdots \ \beta_{1,\bar{d}} \ \cdots \ \beta_{\bar{d},\bar{d}})^\top \\ \tilde{\mathbf{y}}_{i,g} &= (0 \ \cdots \ 0 \ \tilde{y}_{i,1,g} \ \cdots \ \tilde{y}_{i,d_i,g} \ 0 \ \cdots \ 0)^\top \end{aligned}$$

and let $\mathbf{w}_{i,g} = (\tilde{\mathbf{y}}_{i,g}^\top, \mathbf{x}_{i,g}^\top, \bar{\mathbf{x}}_{i,g}^\top)^\top \in \mathbb{R}^{\frac{\bar{d}(\bar{d}+1)}{2}+2l}$. We also denote by $\theta = (\beta^\top, \gamma^\top, \delta^\top)^\top$ the vector of all parameters. The rank-dependent peer effect model (1) can be rewritten as follows:

$$y_{i,g} = \tilde{\mathbf{y}}_{i,g}^\top \beta + \mathbf{x}_{i,g}^\top \gamma + \bar{\mathbf{x}}_{i,g}^\top \delta + \varepsilon_{i,g} = \mathbf{w}_{i,g}^\top \theta + \varepsilon_i \quad (5)$$

for $i = 1, \dots, n_g$ and in matrix form:

$$\mathbf{y}_g = \tilde{\mathbf{Y}}_g \beta + \mathbf{X}_g \gamma + \bar{\mathbf{X}}_g \delta + \boldsymbol{\varepsilon}_g = \mathbf{W}_g \begin{pmatrix} \beta^\top & \gamma^\top & \delta^\top \end{pmatrix}^\top + \boldsymbol{\varepsilon}_g$$

with $\tilde{\mathbf{Y}}_g = \begin{pmatrix} \tilde{\mathbf{y}}_{1,g} & \dots & \tilde{\mathbf{y}}_{n_g,g} \end{pmatrix}^\top$, $\bar{\mathbf{X}}_g = \begin{pmatrix} \bar{\mathbf{x}}_{1,g} & \dots & \bar{\mathbf{x}}_{n_g,g} \end{pmatrix}^\top$, $\mathbf{W}_g = \begin{pmatrix} \mathbf{w}_{1,g} & \dots & \mathbf{w}_{n_g,g} \end{pmatrix}^\top$ and $\boldsymbol{\varepsilon}_g = \begin{pmatrix} \varepsilon_{1,g} & \dots & \varepsilon_{n_g,g} \end{pmatrix}^\top$.

Assumption 2. (EXOGENOUS NETWORK AND COVARIATES) *For each $g = 1, \dots, m$ and $i = 1, \dots, n_g$*

$$\mathbb{E}_g[\varepsilon_{i,g}] = \mathbb{E}[\varepsilon_{i,g} | \mathbf{X}_g, \mathbf{A}_g] = 0.$$

Assumption 2 rules out endogenous network formation and treats the network structure as fixed, meaning it does not depend on the randomness in $\{\varepsilon_{i,g}\}_{i=1}^{n_g}$. While there have been advances in analyzing peer effect models with endogenous network formation (Goldsmith-Pinkham and Imbens, 2013; Hsieh and Lee, 2016; Johnsson and Moon, 2021; Jochmans, 2023), we instead focus on endogeneity in peer effects conditional on the network. To see this, note that we can view our model as a second step of a two-stage process where firstly, the network is formed and secondly, the peer effect is determined based on the connections established in the first step. Models of endogenous network formation address endogeneity in the first step, whereas our model addresses endogeneity in the second step by allowing peer effect coefficients to depend on $\{\varepsilon_i\}_{i=1}^n$.

This endogeneity means that, although $\mathbb{E}_g[\mathbf{y}_g]$ admits a linear representation, the quantities entering the representation are not known functions of $(\mathbf{X}_g, \mathbf{A}_g)$. This is in contrast to the LiM model, where $\mathbb{E}_g[\mathbf{y}_g]$ is linear in $\mathbf{X}_g, \mathbf{G}_g \mathbf{X}_g, \mathbf{G}_g^2 \mathbf{X}_g, \dots$ with $\mathbf{G}_g = [a_{i,j,g}/d_{i,g}]$.¹⁴ This is due to the rank-dependence of the peer-effect coefficients. Consider a simple network with three individuals where individual 1 is friends with both individuals 2 and 3, but individuals 2 and 3 are only friends with individual 1.¹⁵ From Assumption 2, we obtain the following:

$$\begin{aligned} \mathbb{E}_g[y_1] &= \beta_{1,2} \mathbb{E}_g[y_2 \mathbf{1}\{y_2 \leq y_3\} + y_3 \mathbf{1}\{y_2 > y_3\}] \\ &\quad + \beta_{2,2} \mathbb{E}_g[y_3 \mathbf{1}\{y_2 \leq y_3\} + y_2 \mathbf{1}\{y_2 > y_3\}] + \mathbf{x}_1^\top \gamma + \bar{\mathbf{x}}_1^\top \delta. \end{aligned}$$

¹⁴When $d_{i,g} = 0$, we follow convention and set $g_{i,g} = 0$.

¹⁵The network adjacency matrix is

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Substituting $\beta_{1,1}y_1 + \mathbf{x}_2^\top \gamma + \bar{\mathbf{x}}_2^\top \delta + \varepsilon_2$ for y_2 and similarly for y_3 , we get:

$$\begin{aligned} \tilde{\beta} \mathbb{E}_g[y_1] &= \beta_{1,2} \mathbb{E}_g[\varepsilon_2 \mathbf{1}\{y_2 \leq y_3\} + \varepsilon_3 \mathbf{1}\{y_2 > y_3\}] \\ &\quad + \beta_{2,2} \mathbb{E}_g[\varepsilon_3 \mathbf{1}\{y_2 \leq y_3\} + \varepsilon_2 \mathbf{1}\{y_2 > y_3\}] + \sum_{i=1}^3 \left(\mathbf{x}_i^\top \tilde{\gamma}_i + \bar{\mathbf{x}}_i^\top \tilde{\delta}_i \right) \end{aligned}$$

for some $\tilde{\beta}$, $\{\tilde{\gamma}_i\}_{i=1}^3$ and $\{\tilde{\delta}_i\}_{i=1}^3$. Without further assumptions on the distribution of $\{\varepsilon_i\}_{i=1}^n$, the quantities such as $\mathbb{E}_g[\varepsilon_2 \mathbf{1}\{y_2 \leq y_3\}]$ cannot be written as a function of observables $(\mathbf{X}_g, \mathbf{A}_g)$. An exception occurs when $\beta_{1,2} = \beta_{2,2}$, in which case the conditional expectations cancel out, reverting back to the LiM or LiS model. By the same argument, the quantities entering the reduced-form representation of $\mathbb{E}_g[\mathbf{W}_g]$ in our model depend on the distribution of $\varepsilon_{i,g}$ in a way that resists closed-form characterization without further distributional assumptions.¹⁶ For this reason, Assumption 3 is stated directly in terms of $\mathbb{E}_g[\mathbf{W}_g]$ rather than in terms of low-level conditions on the primitives.

The structural parameter θ of our model (β, γ, δ) , or $(\{\beta^k\}_{k=1}^M, \gamma, \delta)$ in a quantile model, is identified under the following conditions.

Assumption 3.

- a. $\varepsilon_g \mid (n_g, \mathbf{X}_g, \mathbf{A}_g) \sim iid$
- b. $\text{plim}_{m \rightarrow \infty} \frac{1}{m} \sum_{g=1}^m \mathbb{E}_g[\mathbf{W}_g]^\top \mathbb{E}_g[\mathbf{W}_g]$ is nonsingular.

Theorem 1 formally establishes the identification result.

Theorem 1. *Suppose that Assumptions 1-3 hold. Then, β, γ and δ are identified from the moment condition below:*

$$\text{plim}_{m \rightarrow \infty} \frac{1}{m} \sum_{g=1}^m \mathbb{E}_g[\mathbf{W}_g]^\top \mathbb{E}_g[\mathbf{W}_g] \theta = \text{plim}_{m \rightarrow \infty} \frac{1}{m} \sum_{g=1}^m \mathbb{E}_g[\mathbf{W}_g]^\top \mathbb{E}_g[\mathbf{y}_g] \quad (6)$$

Proof. See Appendix A.2. □

Part b. of Assumption 3 provides a high-level condition for the identification of the rank-dependent peer effect model. Deriving a low-level sufficient condition for Part b. of Assumption 3 is nontrivial due to the rank dependence of the peer-effect coefficients. However, we believe it will still hold in many empirically relevant cases. Note that for this

¹⁶We show and discuss the closed-form characterization of our model in Appendix B.4.

assumption to not hold means that $\mathbb{E}_g[\tilde{\mathbf{y}}_{i,g}]$, $\mathbf{x}_{i,g}$, and $\bar{\mathbf{x}}_{i,g}$ are linearly dependent. If $\mathbf{x}_{i,g}$ and $\bar{\mathbf{x}}_{i,g}$ are not linearly dependent, which can be easily verified, then this must mean that

$$\mathbb{E}_g[\tilde{\mathbf{y}}_{i,g}]^\top \check{\beta} + \mathbf{x}_{i,g} \check{\gamma} + \bar{\mathbf{x}}_{i,g} \check{\delta} = 0, \quad (7)$$

for some parameters $\check{\beta}$, $\check{\gamma}$, and $\check{\delta}$ such that $\check{\beta} \neq 0$. To see when Equation (7) does not hold for all i and g , assume that i has friends of friends, denoted by l , who are not direct friends of i . Changes in the characteristics $\mathbf{x}_{l,g}$ have no influence on $\mathbf{x}_{i,g}$ and $\bar{\mathbf{x}}_{i,g}$ because l is not a direct friend of i . Equation (7) implies that the total variation in $\mathbb{E}_g[\tilde{\mathbf{y}}_{i,g}]^\top \check{\beta}$ following changes in $\mathbf{x}_{l,g}$ is zero. This would not hold in many contexts because individuals whose outcomes determine $\tilde{\mathbf{y}}_{i,g}$ are friends of i , and some of these friends have l as a direct friend. Therefore, if contextual effects matter, changes in $\mathbf{x}_{l,g}$ are expected to influence the outcomes of i 's friends, except for special cases where the influence on each element of the vector $\tilde{\mathbf{y}}_{i,g}$ cancels out in $\mathbb{E}_g[\tilde{\mathbf{y}}_{i,g}]^\top \check{\beta}$. This identification argument is also employed by [Houndetoungan \(2026\)](#) and extends the result in [Bramoullé et al. \(2009\)](#), who use the presence of friends of friends who are not directly friends to nonlinear models.

The presence of friends of friends who are not direct friends provides restriction exclusions to construct instruments for $\tilde{\mathbf{y}}_{i,g}$. The natural instrument, $\mathbb{E}_g[\tilde{\mathbf{y}}_{i,g}]$, cannot be used directly because it is unobserved and cannot be calculated without making distributional assumptions about $\epsilon_{i,g}$.¹⁷ However, as $\mathbb{E}_g[\tilde{\mathbf{y}}_{i,g}]$ is determined by friends' outcomes, we can use the distribution of $\mathbf{x}_{j,g}$ and $\bar{\mathbf{x}}_{j,g}$ for all j who are direct friends of i to construct instruments. The presence of friends of friends who are not direct friends ensures that $\bar{\mathbf{x}}_{j,g}$ captures relevant information excluded from $\mathbf{x}_{i,g}$ and $\bar{\mathbf{x}}_{i,g}$. Given the nonlinear nature of our model, we use the quantiles of these distributions, for example the deciles of $\mathbf{x}_{j,g}$ and $\bar{\mathbf{x}}_{j,g}$ among friends, as instruments. The strength of these instruments can be easily tested in practice using weak instrument tests.

Given the constructed instrument, $\mathbf{Z}_g$, θ can be easily estimated using a two-stage least squares (TSLS) estimator. Consistency and asymptotic normality of the IV estimator in this setting follow from standard arguments for IV estimators. In particular, under Assumption 5, the sample moment matrices formed from $\{\mathbf{Z}_g, \mathbf{W}_g\}_{g=1}^m$ converge to their population counterparts with full rank, and the sample orthogonality condition between instruments and errors holds. As a result, the TSLS estimator is consistent for θ and satisfies a central limit theorem under some regularity conditions. For completeness, Assumption 5 and formal

¹⁷When errors are iid within a network and homoskedastic, $\mathbb{E}_g[\tilde{\mathbf{y}}_{i,g}]$ is the optimal instrument, as shown in [Chamberlain \(1987\)](#). As shown in Appendix B.4, it does not have an easily computable closed-form even when model parameters are known. Furthermore, assuming iid and homoskedastic errors is a very strong assumption in a network setting.

asymptotic results with their proofs are provided in Appendix B.3.

Finally, a popular empirical practice given a network dataset with many networks is to include network fixed-effects α_g , giving us the model:

$$y_{i,g} = \alpha_g + \tilde{\mathbf{y}}_{i,g}^\top \beta + \mathbf{x}_{i,g}^\top \gamma + \bar{\mathbf{x}}_{i,g}^\top \delta + \varepsilon_{i,g}$$

for $i = 1, \dots, n_g$ and $g = 1, \dots, m$. By demeaning at the network level, we obtain

$$y_{i,g} - \bar{y}_g = (\tilde{\mathbf{y}}_{i,g} - \bar{\tilde{\mathbf{y}}}_g)^\top \beta + (\mathbf{x}_{i,g} - \bar{\mathbf{x}}_g)^\top \gamma + (\bar{\mathbf{x}}_{i,g} - \bar{\bar{\mathbf{x}}}_g)^\top \delta + \varepsilon_{i,g} - \bar{\varepsilon}_g$$

for $i = 1, \dots, n_g$ and $g = 1, \dots, m$. Theorem 1 can be directly extended to a setup with network fixed-effects by replacing $\mathbf{W}_g$ with its demeaned counterpart in Part b. of Assumption 3.

3.2 Encompassing Test on Empirical Specifications

In Subsection 2.2, we propose a quantile restriction on the fully saturated rank-dependent peer effect model, for parsimony and tractability. While certain empirical contexts may naturally call for a particular empirical specification, such as the max or min specification, there are many empirical contexts that offer little basis for imposing such *a priori* quantile selections. Therefore, in this subsection, we introduce an encompassing test which provides guidance on the choice of location and number of quantiles, based on Smith (1992), and develop a data-driven specification selection procedure.¹⁸

The encompassing test we propose compares a pair of candidate specifications, treating one as the null hypothesis and the other as the alternative. Importantly, the test does not require the alternative specification to nest the null specification. Moreover, consistency of the test does not depend on the alternative specification being correctly specified. Instead, we use the alternative specification as a lens to evaluate the null specification. Therefore, the procedure can be used to compare non-nested quantile specifications, such as a three-quantile model with quantiles at $\{0, 1/2, 1\}$ and a four-quantile model with quantiles at $\{0, 1/3, 2/3, 1\}$. In this example, the test remains consistent even if the true model is not the four-quantile specification as long as the true model differs from the three-quantile specification from the perspective of the four-quantile specification.

Let h_1 and h_2 denote the two empirical specifications that we compare with the encompassing test, with h_1 being the null specification and h_2 being the alternative. Given h_1 and

¹⁸This test can also be used to test non-quantile variations of our saturated model (1). As we use quantiles in our empirical application, we will focus our discussion on this case.

h_2 , we will have two sets of endogenous variables. Let $\{\mathbf{W}_g^1\}_{g=1}^m$ and $\{\mathbf{W}_g^2\}_{g=1}^m$ denote the explanatory variables and $\{\mathbf{Z}_g^1\}_{g=1}^m$ and $\{\mathbf{Z}_g^2\}_{g=1}^m$ denote the instruments, respectively under h_1 and h_2 . When h_1 and/or h_2 refer to a quantile-based model, the endogenous variables will be constructed from $\{\mathbf{w}_{i,g}\}_{i,g}$ per the quantile construction as discussed in Subsection 2.2.

The encompassing test (Smith, 1992) is implemented as follows:

1. Regress $\mathbf{y}_g$ on $\mathbf{W}_g^1$, using $\mathbf{Z}_g^1$ as instruments:

$$\hat{\theta}^1 = (\mathbf{H}^1)^{-1} \cdot \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{1\top} \mathbf{Z}_g^1 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{1\top} \mathbf{Z}_g^1 \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{1\top} \mathbf{y}_g$$

where $\mathbf{H}^1 = \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{1\top} \mathbf{Z}_g^1 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{1\top} \mathbf{Z}_g^1 \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{1\top} \mathbf{W}_g^1$. $\hat{\theta}^1$ is the TSLS estimator for the model parameter in the null specification.

2. Construct the residuals from the null specification: $\hat{\epsilon}_g^1 = \mathbf{y}_g - \mathbf{W}_g^1 \hat{\theta}^1$, $\forall g = 1, \dots, m$.
3. Regress $\hat{\epsilon}_g^1$ on $\mathbf{W}_g^2$, using $\mathbf{Z}_g^2$ as instruments:

$$\hat{\psi} = (\mathbf{H}^2)^{-1} \cdot \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{2\top} \mathbf{Z}_g^2 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{Z}_g^2 \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \hat{\epsilon}_g^1$$

where $\mathbf{H}^2 = \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{2\top} \mathbf{Z}_g^2 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{Z}_g^2 \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{W}_g^2$.

4. Construct a Wald statistic based on $\hat{\psi}$. Reject the null specification when the resulting Wald statistic exceeds the critical value from an appropriate chi-squared distribution.

$\hat{\epsilon}_g^1$ is the remaining variation in $\mathbf{y}_g$ that is not explained by the patterns of peer effect allowed in the first model specification h_1 . If there are features of the true peer effect model that the first model specification h_1 fails to capture but the second model specification h_2 does capture, $\hat{\epsilon}_g^1$ will be correlated with $\mathbf{W}_g^2$, leading to $\hat{\psi}$ being centered around a nonzero vector.

In Appendix B.5, under some regularity conditions, we formally show that $\hat{\psi}$ is asymptotically normal with a consistently estimable asymptotic variance matrix. This result does not require either of the models h_1 or h_2 to be correctly specified. The asymptotic theory accounts for the estimation error in $\hat{\epsilon}_g^1$ and so does our asymptotic variance estimator. If the true peer effect model cannot be well approximated to the null specification h_1 when assessed through the alternative specification h_2 , $\hat{\psi}$ is centered around a nonzero vector, giving us

consistency of the test.¹⁹

3.3 Comparison to Other Peer Effect Estimands

A natural question, given that our model generalizes the standard LiM model, is to what extent existing estimands recover the key parameters of interest from our model. A commonly accepted minimal standard for such estimands is that they represent weighted averages of the underlying heterogeneity, such as in the analysis of instrument variables (Mogstad and Torgovitsky, 2024) or the analysis of Difference-in-Differences estimands (De Chaisemartin and d'Haultfoeulle, 2020; Goodman-Bacon, 2021). To explore this, define $\tilde{\mathbf{X}}^{\text{LiM}} = \begin{pmatrix} \mathbf{X} & \mathbf{GX} \end{pmatrix}$ and $\tilde{\mathbf{X}}^{\text{LiS}} = \begin{pmatrix} \mathbf{X} & \mathbf{AX} \end{pmatrix}$, with network-level counterparts $\tilde{\mathbf{X}}_g^{\text{LiM}} = \begin{pmatrix} \mathbf{X}_g & \mathbf{G}_g \mathbf{X}_g \end{pmatrix}$ and $\tilde{\mathbf{X}}_g^{\text{LiS}} = \begin{pmatrix} \mathbf{X}_g & \mathbf{A}_g \mathbf{X}_g \end{pmatrix}$, and the population regression coefficients

$$\mathbf{P} = \left(\text{plim}_{m \rightarrow \infty} \frac{1}{m} \sum_{g=1}^m \tilde{\mathbf{X}}_g^{\text{LiM}} \tilde{\mathbf{X}}_g^{\text{LiM} \top} \right)^{-1} \text{plim}_{m \rightarrow \infty} \frac{1}{m} \sum_{g=1}^m \tilde{\mathbf{X}}_g^{\text{LiM}} \mathbf{Z}_g$$

for the LiM and LiS models. We can then define the standard LiM estimand in a just-identified case as follows:

$$\beta^{\text{LiM}} := \left(\frac{1}{m} \sum_{g=1}^m \mathbb{E}_g \left[\left(\tilde{\mathbf{Z}}_g^{\text{LiM}} \right)^\top \mathbf{G}_g \mathbf{y}_g \right] \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbb{E}_g \left[\left(\tilde{\mathbf{Z}}_g^{\text{LiM}} \right)^\top \mathbf{y}_g \right]$$

where

$$\tilde{\mathbf{Z}}_g^{\text{LiM}} = \mathbf{Z}_g - \tilde{\mathbf{X}}_g^{\text{LiM}} \mathbf{P}^{\text{LiM}}$$

given some set of instruments $\{\mathbf{Z}_g\}_{g=1}^m$ and similarly for the LiS model by replacing $\mathbf{G}_g$ with $\mathbf{A}_g$.

Assumption 4.

i There exist instruments $\{\mathbf{Z}_g\}_{g=1}^m$, known functions of $(\mathbf{X}_g, \mathbf{A}_g)$, satisfying

$$\frac{1}{m} \sum_{g=1}^m \mathbb{E}_g \left[\left(\tilde{\mathbf{Z}}_g^{\text{LiM}} \right)^\top \mathbf{G}_g \mathbf{y}_g \right]$$

is nonsingular.

¹⁹The role of the alternative specification h_2 is not to provide a true model, but to provide a criterion to compare the true model with the null specification h_1 .

ii There exist instruments $\{\mathbf{Z}_g\}_{g=1}^m$, known functions of $(\mathbf{X}_g, \mathbf{A}_g)$, satisfying

$$\frac{1}{m} \sum_{g=1}^m \mathbb{E}_g \left[\left(\tilde{\mathbf{Z}}_g^{\text{LiS}} \right)^\top \mathbf{A}_g \mathbf{y}_g \right]$$

is nonsingular.

These conditions are the instrument relevance conditions used for the LiM model and the LiS model, as shown in [Bramoullé et al. \(2009\)](#).

Proposition 2. Under Assumptions [1](#), [2](#) and [4-i](#),

$$\beta^{\text{LiM}} = \sum_{d=1}^{\bar{d}} \sum_{k=1}^d w_{k,d}^{\text{LiM}} d \cdot \beta_{k,d} + o_p(1)$$

where $\sum_{d=1}^{\bar{d}} \sum_{k=1}^d w_{k,d}^{\text{LiM}} = 1$ and $w_{k,d}^{\text{LiM}} \leq 0$. Similarly, under Assumptions [1](#), [2](#) and [4-ii](#),

$$\beta^{\text{LiS}} = \sum_{d=1}^{\bar{d}} \sum_{k=1}^d w_{k,d}^{\text{LiS}} \beta_{k,d} + o_p(1)$$

where $\sum_{d=1}^{\bar{d}} \sum_{k=1}^d w_{k,d}^{\text{LiS}} = 1$ and $w_{k,d}^{\text{LiS}} \leq 0$.

Proof. See Appendix [A.3](#). □

Note that, in the case of LiM misspecification, the weights are applied not directly to $\beta_{k,d}$ but to $d \cdot \beta_{k,d}$. The rank-dependent peer effect coefficient $\beta_{k,d}$ represents the impact of a single peer outcome when there are d peers in total, so it should be rescaled by d to interpret it as a coefficient on a ‘representative’ or ‘average’ peer. For instance, when $\beta_{1,d} = \dots = \beta_{d,d} = \beta$, we have:

$$\sum_{k=1}^d \beta_{k,d} \tilde{y}_{i,k} = d \beta \bar{y}_i,$$

where the coefficient on the average peer outcome is $d\beta_{k,d}$.

Proposition [2](#) demonstrates that both the LiM estimand β^{LiM} and the LiS estimand β^{LiS} are weighted sums of the rank-dependent peer effect coefficients $\{\beta_{k,d}\}_{k,d}$. While the weights sum to one, which is reasonable, there is no guarantee that the weights have the same sign. As a result, it is possible for all $\beta_{k,d}$ to be positive, while the LiM estimand β^{LiM} is negative, and vice versa. The expressions for the weights are provided in the Appendix.

4 Simulation

In this section, we investigate the finite sample properties of our estimators and compare their performance to existing estimators, specifically the CES approach of [Boucher et al. \(2024\)](#), under various data-generating processes. To ensure our approach is as close to real empirical settings as possible, we will design the setup of the simulations to be as close as possible to the empirical application.

As most empirical applications contain many small networks, we generate networks of size $n_g = 50$ and vary the number of networks, m . To generate networks, we generate the links based on a logit model, defining

$$a_{i,j,g} = \mathbf{1}\{-3 + r_{i,j,g} \leq v_{i,j,g}\}$$

with $v_{i,j,g} \stackrel{\text{iid}}{\sim} \text{logit}$ is an unobserved shock and $r_{i,j,g} \sim N(0, 1)$ is a covariate. When we estimate the fully saturated version of our model, we will limit the maximum number of links in the network, $\bar{d}$, to 5. This is achieved by dropping links from individuals with more links than $\bar{d}$ until all individuals have at most $\bar{d}$ links. This procedure generates a network with a roughly uniform degree distribution, with a small share of individuals in each network having zero links. For the results in this section, we fix the networks across simulations.

Our model specification uses the quantile specification from [Section 2.2](#), with equally spaced quantiles. We primarily use four quantiles at $\{0, 1/3, 2/3, 1\}$. This gives the outcome equation

$$y_{i,g} = \sum_{k=1}^4 \beta^k \tilde{y}_{i,g}^k + \gamma_1 x_{1,i,g} + \gamma_2 x_{2,i,g} + \delta_1 \bar{x}_{1,i,g} + \delta_2 \bar{x}_{2,i,g} + \mu_g + \varepsilon_{i,g}$$

where $\tilde{y}_i^k$ is the $(k - 1)/3$ -th quantile of peer outcomes for unit i , $x_{1,i,g}$ and $x_{2,i,g}$ are the two covariates for unit i , and μ_g is the network fixed effect for network g . The fixed effects are drawn from a uniform distribution on the interval $[8, 10]$. This ensures that the outcome is always positive, which is necessary for the CES-estimator to be well defined. The continuous covariate $x_{1,i,g}$ is normally distributed with a mean equal to the fixed effect and a variance equal to one, while the discrete covariate $x_{2,i,g}$ is distributed Poisson with a mean equal to 2. We also include contextual effects $\bar{x}_{1,i,g}$ and $\bar{x}_{2,i,g}$, the average covariate values for all i 's peers, for each of our two covariates. The error term ε_i is generated from a normal distribution with variance equal to 1. The coefficients on the two covariates are $\gamma_1 = 1.5$ and $\gamma_2 = -0.8$, respectively, and the coefficients on the two contextual effects are $\delta_1 = -0.5$ and $\delta_2 = 1.2$,

where the values are set to be similar to the patterns we see in our empirical application.²⁰ To avoid unnecessary repetition in our tables, we will only show results for a subset of β for the saturated model.

We apply both the correctly specified quantile model and the fully saturated rank-dependent model to estimate the peer effect in simulated samples. Since the fully saturated model nests any quantile model, the fully saturated model is also correctly specified. Given a quantile specification, we calculate the corresponding parameter values of $\beta_{k,d}$ as the true values used to calculate bias and MSE for the saturated model.²¹ In the saturated rank model, we use the peer covariates and contextual effects as instruments, ordered by the values of each variable. Since we have two covariates, this yields $4 \sum_{d=1}^{\bar{d}} d$ instruments in total. For the quantile models we instead use the 4 quantiles of the covariates and contextual effects as our instruments, giving a total of 16 instruments.

Table 4.1 gives the results for TSLS for the saturated rank model in Panel A and the quantile model in Panel B.²² Each row gives the values for a given parameter $\beta_{k,d}$ or β^k . Across the columns, we increase the sample by gradually increasing the number of networks from 5 to 100. These sample sizes therefore go from relatively small to approaching larger network datasets, though the networks are still small. For comparison, in our empirical application we will have 141 networks (schools) with an average size of 531.

In Panel A we see the results for the rank model for the parameters $\beta_{k,2}$ and $\beta_{k,5}$. As expected, all of our estimators are consistent, though we do see some small biases for the smaller sample sizes. Interestingly, the parameters $\beta_{k,2}$ are generally better estimated than $\beta_{k,5}$. This is because there are effectively only two endogenous variables to estimate on the subset of $d_i = 2$, while there are five for individuals with $d_i = 5$, but there is roughly a similar number of observations since the degree distribution is roughly uniform. However, for both groups of parameters we see that estimation error, as measured by the MSE, decreases as the number of networks increases.

The results for the saturated model also reveal an interesting dynamic where the precision of the parameters depends on the ranking, with $\beta_{2,5}, \beta_{3,5}, \beta_{4,5}$ having three times the MSE of the endpoints $\beta_{1,5}, \beta_{5,5}$. This is because the identity of an individual's 2nd, 3rd, or 4th ranked friend is more dependent on the shocks compared to the highest or lowest ranked friend. This difference persists even as the sample size increases.

Comparing the quantile to the rank model, we see that the quantile model generally

²⁰Supplementary Appendix C shows all results from this section excluding these contextual effects, i.e., setting $\delta_1 = \delta_2 = 0$.

²¹For $d_i = 1$, $\beta_{1,1} = \beta^0 + \beta^1 + \beta^2 + \beta^3$; for $d_i = 2$ we get $\beta_{1,2} = \beta^0 + 0.66\beta^1 + 0.33\beta^2$ and so on.

²²In Supplementary Appendix C, we show that the OLS estimator is, as expected, biased in both the saturated and quantile versions of our model.

Table 4.1: Finite sample performance of Rank and Quantile models

	Bias					MSE				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Saturated model										
$\beta_{1,2}$	0.007	0.001	0.002	-0.001	0.000	0.010	0.010	0.005	0.002	0.001
$\beta_{2,2}$	0.003	0.005	0.001	0.002	0.000	0.009	0.009	0.004	0.002	0.001
$\beta_{1,5}$	-0.010	0.002	0.010	0.001	-0.001	0.035	0.025	0.040	0.013	0.011
$\beta_{2,5}$	0.023	-0.005	-0.016	-0.000	0.000	0.087	0.083	0.109	0.064	0.072
$\beta_{3,5}$	-0.019	0.008	0.010	-0.006	0.003	0.101	0.130	0.122	0.111	0.110
$\beta_{4,5}$	0.005	-0.003	0.004	0.009	-0.001	0.101	0.103	0.089	0.086	0.061
$\beta_{5,5}$	0.011	0.004	-0.006	-0.003	-0.001	0.047	0.030	0.027	0.018	0.010
Panel B: Restricted model										
β_{τ_1}	0.005	-0.002	0.003	0.000	-0.000	0.019	0.016	0.014	0.006	0.003
β_{τ_2}	-0.006	0.004	-0.008	-0.001	0.001	0.089	0.109	0.090	0.041	0.021
β_{τ_3}	0.004	0.004	0.007	0.002	-0.000	0.104	0.113	0.078	0.038	0.023
β_{τ_4}	-0.000	-0.003	-0.001	-0.001	-0.000	0.024	0.018	0.011	0.006	0.004
n_g	5	10	20	50	100	5	10	20	50	100
$\bar{d}$	5	5	5	5	5	5	5	5	5	5

Notes: The data generating process is the quantile model with $\beta_\tau = (-0.5, 0.35, 0.15, 0.1)$. n_g gives the number of networks used in the simulation and $\bar{d}$ is the maximum number of peers for each individual. All the simulations include network fixed effects. The instruments used is the distribution of the peers covariate values, as measured either by the ranked peer covariate values or the quantiles of peer covariate values. The results are based on 10,000 draws.

performs better in terms of MSE, and improves faster as the sample grows. We see a similar pattern as before when comparing the middle to the extreme quantiles, with β^2, β^3 having a relatively lower precision compared to β^1, β^4 , especially at smaller sample sizes.

4.1 Model Selection

Implementing our estimator requires the researcher to select how many quantiles they want to include in the model. Table 4.2 shows the rejection rates for the encompassing test from Section 3.2. We run these tests for 6 different DGPs, four DGPs with four quantiles as the true model, one with only the min and max peer outcome mattering, and the Linear-in-Mean model.

Focusing first on the first three rows, we see that the model always rejects the null hypothesis of 3 quantiles in the true DGP. Similarly, we almost never reject the null of 4 quantiles, with rejection rates being fairly steady at around 0.2α . This seeming contradiction of conservative coverage but high power likely stems from the large difference between our two hypotheses. This result is not driven by the structure of the peer effects, as can be seen from the remarkably stable results across DGP’s A, B, C, and E. It is also not driven by having 4 as the true number of quantiles in the DGP, as DGP D shows that when the true number of quantiles is 2, the test does not reject keeping 3 quantiles over 4.

In total, these results indicate that as long as the true DGP is some number of equi-spaced quantiles, this test is capable of distinguishing the correct number with a large degree of power. However, this is not the case when the model is misspecified, as is the case in DGP F. To capture the mean peer outcome we would need an infinite amount of quantiles, which

Table 4.2: Monte Carlo Simulations — Encompassing tests

2 qtls. vs. 3 qtls.		3 qtls. vs. 4 qtls.		4 qtls. vs. 5 qtls.	
5%	10%	5%	10%	5%	10%
DGP A: $\beta = (0, 0.05, 0.2, 0.3)$					
1.000	1.000	0.105	0.180	0.011	0.020
DGP B: $\beta = (0.3, 0.2, 0.05, 0)$					
1.000	1.000	0.914	0.953	0.013	0.026
DGP C: $\beta = (0, 0.275, 0.275, 0)$					
1.000	1.000	0.994	0.998	0.009	0.021
DGP D: $\beta = (0.275, 0, 0, 0.275)$					
0.011	0.023	0.008	0.017	0.013	0.025
DGP E: $\beta = (-0.05, 0.35, 0.15, 0.1)$					
1.000	1.000	1.000	1.000	0.009	0.021
DGP F (LIM model): $\beta = 0.55$					
1.000	1.000	0.961	0.984	0.442	0.574

The columns labeled “ a qtls. vs. b qtls.,” for integers a and b , report the share of rejections of the null hypothesis that the model with a quantile levels does not perform worse than the model with b quantile levels, at the significance levels indicated in the second row.

is clearly infeasible. As one may therefore expect the test consistently prefers more quantiles.

4.2 Comparison to the CES Estimator

Our next simulation compares the ability of our estimator to capture different patterns of peer effects in comparison to the other two main estimators in the literature, the LiM model and the CES estimator of [Boucher et al. \(2024\)](#). We will do this by simulating data from our DGP and then estimating the quantile version of our model, the CES model, and the LiM model on the simulated data. Since we do not estimate the saturated model in this exercise, we no longer limit the maximum number of links, $\bar{d}$, to 5. Other than that, the network formation model is the same as in the rest of this section. In most simulations, most individuals have between 0 and 3 links, with the maximum number of links being 12. This is similar to the degree distributions we see in our empirical application.

Table 4.3 shows the results from estimating our quantile estimator and the CES estimator on the same data. DGP A and B both have monotonically increasing or decreasing peer effects. Our estimator fully captures these effects. These effects are also well captured by the CES aggregator, which is able to correctly assign $\rho > 1$ for the increasing case and $\rho < 1$ for the decreasing case. In both cases, however, it still is not able to fully capture the total effect coming from the peer effects, which equals 0.55 across all our DGPs. The LiM model similarly struggles for both these DGPs, either underestimating or overestimating the total spillover in the data.

DGP C and D instead put all the weight on either the middle or extremes of the distribution. Importantly, our estimator is fully able to estimate the zeros in the parameter space, correctly identifying the parts of the peer outcome distributions that have no effect on an individual’s own outcome. Interestingly, the LiM model performs much better for these symmetric cases, implying the LiM model is more able to approximate the true peer effects when they are symmetric, with the peers around the median having the largest weight. The CES estimator, however, struggles with these cases. Since neither of these cases has concave or convex peer effects, the assumptions of the model do not hold. This leads to the CES estimator wrongly estimating that higher peer outcomes are more important in DGP C ($\rho > 1$), and that lower peer outcomes are more important in DGP D ($\rho < 1$), despite this not being the case for either DGP.

DGP E introduces peer effects of different signs, by having the lowest peer outcome have a negative effect while the other quantiles have a positive effect. Our estimator is fully able to capture this pattern, and since the majority of the weight is on the middle of the distribution, the LiM model also performs decently. The CES estimator wrongly fits a positive ρ , implying

Table 4.3: Monte Carlo Simulations — Estimation of the Quantile Specification

Quantiles				LIM	CES	
β^1	β^2	β^3	β^4	β	ρ	β
DGP A: $\beta = (0, 0.05, 0.2, 0.3)$						
-0.000 (0.005)	0.050 (0.023)	0.200 (0.044)	0.300 (0.028)	0.386 (0.020)	30.057 (8.914)	0.562 (0.011)
DGP B: $\beta = (0.3, 0.2, 0.05, 0)$						
0.300 (0.008)	0.200 (0.033)	0.050 (0.058)	-0.000 (0.033)	0.715 (0.022)	-6.383 (9.338)	0.502 (0.029)
DGP C: $\beta = (0, 0.275, 0.275, 0)$						
0.000 (0.006)	0.275 (0.025)	0.275 (0.048)	0.000 (0.030)	0.476 (0.013)	3.057 (0.275)	0.554 (0.010)
DGP D: $\beta = (0.275, 0, 0, 0.275)$						
0.275 (0.006)	-0.000 (0.028)	-0.000 (0.052)	0.275 (0.031)	0.635 (0.014)	-1.714 (0.253)	0.511 (0.014)
DGP E: $\beta = (-0.05, 0.35, 0.15, 0.1)$						
-0.050 (0.006)	0.350 (0.024)	0.150 (0.047)	0.100 (0.029)	0.434 (0.016)	5.004 (0.575)	0.561 (0.010)
DGP F (LIM model): $\beta = 0.55$						
0.112 (0.006)	0.196 (0.028)	0.098 (0.052)	0.144 (0.031)	0.550 (0.008)	1.002 (0.130)	0.550 (0.010)

The models are simulated and estimated 1,000 times. Values without parentheses represent average peer effect estimates, while those in parentheses correspond to standard errors. The instrument matrix includes the quantiles of $\mathbf{x}$ and $\bar{\mathbf{x}}$ among friends, computed at ten levels uniformly spaced between 0 and 1. DGPs A–E are generated from the proposed quantile-based model, with four quantiles at $\{0, 1/3, 2/3, 1\}$, and $\beta = (\beta^0, \beta^1, \beta^2, \beta^3)$ is the vector of peer effects at each quantile. DGP F follows the standard LIM model with only spillover effects, where $\beta = 0.55$. All estimations account for unobserved sub-network heterogeneity using fixed effects.

peers with higher outcomes are more important for the spillover. This is likely due to the sharp increase in the peer effect from the lowest outcome friend (β^1) to the 33% quantile (β^2). However, it completely contradicts the true DGP where the most influential peers are those at the second quantile.

Finally, we simulate the data from a LiM model. As expected, both the LiM model and the CES model are able to fully capture the spillover, and the CES correctly identifies that $\rho = 1$, i.e., a LiM model. Though the quantile version of our model does not nest the LiM model, it still estimates a roughly uniform effect of peers. This is an important reminder

that restrictions can reduce the performance of our model if the added restrictions are not correct, though the model can still recover meaningful results under misspecifications.

5 Empirical Application

In this section, we present an empirical application using the Wave I dataset from the National Longitudinal Study of Adolescent to Adult Health (Add Health). We examine several outcomes and show that the quantile model captures diverse patterns of peer effects that challenge existing specifications. To illustrate the effectiveness of our approach, we also conduct a counterfactual analysis, demonstrating that ignoring these patterns can lead to the misidentification of key players and reduce the impact of targeted policies.

5.1 Add Health Data

Wave I of the Add Health survey provides nationally representative and detailed information on 7th–12th graders from 144 schools during the 1994–1995 school year in the United States (US). Approximately 90,000 students completed a questionnaire covering demographics, family background, academic performance, health-related behaviors, and friendship links. Each respondent could nominate up to five male and five female best friends within the same school.²³

We study 11 outcomes, including grade point average (GPA), academic effort, participation in extracurricular activities, future expectations, trouble at school, smoking, drinking, risky behaviors, self-esteem, physical exercise, and fighting. All of these outcomes, except for future expectations, have been studied by [Boucher et al. \(2024\)](#) using the CES model. The future expectations outcome is constructed as the sum of binary indicators for whether students believe they will live to age 35, avoid HIV/AIDS, graduate from college, and have a middle-class family income by age 30. We control for several exogenous variables, including student age, grade, sex, race, Hispanic ethnicity (Spanish-speaking), and mother’s education and employment. We also control for contextual variables, defined as the average of the exogenous variables among a student’s friends.

²³The dataset has some limitations that merit discussion. First, some referred friend identifiers are missing and are therefore removed from the network. Recent literature proposes methods to address this issue; however, most of these approaches focus on linear-in-means models and cannot be easily extended to nonlinear settings (e.g., [Lewbel et al., 2023](#); [Herstad, 2023](#); [Boucher and Houndetoungan, 2025](#)). Second, the observed degree is censored, as students can nominate up to ten friends. However, only 1.12% of students reach this maximum. To maintain focus on the main objective of the paper, we do not address this censoring issue. Despite these limitations, it is worth noting that the Add Health dataset remains the most comprehensive network dataset that is currently available for studying peer effects.

After removing observations with missing values, the final sample comprises approximately 75,000 students from 141 schools. The average number of friends per student is 3.47; 22% of students have no friends, and approximately 64% have four friends or fewer.

5.2 Empirical Results

The estimation results for the quantile, LIM, and CES models are summarized in Table 5.1, and the specification tests are presented in Table 5.2.²⁴ The quantile levels are evenly spaced from 0 to 1. The p -values for the encompassing test in Table 5.2 indicate that the specification with three quantiles is often rejected in favor of the one with four quantiles, whereas the specification with four quantiles is generally not rejected against the one with five quantiles. Two exceptions, however, are the outcomes *trouble at school* and *risky behaviour*, for which the specification with four quantiles is still rejected in favor of the one with five quantiles at the 5% level. We consider a specification with four quantile levels throughout the main empirical analysis and present results based on the five-quantile specification in Table D.2 in Appendix D.²⁵

The CES specification restricts peer effects to be monotone across quantile levels. To compare the quantile specification results to those of the LIM and CES models, we perform monotonicity tests (see Table 5.2). Specifically, we test whether peer effects are uniform, increasing, or decreasing across quantile levels. For the increasing (respectively decreasing) test, the null hypothesis is that $\beta^1 \leq \dots \leq \beta^4$ (respectively $\beta^1 \geq \dots \geq \beta^4$). We test these inequality restrictions using a Wald-type criterion (see [Kodde and Palm, 1986](#)).

Our instruments include the quantiles of $\mathbf{x}$ and $\bar{\mathbf{x}}$ among friends and friends of friends, computed at ten levels uniformly spaced between 0 and 1.²⁶ Given that we have several endogenous variables, we assess instrument strength using the rank Wald test proposed by [Kleibergen and Paap \(2006\)](#). This test yields large test statistics, suggesting that the model does not suffer from weak instrument problems.

The decomposition of peer effects across peer outcome quantiles reveals several patterns. A prevalent pattern is one in which mid-level outcomes tend to be the most influential, while

²⁴Note that the estimates for the CES model differ from those reported by [Boucher et al. \(2024\)](#) because their model does not control for contextual effects.

²⁵As a robustness check, we also estimate a four-quantile model on the subsample of nonisolated students; that is, students who have at least one friend (see Table D.2 in Appendix D). Our results are robust to the removal of isolated students.

²⁶This procedure generates approximately 600 instruments. This is less than 1% of the size of our dataset in every specification considered, meaning the many-instrument bias discussed in the literature ([Bekker, 1994](#); [Newey and Smith, 2004](#)) should not be present. As a sanity check, we also compare our estimates to the OLS estimates in the Supplementary Appendix Table D.3 and find that they are statistically significantly different from the IV estimates for every outcome.

Table 5.1: Empirical Results

β^1	Quantile β^2	β^3	β^4	LIM β	CES ρ	β
Academic achievements (GPA)						
0.071 (0.045)	0.147 (0.081)	0.637 (0.089)	-0.122 (0.054)	0.804 (0.048)	0.552 (0.683)	0.801 (0.042)
Academic effort						
0.143 (0.032)	0.158 (0.061)	0.140 (0.054)	0.093 (0.046)	0.643 (0.116)	-4.630 (4.866)	0.500 (0.114)
Extracurricular activities						
-0.081 (0.082)	0.551 (0.121)	0.239 (0.080)	-0.007 (0.021)	0.805 (0.053)	-0.171 (0.397)	0.699 (0.035)
Future perception						
0.155 (0.033)	0.116 (0.087)	0.133 (0.115)	0.078 (0.068)	0.590 (0.061)	0.621 (0.776)	0.577 (0.059)
Trouble at school						
0.028 (0.075)	0.249 (0.112)	0.265 (0.073)	0.031 (0.041)	0.780 (0.109)	0.335 (0.361)	0.835 (0.099)
Smoking						
-0.129 (0.084)	0.384 (0.095)	0.352 (0.055)	0.115 (0.019)	0.783 (0.051)	1.588 (0.667)	0.684 (0.097)
Drinking						
0.110 (0.134)	0.042 (0.177)	0.223 (0.084)	0.081 (0.015)	0.573 (0.087)	0.460 (0.387)	0.638 (0.109)
Risky behaviors						
-0.091 (0.096)	0.363 (0.150)	0.243 (0.082)	0.123 (0.021)	0.672 (0.060)	0.743 (0.364)	0.708 (0.071)
Self-esteem						
0.110 (0.050)	0.143 (0.102)	0.229 (0.081)	-0.023 (0.026)	0.509 (0.169)	-9.190 (12.170)	0.380 (0.106)
Physical exercise						
0.086 (0.048)	0.155 (0.068)	0.193 (0.079)	-0.004 (0.047)	0.614 (0.120)	0.841 (0.372)	0.596 (0.077)
Fighting						
0.228 (0.078)	-0.019 (0.111)	0.185 (0.068)	0.184 (0.027)	0.616 (0.071)	2.124 (0.809)	0.547 (0.079)

Estimates are reported without parentheses, with standard errors (clustered at the subnetwork level) shown in parentheses. The first row indicates the model used: quantile, LIM, or CES. The full table, including coefficients for the control variables, is available upon request.

Table 5.2: Empirical Results — Specification Tests

Encompassing test 3 vs. 4	<i>p</i> -value 4 vs. 5	KP LM test P-value	Monotonicity test <i>p</i> -value		
			Uniform	Increasing	Decreasing
Academic achievements (GPA)					
0.007	0.102	0.000	0.000	0.000	0.000
Academic effort					
0.442	0.113	0.000	0.334	0.003	0.920
Extracurricular activities					
0.023	0.584	0.000	0.000	0.000	0.000
Future perception					
0.884	0.976	0.000	0.047	0.000	0.975
Trouble at school					
0.859	0.021	0.000	0.007	0.000	0.000
Smoking					
0.000	0.103	0.000	0.000	0.000	0.000
Drinking					
0.020	0.839	0.000	0.252	0.001	0.001
Risky behaviors					
0.001	0.007	0.000	0.000	0.000	0.000
Self-esteem					
0.337	0.983	0.000	0.000	0.000	0.004
Physical exercise					
0.983	0.232	0.000	0.005	0.000	0.021
Fighting					
0.037	0.777	0.000	0.108	0.019	0.000

For the encompassing test, the columns labeled ‘ a vs. b ,’ for integers a and b , test the null hypothesis that the specification with a quantile levels does not perform worse than the specification with b quantile levels. The KP test refers to the p -value of the KP LM rank Wald test for weak instruments (Kleibergen and Paap, 2006). The columns labeled ‘monotonicity test p -value’ report the p -values for tests where the null hypothesis is that peer effects are uniform, increasing, or decreasing across quantile levels (see Kodde and Palm, 1986, for Wald-type criteria for monotonicity tests).

students appear largely insensitive to peers at the lowest or highest ends of the outcome distribution. This pattern is observed for GPA, extracurricular activities, trouble at school, physical exercise, and self-esteem. Unsurprisingly, the monotonicity hypothesis is rejected for all of these outcomes, suggesting that the LIM and CES specifications are not suitable for studying them. Specifically, the LIM and CES specifications tend to overestimate peer effects in these cases because they cannot isolate the influence of mid-level outcome peers without implicitly assuming that peers with the lowest or highest outcomes are also influential. Additionally, the CES specification fails to identify which part of the peer-outcome distribution is relevant for shaping individual behavior in this context.

In the case of extracurricular activities, the CES specification identifies a negative substitution parameter, with a standard error indicating that it is statistically different from one. This result suggests that peers with lower outcomes are more influential. Because the influence is monotonic under this specification, one might conclude that the peers at the bottom of the outcome distribution are the most influential. However, the quantile model provides a more nuanced insight: the most influential peers are those around the second quantile, while peers with the lowest outcomes exert little to no effect.

A comparable identification issue arises for self-esteem, where the CES specification yields a highly negative substitution parameter ($\rho = -9.940$) that is not statistically different from one due to the large variance of the estimate. This imprecision indicates that the specification cannot clearly identify the relevant part of the peer-outcome distribution. Our quantile specification sheds light on this imprecision, showing that peers at neither the top nor the bottom of the distribution appear to be the most influential.

For the other outcomes with this peer effect structure (GPA, trouble at school, and physical exercise), the substitution parameter estimates do not indicate a significant difference relative to the LIM model, whereas the monotonicity tests performed with the quantile specification reject the hypothesis that peer effects are uniform.

These results are consistent with theories of social comparison, which suggest that aligning with moderately performing peers, rather than with outliers, may be socially or psychologically optimal (Festinger, 1954). However, the mechanisms underlying these results may vary across outcomes. One explanation is that, because homophily in exogenous characteristics drives friendship formation, peers at the top or bottom of the distribution are more likely to differ substantially from the individual, whereas peers with intermediate values tend to be more similar and more closely connected. Moreover, friendships with high-achieving students may not improve learning if such peers are unwilling to collaborate with friends they perceive as weaker; in this case, these friendships may affect other outcomes rather than academic achievement. Our results are consistent with this interpretation, as we find

a negative (and nearly significant) effect of peers at the top of the distribution for the outcome academic achievement. A related pattern emerges for extracurricular activities, where highly involved peers are not necessarily more influential, as very high involvement may be negatively perceived in an academic context.

For some outcomes, we also find that, in addition to the influence of peers with intermediate outcomes, peers located at the top of the distribution exert significant effects. This pattern arises for outcomes related to potentially harmful behaviors, such as smoking, drinking, and risky behaviors. Here again, the LIM and CES specifications are unable to capture such nuanced structures. For these outcomes, the estimated substitution parameter indicates that the CES specification is not statistically different from the LIM specification.

Furthermore, the results reveal cases in which peers with extreme outcomes are as influential as, or even more influential than, those with intermediate outcomes. This pattern is observed for future expectations, academic effort, and fighting. For these outcomes, the monotonicity assumption is not rejected. For academic effort, we observe a negative ρ , which is not significantly different from one. As with self-esteem, this arises because peers at the top of the distribution are less influential, but the large variance of the estimate indicates that the most influential peers are not those at the bottom of the distribution.

The differences in peer effects across quantiles suggest that not all peers shape individual behaviour. These results carry significant policy implications, particularly for identifying key players within a network. Such identification may vary substantially across the three models, especially in cases where peers with mid-level outcomes are the most influential.

5.3 Measuring Student Influence

This section studies the influence of students within their school. Influence is measured by the change in the school’s average outcome when the student’s outgoing and incoming links are removed. This corresponds to the change in the school’s average outcome in a scenario in which the student becomes *fully isolated*—that is, they have no friends and are not nominated by others. Students who are already fully isolated in the observed network have no influence, as their influence measure is zero. For students who are not fully isolated, removing their links can affect the outcome distribution at the game equilibrium because the peer set for those who nominated them will change. This measure of influence is also considered by [Ballester et al. \(2006\)](#) and [Lee et al. \(2021\)](#), who define the key player as the student with the greatest influence.

We set the model parameters to their estimated values and compute the influence for each student. Within each school, we rank students by assigning the highest rank to the student

with the largest influence. We then compare the rankings obtained from the quantile model to those from the LIM and CES models. Since the effect of removing a single student’s links can be negligible in large networks, we focus on schools with fewer than 50 students.²⁷

Let $\mathbf{G}_g$ be the $n_g \times n_g$ adjacency matrix of school g , and let $\mathbf{G}_g^{(i)}$ denote the matrix obtained by setting the i -th row and column of $\mathbf{G}_g$ to zero. Let $y_{j,g}^{(i)}$ be the outcome of student j when the school network is $\mathbf{G}_g^{(i)}$. The influence of student i is measured by

$$P_{i,g} = \frac{1}{n_g} \sum_{j=1}^{n_s} \left(y_{j,g} - y_{j,g}^{(i)} \right).$$

Student ranks, normalized between 0 and 100, are presented in Figure 5.1 for selected outcomes. Rank gaps are substantial, particularly for smoking, extracurricular activities, and drinking. For these outcomes, some students who receive the highest influence scores under the CES and LIM models are assigned relatively low scores under the quantile model, and vice versa. These discrepancies reflect the non-monotonic pattern of peer effects associated with these outcomes. Under the quantile specification, peers at the top or bottom of the outcome distribution tend to receive lower ranks than under the other models, because the most influential peers are those with intermediate outcome levels. In contrast, the LIM and CES specifications cannot isolate the influence of mid-level outcome peers without simultaneously attributing substantial weight to peers at the extremes. As a result, students connected to peers with extreme outcome values are not assigned low ranks under these models; in the CES specification in particular, the most influential peers are necessarily located at the top or bottom of the peer-outcome distribution.

For the other outcomes, rank gaps are less pronounced because the pattern of peer effects is closer to a monotonic or uniform structure. In particular, for academic effort and future perception, the monotonicity assumption is clearly not rejected. For these outcomes, tests of the null hypothesis that peer effects are decreasing yield p -values above 0.90 (see Table 5.2). The hypothesis of a uniform coefficient is also not rejected, although the corresponding p -values are weaker. Consistent with these results, the rank predictions in Figure 5.1 are similar across the quantile, LIM, and CES specifications.

This counterfactual analysis reveals that key player status varies across the three models. In the standard LIM model, a high outcome for a student who is not fully isolated increases their influence (see Ballester et al., 2006), because the student is nominated as a friend and therefore contributes substantially to the average friend outcome. Since this average is the only aggregated measure of the peer-outcome distribution that affects individuals,

²⁷In larger schools, a similar simulation exercise can be conducted by removing the links of a group of students rather than just one.

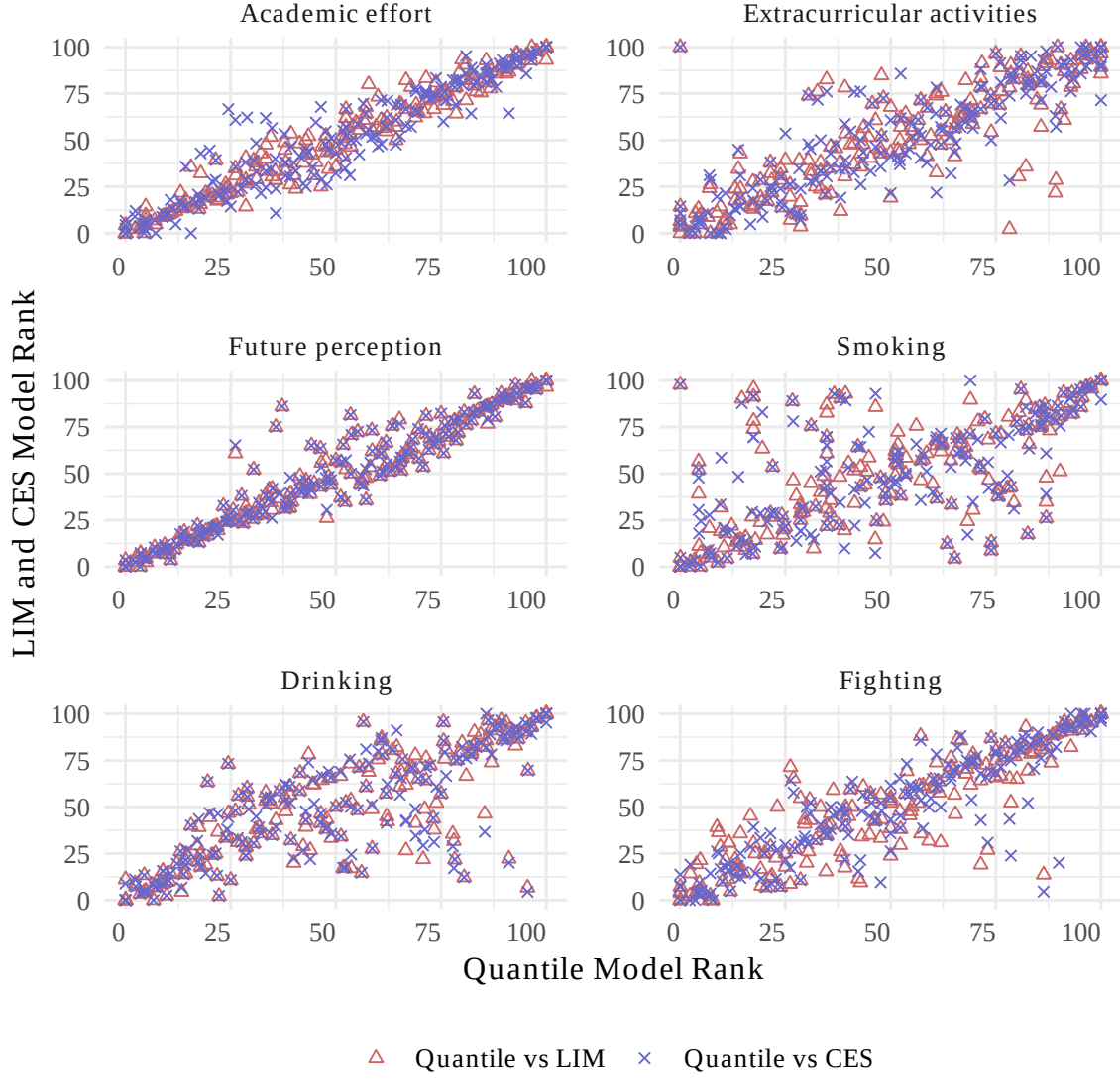


Figure 5.1: Influence Measure

The x-axis reports student ranks based on the influence measure in the quantile model, while the y-axis reports ranks based on the LIM and CES models. Each red triangle represents a student's LIM model rank (on the y-axis) against their quantile model rank (on the x-axis). Each blue "x" marker represents a student's CES model rank (on the y-axis) versus their quantile model rank (on the x-axis).

these peers are systematically influential. In contrast, in the quantile or CES models, a high outcome does not necessarily imply greater influence. Nevertheless, the CES model still imposes restrictions on whether these peers are influential or not. The analysis above shows that when peer effects are non-monotonic, the CES model does not fully overcome the limitations of the LIM model and may misidentify the most influential students and key players. Identifying such agents is crucial in social interactions, for example, to implement targeted policies or to assign individuals to groups, such as students in dormitories. The

quantile model offers a more flexible approach for conducting such analyses.

6 Conclusion

This paper introduced a model that allows for flexible patterns of peer effects. In this model, how you affect your peer depends not only on your own outcome level but also on the outcomes of the other peers that they have. The construction of the peer effect in the model is flexible enough to allow us to discuss many interesting questions regarding how the composition of a peer group affects the peer effect: complementarity v. substitutability, non-monotonic marginal peer effects, etc. Existing peer effect models, such as the linear-in-means (LiM) and the CES model, tend to oversimplify the dynamics of peer influence by focusing on a single scalar parameter or assuming monotonically increasing or decreasing marginal peer effects. This oversimplification neglects the heterogeneity within peer effects, where individuals may experience varying degrees of influence depending on the structure and nature of their interactions, across different empirical contexts. This paper allows researchers to investigate richer patterns of peer interactions, with specific peers—such as the highest or lowest performers—having disproportionately larger effects on outcomes in certain contexts.

As an operational framework, we propose a quantile-based peer effect model to solve the dimensionality problem. There may exist alternative approaches with better finite sample properties, which warrants future research. For example, the model can benefit from having an estimator that has data-driven smoothness on the saturated peer effect coefficients.

In addition, this paper has only considered settings with exogenous networks. As work on incorporating endogenous networks into peer effect models continues to develop, it would be interesting to see if similar approaches can be applied to our model.

References

- BALLESTER, C., A. CALVÓ-ARMENGOL, AND Y. ZENOU (2006): “Who’s who in networks. Wanted: The key player,” Econometrica, 74, 1403–1417.
- BEKKER, P. A. (1994): “Alternative approximations to the distributions of instrumental variable estimators,” Econometrica: Journal of the Econometric Society, 657–681.
- BLUME, L. E., W. A. BROCK, S. N. DURLAUF, AND R. JAYARAMAN (2015): “Linear social interactions models,” Journal of Political Economy, 123, 444–496.
- BOUCHER, V., F. A. DEDEWANOU, AND A. DUFAYS (2022): “Peer-induced beliefs regarding college participation,” Economics of Education Review, 90, 102307.

- BOUCHER, V. AND B. FORTIN (2016): “Some challenges in the empirics of the effects of networks,” Handbook on the Economics of Networks, 45, 48.
- BOUCHER, V. AND A. HOUNDETOUNGAN (2025): “Estimating peer effects using partial network data,” Review of Economics and Statistics, 1–48.
- BOUCHER, V., M. RENDALL, P. USHCHEV, AND Y. ZENOU (2024): “Toward a general theory of peer effects,” Econometrica, 92, 543–565.
- BRAMOULLÉ, Y., H. DJEBBARI, AND B. FORTIN (2009): “Identification of peer effects through social networks,” Journal of Econometrics, 150, 41–55.
- (2020): “Peer effects in networks: A survey,” Annual Review of Economics, 12, 603–629.
- BROCK, W. A. AND S. N. DURLAUF (2007): “Identification of binary choice models with social interactions,” Journal of Econometrics, 140, 52–75.
- CALVÓ-ARMENGOL, A., E. PATACCHINI, AND Y. ZENOU (2009): “Peer effects and social networks in education,” The Review of Economic Studies, 76, 1239–1267.
- CARRELL, S. E., B. I. SACERDOTE, AND J. E. WEST (2013): “From natural variation to optimal policy? The importance of endogenous peer group formation,” Econometrica, 81, 855–882.
- CHAMBERLAIN, G. (1987): “Asymptotic efficiency in estimation with conditional moment restrictions,” Journal of econometrics, 34, 305–334.
- COMOLA, M., R. DIEYE, AND B. FORTIN (2025): “Heterogeneous peer effects and gender-based interventions for teenage obesity,” Journal of Health Economics, 102, 103023.
- DE CHAISEMARTIN, C. AND X. D’HAULTFOEUILLE (2020): “Two-way fixed effects estimators with heterogeneous treatment effects,” American Economic Review, 110, 2964–2996.
- DE PAULA, Á. (2020): “Econometric models of network formation,” Annual Review of Economics, 12, 775–799.
- DIEL, K. AND W. HOFMANN (2019): “Inspired to perspire: The interplay of social comparison direction and standard extremity in the context of challenging exercising goals,” Social Cognition, 37, 247–265.
- FESTINGER, L. (1954): “A theory of social comparison processes,” Human relations, 7, 117–140.
- GALEOTTI, A., B. GOLUB, AND S. GOYAL (2020): “Targeting interventions in networks,” Econometrica, 88, 2445–2471.
- GOLDSMITH-PINKHAM, P. AND G. W. IMBENS (2013): “Social networks and the identification of peer effects,” Journal of Business & Economic Statistics, 31, 253–264.

- GOODMAN-BACON, A. (2021): “Difference-in-differences with variation in treatment timing,” Journal of Econometrics, 225, 254–277.
- GURYAN, J., K. KROFT, AND M. J. NOTOWIDIGDO (2009): “Peer effects in the workplace: Evidence from random groupings in professional golf tournaments,” American Economic Journal: Applied Economics, 1, 34–68.
- HALL, A. R. AND A. INOUE (2003): “The large sample behaviour of the generalized method of moments estimator in misspecified models,” Journal of Econometrics, 114, 361–394.
- HERSTAD, E. (2023): “Estimating peer effects and network formation models with missing links,” Available via www.eyoherstad.com.
- HOUNDETOUNGAN, A. (2026): “Count Data Models with Heterogeneous Peer Effects under Rational Expectations,” Journal of Applied Econometrics.
- HSIEH, C.-S. AND L. F. LEE (2016): “A social interactions model with endogenous friendship formation and selectivity,” Journal of Applied Econometrics, 31, 301–319.
- HYNDMAN, R. J. AND Y. FAN (1996): “Sample quantiles in statistical packages,” The American Statistician, 50, 361–365.
- JOCHMANS, K. (2023): “Peer effects and endogenous social interactions,” Journal of Econometrics, 235, 1203–1214.
- JOHNSON, I. AND H. R. MOON (2021): “Estimation of peer effects in endogenous social networks: Control function approach,” Review of Economics and Statistics, 103, 328–345.
- KANG, C. (2007): “Classroom peer effects and academic achievement: Quasi-randomization evidence from South Korea,” Journal of Urban Economics, 61, 458–495.
- KLEIBERGEN, F. AND R. PAAP (2006): “Generalized reduced rank tests using the singular value decomposition,” Journal of Econometrics, 133, 97–126.
- KODDE, D. A. AND F. C. PALM (1986): “Wald criteria for jointly testing equality and inequality restrictions,” Econometrica: journal of the Econometric Society, 1243–1248.
- LEE, L.-F., X. LIU, E. PATACCHINI, AND Y. ZENOU (2021): “Who is the key player? A network analysis of juvenile delinquency,” Journal of Business & Economic Statistics, 39, 849–857.
- LEWBEL, A., X. QU, AND X. TANG (2023): “Social networks with unobserved links,” Journal of Political Economy, 131, 898–946.
- MANSKI, C. F. (1993): “Identification of endogenous social effects: The reflection problem,” The Review of Economic Studies, 60, 531–542.
- MASTEN, M. A. (2018): “Random coefficients on endogenous variables in simultaneous equations models,” The Review of Economic Studies, 85, 1193–1250.

- MOGSTAD, M. AND A. TORGOVITSKY (2024): “Instrumental Variables with Unobserved Heterogeneity in Treatment Effects,” Tech. rep., National Bureau of Economic Research.
- NEWKEY, W. K. AND R. J. SMITH (2004): “Higher order properties of GMM and generalized empirical likelihood estimators,” Econometrica, 72, 219–255.
- ROGERS, T. AND A. FELLER (2016): “Discouraged by peer excellence: Exposure to exemplary peer performance causes quitting,” Psychological science, 27, 365–374.
- SACERDOTE, B. (2001): “Peer effects with random assignment: Results for Dartmouth roommates,” The Quarterly Journal of Economics, 116, 681–704.
- (2011): “Peer effects in education: How might they work, how big are they and how much do we know thus far?” in Handbook of the Economics of Education, Elsevier, vol. 3, 249–277.
- SMITH, R. J. (1992): “Non-nested tests for competing models estimated by generalized method of moments,” Econometrica: Journal of the Econometric Society, 60, 973–980.
- TAO, J. AND L.-F. LEE (2014): “A social interaction model with an extreme order statistic,” The Econometrics Journal, 17, 197–240.
- TATSI, E. (2015): “Endogenous Social Interactions: Which Peers Matter?” in VfS Annual Conference 2015 (Muenster): Economic Development-Theory and Policy, Verein für Socialpolitik/German Economic Association, 113168.
- YANG, C. AND L.-F. LEE (2017): “Social interactions under incomplete information with heterogeneous expectations,” Journal of Econometrics, 198, 65–83.

APPENDIX

A Proofs

A.1 Proof for Proposition 1

We want to show that the game described by the utility function (2) has a unique Nash equilibrium $\mathbf{y}^* = (y_1^*, \dots, y_n^*)'$ such that $y_i^* = BR_i(\mathbf{y}_{-i}^*)$. To achieve this result, it is sufficient to show that the mapping $BR(\mathbf{y}) = (BR_1(\mathbf{y}_{-1}), \dots, BR_n(\mathbf{y}_{-n}))'$ is a contraction. In other words, there exists some constant $\bar{\beta} < 1$ such that for any two strategy profiles $\mathbf{y}, \mathbf{y}' \in \mathbb{R}^n$, $|BR_i(\mathbf{y}_{-i}) - BR_i(\mathbf{y}'_{-i})| \leq \bar{\beta} \|\mathbf{y} - \mathbf{y}'\|_\infty$ for every $i = 1, \dots, n$. Let $\bar{\beta} = \max_{d \leq \bar{d}} \sum_{k=1}^d |\beta_{k,d}|$. From Assumption 1, $0 \leq \bar{\beta} < 1$.

Let π be the ordering of $\mathbf{y}$, that is, a function that maps i to the rank of y_i in $\mathbf{y}$ under some tie-breaking rule. For example if $y_i = \max_{k=1, \dots, n} y_k$ then $\pi(i) = n$. Note that π , together with the network $\mathbf{A}$, therefore fully determines the mapping from $\mathbf{y}$ to each $\tilde{y}_{i,k}$. We define π' equivalently as the ordering of $\mathbf{y}'$.

Consider first the case where $\pi = \pi'$. The inequality holds trivially: for each i ,

$$\begin{aligned} |BR_i(\mathbf{y}_{-i}) - BR_i(\mathbf{y}'_{-i})| &= \left| \sum_{k=1}^{d_i} (\tilde{y}_{i,k} - \tilde{y}'_{i,k}) \beta_{k,d_i} \right| \\ &\leq \sum_{k=1}^{d_i} |\beta_{k,d_i}| \cdot \|\mathbf{y} - \mathbf{y}'\|_\infty \leq \bar{\beta} \|\mathbf{y} - \mathbf{y}'\|_\infty \end{aligned}$$

with $\bar{\beta} = \max_d \sum_{k=1}^d |\beta_{k,d}| < 1$. The first inequality holds since $\mathbf{y}$ and $\mathbf{y}'$ have the same order and therefore $\tilde{y}_{i,k}$ and $\tilde{y}'_{i,k}$ are the outcome of the same individual: $|\tilde{y}_{i,k} - \tilde{y}'_{i,k}|$ is bounded by $\|\mathbf{y} - \mathbf{y}'\|_\infty$. The second equality is from Assumption 1.

The non-trivial case is when $\pi \neq \pi'$. To proceed, fix i and k and let j be the identity of person i 's k th lowest friend under π , and j' be the identity under π' . Consider the object

$$|(\tilde{y}_{i,k} - \tilde{y}'_{i,k}) \beta_{k,d_i}|.$$

If $j = j'$, this quantity is bounded by $|\beta_{k,d_i}| \|\mathbf{y} - \mathbf{y}'\|_\infty$ by a similar argument as above. If not, we have

$$|(\tilde{y}_{i,k} - \tilde{y}'_{i,k}) \beta_{k,d_i}| \leq |\beta_{k,d_i}| |y_j - y_{j'}|.$$

Let k' be the ranking of j among i 's friends under π' . Then there are two possible cases,

either $k' > k$ or $k' < k$.

Case 1: $k' > k$. Then we have that

$$y'_j = \tilde{y}'_{i,k'} \geq \tilde{y}'_{i,k} = y'_{j'}$$

meaning $y'_j \geq y'_{j'}$. Therefore

$$y_j - y'_{j'} \geq y'_j - \|\mathbf{y} - \mathbf{y}'\|_\infty - y'_{j'} \geq -\|\mathbf{y} - \mathbf{y}'\|_\infty$$

Where the first inequality uses $y'_j - y_j \leq \|\mathbf{y} - \mathbf{y}'\|_\infty$ and the second that $y'_j - y'_{j'} \geq 0$.

To get the inequality for the other sign, note that, by the pigeonhole principle, there must be another friend of i , p , whose rank among i 's friends was above k under π and is lower than or equal to k under π' . That is

$$y'_p \leq \tilde{y}'_{i,k} = y'_{j'} \quad \text{and} \quad y_p \geq \tilde{y}_{i,k} = y_j,$$

and therefore:

$$\begin{aligned} y_j - y'_{j'} &\leq y_p - y'_{j'} \\ &\leq y'_p + \|\mathbf{y} - \mathbf{y}'\|_\infty - y'_{j'} \\ &\leq \|\mathbf{y} - \mathbf{y}'\|_\infty \end{aligned}$$

Combining this with the result above gives the desired inequality: $|y_j - y'_{j'}| \leq \|\mathbf{y} - \mathbf{y}'\|_\infty$

Case 2: $k' < k$. As before, this implies

$$y'_j = \tilde{y}'_{i,k'} \leq \tilde{y}'_{i,k} = y'_{j'}$$

and therefore

$$y_j - y'_{j'} \leq y'_j + \|\mathbf{y} - \mathbf{y}'\|_\infty - y'_{j'} \leq \|\mathbf{y} - \mathbf{y}'\|_\infty.$$

By the pigeonhole principle, we know that as $k' < k$, there must be some other individual p whose rank is higher than or equal to k under π' but lower than k under π . That is

$$y'_p \geq \tilde{y}'_{i,k} = y'_{j'} \quad \text{and} \quad y_p \leq \tilde{y}_{i,k} = y_j$$

and therefore

$$y_j - y'_{j'} \geq y_p - y'_{j'} \geq y'_p - \|\mathbf{y} - \mathbf{y}'\|_\infty - y'_{j'} \geq -\|\mathbf{y} - \mathbf{y}'\|_\infty.$$

Combining the two inequalities we have our desired result: $|y_j - y'_{j'}| \geq \|\mathbf{y} - \mathbf{y}'\|_\infty$.

Finally, by aggregating across k for the fixed i and then aggregating across i , we get

$$\begin{aligned} |BR_i(\mathbf{y}_{-i}) - BR_i(\mathbf{y}'_{-i})| &\leq \left| \sum_{k=1}^{d_i} (\tilde{y}_{i,k} - \tilde{y}'_{i,k}) \beta_{k,d_i} \right| \leq \sum_{k=1}^{d_i} |\beta_{k,d_i}| \cdot \|\mathbf{y} - \mathbf{y}'\|_\infty \leq \bar{\beta} \|\mathbf{y} - \mathbf{y}'\|_\infty \\ \|BR(\mathbf{y}) - BR(\mathbf{y}')\|_\infty &\leq \bar{\beta} \|\mathbf{y} - \mathbf{y}'\|_\infty. \end{aligned}$$

Meaning $BR(\mathbf{y})$ is a contraction mapping for any $\mathbf{y}, \mathbf{y}'$. □

A.2 Proof for Theorem 1

From Assumption 2, we have

$$\mathbb{E}_g[\mathbf{y}_g] = \mathbb{E}_g[\mathbf{W}_g]\theta.$$

Premultiplying by $\mathbb{E}_g[\mathbf{W}_g]^\top$ and averaging over g implies that

$$\text{plim}_{m \rightarrow \infty} \frac{1}{m} \sum_{g=1}^m \mathbb{E}_g[\mathbf{W}_g]^\top \mathbb{E}_g[\mathbf{W}_g] \theta = \text{plim}_{m \rightarrow \infty} \frac{1}{m} \sum_{g=1}^m \mathbb{E}_g[\mathbf{W}_g]^\top \mathbb{E}_g[\mathbf{y}_g].$$

Under Assumption 3, the matrix $\text{plim}_{m \rightarrow \infty} \frac{1}{m} \sum_{g=1}^m \mathbb{E}_g[\mathbf{W}_g]^\top \mathbb{E}_g[\mathbf{W}_g]$ is nonsingular. Thus, θ is uniquely determined, and identification follows. □

A.3 Proof for Proposition 2

Using Assumption 4.i and plugging in

$$\mathbb{E}_g[y_{i,g}] = \sum_{k=1}^{d_{i,g}} \beta_{k,d} \mathbb{E}_g[\tilde{y}_{i,k,g}] + \mathbf{x}_{i,g}^\top \gamma + \bar{\mathbf{x}}_{i,g}^\top \delta,$$

we get

$$\begin{aligned}\beta^{\text{LiM}} &:= \left(\frac{1}{m} \sum_{g=1}^m \left(\tilde{\mathbf{Z}}_g^{\text{LiM}} \right)^\top G_g \mathbb{E}_g[\mathbf{y}_g] \right)^{-1} \frac{1}{m} \sum_{g=1}^m \tilde{\mathbf{Z}}_g^\top \mathbb{E}_g[\mathbf{y}_g] \\ &= \left(\frac{1}{m} \sum_{g=1}^m \tilde{\mathbf{Z}}_g^\top G_g \mathbb{E}_g[\mathbf{y}_g] \right)^{-1} \frac{1}{m} \sum_{g=1}^m \tilde{\mathbf{Z}}_g^\top \begin{pmatrix} \sum_{k=1}^{d_{1,g}} \beta_{k,d} \mathbb{E}_g[\tilde{y}_{1,k,g}] \\ \vdots \\ \sum_{k=1}^{d_{n_g,g}} \beta_{k,d} \mathbb{E}_g[\tilde{y}_{n_g,k,g}] \end{pmatrix} + o_p(1).\end{aligned}$$

$o_p(1)$ term appears since

$$\begin{aligned}\frac{1}{m} \sum_{g=1}^m \left(\tilde{\mathbf{Z}}_g^{\text{LiM}} \right)^\top \tilde{\mathbf{X}}_g^{\text{LiM}} &= \frac{1}{m} \sum_{g=1}^m \left(\mathbf{z}_g^\top - \mathbf{P}^{\text{LiM}\top} \left(\tilde{\mathbf{X}}_g^{\text{LiM}} \right)^\top \right) \tilde{\mathbf{X}}_g^{\text{LiM}} \\ &= \frac{1}{m} \sum_{g=1}^m \mathbf{z}_g^\top \tilde{\mathbf{X}}_g^{\text{LiM}} - \mathbf{P}^{\text{LiM}\top} \cdot \frac{1}{m} \sum_{g=1}^m \left(\tilde{\mathbf{X}}_g^{\text{LiM}} \right)^\top \tilde{\mathbf{X}}_g^{\text{LiM}} = o_p(1).\end{aligned}$$

Suppose that $\mathbf{z}_{i,g}$ and therefore its residual $\tilde{\mathbf{z}}_{i,g}^{\text{LiM}}$ are scalar variables. Then,

$$\begin{aligned}\beta^{\text{LiM}} &= \frac{\frac{1}{m} \sum_{g=1}^m \sum_{i=1}^{n_g} \tilde{\mathbf{z}}_{i,g}^{\text{LiM}} \sum_{k=1}^{d_{i,g}} \mathbb{E}_g[\tilde{y}_{i,k,g}] \beta_{k,d}}{\frac{1}{m} \sum_{g=1}^m \sum_{i=1}^{n_g} \tilde{\mathbf{z}}_{i,g}^{\text{LiM}} \frac{1}{d_{i,g}} \sum_{k=1}^{d_{i,g}} \mathbb{E}_g[\tilde{y}_{i,k,g}]} + o_p(1) \\ &= \frac{\sum_{d=1}^{\bar{d}} \sum_{k=1}^d \frac{1}{m} \sum_{g=1}^m \sum_{i=1}^{n_g} \tilde{\mathbf{z}}_{i,g}^{\text{LiM}} \mathbb{E}_g[\tilde{y}_{i,k,g} \mathbf{1}\{d_{i,g} = d\}] \beta_{k,d}}{\sum_{d=1}^{\bar{d}} \frac{1}{m} \sum_{g=1}^m \sum_{i=1}^{n_g} \frac{1}{d} \sum_{k=1}^d \tilde{\mathbf{z}}_{i,g}^{\text{LiM}} \mathbb{E}_g[\tilde{y}_{i,k,g} \mathbf{1}\{d_{i,g} = d\}]} + o_p(1).\end{aligned}$$

By letting

$$w_{k,d}^{\text{LiM}} = \frac{\frac{1}{m} \sum_{g=1}^m \sum_{i=1}^{n_g} \frac{1}{d} \tilde{\mathbf{z}}_{i,g}^{\text{LiM}} \mathbb{E}_g[\tilde{y}_{i,k,g} \mathbf{1}\{d_{i,g} = d\}]}{\sum_{d=1}^{\bar{d}} \sum_{k=1}^d \frac{1}{m} \sum_{g=1}^m \sum_{i=1}^{n_g} \frac{1}{d} \tilde{\mathbf{z}}_{i,g}^{\text{LiM}} \mathbb{E}_g[\tilde{y}_{i,k,g} \mathbf{1}\{d_{i,g} = d\}]},$$

β^{LiM} becomes

$$\beta^{\text{LiM}} = \sum_{d=1}^{\bar{d}} \sum_{k=1}^d w_{k,d}^{\text{LiM}} d \beta_{k,d} + o_p(1).$$

The weights satisfy the sum-to-one constraint: $\sum_{d=1}^{\bar{d}} \sum_{k=1}^d w_{k,d}^{\text{LiM}} = 1$.

Likewise, for the LiS peer effect parameter, we get

$$\begin{aligned}\beta^{\text{LiS}} &= \frac{\sum_{d=1}^{\bar{d}} \sum_{k=1}^d \frac{1}{m} \sum_{g=1}^m \sum_{i=1}^{n_g} \tilde{\mathbf{z}}_{i,g}^{\text{LiS}} \mathbb{E}_g [\tilde{y}_{i,k,g} \mathbf{1}\{d_{i,g} = d\}] \beta_{k,d}}{\sum_{d=1}^{\bar{d}} \frac{1}{m} \sum_{g=1}^m \sum_{i=1}^{n_g} \tilde{\mathbf{z}}_{i,g}^{\text{LiS}} \mathbb{E}_g \left[\sum_{k=1}^d \tilde{y}_{i,k,g} \mathbf{1}\{d_{i,g} = d\} \right]} + o_p(1) \\ &= \sum_{d=1}^{\bar{d}} \sum_{k=1}^d w_{k,d}^{\text{LiS}} \beta_{k,d} + o_p(1),\end{aligned}$$

where

$$w_{k,d}^{\text{LiS}} = \frac{\frac{1}{m} \sum_{g=1}^m \sum_{i=1}^{n_g} \tilde{\mathbf{z}}_{i,g}^{\text{LiS}} \mathbb{E}_g [\tilde{y}_{i,k,g} \mathbf{1}\{d_{i,g} = d\}]}{\sum_{d=1}^{\bar{d}} \sum_{k=1}^d \frac{1}{m} \sum_{g=1}^m \sum_{i=1}^{n_g} \tilde{\mathbf{z}}_{i,g}^{\text{LiS}} \mathbb{E}_g [\tilde{y}_{i,k,g} \mathbf{1}\{d_{i,g} = d\}]}$$

and $\tilde{\mathbf{z}}_{i,g}^{\text{LiS}}$ is the residual from the linear-in-sum projection $\mathbf{P}^{\text{LiS}}$. Again, the weights sum to one: $\sum_{d=1}^{\bar{d}} \sum_{k=1}^d w_{k,d}^{\text{LiS}} = 1$.

When $\mathbf{z}_{i,g}$ is not a scalar, we can modify the argument above and get similar expressions by replacing $\tilde{\mathbf{z}}_{i,g}$ with $c^\top \tilde{\mathbf{z}}_{i,g}$ with some weighting c . The weighting depends on how we weight over-identifying moments. For example, the TSLS principle uses the weighting vector

$$c^\top = \frac{1}{m} \sum_{g=1}^m \mathbb{E}_g [\mathbf{y}_g^\top] \mathbf{G}_g^\top \tilde{\mathbf{Z}}_g \left(\frac{1}{m} \sum_{g=1}^m \tilde{\mathbf{Z}}_g^\top \tilde{\mathbf{Z}}_g \right)^{-1};$$

misspecification weights depend on the weighting matrix when the LiM or LiS is overidentified.

□

B Supplementary Theoretical Results

B.1 Spillover and Conformity in Preferences

We adopt the approach that was proposed by [Boucher et al. \(2024\)](#), which incorporates both spillover and conformity into the model, allowing their relative importance to be empirically identified from the data. The utility function is given by:

$$U_i(y_i, \mathbf{y}_{-i}) = \underbrace{\alpha_i y_i - \frac{1}{2} y_i^2}_{\text{private benefit}} + \underbrace{\sum_{k=1}^{d_i} \beta_{k,d_i}^s y_i \tilde{y}_{i,k} - \frac{1}{2} \sum_{k=1}^{d_i} \beta_{k,d_i}^c (y_i - \tilde{y}_{i,k})^2}_{\text{social benefit}}, \quad (8)$$

The social benefit reflects both spillover effects and conformity between an agent's effort

and the efforts of their peers.²⁸ The term $\beta_{k,d_i}^s y_i \tilde{y}_{i,k}$ implies strategic complementarity between agent i 's effort and the effort of the peer with the k -th smallest effort when $\beta_{k,d_i}^s > 0$. Conversely, if $\beta_{k,d_i}^s < 0$, then $\beta_{k,d_i}^s y_i \tilde{y}_{i,k}$ reflects strategic substitution (i.e., negative spillover) between efforts. The term $\beta_{k,d_i}^c (y_i - \tilde{y}_{i,k})^2$ captures conformist preferences. If β_{k,d_i}^c is non-negative, agent i incurs a social cost when their effort deviates from the k -th smallest effort among their peers. The case where β_{k,d_i}^c is negative (indicating anticonformity) is rare and is not considered in this paper, as it introduces complications in the analysis.²⁹

For ease of exposition, we introduce the following notation:

$$\lambda_{k,d_i}^s = \frac{\beta_{k,d_i}^s}{1 + \sum_{k=1}^{d_i} \beta_{k,d_i}^c}, \quad \lambda_{k,d_i}^c = \frac{\beta_{k,d_i}^c}{1 + \sum_{k=1}^{d_i} \beta_{k,d_i}^c}, \quad \text{and} \quad \lambda_{k,d_i} = \lambda_{k,d_i}^s + \lambda_{k,d_i}^c \text{ for all } k.$$

$$\lambda_{d_i}^s = \sum_{k=1}^{d_i} \lambda_{k,d_i}^s \quad \text{and} \quad \lambda_{d_i}^c = \sum_{k=1}^{d_i} \lambda_{k,d_i}^c.$$

Agent i chooses their effort y_i by maximizing the utility function $U_i(y_i, \mathbf{y}_{-i})$. By solving the first-order condition of this maximization problem, I obtain the best response function:

$$BR_i(\mathbf{y}_{-i}) = (1 - \lambda_{d_i}^c) \alpha_i + \sum_{k=1}^{d_i} \lambda_{k,d_i} \tilde{y}_{i,k}. \quad (9)$$

The parameter λ_{k,d_i} measures the total peer effect at the k -th smallest peer effort. This total effect is decomposed as $\lambda_{k,d_i} = \lambda_{k,d_i}^s + \lambda_{k,d_i}^c$, where λ_{k,d_i}^s captures the spillover effect and λ_{k,d_i}^c captures the conformity component. The total spillover effect across all peer efforts is given by $\lambda_{d_i}^s = \sum_{k=1}^{d_i} \lambda_{k,d_i}^s$, while the total conformity parameter is $\lambda_{d_i}^c = \sum_{k=1}^{d_i} \lambda_{k,d_i}^c$.

By defining $\tilde{\alpha}_i = (1 - \lambda_{d_i}^c) \alpha_i$, the best response function given by Equation (9) is similar to that in Equation (3), where $\tilde{\alpha}_i$ represents the new productivity level. Therefore, Proposition 1, which establishes the existence and uniqueness of the Nash equilibrium, readily extends to the game featuring both spillover and conformity under a condition analogous to Assumption 1, namely that $\sum_{k=1}^{d_i} |\lambda_{k,d_i}| < 1$.

B.2 Identification under Spillover and Conformity

In this section, we discuss the identification of the model in the case where it includes both spillover and conformity effects. As $\alpha_i = \mathbf{x}_i^\top \boldsymbol{\gamma} + \varepsilon_i$, the reduced-form specification of

²⁸When i has no friends (i.e., no outgoing links or $d_i = 0$), the utility function excludes the social benefit term and is given by $U_i(y_i, \mathbf{y}_{-i}) = \alpha_i y_i - \frac{1}{2} y_i^2$. The corresponding optimal strategy is therefore $y_i = \alpha_i$. An individual without friends is referred to as *isolated*.

²⁹Specifically, when $\beta_{k,d_i}^c < 0$, the utility function may become convex and lack a maximizer unless the strategy space is compact

the model for a non-isolated individual (i.e., one who has friends) is:

$$y_i = \sum_{k=1}^{d_i} \lambda_{k,d_i} \tilde{y}_{i,k} + (1 - \lambda_{d_i}^c) \mathbf{x}_i^\top \boldsymbol{\gamma} + (1 - \lambda_{d_i}^c) \varepsilon_i.$$

Let $\mathbf{x}_i^{iso}$ be the vector including the components of $\mathbf{x}_i$ that are defined for isolated agents (who has no friends). Specifically, $\mathbf{x}_i^{iso}$ excludes the contextual variables. Let $\boldsymbol{\gamma}^{iso}$ denotes the coefficient associated with $\mathbf{x}_i^{iso}$ in $\boldsymbol{\gamma}$. For isolated agents, the outcome is simply given by:

$$y_i = (\mathbf{x}_i^{iso})^\top \boldsymbol{\gamma}^{iso} + \varepsilon_i.$$

As in [Boucher et al. \(2024\)](#), identifying the model's parameters requires the presence of isolated nodes. Given that our inference assumes many independent subnetworks, this implies that the share of subnetworks with isolated nodes is strictly positive asymptotically. Under the exogeneity condition in Assumption 2, we have the moment condition $\mathbb{E}_g(\varepsilon_i) = 0$, conditional on whether i is isolated or not, where $\mathbb{E}_g$ denotes the conditional expectation given the network and exogenous characteristics.

Let n^{iso} be the number of isolated agents. The moment condition for isolated agents helps identify:

$$\boldsymbol{\gamma} = \left(\text{plim} \frac{1}{n} \sum_{i=1}^n \mathbf{1}\{d_i = 0\} \mathbf{x}_i \mathbf{x}_i^\top \right)^{-1} \left(\text{plim} \frac{1}{n} \sum_{i=1}^n \mathbf{1}\{d_i = 0\} \mathbf{x}_i y_i \right).$$

Given the identification of $\boldsymbol{\gamma}$, we can also identify $\lambda_{d_i}^c$ and λ_{k,d_i} for all k and d_i using the moment conditions for non-isolated agents. Indeed, this identification problem is similar to the one studied in the main text. The only difference is that it focuses on non-isolated agents.

Overall, we can then identify λ_{k,d_i} , which represent the total peer effect at each rank, and the total conformity effect $\lambda_{d_i}^c$. The total spillover effect is therefore given by $\lambda_{d_i}^s = \sum_{k=1}^{d_i} \lambda_{k,d_i} - \lambda_{d_i}^c$. However, we cannot disentangle spillover and conformity effects at each rank separately; we can only identify their total effects across all ranks.

B.3 Consistency and Asymptotic Normality

Let $\mathbf{Z}_g$ be the matrix of instruments in the g th subnetwork. As discussed in Section 3, $\mathbf{Z}_g$ includes the quantiles of $\mathbf{x}_{j,g}$ and $\bar{\mathbf{x}}_{j,g}$ among friends, which we use to instrument $\tilde{\mathbf{y}}_{i,g}$. As usual, $\mathbf{Z}_g$ also includes exogenous regressors $\mathbf{x}_{j,g}$ and $\bar{\mathbf{x}}_{j,g}$. To establish consistency and asymptotic normality, we impose the following regularity conditions.

Assumption 5.

- a.* The matrix $\text{plim}_{m \rightarrow \infty} \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^\top \mathbf{Z}_g$ and $\text{plim}_{m \rightarrow \infty} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^\top \mathbf{Z}_g$ are of column full rank.
- b.* $\text{plim}_{m \rightarrow \infty} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^\top \boldsymbol{\varepsilon}_g = 0$.
- c.* $\frac{1}{\sqrt{m}} \sum_{g=1}^m \mathbf{Z}_g^\top \boldsymbol{\varepsilon}_g \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Omega}_\varepsilon)$.

We establish the following result.

Proposition 3. Under Assumption 5, the TSLS estimator $\hat{\theta}^{\text{TSLS}}$ using $\{\mathbf{Z}_g\}_{g=1}^m$ as instruments converges in probability to θ and

$$\sqrt{m} \left(\hat{\theta}^{\text{TSLS}} - \theta \right) \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Omega})$$

with some consistently estimable covariance matrix $\boldsymbol{\Omega}$, as $m \rightarrow \infty$.

Proof. Define the sample moments

$$\hat{\mathbf{Q}}_{WZ} = \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^\top \mathbf{Z}_g, \quad \hat{\mathbf{Q}}_{ZZ} = \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^\top \mathbf{Z}_g, \quad \hat{\mathbf{Q}}_{ZY} = \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^\top \mathbf{y}_g.$$

The TSLS estimator is

$$\hat{\theta}^{\text{TSLS}} = \left(\hat{\mathbf{Q}}_{WZ} \hat{\mathbf{Q}}_{ZZ}^{-1} \hat{\mathbf{Q}}_{ZW} \right)^{-1} \left(\hat{\mathbf{Q}}_{WZ} \hat{\mathbf{Q}}_{ZZ}^{-1} \hat{\mathbf{Q}}_{ZY} \right),$$

where $\hat{\mathbf{Q}}_{ZW} = \hat{\mathbf{Q}}_{WZ}^\top$. Since $\mathbf{y}_g = \mathbf{W}_g \theta + \boldsymbol{\varepsilon}_g$, we have

$$\hat{\theta}^{\text{TSLS}} - \theta = \left(\hat{\mathbf{Q}}_{WZ} \hat{\mathbf{Q}}_{ZZ}^{-1} \hat{\mathbf{Q}}_{ZW} \right)^{-1} \hat{\mathbf{Q}}_{WZ} \hat{\mathbf{Q}}_{ZZ}^{-1} \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^\top \boldsymbol{\varepsilon}_g \right).$$

Under Part a of Assumption 5, $\text{plim} \hat{\mathbf{Q}}_{WZ} = \mathbf{Q}_{WZ}$, and $\text{plim} \hat{\mathbf{Q}}_{ZZ} = \mathbf{Q}_{ZZ}$ with $\mathbf{Q}_{ZZ}$ not singular. Consequently, $\text{plim} \hat{\mathbf{Q}}_{WZ} \hat{\mathbf{Q}}_{ZZ}^{-1} \hat{\mathbf{Q}}_{ZW} = \mathbf{Q}_{WZ} \mathbf{Q}_{ZZ}^{-1} \mathbf{Q}_{ZW}$. By Part b of Assumption 5, $\text{plim} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^\top \boldsymbol{\varepsilon}_g = 0$. Hence $\hat{\theta}^{\text{TSLS}}$ converges in probability to θ , as $m \rightarrow \infty$.

Moreover,

$$\sqrt{m}(\hat{\theta}^{\text{TSLS}} - \theta) = \left(\hat{\mathbf{Q}}_{WZ} \hat{\mathbf{Q}}_{ZZ}^{-1} \hat{\mathbf{Q}}_{ZW} \right)^{-1} \hat{\mathbf{Q}}_{WZ} \hat{\mathbf{Q}}_{ZZ}^{-1} \left(\frac{1}{\sqrt{m}} \sum_{g=1}^m \mathbf{Z}_g^\top \boldsymbol{\varepsilon}_g \right).$$

Using Assumption 5(a)–(c), $\sqrt{m}(\hat{\theta}^{\text{TSLs}} - \theta) \xrightarrow{d} \mathcal{N}(0, \mathbf{\Omega})$, where

$$\mathbf{\Omega} = (\mathbf{Q}_{WZ} \mathbf{Q}_{ZZ}^{-1} \mathbf{Q}_{ZW})^{-1} \mathbf{Q}_{WZ} \mathbf{Q}_{ZZ}^{-1} \mathbf{\Omega}_{\varepsilon} \mathbf{Q}_{ZZ}^{-1} \mathbf{Q}_{ZW} (\mathbf{Q}_{WZ} \mathbf{Q}_{ZZ}^{-1} \mathbf{Q}_{ZW})^{-1}.$$

□

B.4 Reduced-form Representation of the Model

To illustrate why it is not possible to find lower-level conditions for Part b. of Assumption 3, such as those found for the LiM model, as well as motivate our choice of instruments, we will next derive the reduced form of our model. To begin, note that Proposition 1 guarantees there is a unique equilibrium $\mathbf{y}_g$ given a realization of ε_g . Let π_g denote an ordering on individuals in network g , $\{1, \dots, n_g\}$, in terms of their outcomes. For example, $\pi_g(1) = 3$ means that individual 1's outcome $y_{1,g}$ is the third smallest among $\mathbf{y}_g$. Let $\mathbb{B}_g(\pi_g)$ denote a $n_g \times n_g$ peer effect coefficient matrix such that its i -th row j -th column component corresponds to the peer effect coefficient for individual j 's outcome on individual i 's outcome in network g , given the ordering π_g . Then, Proposition 1 implies that there is a unique π such that

$$\mathbf{y}_g = \mathbb{B}_g(\pi_g) \mathbf{y}_g + \mathbf{X}_g \gamma + \bar{\mathbf{X}}_g \delta + \boldsymbol{\epsilon}_g$$

where $\boldsymbol{\epsilon}_g = (\epsilon_{1,g} \ \dots \ \epsilon_{n_g,g})^\top$, and therefore

$$\mathbf{y}_g = (I_{n_g} - \mathbb{B}_g(\pi_g))^{-1} (\mathbf{X}_g \gamma + \bar{\mathbf{X}}_g \delta + \boldsymbol{\epsilon}_g).$$

By taking expectation over the set of every possible ordering, Π , we obtain a reduced-form representation of $\mathbb{E}_g[\mathbf{y}_g]$. The following corollary gives the result of this procedure.

Corollary 1. *Let Assumptions 1 and 2 hold. Then, the peer effects model (1) admits a reduced-form representation with residual terms η_i :*

$$\mathbb{E}_g[y_i] = \sum_{j=1}^n \theta_{i,j} (\mathbf{x}_j^\top \gamma + \bar{\mathbf{x}}_j^\top \delta) + \eta_i$$

where

$$\begin{aligned} \theta_{i,j} &= \sum_{\pi \in \Pi} \theta_{i,j}(\pi) \Pr_n \{\pi\} \\ \eta_i &= \sum_{\pi \in \Pi} \sum_{j=1}^n \theta_{i,j}(\pi) \mathbb{E}_g[\varepsilon_j | \pi] \cdot \Pr_n \{\pi\} \end{aligned}$$

and $(\theta_{i,1}(\pi), \dots, \theta_{i,n}(\pi))$ is the i -th row of the $n \times n$ matrix $(I_n - \mathbb{B}(\pi))^{-1}$. Moreover, there also exists a reduced-form relationship between the ordered peer outcome $\tilde{y}_{i,k}$ and $\{\mathbf{x}_j\}_{j=1}^n$, with residual terms $\{\tilde{\eta}_{j,k}\}_{j=1}^n$:

$$\mathbb{E}_g [\tilde{y}_{i,k}] = \sum_{j=1}^n \tilde{\theta}_{i,k,j}(\mathbf{x}_j^\top \gamma + \bar{\mathbf{x}}_j^\top \delta) + \tilde{\eta}_{i,k}$$

where

$$\begin{aligned} \tilde{\theta}_{i,k,j} &= \sum_{\pi \in \Pi} \tilde{\theta}_{i,k,j}(\pi) \Pr_n \{\pi\} \\ \tilde{\eta}_{i,k} &= \sum_{\pi \in \Pi} \sum_{j=1}^n \tilde{\theta}_{i,k,j}(\pi) \mathbb{E}_g [\varepsilon_j | \pi] \cdot \Pr_n \{\pi\}. \end{aligned}$$

and $(\tilde{\theta}_{i,k,1}(\pi), \dots, \tilde{\theta}_{i,k,n}(\pi))$ is one row of the $n \times n$ matrix $(I_n - \mathbb{B}(\pi))^{-1}$, which corresponds to the k -th lowest performing peer of individual i .

Proof. Let π denote an ordering on $\{1, \dots, n\}$ with some tiebreaking rule, where the ordering comes from $\mathbf{y}$:

$$\pi : \{1, \dots, n\} \mapsto \{1, \dots, n\}$$

and π satisfies $y_{\pi(1)} \leq \dots \leq y_{\pi(n)}$. Note that π is a function of $\{\varepsilon_i\}_{i=1}^n$; from Proposition 1, we have shown that there exists a unique equilibrium and thus a unique ordering for each realization $\{\varepsilon_i\}_{i=1}^n$. To simplify notation, let $\mathbf{x}_i$ denote both the covariates and the contextual effects $\bar{\mathbf{x}}_i$, and equivalently redefine γ to contain the contextual effect parameters δ . Let Π denote the set of all possible orderings on $\{y_1, \dots, y_n\}$. Π is a function of $\{\mathbf{x}_i, a_{ij}\}_{i,j}$ and the distribution of $\{\varepsilon_i\}_i$.

For any realization of $\{\varepsilon_i\}_{i=1}^n$ and the corresponding ordering π , there is a reduced-form linear relationship between $\{y_i\}_{i=1}^n$ and $\{\mathbf{x}_i\}_{i=1}^n$: construct a $n \times n$ matrix $\mathbb{B}(\pi)$ such that

$$\begin{aligned} \mathbb{B}(\pi) &= \left(\beta_{ij}(\pi) \right)_{i,j} \\ \beta_{ij}(\pi) &= \begin{cases} 0 & \text{if } a_{ij} = 0 \\ \beta_{k,d_i} \text{ for some } k \text{ s.t. } \sum_{j'=1}^{\pi^{-1}(j)} a_{i\pi(j')} = k & \text{if } a_{ij} = 1 \end{cases} \end{aligned}$$

$\mathbb{B}(\pi)$ takes the ordering π as fixed and finds the corresponding rank-dependent coefficient $\beta_{k,d}$ for each of the peer outcomes. To show that $(I_n - \mathbb{B}(\pi))$ is invertible for any π : suppose that

there is some nonzero $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{x} = \mathbb{B}(\pi)\mathbf{x}$. Then, for some i such that $|x_i| = \|\mathbf{x}\|_\infty$,

$$|x_i| = \left| \sum_{j=1}^n \beta_{ij}(\pi) x_j \right| \leq \sum_{j=1}^j |\beta_{ij}(\pi)| \cdot |x_j| \leq \bar{\beta} |x_i| < |x_i|,$$

leading to a contradiction; $I - \mathbb{B}(\pi)$ is full rank. Then,

$$\begin{aligned} \mathbf{Y} &= \mathbb{B}(\pi)\mathbf{Y} + \mathbf{X}\gamma + \mathbb{E} \\ \mathbf{Y} &= (\mathbf{I} - \mathbb{B}(\pi))^{-1} \mathbf{X}\gamma + (I_n - \mathbb{B}(\pi))^{-1} \mathbb{E}. \end{aligned}$$

We have a reduced-form linear relationship between $\{y_i\}_{i=1}^n$ and $\{\mathbf{x}_i\}_{i=1}^n$.

Since the reduced-form relationship holds for every realization of $\{\epsilon_i\}_{i=1}^n$ such that the ordering π stays the same, we can consider a conditional expectation of the linear relationship given the event that $\{\epsilon_i\}_{i=1}^n$ induces the ordering π . Then

$$\begin{aligned} \mathbb{E}_g[y_i] &= \sum_{\pi \in \Pi} \sum_{j=1}^n \theta_{ij}(\pi) (\mathbf{x}_j^\top \gamma + \mathbb{E}_g[\epsilon_j|\pi]) \cdot \Pr_n\{\pi\} \\ &= \sum_{j=1}^n \theta_{ij} \mathbf{x}_j^\top \gamma + \eta_i \end{aligned}$$

where $(\theta_{i1}(\pi), \dots, \theta_{in}(\pi))$ is the i -th row of the $n \times n$ matrix $(I - \mathbb{B}(\pi))^{-1}$ and

$$\begin{aligned} \theta_{ij} &= \sum_{\pi \in \Pi} \theta_{ij}(\pi) \Pr_n\{\pi\} \\ \eta_i &= \sum_{\pi \in \Pi} \sum_{j=1}^n \theta_{ij}(\pi) \mathbb{E}_g[\epsilon_j|\pi] \cdot \Pr_n\{\pi\}. \end{aligned}$$

Likewise, by taking a different student while summing over the ordering π , we get

$$\begin{aligned} \mathbb{E}_g[\tilde{y}_{i,k}] &= \sum_{\pi \in \Pi} \sum_{j=1}^n \tilde{\theta}_{i,k,j}(\pi) (\mathbf{x}_j^\top \gamma + \mathbb{E}_g[\epsilon_j|\pi]) \cdot \Pr_n\{\pi\} \\ &= \sum_{j=1}^n \tilde{\theta}_{i,k,j} \mathbf{x}_j^\top \gamma + \tilde{\eta}_{i,k} \end{aligned}$$

where $(\tilde{\theta}_{i,k,1}(\pi), \dots, \tilde{\theta}_{i,k,n}(\pi))$ is one row of the $n \times n$ matrix $(I_n - \mathbb{B}(\pi))^{-1}$, which corresponds

to the k -th lowest performing peer of student i , and

$$\begin{aligned}\tilde{\theta}_{i,k,j} &= \sum_{\pi \in \Pi} \tilde{\theta}_{i,k,j}(\pi) \Pr_n \{\pi\} \\ \tilde{\eta}_{i,k} &= \sum_{\pi \in \Pi} \sum_{j=1}^n \tilde{\theta}_{i,k,j}(\pi) \mathbb{E}_g [\varepsilon_j | \pi] \cdot \Pr_n \{\pi\}.\end{aligned}$$

□

B.5 Encompassing Test on Empirical Specifications

The quantile specification discussed in Subsection 2.2 has the merits of parsimony and tractability. However, an important prerequisite for using such a restricted model specification is determining whether the model restriction sufficiently captures the heterogeneous pattern in the construction of peer effects, since misspecifying the model restriction can lead to incorrect counterfactual analyses and misguided policy recommendations. To check this empirically, we make use of the encompassing test proposed by Smith (1992), which compares two possibly misspecified model restrictions. Thus, the encompassing test provides us with data-driven guidance on the choice of the model specification.

For notational brevity, let h denote a true model, which potentially follows a tractable specification from Subsection 2.2

$$y_{i,g} = \sum_{k=1}^M \beta^k \tilde{y}_{i,g}^k + \mathbf{x}_{i,g}^\top \gamma + \bar{\mathbf{x}}_{i,g}^\top \delta + \varepsilon_i.$$

By stacking at the network level, we can rewrite the model as

$$\mathbf{y}_g = \mathbf{W}_g \theta + \boldsymbol{\varepsilon}_g \tag{10}$$

where $\mathbf{W}_g$ is the subnetwork-level matrix stacking of $\mathbf{w}_{i,g}$ and θ is the model parameters $(\beta^0, \dots, \beta^M, \gamma, \delta)^\top$.

Suppose that we are given two working model restrictions h_1 and h_2 and are given two sets of instruments specific to model restrictions $\{\mathbf{Z}_g^1\}_{g=1}^m$ and $\{\mathbf{Z}_g^2\}_{g=1}^m$, while allowing (true) h to potentially differ from both h_1 and h_2 . The encompassing test is constructed for a pair of model restrictions. In practice, we consider a sequence of model restrictions with growing complexity. For example, in the quantile peer effect model, we gradually increase the number of quantile levels to include in the model restriction whenever a given specification is rejected in favor of another specification that includes more quantile levels.

The TSLS estimators from the respective model restrictions are

$$\hat{\theta}^s = (\mathbf{H}^s)^{-1} \cdot \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{s\top} \mathbf{Z}_g^s \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{s\top} \mathbf{Z}_g^s \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{s\top} \mathbf{y}_g$$

where $\mathbf{H}^s = \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{s\top} \mathbf{Z}_g^s \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{s\top} \mathbf{Z}_g^s \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{s\top} \mathbf{W}_g^s$ for $s = 1, 2$. To discuss the behaviors of the two TSLS estimators under misspecification, we assume regularity on the instruments.

Assumption 6. For any $s \in \{1, 2\}$,

- a. $\text{plim} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{s\top} \mathbf{Z}_g^s$ is a nonstochastic full-rank matrix
- b. $\text{plim} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{s\top} \mathbf{W}_g^s$ is a nonstochastic full-column rank matrix.
- c. $\text{plim} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{s\top} \boldsymbol{\varepsilon}_g = \mathbf{0}$.

Then, we can define pseudo-true coefficient parameters

$$\theta^s = \text{plim}_{m \rightarrow \infty} (\mathbf{H}^s)^{-1} \cdot \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{s\top} \mathbf{Z}_g^s \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{s\top} \mathbf{Z}_g^s \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{s\top} \mathbf{W}_g^s \theta \quad \forall s = 1, 2.$$

Importantly, whenever $\mathbf{W}_g \theta$ lies in the column space of $\mathbf{W}_g^s$ for every g , $\mathbf{W}_g^s \theta^s = \mathbf{W}_g \theta$; even though the working model h_s differs from the true model h , the misspecified model still retrieves the systemic variation of the true model h . Using θ^1 and θ^2 , we rewrite Equation (10) as:

$$\mathbf{y}_g = \mathbf{W}_g^s \theta^s + \boldsymbol{\varepsilon}_g^s \quad \forall s = 1, 2 \quad (11)$$

where $\boldsymbol{\varepsilon}_g^s = \boldsymbol{\varepsilon}_g + \mathbf{W}_g \theta - \mathbf{W}_g^s \theta^s$.

The encompassing test (Smith, 1992) assesses whether the first model specified with h_1 and therefore θ^1 replicates the features of the second model specified with h_2 . Let

$$\psi = \text{plim}_{m \rightarrow \infty} (\mathbf{H}^2)^{-1} \cdot \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{2\top} \mathbf{Z}_g^2 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{Z}_g^2 \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \boldsymbol{\varepsilon}_g^1.$$

ψ is the TSLS estimand of regressing $\{\boldsymbol{\varepsilon}_g^1\}_{g=1}^m$, the residual vectors under the first model restriction, on the second model restriction regressors $\{\mathbf{W}_g^2\}_{g=1}^m$ using the second model restriction instruments $\{\mathbf{Z}_g^2\}_{g=1}^m$. $\psi \neq 0$ suggests that features of the observed data relevant for the specification with h_2 are not completely captured by the model with h_1 , leading to instrumented variation of $\mathbf{W}_g^2$ being correlated with $\boldsymbol{\varepsilon}_g^1$. In that case, the first model restriction must be rejected in favor of the second.

The test statistic of the encompassing test is

$$\hat{\psi} = (\mathbf{H}^2)^{-1} \cdot \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{2\top} \mathbf{Z}_g^2 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{Z}_g^2 \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \hat{\epsilon}_g^1$$

where $\hat{\epsilon}_g^1 = \mathbf{y}_g - \mathbf{W}_g^1 \hat{\theta}^1$. Proposition 4 is the formal asymptotic normality result for the test statistic.

Proposition 4. *Under Assumptions 2 and 6, and the regularity conditions of Assumption 7 (see Appendix B.6):*

- a. $\hat{\psi}$ is a consistent estimator of ψ .
- b. $\sqrt{m}(\hat{\psi} - \psi)$ converges weakly to $N(\mathbf{0}, \mathbf{\Omega}_\psi)$. See Appendix B.6 for details on the asymptotic variance $\mathbf{\Omega}_\psi$.

Proof. See Appendix B.6. □

Since the $\hat{\epsilon}_g^1$ is obtained from an initial regression, the asymptotic variance of $\sqrt{m}(\hat{\psi} - \psi)$ depends on the sampling variability from this initial regression as well; a naive IV standard error obtained from treating $\hat{\epsilon}_g^1$ as ϵ_g^1 would be invalid. Another important complication in the asymptotic theory is that both h_1 and h_2 may differ from h , neither being correctly specified. We therefore use an approach similar to those employed under misspecified moment functions to derive the asymptotic distribution (Hall and Inoue, 2003).

A merit of the encompassing test is that it remains consistent even when both models are misspecified, under some conditions. Note that the test statistic is asymptotically centered around ψ . Thus, the test is consistent as long as $\psi \neq 0$ under an alternative. Since

$$\begin{aligned} \psi &= (\mathbf{H}^2)^{-1} \cdot \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{2\top} \mathbf{Z}_g^2 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{Z}_g^2 \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \epsilon_g^1 \\ &= (\mathbf{H}^2)^{-1} \cdot \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{2\top} \mathbf{Z}_g^2 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{Z}_g^2 \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} (\mathbf{W}_g \theta - \mathbf{W}_g^1 \theta^1), \end{aligned}$$

ψ equals to a zero vector if and only if the structural part of the true model, $\mathbf{W}_g \theta$, and its approximation under the first model restriction, $\mathbf{W}_g^1 \theta^1$, are equivalent from the perspective of

$$(\mathbf{H}^2)^{-1} \cdot \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{2\top} \mathbf{Z}_g^2 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{Z}_g^2 \right)^{-1} \mathbf{Z}_g^{2\top}.$$

Thus, even though the test seemingly tests the null hypothesis of the first model restriction against the alternative hypothesis of the second model restriction, the test still has power against an alternative hypothesis which is not the second model restriction, as long as $(\mathbf{H}^2)^{-1} \cdot \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{2\top} \mathbf{Z}_g^2 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{Z}_g^2 \right)^{-1} \mathbf{Z}_g^{2\top}$ preserves the nonzero directions of

$$\{\mathbf{W}_g \theta - \mathbf{W}_g^1 \theta^1\}_{g=1}^m.$$

B.6 Proof for Proposition 4

We start by showing that $\hat{\psi}$ is consistent for ψ :

$$\begin{aligned} \text{plim}_{m \rightarrow \infty} \hat{\psi} - \psi &= \text{plim}_{m \rightarrow \infty} (\mathbf{H}^2)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{2\top} \mathbf{Z}_g^2 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{Z}_g^2 \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} (\hat{\boldsymbol{\epsilon}}_g^1 - \boldsymbol{\epsilon}_g^1) \\ &= \text{plim}_{m \rightarrow \infty} (\mathbf{H}^2)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{2\top} \mathbf{Z}_g^2 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{Z}_g^2 \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} (\mathbf{y}_g - \mathbf{W}_g^1 \hat{\theta}^1 - \boldsymbol{\epsilon}_g^1) \\ &= \text{plim}_{m \rightarrow \infty} (\mathbf{H}^2)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{2\top} \mathbf{Z}_g^2 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{Z}_g^2 \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{W}_g^1 (\theta^1 - \hat{\theta}^1) = 0. \end{aligned}$$

The second and third equalities use $\mathbf{y}_g = \mathbf{W}_g^1 \theta^1 + \boldsymbol{\epsilon}_g^1$. The last equality holds from Assumption 6. This completes the proof of the first statement of the proposition.

To show that $\sqrt{m}(\hat{\psi} - \psi)$ is asymptotically normally distributed, we define

$$\tilde{\boldsymbol{\epsilon}}_g := \boldsymbol{\epsilon}_g^1 - \mathbf{W}_g^2 \psi,$$

where $\tilde{\boldsymbol{\epsilon}}_g$ is the error term of the linear approximation of $\boldsymbol{\epsilon}_1$ onto $\mathbf{W}_g^2$. Recall that $\psi = \mathbf{0}$ when h_1 is true. The main challenge in deriving the asymptotic distribution is that h_1 may be misspecified, in which case both $\mathbb{E}_g[\boldsymbol{\epsilon}_g^1]$ and $\mathbb{E}_g[\tilde{\boldsymbol{\epsilon}}_g]$ may differ from zero. Let $\mathbf{u}_1 = \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{1\top} \boldsymbol{\epsilon}_g^1$, $\mathbf{u}_2 = \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \tilde{\boldsymbol{\epsilon}}_g$. For $s = 1, 2$, let $\mathbf{u}_{s,3} = \frac{1}{m} \sum_{g=1}^m \text{vec}(\mathbf{W}_g^{s\top} \mathbf{Z}_g^s)$ and $\mathbf{u}_{s,4} = \frac{1}{m} \sum_{g=1}^m \text{vec}(\mathbf{Z}_g^{s\top} \mathbf{Z}_g^s)$, where the operator vec vectorizes a matrix by stacking its columns sequentially. Lastly, let $\mathbf{u} = (\mathbf{u}_1^\top, \mathbf{u}_2^\top, \mathbf{u}_{1,3}^\top, \mathbf{u}_{1,4}^\top, \mathbf{u}_{2,3}^\top, \mathbf{u}_{2,4}^\top)^\top$. We introduce the following conditions.

Assumption 7.

- a.* For any $s \in \{1, 2\}$, $\text{plim}_{m \rightarrow \infty} (\mathbf{u}_{s,3} - \mathbb{E}[\mathbf{u}_{s,3}]) = 0$ and $\text{plim}_{m \rightarrow \infty} (\mathbf{u}_{s,4} - \mathbb{E}[\mathbf{u}_{s,4}]) = 0$.

b. $\text{plim}_{m \rightarrow \infty} (\mathbf{u}_1 - \mathbb{E}[\mathbf{u}_1]) = 0$ and $\text{plim}_{m \rightarrow \infty} (\mathbf{u}_2 - \mathbb{E}[\mathbf{u}_2]) = 0$.

c. $\sqrt{m}(\mathbf{u} - \mathbb{E}[\mathbf{u}])$ converges in distribution to $\mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}_{\mathbf{u}})$.

Assumption 7 imposes weak regularity conditions to establish the asymptotic distribution of $\sqrt{m}(\hat{\psi} - \psi)$. Conditions **a.** and **b.** follow from the weak law of large numbers applied to $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_{s,3}$, and $\mathbf{u}_{s,4}$, which are sample means. Condition **c.** follows from standard central limit theorems (CLT) because $\mathbf{u}$ is a sample mean. The reason for also requiring the CLT to apply to $\mathbf{u}_{s,3}$ and $\mathbf{u}_{s,4}$ is that $\mathbb{E}[\mathbf{u}_1]$ and $\mathbb{E}[\mathbf{u}_2]$ may differ from zero when both h_1 and h_2 are misspecified (Hall and Inoue, 2003). We have

$$\begin{aligned}
& \sqrt{m}(\hat{\psi} - \psi) \\
&= \sqrt{m}(\mathbf{H}^2)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{2\top} \mathbf{Z}_g^2 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{Z}_g^2 \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \hat{\boldsymbol{\varepsilon}}_g^1 - \sqrt{m}\psi \\
&= \sqrt{m}(\mathbf{H}^2)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{2\top} \mathbf{Z}_g^2 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{Z}_g^2 \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} (\mathbf{y}_g - \mathbf{W}_g^1 \hat{\theta}^1) - \sqrt{m}\psi \\
&= \sqrt{m}(\mathbf{H}^2)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{2\top} \mathbf{Z}_g^2 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{Z}_g^2 \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} (\mathbf{W}_g^1 (\theta^1 - \hat{\theta}^1) + \boldsymbol{\varepsilon}_g^1) - \sqrt{m}\psi \\
&= \sqrt{m}(\mathbf{H}^2)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{2\top} \mathbf{Z}_g^2 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{Z}_g^2 \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} (\mathbf{W}_g^1 (\theta^1 - \hat{\theta}^1) + \tilde{\boldsymbol{\varepsilon}}_g) \\
&= \sqrt{m}(\mathbf{H}^2)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{2\top} \mathbf{Z}_g^2 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{Z}_g^2 \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \tilde{\boldsymbol{\varepsilon}}_g \\
&\quad - \sqrt{m}(\mathbf{H}^2)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{2\top} \mathbf{Z}_g^2 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{Z}_g^2 \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{2\top} \mathbf{W}_g^1 \\
&\quad \cdot (\mathbf{H}^1)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{W}_g^{1\top} \mathbf{Z}_g^1 \left(\frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{1\top} \mathbf{Z}_g^1 \right)^{-1} \frac{1}{m} \sum_{g=1}^m \mathbf{Z}_g^{1\top} \boldsymbol{\varepsilon}_g^1
\end{aligned}$$

From the last equation, even though $\mathbb{E}[\mathbf{Z}_g^{1\top} \boldsymbol{\varepsilon}_g^1]$ and $\mathbb{E}[\mathbf{Z}_g^{2\top} \tilde{\boldsymbol{\varepsilon}}_g]$ are not zero, we can adapt results from the misspecified moment condition framework (see Hall and Inoue, 2003). Specifically, Assumption 7 guarantees that $\sqrt{m}(\hat{\psi} - \psi)$ weakly converges to a normal distribution with mean zero. The exact expression of the asymptotic variance depends on $\boldsymbol{\Omega}_{\mathbf{u}}$ and is given in (Hall and Inoue, 2003). In practice, we estimate this asymptotic variance using a paired bootstrap method.

C Supplementary Simulation Results

C.1 Comparing OLS and TSLS Estimates

Table C.1 compares the OLS estimator and the TSLS estimator using ordered peer covariates as instruments in a sparse network setting, setting $\bar{d} = 2$. We estimate the models with small samples, consisting of either 2 or 5 networks, implying either 100 or 250 total observations. All estimators include network fixed effects.

As expected, OLS is biased, and the bias is unaffected by the sample size. However, when we increase the strength of the covariates by increasing their coefficients in absolute value, the bias in the OLS estimator decreases. This is caused by the increased covariate strength reducing the relative share of endogeneity in the outcomes, which then reduces the impact of the reflection bias.

C.2 Simulations Without Contextual Effects

For completeness, we here present our simulation results from the same DGP as the main text, except we have removed the contextual effects. The results are qualitatively similar, with any differences likely stemming from the relative increase in ε 's share of the variance in the outcomes.

Table C.1: Comparison between the OLS estimator and the TSLS estimator

	Bias				MSE			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Rank Models								
γ_1^{OLS}	-0.021	-0.026	-0.022	-0.038	0.009	0.010	0.005	0.007
$\beta_{1,2}^{OLS}$	0.044	0.021	0.033	0.017	0.006	0.002	0.003	0.001
$\beta_{2,2}^{OLS}$	-0.013	-0.007	0.006	0.004	0.004	0.001	0.002	0.001
γ_1^{IV}	0.001	0.001	0.001	-0.002	0.010	0.010	0.005	0.006
$\beta_{1,2}^{IV}$	0.012	0.003	-0.000	0.001	0.012	0.002	0.007	0.002
$\beta_{2,2}^{IV}$	-0.008	-0.002	0.002	0.000	0.011	0.002	0.007	0.001
Panel A: Quantile Models								
γ_1^{OLS}	-0.023	-0.027	-0.023	-0.038	0.009	0.010	0.005	0.007
β_{min}^{OLS}	0.034	0.016	0.027	0.015	0.004	0.001	0.002	0.001
β_{max}^{OLS}	-0.004	-0.002	0.012	0.006	0.003	0.001	0.001	0.000
γ_1^{IV}	0.001	0.001	0.001	-0.002	0.009	0.010	0.005	0.006
β_{min}^{IV}	0.003	0.001	-0.001	0.000	0.006	0.001	0.002	0.001
β_{max}^{IV}	-0.001	-0.001	0.001	0.000	0.005	0.001	0.002	0.001
γ	1	2	1	2	1	2	1	2
<i>networks</i>	2	2	5	5	2	2	5	5

Notes: The values of the coefficients used in DGP are as follows: $\gamma_0 = \gamma \times (1, 0.25)$ and $\beta = (0.2, 0.5)$. The instruments used in TSLS estimation for the rank models are peers' covariates ordered by their own values. The quantile models use the quantiles of the distribution of each covariate amongst a persons peers.

Table C.2: Comparison between the OLS estimator and the TSLS estimator

	Bias				MSE			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Rank Models								
γ_1^{OLS}	-0.005	-0.003	-0.001	-0.001	0.009	0.009	0.005	0.005
$\beta_{1,2}^{OLS}$	0.027	0.008	0.030	0.010	0.003	0.001	0.002	0.001
$\beta_{2,2}^{OLS}$	-0.025	-0.007	-0.027	-0.009	0.003	0.001	0.002	0.001
γ_1^{IV}	0.003	0.003	-0.001	-0.001	0.009	0.009	0.005	0.005
$\beta_{1,2}^{IV}$	0.005	0.001	0.002	0.001	0.010	0.002	0.009	0.002
$\beta_{2,2}^{IV}$	-0.005	-0.001	-0.002	-0.001	0.009	0.002	0.008	0.002
Panel A: Quantile Models								
γ_1^{OLS}	-0.005	-0.002	-0.003	-0.002	0.009	0.009	0.005	0.005
β_{min}^{OLS}	0.017	0.005	0.018	0.006	0.002	0.001	0.001	0.000
β_{max}^{OLS}	-0.016	-0.004	-0.016	-0.005	0.002	0.000	0.001	0.000
γ_1^{IV}	0.002	0.002	-0.001	-0.001	0.009	0.009	0.005	0.005
β_{min}^{IV}	0.000	-0.000	0.001	0.000	0.003	0.001	0.002	0.001
β_{max}^{IV}	-0.000	0.000	-0.000	-0.000	0.003	0.001	0.002	0.001
γ	1	2	1	2	1	2	1	2
<i>networks</i>	2	2	5	5	2	2	5	5

Notes: The values of the coefficients used in DGP are as follows: $\gamma_0 = \gamma \times (1, 0.25)$ and $\beta = (0.2, 0.5)$. The instruments used in TSLS estimation for the rank models are peers' covariates ordered by their own values. The quantile models use the quantiles of the distribution of each covariate amongst a persons peers.

Table C.3: Finite sample performance of Rank and Quantile models

	Bias					MSE				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Saturated model										
$\beta_{1,2}$	-0.001	-0.001	-0.001	-0.000	0.000	0.007	0.007	0.004	0.002	0.001
$\beta_{2,2}$	0.001	0.001	0.001	0.000	-0.000	0.006	0.006	0.004	0.002	0.001
$\beta_{1,5}$	-0.007	0.012	-0.001	0.000	0.000	0.088	0.067	0.060	0.015	0.015
$\beta_{2,5}$	0.012	-0.022	0.003	0.003	-0.003	0.323	0.288	0.260	0.092	0.135
$\beta_{3,5}$	-0.008	0.005	-0.004	-0.011	0.005	0.352	0.312	0.281	0.240	0.271
$\beta_{4,5}$	0.014	0.010	0.007	0.012	-0.003	0.253	0.271	0.161	0.169	0.147
$\beta_{5,5}$	-0.010	-0.004	-0.005	-0.005	0.001	0.084	0.072	0.045	0.027	0.015
Panel B: Restricted model										
β_{τ_1}	0.008	0.001	-0.003	-0.000	-0.001	0.035	0.047	0.022	0.005	0.003
β_{τ_2}	-0.024	0.001	0.008	-0.002	0.001	0.227	0.335	0.160	0.047	0.027
β_{τ_3}	0.022	-0.002	-0.005	0.004	-0.000	0.224	0.289	0.138	0.056	0.030
β_{τ_4}	-0.006	0.001	0.000	-0.002	-0.000	0.034	0.033	0.016	0.008	0.004
n_g	5	10	20	50	100	5	10	20	50	100
$\bar{d}$	5	5	5	5	5	5	5	5	5	5

Notes: The data generating process is the quantile model with $\beta_\tau = (-0.05, 0.35, 0.15, 0.1)$. n_g gives the number of networks used in the simulation and $\bar{d}$ is the maximum number of peers for each individual. All the simulations include network fixed effects. The instruments used is the distribution of the peers covariate values, as measured either by the ranked peer covariate values or the quantiles of peer covariate values. The results are based on 10,000 draws.

Table C.4: Monte Carlo Simulations — Encompassing tests

2 qtls. vs. 3 qtls.		3 qtls. vs. 4 qtls.		4 qtls. vs. 5 qtls.	
5%	10%	5%	10%	5%	10%
DGP A: $\beta = (0, 0.05, 0.2, 0.3)$					
1.000	1.000	0.201	0.307	0.008	0.020
DGP B: $\beta = (0.3, 0.2, 0.05, 0)$					
1.000	1.000	0.987	0.994	0.013	0.027
DGP C: $\beta = (0, 0.275, 0.275, 0)$					
1.000	1.000	0.999	1.000	0.008	0.021
DGP D: $\beta = (0.275, 0, 0, 0.275)$					
0.011	0.023	0.008	0.016	0.012	0.026
DGP E: $\beta = (-0.05, 0.35, 0.15, 0.1)$					
1.000	1.000	1.000	1.000	0.010	0.023
DGP F (LIM model): $\beta = 0.55$					
1.000	1.000	0.991	0.997	0.703	0.815

The columns labeled “ a qtls. vs. b qtls.”, for integers a and b , report the share of rejections of the null hypothesis that the model with a quantile levels does not perform worse than the model with b quantile levels, at the significance levels indicated in the second row.

Table C.5: Monte Carlo Simulations — Estimation of the Quantile Specification

Quantiles				LiM	CES	
β^1	β^2	β^3	β^4	β	ρ	β
DGP A: $\beta = (0, 0.05, 0.2, 0.3)$						
-0.000 (0.004)	0.050 (0.018)	0.200 (0.037)	0.300 (0.025)	0.544 (0.003)	34.759 (12.301)	0.554 (0.004)
DGP B: $\beta = (0.3, 0.2, 0.05, 0)$						
0.300 (0.006)	0.200 (0.027)	0.050 (0.049)	0.000 (0.030)	0.544 (0.004)	-3.165 (0.165)	0.530 (0.003)
DGP C: $\beta = (0, 0.275, 0.275, 0)$						
0.000 (0.004)	0.275 (0.020)	0.275 (0.040)	0.000 (0.027)	0.544 (0.003)	2.918 (0.243)	0.552 (0.003)
DGP D: $\beta = (0.275, 0, 0, 0.275)$						
0.275 (0.005)	-0.000 (0.023)	0.000 (0.044)	0.275 (0.028)	0.555 (0.003)	-1.102 (0.132)	0.545 (0.003)
DGP E: $\beta = (-0.05, 0.35, 0.15, 0.1)$						
-0.050 (0.004)	0.350 (0.020)	0.150 (0.039)	0.100 (0.026)	0.543 (0.003)	4.696 (0.553)	0.556 (0.003)
DGP F (LIM model): $\beta = 0.55$						
0.112 (0.005)	0.198 (0.024)	0.093 (0.046)	0.147 (0.029)	0.550 (0.003)	1.001 (0.078)	0.550 (0.003)

The models are simulated and estimated 1,000 times. Values without parentheses represent average peer effect estimates, while those in parentheses correspond to standard errors. The instrument matrix includes the quantiles of $\mathbf{x}$ and $\bar{\mathbf{x}}$ among friends, computed at ten levels uniformly spaced between 0 and 1. DGPs A–E are generated from the proposed quantile-based model, with four quantiles at $\{0, 1/3, 2/3, 1\}$, and $\beta = (\beta^0, \beta^1, \beta^2, \beta^3)$ is the vector of peer effects at each quantile. DGP F follows the standard LIM model with only spillover effects, where $\beta = 0.55$. All estimations account for unobserved subnetwork heterogeneity using fixed effects.

D Supplemental Results on the Empirical Application

Table D.1: Sample size

Outcome	<i>n</i>
Academic achievements (GPA)	70 022
Academic effort	76 261
Extracurricular activities	79 694
Future perception	75 822
Trouble at school	76 206
Smoking	74 904
Drinking	74 741
Risky behaviors	75 545
Self-esteem	71 649
Physical exercise	71 683
Fighting	71 606

Table D.2: Empirical Results with five quantile levels or without isolated students (IS)

Five quantile levels					Four quantile levels without IS			
β^1	β^2	β^3	β^4	β^5	β^1	β^2	β^3	β^4
Academic achievements (GPA)								
0.011 (0.054)	0.225 (0.104)	0.174 (0.139)	0.464 (0.135)	-0.129 (0.073)	0.061 (0.045)	0.176 (0.082)	0.641 (0.090)	-0.115 (0.054)
Academic effort								
0.110 (0.038)	0.157 (0.068)	0.110 (0.075)	0.073 (0.065)	0.106 (0.051)	0.144 (0.032)	0.151 (0.061)	0.147 (0.055)	0.099 (0.047)
Extracurricular activities								
-0.192 (0.097)	0.649 (0.155)	0.046 (0.139)	0.209 (0.076)	-0.015 (0.022)	-0.085 (0.080)	0.552 (0.118)	0.249 (0.079)	-0.004 (0.021)
Future perception								
0.157 (0.035)	0.030 (0.085)	0.166 (0.131)	0.073 (0.139)	0.068 (0.082)	0.149 (0.033)	0.138 (0.086)	0.167 (0.115)	0.073 (0.068)
Trouble at school								
0.122 (0.097)	-0.147 (0.153)	0.501 (0.122)	0.043 (0.088)	0.037 (0.046)	0.035 (0.075)	0.242 (0.113)	0.275 (0.074)	0.028 (0.041)
Smoking								
-0.082 (0.107)	0.227 (0.139)	0.245 (0.092)	0.279 (0.048)	0.085 (0.022)	-0.107 (0.083)	0.365 (0.095)	0.373 (0.055)	0.122 (0.019)
Drinking								
0.018 (0.204)	0.205 (0.333)	0.045 (0.195)	0.120 (0.068)	0.081 (0.016)	0.127 (0.137)	0.043 (0.180)	0.235 (0.083)	0.080 (0.015)
Risky behaviors								
-0.089 (0.130)	0.441 (0.219)	-0.065 (0.158)	0.321 (0.083)	0.085 (0.024)	-0.082 (0.095)	0.383 (0.150)	0.251 (0.082)	0.123 (0.021)
Self-esteem								
0.126 (0.058)	0.038 (0.134)	0.202 (0.181)	0.121 (0.106)	-0.026 (0.026)	0.112 (0.050)	0.157 (0.101)	0.239 (0.082)	-0.021 (0.025)
Physical exercise								
0.113 (0.055)	0.001 (0.077)	0.232 (0.085)	0.130 (0.096)	-0.040 (0.059)	0.092 (0.048)	0.164 (0.070)	0.193 (0.081)	-0.003 (0.047)
Fighting								
0.287 (0.104)	-0.088 (0.157)	0.067 (0.112)	0.160 (0.067)	0.165 (0.030)	0.236 (0.079)	-0.012 (0.112)	0.203 (0.068)	0.183 (0.027)

This table reports supplemental results from the empirical application using a model with five quantile levels (from 0 to 1, spaced by $\frac{1}{5}$) and a model with four quantile levels, as in the paper, but excluding isolated students (i.e., those with no friends).

Table D.3: Comparison of OLS and IV Estimates

OLS				IV				Hausman
β^1	β^2	β^3	β^4	β^1	β^2	β^3	β^4	(p -value)
Academic achievements (GPA)								
0.051 (0.009)	0.144 (0.016)	0.220 (0.017)	-0.027 (0.012)	0.061 (0.045)	0.176 (0.082)	0.641 (0.091)	-0.115 (0.055)	0.000
Academic effort								
0.027 (0.005)	0.045 (0.010)	0.075 (0.008)	0.004 (0.008)	0.144 (0.032)	0.151 (0.062)	0.147 (0.055)	0.099 (0.047)	0.000
Extracurricular activities								
0.008 (0.019)	0.205 (0.019)	0.083 (0.016)	0.004 (0.005)	-0.085 (0.081)	0.552 (0.118)	0.249 (0.079)	-0.004 (0.021)	0.000
Future perception								
0.028 (0.006)	0.064 (0.015)	0.065 (0.019)	-0.032 (0.014)	0.149 (0.032)	0.138 (0.087)	0.167 (0.114)	0.073 (0.067)	0.000
Trouble at school								
-0.050 (0.016)	0.083 (0.017)	0.048 (0.012)	-0.007 (0.006)	0.035 (0.075)	0.242 (0.114)	0.275 (0.074)	0.028 (0.042)	0.000
Smoking								
-0.106 (0.023)	0.239 (0.023)	0.299 (0.012)	0.055 (0.004)	-0.107 (0.084)	0.365 (0.095)	0.373 (0.055)	0.122 (0.019)	0.000
Drinking								
-0.141 (0.040)	0.126 (0.057)	0.173 (0.023)	0.020 (0.004)	0.127 (0.135)	0.043 (0.178)	0.235 (0.084)	0.080 (0.015)	0.000
Risky behaviors								
-0.176 (0.026)	0.290 (0.033)	0.125 (0.016)	0.030 (0.005)	-0.082 (0.094)	0.383 (0.149)	0.251 (0.082)	0.123 (0.021)	0.000
Self-esteem								
0.038 (0.008)	0.063 (0.017)	0.018 (0.014)	-0.015 (0.003)	0.112 (0.050)	0.157 (0.101)	0.239 (0.082)	-0.021 (0.025)	0.000
Physical exercise								
0.038 (0.007)	0.088 (0.011)	0.080 (0.013)	-0.029 (0.009)	0.092 (0.048)	0.164 (0.068)	0.193 (0.081)	-0.003 (0.048)	0.000
Fighting								
-0.045 (0.018)	0.098 (0.020)	0.069 (0.010)	0.031 (0.005)	0.236 (0.079)	-0.012 (0.112)	0.203 (0.068)	0.183 (0.027)	0.000

This table reports supplemental results from the empirical application where the model is estimated by OLS. The IV results correspond to the main estimates reported in the paper.

E Constructing Sample Quantiles

When $d_i = 0$, we set $\tilde{y}_i^k = 0$ for $k = 1, \dots, M$. This decision is innocuous if the model includes separate intercepts for isolated and non-isolated individuals. When $d_i = M > 0$, we simply let $\tilde{y}_i^k = \tilde{y}_{i,k}$ for $k = 1, \dots, M$.

When $d_i > 0$ and $d_i \neq M$, we compute sample quantiles using linear interpolation. For a given quantile index k and sample size d_i , we first compute the (possibly non-integer) position $h(k, d_i) = \frac{k-1}{M-1}(d_i - 1) + 1$, which maps the quantile level $(k - 1)/(M - 1)$ to a location within the ordered peer outcomes $\{\tilde{y}_{i,k}\}_{k=1}^{d_i}$. When $h(k, d_i)$ is not an integer, the k -th sample quantile $\tilde{y}_i^k$ is obtained by linearly interpolating between the two peer outcomes $\lfloor h(k, d_i) \rfloor$ and $\lfloor h(k, d_i) \rfloor + 1$, weighted by their respective distances from $h(k, d_i)$:

$$\tilde{y}_i^k = \begin{cases} \tilde{y}_{i, \lfloor h(k, d_i) \rfloor} & \text{if } h(k, d_i) = \lfloor h(k, d_i) \rfloor \\ \left(h(k, d_i) - \lfloor h(k, d_i) \rfloor \right) \tilde{y}_{i, \lfloor h(k, d_i) \rfloor + 1} \\ \quad + \left(\lfloor h(k, d_i) \rfloor + 1 - h(k, d_i) \right) \tilde{y}_{i, \lfloor h(k, d_i) \rfloor} & \text{if } h(k, d_i) \neq \lfloor h(k, d_i) \rfloor \end{cases}$$

This definition of the sample quantile is popular and is known as the Type 7 sample quantile (see [Hyndman and Fan, 1996](#)).