Submitted to *Econometrica*

TOWARD A GENERAL THEORY OF PEER EFFECTS

VINCENT BOUCHER
Department of Economics, University of Laval, Canada, CRREP and CREATE

MICHELLE RENDALL
Department of Economics, Monash University, Australia, and CEPR

PHILIP USHCHEV
ECARES, Universite libre de Bruxelles, Belgium

YVES ZENOU
Department of Economics, Monash University, Australia, CEPR, and IZA

There is substantial empirical evidence showing that peer effects matter in many activities. The workhorse model in empirical work on peer effects is the linear-in-means (LIM) model, whereby it is assumed that agents are linearly affected by the mean action of their peers. We develop a new general model of peer effects that relaxes the linear assumption of the best-reply functions and the mean peer behavior and that encompasses the spillover, conformist model, and LIM model as special cases. Then, using data on adolescent activities in the U.S., we structurally estimate this model. We find that for many activities, individuals do not behave according to the LIM model. We run some counterfactual policies and show that imposing the mean action as an individual social norm is misleading and leads to incorrect policy implications.

We thank Yannis Ioannides, Matthew O. Jackson, Efthymia Kyriakopoulou, Nicolás Salamanca, and the participants at the research seminar, Northwestern University, Johns Hopkins University, the Labour Economics Seminar, SOFI, the economics seminar, Nantes University, Tinbergen Institute, and the labour and public economics seminar, PSE, for all the very helpful comments. The financial support from the Australian Research Council Discovery Project DP240100158 is gratefully acknowledged by Yves Zenou. This research uses data from AddHealth, a program 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 is funded by Grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other 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 AddHealth data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from Grant P01-HD31921 for this research.

KEYWORDS: spillovers, conformism, structural estimation, policies.

1. INTRODUCTION

Individuals interact in all kinds of ways. In particular, they imitate and learn from the behavior of others, especially those close to them, such as their friends, neighbors, and colleagues. The impact of these interactions on individual behavior is referred to as peer effects. The decisions individuals take in the presence of peer effects generate externalities. There is substantial empirical evidence showing that peer effects matter in many contexts, such as education, crime, and program participation.[1] The overwhelming majority of empirical research assumes that individuals are affected by a linear function of the mean behavior of their peers and are silent about the underlying behavioral foundation generating the estimated peer effects.

Indeed, most peer-effect studies use the standard linear-in-means (LIM) model.[2] For example, the criminal activity of an individual is assumed to depend on the average criminal activity of the neighborhood where she lives. In education, each student compares herself with the average performance of students in her classroom, and so forth. The game theoretic foundation of the LIM model is a network model such that the best-reply function of each agent is linear in the mean action of her peers.[3] Moreover, it is now well recognized that the LIM model can be equivalently microfounded by games assuming either conformist preferences or positive spillovers.[4] With conformist preferences, individuals pay a cost for deviating from the average effort of their peers (their social norm) while, with spillover preferences, agents always benefit from a higher social norm.

In this paper, we develop a new general model of peer effects that relaxes the assumptions of linearity of the best-reply functions and the mean peer behavior of the LIM model and encompasses the spillover and conformist models as special cases. Instead of assuming

[1]See, e.g., Calvó-Armengol et al. (2009), Sacerdote (2011), and Dahl et al. (2014), and Lee et al. (2021).

[2]Manski (1993) was among the first to highlight the identification issues in estimating the LIM model, in particular, the reflection problem.

[3]See, e.g., Patacchini and Zenou (2012), Blume et al. (2015), Boucher (2016), Kline and Tamer (2020) and Ushchev and Zenou (2020).

[4]See Blume et al. (2015) and Boucher and Fortin (2016). In particular, Boucher and Fortin (2016) have highlighted the fact that both the conformist and spillover models can microfound the LIM model; they also have suggested, but not implemented, a way to identify them separately using isolated individuals.

that the social norm of each agent is given by the average action of her peers, we allow for more flexibility and define the social norm using a CES function with elasticity parameter β. In this model, the individual outcome is referred to as "effort", whether it is a positive (e.g., GPA) or a negative outcome (e.g., risky behavior). When β is equal to 1, we revert to the LIM model. When β is very large, we obtain the "max" model in which agents only care about the agents who exert the most effort (e.g., the highest-ability students in the case of GPA) in their network, while when β is very negative, the "min" model prevails in which agents only care about the agents who exert the least effort (e.g., the lowest-ability students) in their network. The main advantage of providing a model in which the individual action/outcome is a function of a CES of peer actions is that it contains the LIM model, as well as the min and the max model, as special cases but provides flexibility in modeling, since the relevant peer group is estimated from the data.

We show that, contrary to the linear case, the best-response function for the general model with flexible β is not necessarily contracting. However, by relying on the literature on supermodular games (Milgrom and Roberts, 1990) and by studying the structure of the smallest and largest equilibria, we show that there always exists a unique Nash equilibrium. Our proof of existence and uniqueness of a Nash equilibrium applies to any social norm function that is homogeneous of degree one, increasing in individual action, and satisfies either global convexity or global concavity; it includes the CES function as a special case.

Then, using data on teenagers in the United States from the National Longitudinal Survey of Adolescent Health (AddHealth), we structurally estimate this general model. We estimate the value of β to determine the relevant peer reference group and estimate which model matches best the data. We find that for GPA, social clubs, self-esteem, and exercise, the spillover effect strongly dominates, while for risky behavior, study effort, fighting, smoking, and drinking, conformism plays a stronger role. We also find that for many activities, individuals do not behave according to the LIM model. Indeed, for GPA, self-esteem, exercise, and study effort, individuals have peer preferences skewed towards the right-hand side of the distribution (that is, agents who exert relatively high effort), while for trouble behavior at school, fighting, and drinking, the peers that matter are those who exert little effort. This confirms the fact that imposing the mean action as an individual social norm is misleading and may lead to incorrect policy implications.

In order to quantitatively evaluate the policy implications in the context of our data, for each activity, we simulate a counterfactual tax/subsidy policy that restores the first best. In particular, we contrast the differences between a planner that uses the LIM model and one that uses the general results obtained in our structural estimations. We show that the differences are large. In general, with the LIM model, the planner tends to uniformly tax or subsidize all agents in the network. In contrast, with the general model, it targets some key agents; the choice of which agents are targeted depends on whether the spillover or the conformist model dominates, and on the value of β. Consider, for example, GPA, which is a spillover model for which β is much greater than 1; this means that peer preferences are skewed toward students with the highest GPA. In contrast to the LIM model, we find in our policy simulations that in the general model, there is a large mass at zero because these individuals do not provide any positive spillover to their neighbors (they are not the highest-ability friends), and there is therefore little social value in subsidizing them. We also show that some individuals obtain very large positive subsidies; this is when the social norm is made up of very low-performing students, and thus it becomes valuable to give large subsidies to the highest-ability peers because they will generate large spillover effects.

We consider the baseline situation in which the network is exogenous and not affected by policy shocks. This allows us to focus on the impact of the behavioral foundations (conformism or spillovers) and non-linearity for public policies when the network is fixed. We find that the policy recommendations resulting from conformism or spillover effects differ wildly and that non-linearities have a huge influence on who should be targeted by the policies. We provide an estimator that allows us to identify the behavioral foundation and the non-linearity from the data.

Our main contribution is to provide a general structural framework to study peer effects in a context in which peers do not necessarily react to the average of their peers' behavior, and that enables identification of the behavioral foundation of the estimated peer effects. Even though the vast majority of papers have used the LIM model to estimate peer effects, some have considered the maximum peers, namely the leaders, shining lights, or high achievers (Carrell et al., 2010, Tao and Lee, 2014, Diaz et al., 2021, Islam et al., 2021), some have included the minimum peers, namely the bad apples or low-ability individuals (Bietenbeck, 2020, Hahn et al., 2020), and some have incorporated both (Hoxby and

Weingarth, 2005, Tatsi, 2015).[5] However, none of these papers have developed a general theoretical framework with different mechanisms (spillover or conformism) and different peer-group references. In contrast, our proposed framework allows us to identify which mechanism and which peer group matter most. Our main conclusion confirms the fact that ex ante imposing the mean (max or min) action as an individual social norm is misleading and leads to incorrect policy recommendations.

2. THEORY

2.1. *Linear-in-means models*

Before presenting our general model, we describe our setup using the well known linear-in-means model. Our main specification is presented in Section 2.2. Consider $n \geq 2$ individuals who are embedded in a network $\mathbf{g}$. The adjacency matrix $\mathbf{G} = [g_{ij}]$ is an $(n \times n)$-matrix with $\{0, 1\}$ entries that keeps track of the direct connections in the network. By definition, agents i and j are directly connected if and only if $g_{ij} = 1$; otherwise, $g_{ij} = 0$. We assume that the network is directed (i.e., g_{ij} and g_{ji} are potentially different)[6] and has no self-loops (i.e. $g_{ii} = 0$). $\mathcal{N}_i = \{j \mid g_{ij} \in \mathbf{g}\}$ denotes the set of i's neighbors. The cardinal of $\mathcal{N}_i$ is d_i, the degree or the number of direct neighbors of individual i, so that $d_i := \sum_{j=1}^{n} g_{ij} = |\mathcal{N}_i|$. Finally, $\widetilde{\mathbf{G}} = [\widetilde{g}_{ij}]$ denotes the $(n \times n)$ row-normalized adjacency matrix defined by $\widetilde{g}_{ij} := g_{ij}/d_i$ if $d_i > 0$ and $\widetilde{g}_{ij} := 0$ otherwise.

Assume that each individual has at least one neighbor, namely $d_i > 0$ for all i. Consider the following class of best-response functions:

$$y_i = \alpha_i + \lambda \overline{y}_{-i}, \tag{1}$$

where y_i is the effort or outcome in some activity (such as GPA in education); $\alpha_i = \mathbf{x}_i \boldsymbol{\gamma} + \epsilon_i$ is a vector of individual characteristics that includes both the observable $(\mathbf{x}_i)$[7] and unobservable (ϵ_i) characteristics of individual i; λ is the peer-effect propagation rate (common

[5] See also Brock and Durlauf (2001b) who look at a variety of models of peer behavior and suggest some nonlinearities in some applications, and Blume et al. (2011) who classify these different peer effects.

[6] We can easily generalize our results to undirected and weighted networks.

[7] $\mathbf{x}_i$ is a $(1 \times k)$ vector of k observable characteristics, and $\boldsymbol{\gamma}$ is a $(k \times 1)$ vector, so that $\mathbf{x}_i \boldsymbol{\gamma} = \sum_{l=1}^{k} x_l \gamma_l$.

for all individuals in the group); and

$$\overline{y}_{-i} = \sum_{j=1}^{n} \widetilde{g}_{ij} y_j. \tag{2}$$

is the average eort of is neighbors (excluding i). In Equation (1), individuals choose their eort y_i as a function of their own characteristics α_i and also as a function of the eort of the other individuals $\overline{y}_{-i}$ in the population. The model in (1) with the norm $\overline{y}_{-i}$ defined in (2) is referred to as the linear-in-means (LIM) model and can thus be written as

$$y_i = \mathbf{x}_i \boldsymbol{\gamma} + \lambda \sum_{j=1}^{n} \widetilde{g}_{ij} y_j + \epsilon_i. \tag{3}$$

2.2. *A model with general social norms*

So far, following the LIM model, we assumed that peers operate through a linear and an average effect, that is, the social norm $\overline{y}_{-i}$ is the average effort of i's peers. This is a strong assumption, especially for empirical applications. The empirical literature has been partially addressing this issue by not only looking at the effect of the average peer, but instead the minimum or maximum. In this section, we relax the linearity assumption and provide a more general and flexible structure of peer preferences.

2.2.1. *A general social norm*

For each individual i, we generalize the social norm $\overline{y}_{-i}$ given in (2) by considering the following CES social norm:[8]

$$\widetilde{y}_{-i}(\beta) = \left(\sum_{j=1}^{n} \widetilde{g}_{ij} y_j^{\beta} \right)^{\frac{1}{\beta}}. \tag{4}$$

We can easily see that the social norm in the LIM model defined in (2) is a special case of (4) when $\beta = 1$, that is, $\overline{y}_{-i} \equiv \widetilde{y}_{-i}(1) = \sum_{j=1}^{n} \widetilde{g}_{ij} y_j$. Our social norm is general, since (4) allows for any β, that is, $\beta \in [-\infty, +\infty]$. We argue that this parameter, β, captures peer

[8]Observe that, in the theory, we focus (and provide conditions) on equilibria for which $y_i > 0$, for all i. Thus, Equation (4) is well-defined for all $y_i > 0$.

preference. Consider, for example, the study effort y_i of each student i. Then, if we set $\beta = +\infty$, (4) becomes

$$\lim_{\beta \to +\infty} \widetilde{y}_{-i}(\beta) = \max_{j \in \mathcal{N}_i} \{y_j\},$$

that is, the social norm is defined with respect to the most studious agents in the network. Under $\beta = -\infty$, (4) becomes

$$\lim_{\beta \to -\infty} \widetilde{y}_{-i}(\beta) = \min_{j \in \mathcal{N}_i} \{y_j\},$$

that is, the social norm is defined with respect to the least studious agents. Other possible values of $\beta \in \mathbb{R}$ capture a rich spectrum of intermediate cases.[9] We have:

$$\frac{\partial \widetilde{y}_{-i}(\beta)}{\partial y_j} = \widetilde{g}_{ij} \left(\sum_{j=1}^{n} \widetilde{g}_{ij} y_j^{\beta} \right)^{(\frac{1}{\beta} - 1)} y_j^{\beta - 1},$$

while, in the LIM with the social norm given by Equation (2), we have: $\frac{\partial \overline{y}_{-i}}{\partial y_j} = \widetilde{g}_{ij}$.

The main advantage of providing a model in which the individual action is a function of a CES function of peer actions is that it contains the LIM, the min and max models and any combination of them as a special case.

2.2.2. *A general model*

Define $\mathbf{y}_{-i} := (y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n)^T$ the vector of effort without the effort of agent i. The utility function for each non-isolated individual i is given by:[10]

$$U_i(y_i, \mathbf{y}_{-i}, \mathbf{g}) = \underbrace{\alpha_i y_i + \theta_1 y_i \widetilde{y}_{-i}(\beta)}_{\text{benefit}} - \underbrace{\frac{1}{2} \left[y_i^2 + \theta_2 \left(y_i - \widetilde{y}_{-i}(\beta) \right)^2 \right]}_{\text{cost}}, \tag{6}$$

[9] When $\beta < 0$, expression (4) reads as follows:

$$\widetilde{y}_{-i}(\beta) = \begin{cases} \left(\sum_{j=1}^{n} \widetilde{g}_{ij} y_j^{\beta} \right)^{\frac{1}{\beta}}, & y_j > 0 \text{ for all } j \in \mathcal{N}_i; \\ 0, & y_j = 0 \text{ for some } j \in \mathcal{N}_i. \end{cases} \tag{5}$$

[10] For an isolated individual i, her utility is equal to $U_i = \alpha_i y_i - \frac{1}{2} y_i^2$, so that her action is $y_i = \alpha_i$.

where the social norm $\widetilde{y}_{-i}(\beta)$ is defined in (4) and $\alpha_i > 0$. The term $\theta_1 y_i \widetilde{y}_{-i}(\beta)$ corresponds to the spillover model (Brock and Durlauf, 2001a, Glaeser and Scheinkman, 2002, Boucher and Fortin, 2016, Reif, 2019) while the term $-\frac{1}{2}\theta_2 \left(y_i - \widetilde{y}_{-i}(\beta)\right)^2$ corresponds to the conformism model (Akerlof, 1997, Bernheim, 1994, Patacchini and Zenou, 2012, Boucher, 2016, Ushchev and Zenou, 2020). Thus, this utility function has the conformist model (when $\theta_1 = 0$) and the spillover model (when $\theta_2 = 0$) as special cases.[11]

Denote $\lambda_1 := \theta_1/(1+\theta_2)$ and $\lambda_2 := \theta_2/(1+\theta_2)$. Then, the first-order condition can be written as

$$y_i = (1-\lambda_2)\alpha_i + (\lambda_1+\lambda_2)\widetilde{y}_{-i}(\beta), \tag{7}$$

or equivalently

$$y_i = (1-\lambda_2)\alpha_i + (\lambda_1+\lambda_2)\left(\sum_{j=1}^{n} \widetilde{g}_{ij} y_j^{\beta}\right)^{\frac{1}{\beta}}. \tag{8}$$

The main difference with the LIM model (where $\beta = 1$) is that the first-order conditions (8) are not linear anymore. Thus, when estimating (8), in particular, β, we can determine whether the correct model is the LIM (i.e., $\beta = 1$) and, if not, which peers matter. We have the following result.

PROPOSITION 1: *Assume that the utility function of each individual $i = 1, \ldots, n$ is given by (6), with $0 < \lambda_1 + \lambda_2 < 1$ and $0 < \lambda_1 < 1$, and her social norm $\widetilde{y}_{-i}(\beta)$ has the CES functional form (4). Then, there exists a unique Nash equilibrium.*

The proof of Proposition 1 is given in Appendix A. It is is not obvious because, contrary to the LIM model, the best-reply mapping is neither linear nor a contraction. First, for the existence of equilibrium, we use the fact that the game is supermodular and solve for a fixed point theorem. To prove uniqueness, we use the fact that there always exist a maximum and a minimum equilibrium and show that they are equal. To achieve this, we need to differentiate between concave and convex norms and demonstrate this equality; thus, the equilibrium is unique. In fact, we show that the existence and uniqueness of the equilibrium

[11] As highlighted by Boucher and Fortin (2016), when $\overline{y}_{-i} \equiv \widetilde{y}_{-i}(1)$, the LIM model corresponds to the best-response function of two, observationally equivalent, types of social preferences: spillover or conformism.

of this game is true for more general norms than the CES one, since we only need to assume that the social norm is monotone increasing, continuous, linear homogeneous, and normalized (see Assumption 1 and the proof of Proposition 1 in Appendix A).

3. STRUCTURAL ESTIMATION

3.1. *Empirical Strategy*

The model has two main components: the non-linearity of the social norm, and the nesting of conformity and spillover effects. We are interested in structurally estimating (i) the intensity of the spillover effect, λ_1; (ii) the taste for conformity, λ_2; and (iii) the peer preference, β. We can formulate the equilibrium effort of individual i by

$$y_{is} = (1-\lambda_2)\mathbf{x}_{is}\boldsymbol{\gamma} + (\lambda_1+\lambda_2)\widetilde{y}_{-is}(\beta) + \zeta_s + \varepsilon_{is}. \tag{9}$$

Equation (9) is the equivalent of the first-order condition (7), where, as above, $\alpha_{is} := \mathbf{x}_{is}\boldsymbol{\gamma} + \xi_s + \epsilon_{is}$ captures the observable and unobservable characteristics of i as well as the school fixed effects ξ_s, where $\zeta_s := (1-\lambda_2)\xi_s$, and $\varepsilon_i := (1-\lambda_2)\epsilon_i$. Indeed, as we discuss in Section 3.3 below, students in the data are assumed to interact within their school. We therefore added the subscript s to denote each school s in our data. Thus, compared to (7), we control for school fixed effects, ξ_s, which will absorb any factor that is common to all students within a given school, including the effect of the school itself. We assume that ε_{is}, the error term, is such that $\mathbb{E}(\varepsilon_{is}|\mathbf{X},\mathbf{G}) = 0$ for all i, implying an exogenous network.

For comparison purposes, we will also provide results for the reduced-form LIM model (Equation (3)),

$$y_{is} = \boldsymbol{x}_{is}\boldsymbol{\gamma} + \lambda\overline{y}_{-is} + \epsilon_{is}, \tag{10}$$

where y_{is} is the effort or outcome in some activity (e.g., GPA), and $\overline{y}_{-is}$ is the average effort of i's neighbors (excluding i), instrumented by their friends' characteristics, x_{-is} (Bramoullé et al., 2009).

3.2. *Identification*

We show in Appendix B how to formally estimate $\boldsymbol{\theta} = [\boldsymbol{\gamma}', \lambda_1, \lambda_2, \beta]'$ by deriving the appropriate generalized method of moment (GMM) estimator. Let us provide some intuition for the estimation procedure. Equation (9) does not allow us to separately identify λ_2 from $\boldsymbol{\gamma}$ or λ_1. However, we can consider two types of individuals in the data: (a) isolated and (b) non-isolated individuals. Isolated individuals are individuals without friends. This separation allows us to break the estimation problem into two parts and consequently identify λ_2 and $\boldsymbol{\gamma}$ separately. This strategy rests on the assumption that isolated and non-isolated individuals are similar in that the impact of their individual characteristics (e.g. age, gender) on the private benefit of effort is similar. This is a classical assumption, even in models with endogenous network formation (e.g. Lee et al. (2021)).

First, note that isolated individuals have a simplified version of the general first-order condition (9), given by

$$y_{is} = \mathbf{x}_{is}\boldsymbol{\gamma} + \xi_s^{iso} + \varepsilon_{is}, \tag{11}$$

where ξ_s^{iso}, the school fixed effect specific to isolated individuals, has been added. This equation is independent of any social norm and, therefore, of β, λ_1 and λ_2. Thus, in our specification, the identification of $\boldsymbol{\gamma}$ can be obtained from isolated individuals, under the independence assumption of the error term, $\mathbb{E}\left(\varepsilon_{is}\mathbf{x}_{is}\right) = \mathbf{0}$. Note, identification does not allow us to estimate separate $\boldsymbol{\gamma}$s for isolated and non-isolated students. However, we allow the school fixed effect to differ between the two types of students. We acknowledge that assuming $\boldsymbol{\gamma}$ to be identical for isolated and non-isolated individuals can be a strong assumption. In Section 3.4, we consider a robustness exercise by estimating a model that does not identify separately the conformist and spillover model, so that the peer preference parameter β can be estimated without isolated individuals.

Second, to identify $\tilde{\boldsymbol{\theta}} = [\lambda_1, \lambda_2, \beta]$, we require three further moment conditions. Let us define three instruments, $\mathbf{z}_{is}$ for non-isolated individuals that satisfy the moment conditions, $\mathbb{E}\left(\varepsilon_{is}\mathbf{z}_{is}\right) = \mathbf{0}$. First, we can identify $(1-\lambda_2)\boldsymbol{\gamma}$, and thus λ_2, given the result of $\boldsymbol{\gamma}$ from the solution of isolated individuals. Consequently, our first instrument is the set of covariates, $\mathbf{x}_{is}$. Secondly, if $\hat{\mathbf{y}}_s$ is the OLS predictor of $\mathbf{y}_s$, on the covariates, $\mathbf{x}_{is}$,[12] then, given λ_2, the identification of λ_1 comes from the moment $\widetilde{y}_{-is}(\hat{\mathbf{y}}_s, \beta)$ for non-isolated indi-

[12]Because y_{is} is potentially endogenous, we estimate the reduced form of y_{is} on $\mathbf{x}_{is}$ to obtain the predicted $\hat{\mathbf{y}}_s$ for each i. Thus, $\hat{\mathbf{y}}_s$ is an exogenous predictor and independent of β. While we use a simple OLS predictor for simplicity, any predictor of y_i (potentially non-parametric) that is based on the exogeneous variables $\mathbf{X}$ would be admissible.

viduals—our second instrument. Finally, the identification of β comes from the derivative of the social norm with respect to β, $\frac{\partial \widetilde{y}_{-is}(\hat{\mathbf{y}}_s,\beta)}{\partial \beta} = \widetilde{y}'_{-is}(\hat{\mathbf{y}}_s,\beta)$—our last instrument. The intuition behind this instrument is that the slope of the social norm with respect to β should inform the directional movement of search for the parameter that minimizes the objective function during the numerical simulation. Also note that the choice of $\widetilde{y}'_{-is}(\hat{\mathbf{y}}_s,\beta)$ as a moment condition is justified by the fact that $\widetilde{y}'_{-is}(\hat{\mathbf{y}}_s,\beta)$ is equal to the first-order condition for the nonlinear least-squares estimator of a model in which $\tilde{y}_{is}(\beta)$ is substituted by $\tilde{y}_{is}(\hat{\mathbf{y}}_s,\beta)$.[13] Thus, the set of instruments for all non-isolated individuals can be summarized by $\mathbf{z}_{is} = [\mathbf{x}_{is}, \widetilde{y}_{-is}(\hat{\mathbf{y}}_s,\beta), \widetilde{y}'_{-is}(\hat{\mathbf{y}}_s,\beta)]$, with the assumption that $\mathbb{E}(\varepsilon_{is}\mathbf{z}_{is}) = \mathbf{0}$.

Note that our additional moment conditions are evaluated at $\hat{\mathbf{y}}_s$, the OLS predictor of $\mathbf{y}_s$. While it is standard to use the entire matrix of observable characteristics as instruments, namely $\widetilde{\mathbf{G}}\mathbf{X}$ (Bramoullé et al., 2009) when $\beta = 1$, this approach is not suitable for the general model when β could be substantially different to 1. Indeed, suppose that $\beta = +\infty$, so that $\widetilde{y}_{-is} = \max_{j:g_{ij,s}=1} y_{j,s}$. While $\bar{z}_{i,s} = \max_{j:g_{ij,s}=1} \hat{y}_{j,s}$ is likely a good predictor of $\widetilde{y}_{-i,s}$, it may well be the case that none of the maximum of characteristics of i's friends, taken individually (i.e., $\max_{j:g_{ij,s}=1} x^l_{j,s}$, $l = 1, ..., k$), would be a good predictor of $\widetilde{y}_{-i,s}$. Evaluating the instruments at $\hat{\mathbf{y}}_s$ is therefore a simple and effective way to ensure strong instruments, irrespective of the value of β.

We therefore have four sets of moment conditions, one from isolated individuals (i.e. $\mathbb{E}(\varepsilon_{is}\mathbf{x}_{is}) = \mathbf{0}$) and three from non-isolated individuals (i.e. $\mathbb{E}(\varepsilon_{is}\mathbf{x}_{is}) = \mathbf{0}$, $\mathbb{E}(\varepsilon_{is}\widetilde{y}_{-is}(\hat{\mathbf{y}}_s,\beta)) = \mathbf{0}$, and $\mathbb{E}(\varepsilon_{is}\widetilde{y}'_{-is}(\hat{\mathbf{y}}_s,\beta)) = \mathbf{0}$). Notice that the moment conditions for $\boldsymbol{\gamma}$ and for $(\lambda_1, \lambda_2, \beta)$, are not based on the same number of observations (isolated and non-isolated individuals). Thus, the estimations for isolated and non-isolated individuals are performed jointly using the sum of the GMM objective functions for both sets of moments, leading to an observation-weighted average of the two sets of moment conditions (Angrist and Krueger, 1992, Arellano and Meghir, 1992). Lastly, since the first-order condition is linear in $\boldsymbol{\gamma}$, we can use the model and the objective function to derive $\boldsymbol{\gamma}$ as a function of the three remaining parameters $[\lambda_1, \lambda_2, \beta]$ for both isolated and non-isolated individuals. This allows us to concentrate the objective function around these remaining three param-

[13] For a textbook treatment of the nonlinear least-squares estimator, see Section 5.8.2 in Cameron and Trivedi (2005). For an in-depth discussion of the optimal moment conditions for non-linear GMM, see also Section 6.3.7 in Cameron and Trivedi (2005).

eters, $\tilde{\boldsymbol{\theta}} = [\lambda_1, \lambda_2, \beta]$, which are numerically estimated. For further estimation details, see Appendix B.

3.3. *Data description*

Our analysis is based on a well-known database of friendship networks from the National Longitudinal Survey of Adolescent Health (AddHealth). The AddHealth survey has been designed to study the impact of the social environment (i.e., friends, family, neighborhood, and school) on adolescents behavior in the United States by collecting data on students in grades 7–12 from a nationally representative sample of more than 130 private and public schools in years 1994–1995 (Harris et al., 2019). AddHealth provides a wealth of information regarding students' activities and outcomes. We extracted a large number of the activities available in the in-school interview sample to test our theory. For the purpose of studying peer effects, AddHealth data also record friendship information, which is based upon actual friends' nominations. Pupils were asked to identify their best friends from a school roster (up to five males and five females). Our estimation sample comprised over 70,000 students, from 134 schools.

We use AddHealth data because it is one of the few datasets that provides both the exact network of all students and has information on multiple activities, so we can illustrate our theory with different values of β and consider their policy implications. Nonetheless, we acknowledge that AddHealth poses some limitations in terms of network endogeneity. However, our aim is methodological, as we want to illustrate the importance of having a microfoundation in estimating peer effect models. Thus, we assume that the network $\mathbf{G}$ is conditional exogenous.

In Section 3.4, we report the estimation results for 10 activities: (1) grade point average (GPA), (2) social clubs, (3) self-esteem, (4) risky behavior, (5) exercise, (6) study effort, (7) fighting, (8) smoking, (9) drinking, and (10) trouble behavior. We also use a series of students individual characteristics, such as age, gender, racial group, and mother's education and occupation.

Activities are reported consistently across schools, with summary statistics similar across standard demographic characteristics (e.g., age, gender, race). All activities are based on increasing activity levels; for example, a higher value for self-esteem reflects an increased level of self-esteem. Importantly, 14–15 percent of students did not report any friends, and we labeled them isolated.[14,15]

3.4. *Empirical results*

For each activity in the data, we estimate the value of β to determine the relevant peer reference group, and we test which model (conformist or spillover) is the most appropriate one. The aim here is not to study the quantitative impact of specific peer effects, but rather to illustrate the importance of using a generalised theory of peer effects when trying to estimate their economic impact and, consequently, design adequate policies that improve agents outcomes. Thus, to provide a broad view of our theory, we pick from a wide range of activities documented in AddHealth. That is, we provide estimation results for $\tilde{\boldsymbol{\theta}} = [\lambda_1, \lambda_2, \beta]$ for the 10 activities described in Section 3.3.

Table I shows the GMM results of (9) for the general model with estimated peer preferences, that is, the βs. For comparison purposes, the table also shows the general LIM model, whereby we impose the social norm of the average peer (i.e., $\beta = 1$), and reduced-form estimates from Equation (10). Lastly, the table reports two ways of testing the validity of the general model relative to the LIM model; (1) we report the significance of the hypothesis test of $H_0 : \beta \neq 1$, and (2) we report the objective value of each GMM procedure, as well as the likelihood-ratio statistic, $2N(Q^{LIM}(\lambda_1, \lambda_2) - Q(\lambda_1, \lambda_2))$, comparing it to the general model. This latter measure is a simple way of establishing the relevant peer preference in the estimation, as it compares the goodness of fit for a model with $\beta \neq 1$ and $\beta = 1$.

Just two of the ten activities have marginal cases regarding the most appropriate general model choice (spillover, conformism, or both). As the coefficient on λ_2 for self-esteem and exercise are statistically zero, the general model exclusively estimates spillover effects.

[14] We follow Boucher and Houndetoungan (2023) in dropping individuals who are potentially falsely classified as having no friends. Friendship nominations in the data are not always reciprocal, even when restricting to the set of individuals for whom we can identify all of their nominated friends.

[15] The activity or outcome values often included an outcome of zero (e.g., non-smokers never smoke). Equation (4) is not defined for values of zero if peer preference is skewed to the least active agent, $\beta < 0$. To avoid this computational error, we have added in the estimation a value of 1 for each activity or outcome. For comparability, we did this for all activities. Results without adding 1 in instances where $\beta > 1$ are comparable and available upon request.

As the objective value is, by construction, always lowest for the general model with both spillovers and conformism, we report it as our baseline.[16]

Table I also shows that activities have varying degrees of peer preference. Overall we find a wide range of values for $\beta = [-386, 372]$. Thus, peer preference does not necessarily conform to the average peer, as per the assumption in the commonly used LIM model. For example, GPA, self esteem, exercise, and study effort, have peer preferences skewed towards the right of the distribution (that is, agents who exert relatively high effort). Trouble behavior at school, fighting and drinking are skewed towards the left of the distribution (that is, agents who exert relatively low effort), while social clubs, risky behavior, and smoking are close to the LIM assumption of $\beta = 1$. However, only for risky behavior we cannot reject that $\beta = 1$.

With varying degrees of peer preference, the resulting estimates on total peer effects, as well as the magnitudes of spillover versus conformism effects, change in non-negligible ways across activities. We find that the difference in total peer effects $(\lambda_1 + \lambda_2)$ between the general model (GM) and the general LIM model to be large for (i) trouble behavior at school (36 percent), (ii) GPA and drinking (30 percent), (iii) study effort and fighting (20 percent), (iv) exercise (30 percent), and (v) self-esteem, risky behavior, exercise and smoking (10 percent).[17] Only for social clubs and self-esteem the total peer effects remain unchanged.

Thus, estimating the general model compared to imposing the reduced form model has implications on behavior through changing (i) the shape of the distribution of individuals' social norms, and/or (ii) the distribution of individuals' total peer effect exposure. We illustrate this further with three distinct examples, risky behavior, study effort and GPA outcomes.

As for risky behavior, the estimated peer preference is close to one while the density distribution of social norms is similar in the general model and LIM model. However, given the sizable difference between the estimates on λ in the reduced form and $(\lambda_1 + \lambda_2)$ in the

[16]For the two instances of self-esteem and exercise, the corresponding estimates of λ_1 and β for the general model with only spillovers are very similar to the general model reported here.

[17]Formally, the difference in peer preference reported is $\left|1 - \frac{\lambda_1^{GM} + \lambda_2^{GM}}{\lambda_1^{LIM} + \lambda_2^{LIM}}\right|$.

general model, the distribution of peer effects interacted with individuals' social norms in the reduced form has considerably more mass to the left.

The distribution of the social norm of study effort, for which the coefficient is above one (i.e., $\beta = 3.9$), shows a slight skewedness towards the right for the general model. Moreover, even stronger than for risky behavior, the *total peer effect* $(\lambda \overline{y}_{-i})$ or $(\lambda \tilde{y}_{-i})$, has large variance across the three models, with the general model strongly skewed toward higher peer effect outcomes.

Lastly, GPA outcomes serve as an example where peer preference is highly skewed towards the highest-ability students (i.e., $\beta = 372$). Thus, the distribution of the social norm in the general model is skewed toward the right, with several distinct peaks, but still far from the highest-ability students. Peaks appear because individuals might naturally have peers who do not achieve the highest (4.0) GPA, but something just below it, such as from 3.0 to 4.0. In comparison, the LIM model will have a perfectly hump-shaped distribution following the average peer social norm. The intensity of peer preference across the models greatly exacerbates differences between the general model and the LIM model. As peer preference increases towards the highest-ability agents in the general model, the conformism effect mostly disappears. That is, for GPA, the friends with the best academic success will influence one's outcomes through positive spillovers, while friends with average academic outcomes will have no effect. Moreover, individuals do not try or succeed in conforming to their peers' academic outcomes.

These three different examples highlight the importance of moving towards a general theory of peer effects. There is a large range of peer preference estimates. These differences translate into vastly different social norms that cannot be consistently approximated by either the average, the max or the min. That is, moving from the reduced form to a general LIM model (with both spillover and conformism behavior) but without relaxing the functional form of the social norm is not enough to provide a general theory of peer effects.

TABLE I

STRUCTURAL ESTIMATION: PEER PREFERENCES (β)

	Peer preferences with general β				$\boldsymbol{\beta = 1}$				Reduced form	Non-isolated only	
Activity	$\boldsymbol{\lambda_1}$	$\boldsymbol{\lambda_2}$	$\boldsymbol{\beta}$	**Obj. Value**	$\boldsymbol{\lambda_1}$	$\boldsymbol{\lambda_2}$	**Obj. Value**	**L-R test**	$\boldsymbol{\lambda}$	$\boldsymbol{\lambda}$	$\boldsymbol{\beta}$
GPA	0.32 (0.06)	0.06 (0.05)	371.8 $^{(a)}$ (116.0)	0.0014	0.39 (0.06)	0.19 (0.04)	0.0026	180.5	0.59 (0.02)	0.36 (0.01)	$+\infty$ -
Clubs	0.33 (0.11)	0.33 (0.07)	1.4 $^{(b)}$ (0.2)	0.0022	0.35 (0.10)	0.34 (0.07)	0.0023	18.9	0.64 (0.03)	0.62 (0.04)	1.55 $^{(a)}$ (0.18)
Self esteem	0.28 (0.11)	0.01 (0.07)	22.3 $^{(a)}$ (8.1)	0.0020	0.25 (0.11)	0.03 (0.07)	0.0027	110.1	0.29 (0.04)	0.31 (0.04)	21.78 $^{(a)}$ (7.83)
Risky	-0.16 (0.08)	0.50 (0.03)	0.8 (0.4)	0.0016	-0.19 (0.07)	0.49 (0.03)	0.0017	9.0	0.24 (0.03)	0.22 (0.05)	1.40 (0.56)
Exercise	0.21 (0.06)	0.01 (0.04)	8.1 $^{(b)}$ (3.5)	0.0008	0.18 (0.06)	0.02 (0.04)	0.0017	122.9	0.18 (0.02)	0.19 (0.02)	49.63 (60.85)
Study effort	0.02 (0.09)	0.34 (0.04)	3.9 $^{(b)}$ (1.4)	0.0008	-0.02 (0.08)	0.32 (0.04)	0.0010	48.5	0.22 (0.04)	0.24 (0.05)	5.34 $^{(c)}$ (2.28)
Fight	-0.08 (0.05)	0.12 (0.04)	73.5 $^{(a)}$ (14.8)	0.0011	0.06 (0.07)	0.19 (0.04)	0.0013	18.6	0.25 (0.03)	0.03 (0.01)	$+\infty$ -
Smoke	0.21 (0.09)	0.65 (0.04)	0.7 $^{(b)}$ (0.1)	0.0006	0.13 (0.07)	0.62 (0.04)	0.0009	41.8	0.73 (0.03)	0.74 (0.05)	0.95 (0.16)
Drink	-0.05 (0.10)	0.64 (0.03)	0.4 $^{(a)}$ (0.2)	0.0018	-0.18 (0.07)	0.63 (0.03)	0.0019	13.4	0.16 (0.04)	0.09 (0.05)	4.32 (3.21)
Trouble	0.19 (0.11)	0.11 (0.08)	-386.0 $^{(a)}$ (161.6)	0.0004	0.22 (0.12)	0.25 (0.07)	0.0007	48.5	0.52 (0.04)	0.29 (0.03)	$-\infty$ -

Notes: We report the general model (Equation (9)), the general LIM model ($\beta = 1$), the reduced form (Equation (10)), and the general restricted model (Equation (12)). All results control for school-fixed effects. Standard errors are reported in parentheses. For the general model, we report the significance of the hypothesis test, $H_0 : \beta \neq 1$ with (a) $p < 0.01$, (b) $p < 0.05$, (c) $p < 0.1$. For the case of $\pm\infty$ the hypothesis test is omitted because, up to machine precision, we cannot distinguish between $\pm\infty$ and and also between very large/small values of β. In the case of the LIM model, likelihood ratio tests, $2N(Q^{LIM}(\lambda_1, \lambda_2) - Q(\lambda_1, \lambda_2))$, are also reported comparing the general LIM model with $\beta = 1$ with the *peer preference* outcome, $\beta \neq 1$.

Finally, to illustrate the validity of our assumption of using isolated individuals to identify a common $\boldsymbol{\gamma}$, we estimate a model with general social norms, but with only one type of peer effect (spillover or conformism) without isolated individuals. Without distinguishing between spillover and conformism, Equation (9) simplifies to

$$y_{is} = \mathbf{x}_{is}\boldsymbol{\gamma}^{niso} + \lambda\widetilde{y}_{-is}(\beta) + \xi_s + \epsilon_{is}, \tag{12}$$

where the superscript $niso$ refers to non-isolated individuals. Identification of $\tilde{\boldsymbol{\theta}} = [\lambda, \beta]$ comes exclusively from non-isolated individuals. The last two columns of Table I show the results of estimating Equation (12) by GMM without isolated individuals compared to the benchmark results. The results confirm that, for all activities, the average peer model is rejected, that is β is not equal to one. Moreover, for the majority of activities, the peer preference, β, is similar to the general model, although the precision of estimates is generally weaker. This confirms the fact that relying on both isolated and non-individuals for the estimation of β is not crucial for our results.

4. POLICY IMPLICATIONS

4.1. *Policy implications: Theory*

The decisions individuals make in the presence of peer effects generate externalities and thus inefficiencies. This creates room for a benevolent social planner to intervene. The planner chooses the actions $y_1, y_2, ..., y_n$ of each of the n agents that maximizes the welfare $\mathcal{W}(\mathbf{y}, \mathbf{g}) = \sum_i U_i(y_i, \mathbf{y}_{-i}, \mathbf{g})$, where each agent i's utility $U_i(y_i, \mathbf{y}_{-i}, \mathbf{g})$ is given by (6). In Appendix C, we determine the first-best outcome (Equation (24)) and show that it is unique (Proposition 6). Because of different externalities, the market (Nash) equilibrium is inefficient. For the spillover model ($\lambda_2 = 0$), agents exert too little effort at the Nash equilibrium as compared to the social optimum outcome because each agent ignores the positive impact of her effort on the effort choices of others. For the conformist model ($\lambda_1 = 0$), at the Nash equilibrium, when deciding her individual effort, each agent does not take into account the effect of her effort on the social norm of her peers, which creates an externality that can be positive or negative.

It is then straightforward to determine which subsidies restore the first best; the planner needs to give to each agent i the following subsidy:[18]

$$S_i^G = \frac{y_i^o - y_i^N}{1-\lambda_2} = \frac{1}{1-\lambda_2}\left[\lambda_1 \sum_j y_j^o \frac{\partial \widetilde{y}_{-i}^o(\beta)}{\partial y_i^o} + \lambda_2 \sum_j (y_j^o - \widetilde{y}_{-i}^o(\beta)) \frac{\partial \widetilde{y}_{-i}^o(\beta)}{\partial y_i^o}\right]. \tag{13}$$

4.2. *Policy implications: Simulations*

We now illustrate the importance of microfounding a model of peer effects through our estimated activities. We proceed in two steps. First, following the general estimation procedure, we simulate the Nash equilibrium and the social optimum (first best) for the LIM spillover and conformist model. Second, we simulate the Nash and social optimum for the general LIM model and the general model.

In Figure 1, we display the (kernel) density of the subsidies required for all non-isolated individuals to reach the first best for each model (see Equation (13) and Appendix C for details). We use the same three activities as above for illustration, that is, risky behavior, study effort and GPA outcomes; the results for all the other activities are available upon request. In the left panels of Figure 1, we show the subsidies in the LIM spillover and conformist models, while in the right panels, we show the subsidies in the general LIM model (i.e., $\beta = 1$) and the general model (i.e., β can take any value).

Technically, we simulate the best-response outcomes y_i for all students within each school based on their individual characteristics, estimated school fixed effects, and a truncated normally distributed error term using the parameter estimates from Table I.[19] We then proceed in steps to find the social optimum outcome and subsidy. First, we guess an initial value of the first best subsidy, $\hat{S}_0$, based on the Nash outcome using (13). Second, we compute the *subsidised* Nash outcome y_i with subsidy $\hat{S}_0$. Third, we recompute the first best subsidy $\hat{S}_1$ based on this new *subsidised* Nash outcome. Fourth, we repeat the second and third steps until we have convergence of the first best subsidy. In the spillover model, with positive peer effects, subsidies can potentially become un-

[18]A variable with the superscript o denotes its optimal value while a variable with the superscript N denotes its Nash-equilibrium value.

[19]Errors come from the same normal distribution with mean zero, but truncation is individual specific, based on the natural bounds of each outcome (e.g., the outcomes for GPA lie between 1 and 4.)

bounded if the cost of exerting effort is smaller than the peer effect. To avoid such an issue, we limit the amount of subsidy such that the first best outcomes never exceed the highest outcome observed in the data. Thus, subsidies are bounded for each individual by $S_i \in \left[\min\left\{y_i^{data}\right\}_i - y_i^N, \max\left\{y_i^{data}\right\}_i - y_i^N\right]$.

From the theory (see Section 4.1), we know that the spillover model requires only positive subsidies, while, in the conformist model, the planner can tax or subsidize agents. Consequently, policy prescriptions are vastly different depending on the selected microfoundation. Consider the left panels in Figure 1, where we compare the subsidies/taxes in the LIM spillover and LIM conformist model. As predicted by the theory, to reach the first best (social optimum) in the conformism model, the planner subsidizes some agents and taxes others, while in the spillover model, all agents are subsidized. For both risky behavior and study effort, peer effects are all driven by conformism (dashed line in left panels of Figure 1), while GPA is almost exclusively driven by spillover effects (solid line). However, picking between the conformism (right dashed line) and spillover (solid line), subsidy schedules do not necessarily reflect the correct policy intervention, as we have so far ignored the degree of peer preference.

Let us now focus on the right panels of Figure 1, in which we compare the policy that restores the first best for the general LIM model (imposing $\beta = 1$) and the general model with flexible peer preferences.[20] We observe a wide range of results, which is consistent with the large variation we obtained in the peer-preference estimates of β in Section 3.4. An activity that has peer preferences close to the average peer (i.e., $\beta = 1$) displays similar subsidy schedules between the general model and general LIM model (i.e., risky behavior). As peer preferences skew towards the right or the left of the distribution of agents' efforts, policy prescriptions start to differ more (e.g., study effort, GPA).

More specifically, in panel (b), for risky behavior, since peer preferences are close to 1 ($\beta = 0.8$), the general model policy implications seem to be similar to that of the general LIM model. There is, however, one subtle difference: The general model mostly taxes individuals, while the general LIM model does also subsidize a share of individuals, some of them with relatively large subsidies. Indeed, since for risky behavior, the conformist model is the most prominent one, and since peer preferences are slightly skewed toward the agents

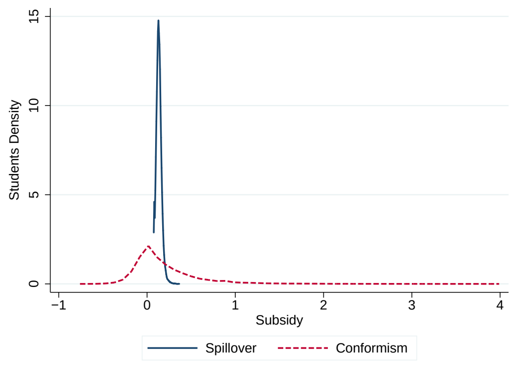
[20]Note the general LIM model represents the full effect (spillover, conformism, or both) from the left panel.

who do not engage in risky activities, there is little use in subsidizing individuals to increase their risky behavior. Therefore, to reach the first best outcomes, it is optimal to tax the most risk-loving agents because it will induce them to decrease their risky behavior. In contrast, in the general LIM model, as the social planner tries to move individuals closer to their *average* peers' risky behavior, it is beneficial to increase some individuals' risky behavior, as each friend's behavior has an equal impact on one's social norm. Thus, the planner finds it optimal to subsidize some individuals to engage in more risky behavior.

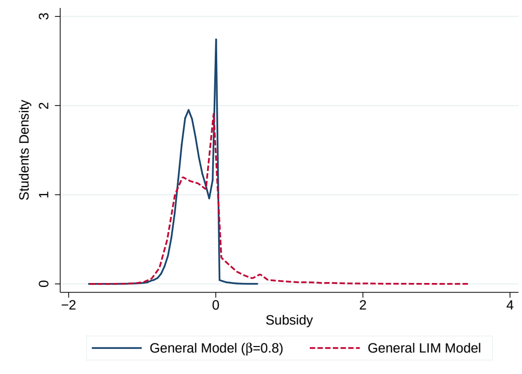
In panel (d), study effort requires, in general, greater levels of subsidy to reach the social optimum (the mean of the distribution of subsidies is larger for the general model). In the LIM model, since study effort is driven by conformism, the social planner needs to tax high-effort students. However, in the general model, since peer preferences are skewed towards the right (i.e., $\beta = 4$), the social planner can implement more targeted policies that require less taxation and overall higher utility outcomes. Thus, in the general model, while the social planner still taxes some of the most studious students, the number of students who need to be taxed to reach the first best are fewer while the number of students receiving a subsidy increases. This results in higher overall study effort and also higher utility.

Lastly, in panel (f), since GPA is driven by spillover effects, for the general LIM model, the social planner gives every individual $\lambda\overline{y}_{-i}$, namely the social norm times the peer effect. In contrast, in the general model, since peer preferences are skewed towards the highest-ability agents (i.e., $\beta = 372$), there is a large mass at zero because these individuals do not have any positive spillover effect (they are not the highest-ability friends), and there is therefore little social value in subsidising them. The general model also has some instance of very large positive subsidies, which are cases where the social norm is made up of very low-performing students. Indeed, from the social planner's perspective, it is valuable to greatly subsidize the highest-ability peers, even if they are low-performing students. This is because, within their friendship group, being the highest performer will generate large spillover effects for their poorly performing peers. For example, take two groups, one where all peers have a GPA between 3.3 and 3.8 and one where all peers have a GPA between 1.2 and 1.5. For both of these groups, the social planner will give most subsidies to the peers

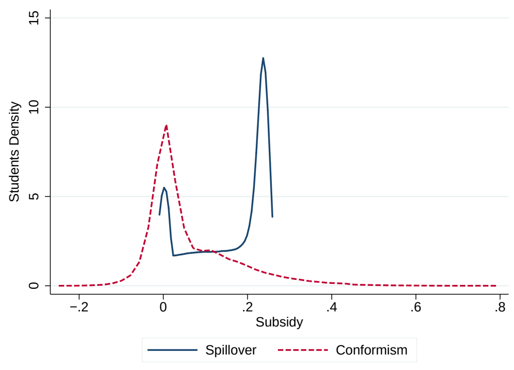
FIGURE 1.—*First-best subsidies (Examples)*



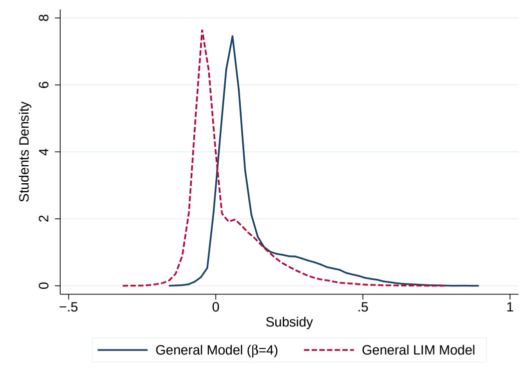
(a) LIM (Risky Behavior)



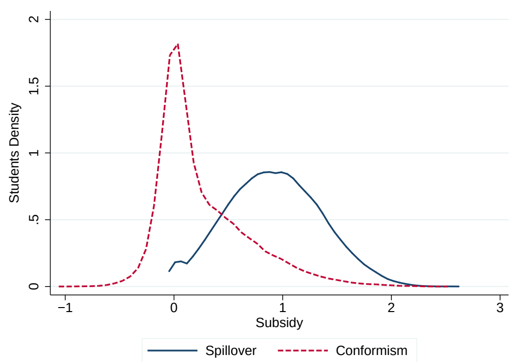
(b) Peer Preference (Risky Behavior)



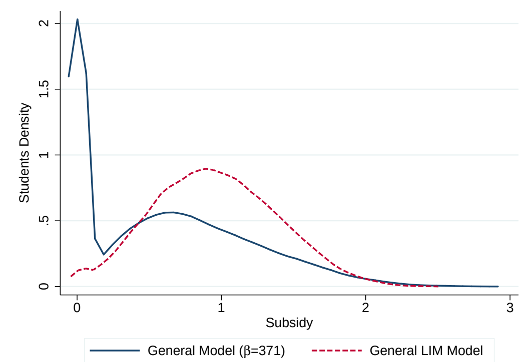
(c) LIM (Study Effort)



(d) Peer Preference (Study Effort)



(e) LIM (GPA)



(f) Peer Preference (GPA)

Notes: Kernel density distribution of non-isolated individuals of the subsidy required to reach the social optimum for (i) the linear-in-means (LIM) spillover and conformist model (estimates available upon request); and (ii) the general model and general LIM model using estimates from the two left-hand panels of Table I.

with the highest GPA because peer preferences are skewed toward the highest-ability agents. A student with a lower GPA (1.5) in the second group will receive a considerably higher subsidy than a student with a higher GPA (3.8) in the first group. This is mechanical since subsidies are capped by the natural limit of 4.0 (the highest achievable GPA). Observe that since the model has mostly spillover effects, no agent is taxed. More generally, policies are very different between the LIM and the general model. In particular, compared to the LIM model, with the general model, the planner gives no subsidy to a large share of agents because they do not have the highest GPA in their peer group but she does give larger subsidies; that is, the curve is flatter but more spread for the general model, compared to the LIM model. In other words, with peer preferences skewed toward high-GPA students, the most effective way of reaching the social optimum in the general model is by subsidizing only a selected number of individuals.

We acknowledge that, as an applied exercise, there are limitations in our policy implications (in particular, potential network endogeneity). We therefore do not recommend to take the exact estimates at face value. However, we urge researchers to acknowledge that individuals do not necessarily respond to the mean of the peer group action and it is likely a function of whether a conformism or spillover game is at play. These have important policy implications that can be tested, since nonlinearity and behavioral foundations can be inferred from data.

5. CONCLUSION

Most papers that estimate peer effects use the LIM model, which assumes that impact on outcomes is linear and that the mean peers' outcomes matter. In this paper, we have argued that to prescribe adequate policies, one needs to determine the correct peer reference group (or social norm) and know which model microfounds the LIM model. We have developed a general model that embeds the spillover and the conformist model and a general social norm for which the LIM model is a special case.

We structurally estimated this model for ten different activities and showed which model mattered the most for each activity. We found that, for most activities, individuals did not behave according to the LIM model; that is, their social norm was not the average outcome of their peers. For example, for GPA, self-esteem, exercise, and study effort, we found that individuals cared mostly about the agents who made a relative high effort among their peers, while for trouble behavior, fighting and drinking, the peers that mattered were the ones who exerted relatively low effort in each activity. We then implemented some counterfactual policies; that is, we determined for each activity the taxes/subsidies that would restore the first best. We found that in most cases, it was optimal to target some individuals in the network. For example, for GPA, the most effective way of reaching the social optimum would be to only subsidize a selected number of individuals while, in the LIM model, the planner should give the same subsidy to all individuals. This implies that by imposing the LIM model, the policy recommendations may be very wrong and lead to inefficient outcomes.

More generally, our aim in this study was mainly methodological, as we wanted to show the potential mistakes made by using the (reduced-form) LIM model. While we considered a tax/subsidy policy that would restore the first best, other policies could be implemented. For example, we could consider a policy for which the planner would either maximize (for positive activities such as GPA or self-esteem) or minimize (for negative activities such as risky behavior or drinking) total outcome (instead of welfare) under a budget constraint. Since in our estimations, we show that the peer reference group greatly varies between different activities and very rarely corresponds to the mean peers' outcomes, the discrepancy between the LIM and our general model in terms of policy recommendations would still be very large.

The takeaway from our study is that a tighter link between theory, econometric methods, and data is necessary to deeply understand how peer effects work and which policy to recommend.

APPENDIX A: EXISTENCE AND UNIQUENESS OF EQUILIBRIUM (PROPOSITION 1)

A.1. *Existence of Nash equilibrium*

Define the social norm mapping $\widetilde{\mathbf{y}}:\mathbb{R}^n_+\to\mathbb{R}^n_+$ as follows:[21]

$$\widetilde{\mathbf{y}}(\mathbf{y}) := \left(\widetilde{y}_1(\mathbf{y}), \widetilde{y}_2(\mathbf{y}), \ldots, \widetilde{y}_n(\mathbf{y})\right). \tag{14}$$

[21]For the ease of the presentation, the social norm of individual i is denoted by $\widetilde{y}_i(\mathbf{y}) \equiv \widetilde{y}_{-i}(\beta) \equiv \widetilde{y}_{-i}(\mathbf{y}_{-i},\beta)$.

ASSUMPTION 1: *For all $i = 1, 2, \dots, n$, agent i's social norm $\widetilde{y}_i(\mathbf{y}) : \mathbb{R}_+^n \to \mathbb{R}_+$, is monotone increasing, continuous, linear homogeneous, and normalized: $\widetilde{y}_i(\mathbf{1}) = 1$.*[22]

Clearly, the CES norm (4) satisfies Assumption 1 for any $\beta \in [-\infty, +\infty]$. The best-reply (BR) mapping is given by (7) or, in matrix form:

$$\mathbf{y} = \mathbf{b}(\mathbf{y}) := (1 - \lambda_2)\boldsymbol{\alpha} + (\lambda_1 + \lambda_2)\widetilde{\mathbf{y}}(\mathbf{y}). \tag{15}$$

Let $\mu := (\lambda_1 + \lambda_2) \in (0, 1)$. For each $i = 1, 2, \dots, n$, define

$$\widetilde{\alpha}_i := \frac{1 - \lambda_2}{1 - \mu}\alpha_i \qquad \text{and} \qquad \widetilde{\boldsymbol{\alpha}} := (\widetilde{\alpha}_1, \widetilde{\alpha}_2, \dots, \widetilde{\alpha}_n)\,. \tag{16}$$

Then, the BR mapping (15) can be written as follows:

$$\mathbf{y} = \mathbf{b}(\mathbf{y}) := (1 - \mu)\widetilde{\boldsymbol{\alpha}} + \mu\,\widetilde{\mathbf{y}}(\mathbf{y}). \tag{17}$$

For any n-dimensional vector $\mathbf{x} = (x_1, x_2, \dots, x_n)$, define:

$$x_{\min} := \min_{1 \le i \le n} \{x_i\}, \quad x_{\max} := \max_{1 \le i \le n} \{x_i\}.$$

LEMMA 2: *Under Assumption 1, the following inequalities hold for any $\mathbf{y} \in \mathbb{R}_+^n$:*

$$y_{\min} \le \widetilde{y}_{\min}(\mathbf{y}) \le \widetilde{y}_{\max}(\mathbf{y}) \le y_{\max}. \tag{18}$$

PROOF: By Assumption 1, $\widetilde{\mathbf{y}}(\cdot)$ is increasing over $\mathbb{R}_+^n$. Furthermore, linear homogeneity and the normalization together imply that $\widetilde{\mathbf{y}}(c\mathbf{1}) = c\mathbf{1}$ for any scalar $c \ge 0$. Hence:

$$y_{\min}(\mathbf{y})\mathbf{1} = \widetilde{\mathbf{y}}\left(y_{\min}(\mathbf{y})\mathbf{1}\right) \le \widetilde{\mathbf{y}}\left(\mathbf{y}\right) \le \widetilde{\mathbf{y}}\left(y_{\max}(\mathbf{y})\mathbf{1}\right) = y_{\max}(\mathbf{y})\mathbf{1},$$

which is equivalent to (18). This completes the proof. *Q.E.D.*

PROPOSITION 3: *For any $\widetilde{\boldsymbol{\alpha}} \in \mathbb{R}_{++}$, any $\mu \in (0, 1)$, any network $\mathbf{g}$, and any norm satisfying Assumption 1, the set of Nash equilibria is non-empty and contains a minimum equilibrium $\mathbf{y}^*$ and a maximum equilibrium $\mathbf{y}^{**}$.*

[22] Here, and everywhere below, $\mathbf{1}$ stands for the n-dimensional vector of ones.

PROOF: Consider a game with payoffs (6), social norms (14), and each agent's strategy space $[\widetilde{\alpha}_{\min} - \varepsilon, \widetilde{\alpha}_{\max} + \varepsilon]$, where $\varepsilon > 0$ is small. Clearly, this is a supermodular game.[23] Hence, the set of Nash equilibria of the restricted game contains a minimum equilibrium $\mathbf{y}^*$ and a maximum equilibrium $\mathbf{y}^{**}$ (Topkis, 1998, Theorem 4.2.1). Also, the set of Nash equilibria in the original game is the same as in the restricted game. To see this, using the BR (17), concavity of the min-function (resp., convexity of the max-function), and Lemma 2, for any agent i, we find that

$$
\begin{aligned}
&b_i(\mathbf{y}^*) \geq \min\nolimits_{1\leq i\leq n} \{(1-\mu)\widetilde{\alpha}_i + \mu\widetilde{y}_i(\mathbf{y}^*)\} \geq (1-\mu)\widetilde{\alpha}_{\min} + \mu\widetilde{y}_{\min}(\mathbf{y}^*) \\
&\geq (1-\mu)\widetilde{\alpha}_{\min} + \mu y^*_{\min} \implies b_i(\mathbf{y}^*) \geq \widetilde{\alpha}_{\min}; \\
&b_i(\mathbf{y}^*) \leq \max\nolimits_{1\leq i\leq n} \{(1-\mu)\widetilde{\alpha}_i + \mu\widetilde{y}_i(\mathbf{y}^*)\} \leq (1-\mu)\widetilde{\alpha}_{\max} + \mu\widetilde{y}_{\max}(\mathbf{y}^*) \\
&\leq (1-\mu)\widetilde{\alpha}_{\min} + \mu y^*_{\max} \implies y^*_{\max} \geq \widetilde{\alpha}_{\max}.
\end{aligned}
$$

Thus, every Nash equilibrium in both the restricted game and the unrestricted game is interior and satisfies $\widetilde{\alpha}_{\min} \leq y^*_i \leq \widetilde{\alpha}_{\max}$ for all $i = 1, 2, \ldots, n$. Hence, restricting the strategy space of each player to $[\widetilde{\alpha}_{\min} - \varepsilon, \widetilde{\alpha}_{\max} + \varepsilon]$ does not change the set of Nash equilibria. This completes the proof. *Q.E.D.*

A.2. *Uniqueness of Nash equilibrium*

A.2.1. *Uniqueness of Nash equilibrium for convex norms*

ASSUMPTION 2: *The social norm mapping $\widetilde{\mathbf{y}} : \mathbb{R}^n_+ \to \mathbb{R}^n_+$ is globally convex, that is, the inequality $\widetilde{\mathbf{y}}((1-\gamma)\mathbf{x} + \gamma\mathbf{z}) \leq (1-\gamma)\widetilde{\mathbf{y}}(\mathbf{x}) + \gamma\widetilde{\mathbf{y}}(\mathbf{z})$ holds for any $\gamma \in [0,1]$ and for any $\mathbf{x}, \mathbf{z} \in \mathbb{R}^n_+$*

PROPOSITION 4: *Let the social norm mapping $\widetilde{\mathbf{y}} : \mathbb{R}^n_+ \to \mathbb{R}^n_+$ satisfy Assumptions 1 and 2. Then, (17), and hence (15), has a unique fixed point.*

PROOF: Let $\|\cdot\|_\infty$ be the standard sup-norm over $\mathbb{R}^n$, that is,

$$
\|\mathbf{z}\|_\infty := \max_{i=1,2,\ldots,n} |z_i|, \qquad \text{for all } \mathbf{z} = (z_1, z_2, \ldots, z_n) \in \mathbb{R}^n.
$$

[23] A game with strategies in $\mathbb{R}$ is a supermodular game if (i) each player's strategy space is compact; (ii) each player's utility is upper semi-continuous; (iii) each player's utility function has increasing differences.

For all $\mathbf{x}, \mathbf{z} \in \mathbb{R}^n_+$, the following relations are readily verified:[24]

$$\|\mathbf{z} - \mathbf{x}\|_\infty = \|\mathbf{x} \vee \mathbf{z} - \mathbf{x} \wedge \mathbf{z}\|_\infty; \tag{19}$$

$$-\mathbf{z} \le \mathbf{x} \le \mathbf{z} \quad \Longrightarrow \quad \|\mathbf{x}\|_\infty \le \|\mathbf{z}\|_\infty; \tag{20}$$

$$\|\widetilde{\mathbf{y}}(\mathbf{z})\|_\infty \le \|\mathbf{z}\|_\infty. \tag{21}$$

From the monotonicity of the BR mapping (17), we have:

$$\mathbf{b}(\mathbf{x} \wedge \mathbf{z}) - \mathbf{b}(\mathbf{x} \vee \mathbf{z}) \le \mathbf{b}(\mathbf{z}) - \mathbf{b}(\mathbf{x}) \le \mathbf{b}(\mathbf{x} \vee \mathbf{z}) - \mathbf{b}(\mathbf{x} \wedge \mathbf{z}).$$

Using consecutively (20), (17), the homogeneity and convexity of $\widetilde{\mathbf{y}}(\cdot)$, (21), and (19), we obtain:

$$\begin{aligned}
\|\mathbf{b}(\mathbf{z}) - \mathbf{b}(\mathbf{x})\|_\infty \le \|\mathbf{b}(\mathbf{x} \vee \mathbf{z}) - \mathbf{b}(\mathbf{x} \wedge \mathbf{z})\|_\infty = \mu \|\widetilde{\mathbf{y}}(\mathbf{x} \vee \mathbf{z}) - \widetilde{\mathbf{y}}(\mathbf{x} \wedge \mathbf{z})\|_\infty \\
= \mu \|\widetilde{\mathbf{y}}(\mathbf{x} \wedge \mathbf{z} + \mathbf{x} \vee \mathbf{z} - \mathbf{x} \wedge \mathbf{z}) - \widetilde{\mathbf{y}}(\mathbf{x} \wedge \mathbf{z})\|_\infty \\
= \mu \left\| 2\widetilde{\mathbf{y}}\left(\tfrac{1}{2}\mathbf{x} \wedge \mathbf{z} + \tfrac{1}{2}(\mathbf{x} \vee \mathbf{z} - \mathbf{x} \wedge \mathbf{z})\right) - \widetilde{\mathbf{y}}(\mathbf{x} \wedge \mathbf{z}) \right\|_\infty \\
\le \mu \|\widetilde{\mathbf{y}}(\mathbf{x} \wedge \mathbf{z}) + \widetilde{\mathbf{y}}(\mathbf{x} \vee \mathbf{z} - \mathbf{x} \wedge \mathbf{z}) - \widetilde{\mathbf{y}}(\mathbf{x} \wedge \mathbf{z})\|_\infty \\
= \mu \|\widetilde{\mathbf{y}}(\mathbf{x} \vee \mathbf{z} - \mathbf{x} \wedge \mathbf{z})\|_\infty \le \mu \|\mathbf{x} \vee \mathbf{z} - \mathbf{x} \wedge \mathbf{z}\|_\infty = \mu \|\mathbf{z} - \mathbf{x}\|_\infty,
\end{aligned}$$

i.e., for all $\mathbf{x}, \mathbf{z} \in \mathbb{R}^n_+$, the following inequality holds: $\|\mathbf{b}(\mathbf{z}) - \mathbf{b}(\mathbf{x})\|_\infty \le \mu \|\mathbf{z} - \mathbf{x}\|_\infty$. Therefore, as $\mu \in (0, 1)$, $\mathbf{b}(\cdot)$ is a contraction over the complete metric space $\left(\mathbb{R}^n_+, \rho\right)$, with the distance $\rho(\mathbf{x}, \mathbf{z}) := \|\mathbf{x} - \mathbf{z}\|_\infty$ for all $\mathbf{x}, \mathbf{z} \in \mathbb{R}^n_+$. By the contraction mapping theorem, $\mathbf{b}(\cdot)$ has a unique fixed point $\mathbf{y}^* \in \mathbb{R}^n_+$. This completes the proof. *Q.E.D.*

[24] We use standard notation of lattice theory: for all $\mathbf{x}, \mathbf{z} \in \mathbb{R}^n$,

$$\begin{aligned}
\mathbf{x} \vee \mathbf{z} := (\max\{x_1, z_1\}, \max\{x_2, z_2\}, \ldots, \max\{x_n, z_n\}), \\
\mathbf{x} \wedge \mathbf{z} := (\min\{x_1, z_1\}, \min\{x_2, z_2\}, \ldots, \min\{x_n, z_n\}).
\end{aligned}$$

Observe that the inequality (21) is an immediate implication of Lemma 2.

A.2.2. *Uniqueness of Nash equilibrium for concave norms*

ASSUMPTION 3: *The social norm mapping* $\widetilde{\mathbf{y}} : \mathbb{R}^n_+ \to \mathbb{R}^n_+$ *is globally concave, that is, the inequality* $\widetilde{\mathbf{y}}((1-\gamma)\mathbf{x} + \gamma\mathbf{z}) \geq (1-\gamma)\widetilde{\mathbf{y}}(\mathbf{x}) + \gamma\widetilde{\mathbf{y}}(\mathbf{z})$ *holds for any* $\gamma \in [0,1]$ *and for any* $\mathbf{x}, \mathbf{z} \in \mathbb{R}^n_+$

PROPOSITION 5: *Let the social norm mapping* $\widetilde{\mathbf{y}} : \mathbb{R}^n_+ \to \mathbb{R}^n_+$ *satisfy Assumptions 1 and 3. Then, (17), and hence (15), has a unique fixed point.*

PROOF: We proceed by contradiction. From Proposition 3, there exist the minimum equilibrium $\mathbf{y}^*$ and the maximum equilibrium $\mathbf{y}^{**}$. Assume that $\mathbf{y}^{**} \neq \mathbf{y}^*$, so that the set $\mathcal{I} := \{i \,|\, y_i^{**} > y_i^*\}$ of agents is non-empty. For each $i \in \mathcal{I}$, define τ_i by

$$\tau_i y_i^{**} + (1-\tau_i) y_i^* = 0 \quad \Longrightarrow \quad \tau_i := 1 - \frac{y_i^{**}}{y_i^*} < 0.$$

Pick $j \in \mathcal{I}$ such that $\tau_j = \max\{\tau_i \,|\, i \in \mathcal{I}\}$. It is readily verified that[25]

$$\tau_j \mathbf{y}^{**} + (1-\tau_j)\mathbf{y}^* \geq \mathbf{0} \implies b_j\left(\tau_j \mathbf{y}^{**} + (1-\tau_j)\mathbf{y}^*\right) \geq b_j(\mathbf{0}) > 0. \tag{22}$$

From (22), and by definition of τ_j, we get:[26]

$$\begin{aligned} &b_j\left(\tau_j \mathbf{y}^{**} + (1-\tau_j)\mathbf{y}^*\right) > 0 = \tau_j y_j^{**} + (1-\tau_j) y_j^* \Longrightarrow \\ &b_j\left(\tau_j \mathbf{y}^{**} + (1-\tau_j)\mathbf{y}^*\right) > \tau_j b_j(\mathbf{y}^{**}) + (1-\tau_j) b_j(\mathbf{y}^*). \end{aligned} \tag{23}$$

However, applying $b_j(\cdot)$ to both parts of the identity, we obtain:

$$\mathbf{y}^* = \frac{-\tau_j}{1-\tau_j}\mathbf{y}^{**} + \frac{1}{1-\tau_j}\left(\tau_j \mathbf{y}^{**} + (1-\tau_j)\mathbf{y}^*\right),$$

and using concavity of $b_j(\cdot)$, we get:

$$b_j(\mathbf{y}^*) \geq \frac{-\tau_j}{1-\tau_j} b_j(\mathbf{y}^{**}) + \frac{1}{1-\tau_j} b_j\left(\tau_j \mathbf{y}^{**} + (1-\tau_j)\mathbf{y}^*\right) \quad \Longrightarrow$$

[25] Here $b_j(\cdot)$ is the jth component of (17). We use monotonicity of $b_j(\cdot)$ and $b_j(\mathbf{0}) = (1-\mu)\widetilde{\alpha}_i > 0$.

[26] Because $\mathbf{y}^*$ and $\mathbf{y}^{**}$ are Nash equilibria, $\tau_j b_j(\mathbf{y}^{**}) + (1-\tau_j) b_j(\mathbf{y}^*) = \tau_j y_j^{**}(1-\tau_j) y_j^*$.

$$b_j\left(\tau_j \mathbf{y}^{**} + (1-\tau_j)\mathbf{y}^{*}\right) \leq (1-\tau_j) b_j\left(\mathbf{y}^{*}\right) + \tau_j b_j\left(\mathbf{y}^{**}\right),$$

which contradicts (23). This completes the proof. *Q.E.D.*

It is now straightforward to prove Proposition 1.

Case 1: $\beta \in (1, +\infty]$. In this case, the social norm mapping $\widetilde{\mathbf{y}} : \mathbb{R}_{+}^{n} \to \mathbb{R}_{+}^{n}$ is convex. The existence and uniqueness result follows from Proposition 4.

Case 2: $\beta \in (1, +\infty]$. In this case, the social norm mapping $\widetilde{\mathbf{y}} : \mathbb{R}_{+}^{n} \to \mathbb{R}_{+}^{n}$ is concave. The existence and uniqueness result follows from Proposition 5.

APPENDIX B: ADDITIONAL DETAILS STRUCTURAL ESTIMATION

Weighted Average GMM

Given the moment conditions are based on two distinct groups, we follow the estimation strategy of Arellano and Meghir (1992). Formally, let $\hat{\mathbf{y}}$ be the OLS predictor of $\mathbf{y}$ (or any exogenous predictor of $\mathbf{y}$, that is, an object that is only a function of $\mathbf{x}$), and let $\widetilde{y}'_{-is}(\hat{\mathbf{y}}_s, \beta)$ denote the derivative of $\widetilde{y}_{-is}(\hat{\mathbf{y}}_s, \beta)$ with respect to β. Further, define the set of instruments as $\mathbf{z}_i = [\mathbf{x}_i, \widetilde{y}_{-is}(\hat{\mathbf{y}}_s, \beta), \widetilde{y}'_{-is}(\hat{\mathbf{y}}_s, \beta)]$. There are two orthogonality assumptions on the error term: (1) $\mathbb{E}\left(\varepsilon_i \mathbf{z}_i\right) = \mathbf{0}$ for all non-isolated individuals, and (2) $\mathbb{E}\left(\varepsilon_i \mathbf{x}_i\right) = \mathbf{0}$ for all isolated individuals. The orthogonality conditions follows directly from the assumption that $\mathbb{E}\left(\varepsilon_i | \mathbf{Z}, \mathbf{G}\right) = 0$ for all i as $\hat{y}$ is only a function of $\mathbf{x}$. Then, the method of moments estimator $\boldsymbol{\theta} = [\boldsymbol{\gamma}', \lambda_1, \lambda_2, \beta]'$ is the solution of,

$$Q(\boldsymbol{\theta}) = h_1(\boldsymbol{\theta}) \mathbf{W}_1 h_1'(\boldsymbol{\theta}) + h_2(\boldsymbol{\theta}) \mathbf{W}_2 h_2'(\boldsymbol{\theta}),$$

where

$$h_1(\boldsymbol{\theta}) = \frac{1}{N_1} \sum_{i=1}^{N_1} [y_i - (1-\lambda_2)\boldsymbol{x}_i \boldsymbol{\gamma} - (\lambda_1 + \lambda_2)\widetilde{y}_{-is}(\mathbf{y}_{-is}, \beta))]\mathbf{z}_i$$

and

$$h_2(\boldsymbol{\theta}) = \frac{1}{N_2} \sum_{i=1}^{N_2} [y_i - \boldsymbol{x}_i \boldsymbol{\gamma}]\mathbf{x}_i$$

for non-isolated and isolated individuals, respectively. Note, N_1 is the number of non-isolated individuals and N_2 is the number of isolated individuals . Note, the identification of $\boldsymbol{\theta}$ relies on both moment conditions so we need to ensure that both are asymptotically not-negligible, i.e. $\lim_{N_1+N_2\to\infty} \frac{N_1}{N_1+N_2} = r_1 \in (0,1)$ (which is equivalent to $\lim_{N_1+N_2\to\infty} \frac{N_2}{N_1+N_2} = r_2 \in (0,1)$).

Concentrated GMM

For estimation purposes, as the moment functions are linear in $\boldsymbol{\gamma}$, we can concentrate the objective function around $[\lambda_1, \lambda_2, \beta]$. Taking the first order condition of $Q(\boldsymbol{\theta})$ with respect to $\boldsymbol{\gamma}$, we obtain (after long, but straightforward algebra):

$$\hat{\boldsymbol{\gamma}}(\lambda_1, \lambda_2, \beta) = \left[\frac{(1-\lambda_2)^2}{N_1^2}\mathbf{X}_1'\mathbf{Z}_1\mathbf{W}_1\mathbf{Z}_1'\mathbf{X}_1 + \frac{1}{N_2^2}\mathbf{X}_2'\mathbf{X}_2\mathbf{W}_2\mathbf{X}_2'\mathbf{X}_2\right]^{-1} \times$$
$$\left[\frac{(1-\lambda_2)}{N_1^2}\mathbf{X}_1'\mathbf{Z}_1\mathbf{W}_1\mathbf{Z}_1'(\mathbf{y}_1 - (\lambda_1+\lambda_2)\boldsymbol{\phi}_1(\mathbf{y}_{-i}, \beta)) + \frac{1}{N_2^2}\mathbf{X}_2'\mathbf{X}_2\mathbf{W}_2\mathbf{X}_2'\mathbf{y}_2\right]^{-1},$$

where for any (row) vector $\mathbf{a}_i = (\mathbf{x}_i, \mathbf{z}_i, \mathbf{w}_i, \mathbf{y}_i)$, the matrix $\mathbf{A}_1 = (\mathbf{X}_1, \mathbf{Z}_1, \mathbf{W}_1)$ is obtained by staking $\mathbf{a}_i$ for all non-isolated individual i, and $\mathbf{A}_2 = (\mathbf{X}_2, \mathbf{Z}_2, \mathbf{W}_2)$ is obtained by staking $\mathbf{a}_i$ for all isolated individual i.

The concentrated objective function is therefore $\tilde{Q}(\tilde{\boldsymbol{\theta}}) = \tilde{Q}([\lambda_1, \lambda_2, \beta]) = Q([\hat{\boldsymbol{\gamma}}'(\lambda_1, \lambda_2, \beta), \lambda_1, \lambda_2, \beta])$, where $\tilde{\boldsymbol{\theta}} = [\lambda_1, \lambda_2, \beta]$. The function is minimized numerically in Section 3.4.

APPENDIX C: SOCIAL OPTIMUM (FIRST BEST)

Consider a standard welfare function $\mathcal{W}(\mathbf{y}, \mathbf{g}) = \sum_i U_i(y_i, \mathbf{y}_{-i}, \mathbf{g})$, where each agent i's utility $U_i(y_i, \mathbf{y}_{-i}, \mathbf{g})$ is given by (6). The planner chooses the actions $y_1, y_2, ..., y_n$ of each of the n agents that maximizes $\mathcal{W}(\mathbf{y}, \mathbf{g})$. The first best is equal to:

$$y_i^o = (1-\lambda_2)\alpha_i + (\lambda_1+\lambda_2)\widetilde{y}_{-i}(\beta) + \lambda_1 \sum_j y_j \frac{\partial \widetilde{y}_{-i}(\beta)}{\partial y_i} + \lambda_2 \sum_j (y_j - \widetilde{y}_{-j}(\beta)) \frac{\partial \widetilde{y}_{-i}(\beta)}{\partial y_i}, \tag{24}$$

PROPOSITION 6: *Assume that $\lambda_1 + \lambda_2 < 1$. Then, the first best outcome is unique.*

PROOF: To render the notations more explicit, denote $\widetilde{y}_{-i}(\beta) \equiv \widetilde{y}_{-i}(\mathbf{z},\beta)$. Let us restate (24) in vector-matrix form:

$$\mathbf{y} = (1-\lambda_2)\boldsymbol{\alpha} + \lambda_1 \mathbf{F}(\mathbf{y}) + \lambda_2 \mathbf{G}(\mathbf{y}), \tag{25}$$

where the mappings $\mathbf{F}(\mathbf{y}) = (F_1(\mathbf{y}), F_2(\mathbf{y}), \ldots, F_n(\mathbf{y}))$ and $\mathbf{G}(\mathbf{y}) = (G_1(\mathbf{y}), G_2(\mathbf{y}), \ldots, G_n(\mathbf{y}))$ are defined, respectively, as follows:

$$F_i(\mathbf{y}) := \widetilde{y}_{-i}(\mathbf{y}_{-i},\beta) + \sum_j y_j(\mathbf{y}_{-j}) \frac{\partial \widetilde{y}_{-i}(\mathbf{y}_{-j},\beta)}{\partial y_i},$$

$$G_i(\mathbf{y}) := \widetilde{y}_{-i}(\mathbf{y}_{-i},\beta) + \sum_j (y_j - \widetilde{y}_{-i}(\mathbf{y}_{-j},\beta)) \frac{\partial \widetilde{y}_{-i}(\mathbf{y}_{-j},\beta)}{\partial y_i}.$$

At the extreme case of $\lambda_1 = \lambda_2 = 0$, the fixed point condition (25) has a unique solution $\mathbf{y}^O = \boldsymbol{\alpha}$. Furthermore, since the right-hand side of (25) is continuously differentiable with respect to λ_1, λ_2, and at $(\lambda_1, \lambda_2, \mathbf{y}) = (0,0,\boldsymbol{\alpha})$, by the implicit function theorem, there exist threshold values $\widehat{\lambda}_1 > 0$ and $\widehat{\lambda}_2 > 0$ of λ_1 and λ_2 respectively, such that (25) defines a single-valued function $\mathbf{y}^O(\lambda_1,\lambda_2)$ for all $(\lambda_1,\lambda_2) \in \Big[(0,0);(\widehat{\lambda}_1,\widehat{\lambda}_2)\Big]$. Moreover, we need to impose that $\lambda_1 + \lambda_2 < 1$ for each model to be well-defined, which implies that $0 < \widehat{\lambda}_1 + \widehat{\lambda}_2 < 1$. It remains to prove that $\mathbf{y}^O(\lambda_1,\lambda_2)$ is a unique solution to (25). We proceed by contradiction. Assume that there exists a sequence $(\lambda_1^k, \lambda_2^k) \to 0$, such that, for any $(\lambda_1^k, \lambda_2^k)$ there exists $\widehat{\mathbf{y}}(\lambda_1^k, \lambda_2^k) \neq \mathbf{y}^O(\lambda_1^k, \lambda_2^k)$. Two cases may arise.

Case 1: the sequence $\widehat{\mathbf{y}}(\lambda_1^k, \lambda_2^k)$ converges to $\boldsymbol{\alpha}$. This case is impossible since it implies the existence of two distinct branches of the fixed-point correspondence defined by (25), which violates the implicit function theorem.

Case 2: the sequence $\widehat{\mathbf{y}}(\lambda_1^k, \lambda_2^k)$ has a subsequence, which does not converge to $\boldsymbol{\alpha}$ but converges to some $\boldsymbol{\xi} \neq \boldsymbol{\alpha}$. This leads to a contradiction since both the left-hand side and the right-hand side of (25) are continuous with respect to $(\lambda_1, \lambda_2, \mathbf{y})$ at $(\lambda_1, \lambda_2, \mathbf{y}) = (0,0,\boldsymbol{\xi})$. Taking the limit on both sides of (25) under $(\lambda_1^k, \lambda_2^k, \widehat{\mathbf{y}}(\lambda_1^k, \lambda_2^k)) \to (0,0,\boldsymbol{\xi})$, we conclude that $\mathbf{y} = \boldsymbol{\xi}$ must be a solution to (25) in the extreme case of $\lambda_1 = \lambda_2 = 0$. But we have assumed $\boldsymbol{\xi} \neq \boldsymbol{\alpha}$, and (25) clearly has no solutions other than $\boldsymbol{\alpha}$, a contradiction.

This completes the proof. *Q.E.D.*

REFERENCES

Akerlof, G. (1997). Social distance and social decisions. *Econometrica 65*, 1005–1027. [8]

Angrist, J. and A. Krueger (1992). The effect of age at school entry on educational attainment: An application of instrumental variables with moments from two samples. *Journal of the American Statistical Association 87*, 328–336. [11]

Arellano, M. and C. Meghir (1992). Female labour supply and on-the-job search: An empirical model estimated using complementary data sets. *Review of Economic Studies 59*, 537–559. [11, 28]

Bernheim, B. (1994). A theory of conformity. *Journal of Political Economy 102*, 841–877. [8]

Bietenbeck, J. (2020). The long-term impacts of low-achieving childhood peers: Evidence from project star. *Journal of the European Economic Association 18*(1), 392–426. [4]

Blume, L., W. Brock, S. Durlauf, and Y. Ioannides (2011). Identification of social interactions. In: J. Benhabib, A. Bisin, and M.O. Jackson (Eds.)*, Handbook of Social Economics,* Amsterdam: North-Holland, 853–964. [5]

Blume, L., W. Brock, S. Durlauf, and R. Jayaraman (2015). Linear social interactions models. *Journal of Political Economy 123*, 444–496. [2]

Boucher, V. (2016). Conformism and self-selection in social networks. *Journal of Public Economics 136*, 30–44. [2, 8]

Boucher, V. and B. Fortin (2016). Some challenges in the empirics of the effects of networks. In: Bramoullé, Y., Galeotti, A., and Rogers, b.W. (Eds.), *The Oxford Handbook on the Economics of Networks,* Oxford: Oxford University Press, 277–302. [2, 8]

Boucher, V. and A. Houndetoungan (2023). Estimating peer effects using partial network data. Unpublished manuscript, University of Laval. [13]

Bramoullé, Y., H. Djebbari, and B. Fortin (2009). Identification of peer effects through social networks. *Journal of Econometrics 150*, 41–55. [9, 11]

Brock, W. and S. Durlauf (2001a). Discrete choice with social interactions. *Review of Economic Studies 68*(2), 235–260. [8]

Brock, W. and S. Durlauf (2001b). Interactions-based models. In: J.J. Heckman and E. Leamer (Eds.)*, Handbook of Econometrics, Vol. 5,* Amsterdam: North-Holland, 3297–3380. [5]

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009). Peer effects and social networks in education. *The Review of Economic Studies 76*(4), 1239–1267. [2]

Cameron, A. and P. Trivedi (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press. [11]

Carrell, S. E., M. E. Page, and J. E. West (2010). Sex and science: How professor gender perpetuates the gender gap. *Quarterly Journal of Economics 125*(3), 1101–1144. [4]

Dahl, G., K. Loken, and M. Mogstad (2014). Peer efects in program participation. *American Economic Review 104*(7), 2049–2074. [2]

Diaz, C., E. Patacchini, T. Verdier, and Y. Zenou (2021). Leaders in juvenile crime. *Journal of Economic Behavior and Organization 192*, 638–667. [4]

Glaeser, E. and J. Scheinkman (2002). Non-market interactions. In: M. Dewatripoint, L., Hansen, P. and S. Turnovsky (Eds.), *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress,* Cambridge: Cambridge University Press, 339–369. [8]

Hahn, Y., A. Islam, E. Patacchini, and Y. Zenou (2020). Friendship and female education: Evidence from a field experiment in Bangladeshi primary schools. *Economic Journal 130*, 740–764. [4]

Harris, K., C. Halpern, E. Whitsel, J. Hussey, L. Killeya-Jones, J. Tabor, and S. Dean (2019). Endogenous social interactions: Which peers matter? *International Journal of Epidemiology 48*(5), 1415–1425. [12]

Hoxby, C. and G. Weingarth (2005). Taking race out of the equation: School reassignment and the structure of peer effects. Unpublished manuscript, Harvard University. [4]

Islam, A., M. Vlassopoulos, Y. Zenou, and X. Zhang (2021). Centrality-based spillover effects. CEPR Discussion Paper No. 16321. [4]

Kline, B. and E. Tamer (2020). Econometric analysis of models with social interactions. In: B. Graham and A. de Paula (Eds.)*, The Econometric Analysis of Network Data,* New York: Academic Press, 149–181. [2]

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*(3), 849–857. [2, 10]

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies 60*(3), 531–542. [2]

Milgrom, P. and J. Roberts (1990). Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica 58*(6), 1255–1277. [3]

Patacchini, E. and Y. Zenou (2012). Juvenile delinquency and conformism. *Journal of Law, Economics, and Organization 28*(1), 1–31. [2, 8]

Reif, J. (2019). A model of addiction and social interactions. *Economic Inquiry 57*(2), 759–773. [8]

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In: Hanushek, E., Machin, S. and L. Woessmann (Eds.)*, The Handbook of the Economics of Education,* Dordrecht: Elsevier, 249–277. [2]

Tao, J. and L. Lee (2014). A social interaction model with an extreme order statistic. *The Econometrics Journal 17*(3), 197–240. [4]

Tatsi, E. (2015). Endogenous social interactions: Which peers matter? Unpublished manuscript, SOFI. [5]

Topkis, D. (1998). *Supermodularity and Complementarity*. Princeton University Press. [25]

Ushchev, P. and Y. Zenou (2020). Social norms in networks. *Journal of Economic Theory 185*, 104969. [2, 8]