Whom to Target under Peer Pressure? A Social Marginal Effects Approach*

Pablo Estrada†

May, 2023

Preliminary and Incomplete.
Please do not cite or circulate.

Abstract

This study investigates how a policymaker can effectively target adolescents with a prevention program to reduce risky behavior such as marijuana consumption. We assume that peer pressure is the mechanism that determines how social interactions influence such behaviors. To determine the most influential individuals, we estimate social marginal effects using observational data on multiple social connections with a generalized method of moments (GMM). Our empirical strategy uses the observed characteristics of distant individuals in a multilayered network space as instruments to address the endogeneity of the networks that arise from homophily. We use the Add Health data to find positive peer effects, for friends and classmates, on both cigarette smoking and marijuana use, with the friends effect having a greater impact. Our findings suggest that policymakers can benefit from using social marginal effects to target high-influence individuals in their prevention programs.

Keywords: Health Behaviors, Peer Effects, Multilayer Networks, Policy Targeting

JEL Codes: I12, C33

*We are grateful to Guido Kuersteiner, Paul Goldsmith-Pinkham, and participants in the Network Science and Economics Conference 2023, and the Econometric Sociesty European Meeting 2023 for their useful feedback. Rachel Clohan provided expert research assistant.

†Emory University. pablo.estrada@emory.edu

1 Introduction

Risky behaviors such as smoking marijuana and vaping can adversely affect the well-being and cognitive development of teenagers, according to the U.S. Center for Disease Control[1]. Previous literature has established that such behaviors can impact education, crime, future earnings, and healthcare costs[2]. The 2021 National Youth Tobacco Survey reported that the most common reason for U.S. adolescents to try e-cigarettes is because "a friend used them." Since peer pressure plays a critical role in adolescents' risky behaviors, designing effective prevention programs that account for social connections remains a pressing concern.

Suppose a policymaker aims to reduce the aggregate level of a risky behavior, such as the total amount of marijuana consumption, but is limited by the number of individuals that can be included in a prevention program. In such a scenario, how should the policymaker decide which adolescents to target? To address this question, the policymaker must be able to identify highly influential adolescents to the impact of the prevention program. We assume that peer pressure is the mechanism through which social interactions influence risky behaviors. Under this assumption, an adolescent's risky behavior tends to align with the average of their connections.

However, different types of social links can create varying levels of social pressure. For example, adolescents may be more influenced by their friends than their neighbors when deciding whether or not to consume marijuana. Therefore, it may be relevant to incorporate a broader range of possible social connections. In this paper, we demonstrate how, under the influence of peer pressure and the potentially heterogeneous effects of different types of connections, a policymaker can use observational social network data to select individuals for a program in a way that maximizes the impact on risky behavior.

We assume that peer pressure is the underlying mechanism through which social interactions affect adolescents' risky behaviors. Ushchev and Zenou (2020) provide a microfoundation for this assumption, linking peer pressure to "social norms." They show that group-based policies are more effective when social norms are the driving force behind individual behavior. Building on their work, we incorporate different types of social connections into the social interactions game. With this social interactions game, we are able to identify highly influential individuals through the development of a novel network centrality measure. Our centrality measure, called the social marginal effects, is related to a heterogeneous Katz-Bonacich centrality, which builds upon the eigenvector centrality for multilayer networks. It incorporates a discount factor for each type of social connection, which is tied to the peer

[1] CDC - Teens Risky Behaviors

[2] For a review of the economic consequences of risky behaviors, see Cawley and Ruhm (2011)

effects specific to that particular connection.

Numerous studies have explored the influence of peers on risky health behaviors, such as marijuana use and smoking. For instance, Card and Giuliano (2013) estimate a structural model that accounts for peer effects and find that peer pressure has a significant impact on marijuana use among adolescents. Similarly, Arduini et al. (2019) use the National Longitudinal Study of Adolescent Health to estimate a dynamic social interactions model that considers addiction effects in smoking. These works, along with others[3], have established the existence of peer effects on risky behaviors. This article contributes to the literature on peer effects by incorporating various types of social connections that impact individuals' behavior.

Estimating peer effects using multiple networks poses several methodological challenges, including potential endogeneity issues stemming from the social networks. Some existing studies (e.g., Goldsmith-Pinkham and Imbens, 2013; Kuersteiner and Prucha, 2020) have tackled these challenges using Bayesian and generalized method of moments (GMM) frameworks. Others (e.g., Comola and Prina, 2021) have focused on identifying treatment effects when the structure of the network changes. In contrast, we adopt the approach proposed by Estrada (2021) and utilize the distance of individuals across different networks to construct instruments for identifying and estimating peer effects in a multilayered linear-in-means model (Manta et al., 2022).

This paper presents a new strategy to identify heterogeneous peer effects using multilayer networks, and shows how to recover individuals' social marginal effects with observational data on multiple social connections. These estimated social effects are then used to identify the group of individuals to target with a particular policy, with the assumption of peer pressure. The main econometric challenge is the endogeneity of the networks that arise from homophily. To address this, our empirical strategy uses the observed characteristics of distant individuals in a multilayered network space as instruments. The resulting estimates are non-linear functions of social effect parameters in a linear model, estimated by a non-linear GMM.

Using data from Add Health, we assess the influence of peers on cigarette smoking and marijuana use, considering two types of social connections: friends and classmates. Our results reveal positive effects of both friends and classmates on smoking and marijuana use, with the former being more dominant in both cases. Notably, the friends effect is twice as large for marijuana use compared to cigarette smoking. Based on our theoretical model, we identify a group of highly influential individuals, who differ from their low-influence

[3] See also Manski (1993), Bramoullé et al. (2009), Patacchini and Zenou (2009), Calvó-Armengol et al. (2009), Lin (2010), Liu et al. (2014), Boucher and Mourifié (2017), Hsieh and Lin (2020), Mele (2021).

counterparts only in terms of their ability test score, age, and race. Policymakers can leverage these findings when dealing with policy interventions that rely on observational data from multiple social connections.

The contribution of this paper is threefold. First, we present a novel approach to identify heterogeneous peer effects using multilayer networks. This approach employs instrumental variables constructed from distances across various types of social connections. Second, we demonstrate that choosing individuals based on social marginal effects, under the assumption of peer pressure, can considerably reduce the aggregate level of risky behavior. Lastly, we show that identifying highly influential individuals through the social marginal effects aligns with a measure of popularity captured by the heterogeneous Katz-Bonacich centrality.

The remainder of this article is organized as follows. Section 2 presents the theoretical framework for a model of heterogeneous social interactions. In section 3, we describe the data and present descriptive statistics. Section 4 discusses the identification strategy and the non-linear GMM estimation. In section 5, we present the empirical results, robustness checks, and discussion of results. Finally, section 6 concludes the paper and provides suggestions for future research.

2 Social Marginal Effects

This section presents a theoretical support for our influence metric: the social marginal effects. To do so, we develop a social interaction model that incorporates peer pressure as a determinant of individual preferences. Our model disaggregates the decision to engage in risky behaviors into two components: individual characteristics and social effects. We demonstrate that the matrix of social effects yields an index of influence over the risky behavior of others. The social marginal effects can be effectively quantified through a novel measure of heterogeneous Katz-Bonacich centrality that we develop.

Building on the work of Ushchev and Zenou (2020), our model of social interactions captures the influence of social conformity, while also acknowledging the heterogeneity of social connections. To account for the endogeneity of social networks, we employ a two-stage bayesian game, following the approach of Blume et al. (2015). In the first stage, agents make decisions regarding their social connections for each type of interaction. Subsequently, they determine the optimal level of risky behavior to engage in. Appendix B provides detailed exposition on the social interactions game. The optimal choice of risky behavior for an individual is given by Eq. (1),

$$\mathbf{y} = \frac{1}{1+\sum_{m=1}^{M}\beta_m^*}\left[\sum_{m=1}^{M}\beta_m^*\mathbf{W}_m\mathbf{y} + \sum_{m=1}^{M}\mathbf{W}_m\mathbf{X}\delta_m^* + \mathbf{X}\gamma_1^* + \mathbf{B}\gamma_2^* + \mathbf{e}\right], \tag{1}$$

which includes social effects stemming from the average outcomes $\mathbf{W}_m\mathbf{y}$ and characteristics $\mathbf{W}_m\mathbf{X}$ of her connections. The outcome variable $\mathbf{y}$ is a $n \times 1$ vector, and $\mathbf{X}$ and $\mathbf{B}$ are $n \times k_1$ and $n \times k_2$ matrices of covariates, respectively, while $\mathbf{e}$ captures the error term.

In the proposed model, we account for the impact of diverse social connections stemming from M networks, each represented by a row-normalized $(n \times n)$ adjacency matrix $\mathbf{W}_m$. The parameters β_m quantify an individual's conformity with the average behavior of her peers of type m. For instance, β_m denotes the social effect of the average marijuana use among friends and classmates when $m = 1, 2$. Equation (1) can be reformulated as

$$\mathbf{y} = \mathbf{S}(\beta^*)\left[\sum_{m=1}^{M}\mathbf{W}_m\mathbf{X}\delta_m^* + \mathbf{X}\gamma_1^* + \mathbf{B}\gamma_2^* + \mathbf{e}\right],$$

where $\mathbf{S}(\beta^*) = \left[\left(1+\sum_{m=1}^{M}\beta_m^*\right)\mathbf{I} - \sum_{m=1}^{M}\beta_m^*\mathbf{W}_m\right]^{-1}$ represents a $n \times n$ matrix capturing social effects.

The matrix $\mathbf{S}(\beta^*)$ captures the marginal effect of a change in j's characteristics on i's outcome, as reflected in the entry s_{ij}. Correct identification of the conformity parameters β_m^* is key to estimating $\mathbf{S}(\beta^*)$. Section 4 outlines a method to identify the model parameters by leveraging the heterogeneity in social connections. The social marginal effect for individual j, computed as $s_j = \sum_{j \neq i} s_{ij}$, can be obtained from $\mathbf{S}(\beta^*)$. Additionally, we can express s_j as a multilayer version of the Katz-Bonacich centrality. We first show that the Katz-Bonacich centrality describes the monolayer counterpart of our definition of social marginal effects s_j. Later in this section, we provide a full description of the heterogeneous Katz-Bonacich centrality.

From the structural equation (1), we can derive the regression equation (2) that we seek to estimate. Equation (2) considers solely the two types of social interactions - friends and classmates - that we employ in the empirical analysis. Specifically, we utilize the vector $\mathbf{W}_1\mathbf{y}$ to represent the average level of risky behavior among friends and the vector $\mathbf{W}_2\mathbf{y}$ to denote the average level of risky behavior among classmates. Analogously, the vectors $\mathbf{W}_1\mathbf{X}$ and $\mathbf{W}_2\mathbf{X}$ convey the average characteristics of friends and classmates, respectively. In section 4, we discuss the underlying assumptions that enable the identification of the coefficients

$\theta = [\beta_1, \beta_2, \delta_1, \delta_2, \gamma_1, \gamma_2]^\top$.

$$\mathbf{y} = \beta_1 \mathbf{W}_1 \mathbf{y} + \beta_2 \mathbf{W}_2 \mathbf{y} + \mathbf{W}_1 \mathbf{X} \delta_1 + \mathbf{W}_2 \mathbf{X} \delta_2 + \mathbf{X} \gamma_1 + \mathbf{B} \gamma_2 + \mathbf{e}_t \tag{2}$$

The coefficients β_1 and β_2 capture the influence of friends' and classmates' risky behavior on their own individual behavior. However, it is important to note that these coefficients should not be conflated with the parameters β_1^* and β_2^*, which represent the peer effects in the proposed model. Proposition 1 formally outlines how to recover the model's parameters under Assumption 1.

Proposition 1. *Assuming invertibility of the matrix $\mathbf{S}(\beta^*)$, the parameters of the model are given by*

$$\beta_m^* = \frac{\beta_m}{1 - \beta_1 - \beta_2}\ , \quad \delta_m^* = \frac{\delta_m}{1 - \beta_1 - \beta_2}\ , \quad \textit{and} \quad \gamma_{1,2}^* = \frac{\gamma_{1,2}}{1 - \beta_1 - \beta_2} \tag{3}$$

for $m = 1, \ldots, M$.

Calvó-Armengol et al. (2009) establish that the Katz-Bonacich centrality is the network index that describes the equilibrium behavior of individuals. Following their work, we show the connection between the Katz-Bonacich centrality and the social marginal effects when we only account for one type of social network. What is more, we propose a heterogeneous Katz-Bonacich centrality that is defined in terms of the social marginal effects s_j. This heterogeneous Katz-Bonacich centrality results from modifying the eigenvector centrality for multilayered networks proposed by Solá et al. (2013). However, we show that the influence parameters to choose on the heterogeneous Katz-Bonacich centrality calculation come from the peer effects β_1 and β_2.

We first establish the connection between the social marginal effects and the Katz-Bonacich centrality for the case of a single network. Calvó-Armengol et al. (2009) shows that the Katz-Bonacich centrality can be written as $\mathbf{b} = (\mathbf{I} - \beta \mathbf{W})^{-1}(\beta \mathbf{W} \cdot \mathbf{1})$, where $\mathbf{1}$ is a vector of ones and the influence parameter β is non-negative discount factor. If the social interaction matrix $\mathbf{W}$ is row-normalized, we can rewrite it as $\mathbf{b} = \beta(\mathbf{I} - \beta \mathbf{W})^{-1} \cdot \mathbf{1}$. Therefore, the social marginal effects can be expressed as $\mathbf{s} = \frac{1-\beta}{\beta} \mathbf{b}$. This result indicates that the social marginal effects are proportional to the Katz-Bonacich centrality measure with influence parameter β.

From the regression equation (2), the matrix of social effects can also be defined as $\mathbf{S}(\beta) = (1 - \beta_1 - \beta_2)(\mathbf{I} - \beta_1 \mathbf{W}_1 - \beta_2 \mathbf{W}_2)^{-1}$. Using this definition, we relate the social marginal effects to a new measure of Katz-Bonacich centrality over multiple networks. Previous work by Solá et al. (2013) show that the eigenvector centrality can be extended to include multilayer networks. They define uniform and heterogeneous types of eigenvector centrality. In Appendix B, we propose a heterogeneous Katz-Bonacich centrality for the case of two networks as $\mathbf{b} = (\beta_1 + \beta_2)(\mathbf{I} - \beta_1\mathbf{W}_1 - \beta_2\mathbf{W}_2)^{-1} \cdot \mathbf{1}$. And we show that the relationship between the social marginal effects and the heterogeneous Katz-Bonacich centrality is $\mathbf{s} = \frac{1-\beta_1-\beta_2}{\beta_1+\beta_2}\mathbf{b}$.

3 Data and Descriptive Statistics

3.1 Add Health Data

The empirical analysis is based on data from the National Longitudinal Study of Adolescent Health (Add Health), which is a comprehensive, nationally representative survey that gathered data from more than 20,000 adolescents through in-home and in-school interviews. The sample comprises adolescents in grades 7-12 during 1994-95, and the study tracks them through five waves to date, the latest of which was conducted in 2016-18. We restrict our attention to the saturated subsample of 16 schools where all students were eligible for the in-home questionnaire. Within this subset, we focus exclusively on high school students and their social connections with both friends and classmates.

We examine the impact of peers through two distinct social connections, which we detail in Table A1 in Appendix A. The first is the friendship network, constructed from the Add Health in-home questionnaire, where respondents were asked to nominate up to five female and male friends. We identify reciprocal nominations and construct the corresponding network, which includes 454 pairs and has a density of 0.0005. Despite the low density, the network has an average shortest path of 10 links, providing sufficient variability in connections to investigate peer influence.

The construction of the classmates network is based on the in-school questionnaire and follows the approach of Arduini et al. (2019). Specifically, the classmates network is formed by aggregating all same-gender students within the same grade-school. As shown by Arduini et al. (2019), the classmates network exhibits strong assortativity by gender, which aligns with the gender-based homophily commonly observed in adolescent social networks in the sociological literature. Due to its inherent structure, the classmates network is highly clustered, with 300 clusters and an average cluster size of 23 adolescents.

The two outcomes of interest are cigarette and marijuana use. To measure these outcomes, we use the answers provided in the questionnaires to the following inquiry: "During the past 30 days, how many days did you smoke *cigarettes/marijuana*?". We also conduct robustness checks using other definitions for the outcome variables, such as a score that takes the value of 0 if the respondent has never smoked, 1 if they have smoked at least once in their lifetime, and a score from 2 to 7 following the National Survey on Drug Use and Health (NSDUH) scales.[4] However, to uphold the assumptions of the model, the analysis primarily employs continuous outcome variables to examine the effects of social connections. Appendix A presents detailed information on the outcomes and the corresponding score definitions.

The individual and parental characteristics of the adolescents are obtained through the in-home questionnaire of the Add Health dataset. Specifically, we extract information related to gender, race, age, PVT test (ability) score, physical activity index, parents' characteristics, and the availability of prevention programs in the school from the first wave. Moreover, we utilize the questions from the second wave that measure risk aversion and future orientation, as these factors are closely associated with potential risky behavior. The inclusion of future orientation is particularly pertinent, as it captures the likelihood that the adolescent believes they will die at 21 years old.

3.2 Descriptive Statistics

In our regression analysis, we incorporate a set of individual and parental characteristics. To distinguish between covariates that reflect social effects and those that only affect the outcome directly, we allow for individual variables such as PVT test score, physical activity index, risk aversion, and future orientation to have effects through friends and classmates. The remaining covariates such as gender, race, age, parent's characteristics, and school prevention program are included in the matrix $\mathbf{B}$ with direct effects γ_2.

Our sample consists of 1,334 high school students, of whom 49% are females. The race variable is constructed using previous work on mixed-race identity (Udry et al., 2003), and the PVT test score is normalized. Future orientation and risk aversion are measured using indices with a 1-5 scale, while physical activity is measured on a 0-3 scale to indicate how many times the adolescent exercised in the past week. The parents' characteristics consist of dummies for a parent college degree, smokers in the household, and a two-parent household. We also include a dummy for drug and tobacco prevention programs in the adolescent's school. Finally, we report descriptive statistics for missing variables.

Table A2 displays the descriptive statistics of the high school students. On average, students in our sample have smoked cigarettes for four days and marijuana for two days in the past month. There is a high variance in the outcomes, particularly for marijuana consumption. The future orientation index is a 1-5 scale measure, with higher values indicating that the adolescent is less future-oriented. In the saturated sample of schools, 32% and 28% have

[4] Specifically, the score is constructed by the following rule: 0 - never smoked, 1 - smoked at least once in lifetime, 2 - smoked 1-2 days past month, 3 - smoked 3-5 days past month, 4 - smoked 6-9 days past month, 5 - smoked 10-19 days past month, 6 - smoked 20-29 days past month, 7 - smoked everyday past month.

a drug and tobacco prevention program, respectively. The dummies for missing values range from 5% to 7%, and we include them when estimating regression (2). Our sample has very similar demographics to Arduini et al. (2019).

4 Empirical Framework

The main goal of the paper is to distinguish individuals that bear high social effects. However, estimating the social effects resulting from peer pressure presents significant challenges in identifying the underlying peer effects. To overcome the endogeneity of social networks that arise from homophily, we adopt Estrada's (2021) innovative IV strategy. We use the observed characteristics of distant individuals across different networks as instruments to identify peer effects. Our approach relies on the assumption that the observed characteristics of my friends' classmates, who are not my classmates, are a good instrument for identifying the peer effect of friends.

To estimate peer effects for friends and classmates, we employ a GMM framework. The resulting estimates are then used to calculate the social effects for each adolescent. The GMM estimator incorporates linear and quadratic instruments based on our main identification strategy. Kuersteiner and Prucha (2020) provide a detailed discussion of our instrument selection and estimation procedure. We derive the social effects from our assumptions on the peer pressure mechanism that governs the preferences of individuals. Additionally, we calculate confidence intervals on the individual social effects by applying the delta method to a function of our parameters of interest.

4.1 Identification Strategy

The literature on peer effects proposes various methods to address the problem of homophily in social interactions, also known as the problem of correlated effects. One approach to identifying the structural parameters of equation (1) is to assume a multilayer network formation process. Alternatively, if we assume that the observable characteristics of individuals are strictly or contemporaneously exogenous, we can adopt the approach proposed by Kuersteiner and Prucha (2020). They suggest using past networks, $\mathbf{W}_{m,t-1}$, as instruments for present networks, $\mathbf{W}_{m,t}$.

Past networks, such as $\mathbf{W}_{m,t-1}$, allow for the creation of instruments in the form of $\mathbf{W}^2_{m,t-1}\mathbf{X_t}$. If the adjacency matrix $\mathbf{W}_m$ represents friendships, then $\mathbf{W}^2_{m,t-1}\mathbf{X_t}$ denotes the average characteristics of my previous friends' friends. In the case of the classmates' network, this is the average of students' characteristics who are not in my classroom in the past period. Under the assumption that the observed characteristics $\mathbf{X_t}$ are sequentially exogenous, these instruments address the problem of homophily. However, this method requires access to at least three time periods.

Instead, we allow observable characteristics to be correlated with the unobserved characteristics of peers (friends and classmates), meaning $\mathbb{E}[\mathbf{x}_j e_i] \neq 0$. Under this condition, we can impose the weak neighborhood dependence assumption from Estrada (2021). We use distant individuals in the multilayer network to set up moment conditions that help us identify and estimate the social parameters of equation (1). This multilayer network encompasses two dimensions: path lengths and edge-type changes. The primary assumption is that individuals connected through different networks are less likely to be dependent than those connected in the same network. For instance, this indicates that individual i is less likely to be similar to her friend's classmates than to her friend's friends.

Proposition 2 formalizes the key identification strategy for setting up moment conditions. We define d^*_{ij} as the minimum path length between individual i and j in the multilayer space (which includes friends and classmates' links), c^*_{ij} as the minimum number of edge-type changes between i and j, and d^c_{ij} as the second shortest path between i and j for which the number of edge-type changes is less than c^*_{ij}. Appendix C details assumptions and results for identification.

Proposition 2. *Under Assumptions 2 and 3, the conditional distribution $\mathcal{F}(\mathbf{X}, \mathbf{e} \mid \mathcal{M})$ is such that*

$$\mathbb{E}\left(\mathbf{x}_j e_i \mid c^*_{ij} \geq 1,\ d^c_{ij} \geq 2\right) = \mathbf{0}, \tag{4}$$

$$\mathbb{E}\left(\mathbf{x}_j e_i \mid c^*_{ij} < 1,\ d^*_{ij} \geq 2\right) = \mathbf{0}, \tag{5}$$

and the matrices $\mathcal{W}_{m,\beta} = [w_{m,\beta;i,j}]$ and $\mathcal{W}_{m,\delta} = [w_{m,\delta;i,j}]$ if equations (4) or (5) are satisfied for agents i and j in layer m, respectively, and $w_{m,\beta;i,j} = w_{m,\delta;i,j} = 0$ otherwise.

Proposition 2 imposes conditional restrictions between the observed and unobserved characteristics of individuals. Intuitively, conditions (4) and (5) state that there is no correlation between the observed characteristics $\mathbf{x}$ of individual j and the the unobserved characteristics $\mathbf{e}$ of individual i if the edge-type changes c^* and the path length d^c is large enough.

4.2 Estimation Strategy

The estimation strategy adopts a GMM framework with linear and quadratic moments. The choice of the functional form of the instruments and the estimation procedure is further explored by Kuersteiner and Prucha (2020). To address the endogeneity of the networks, we use the moment restrictions from Proposition 2 to set up matrices $\mathcal{W}_{m,\beta}$ and $\mathcal{W}_{m,\delta}$ as instruments for $\mathbf{W}_m$. Thus, the linear and quadratic instruments are

$$\mathbf{Z} = [\mathcal{W}_{1\beta}\widetilde{\mathbf{X}},\ \mathcal{W}_{2\beta}\widetilde{\mathbf{X}},\ \mathcal{W}_{1\delta}\widetilde{\mathbf{X}},\ \mathcal{W}_{2\delta}\widetilde{\mathbf{X}},\ \mathbf{X},\ \mathbf{B}]$$
$$\mathbf{A}_r = \mathcal{W}_{m,\lambda}^{\top}\mathcal{W}_{m,\lambda} - \text{diag}\left(\mathcal{W}_{m,\lambda}^{\top}\mathcal{W}_{m,\lambda}\right),\ r = 1, \ldots, q$$

where the matrix $\mathbf{Z}$ contains p linear instruments and there are q quadratic instruments by the matrices $\mathbf{A}_r$. For instance, $\mathcal{W}_{1\beta}\widetilde{\mathbf{X}}$ represents the average characteristics of the distant individuals that support Proposition 2, and we use it as an instrument for $\mathbf{W}_1\mathbf{y}$ which is the average risky behavior of her friends. Using these $p+q$ instruments, we write the moment conditions

$$\overline{\mathbf{m}}_l(\theta) = \frac{1}{\sqrt{n}}\left[\mathbf{Z}^{\top}\mathbf{e}(\theta)\right], \qquad \overline{\mathbf{m}}_q(\theta) = \frac{1}{\sqrt{n}}\begin{bmatrix} \mathbf{e}(\theta)^{\top}\mathbf{A}_1\mathbf{e}(\theta) \\ \vdots \\ \mathbf{e}(\theta)^{\top}\mathbf{A}_q\mathbf{e}(\theta) \end{bmatrix}.$$

We concatenate the linear and quadratic instruments as $\overline{\mathbf{m}}_n(\theta) = \left[\overline{\mathbf{m}}_l(\theta)^{\top},\ \overline{\mathbf{m}}_q(\theta)^{\top}\right]^{\top}$, where we stack the regression parameters $\theta = [\beta_1, \beta_2, \delta_1, \delta_2, \gamma_1, \gamma_2]^{\top}$. Thus, the GMM estimator is defined as

$$\hat{\theta} = \arg\min_{\theta\in\Theta}\ n^{-1}\overline{\mathbf{m}}_n(\theta)^{\top}\widehat{\mathbf{\Omega}}^{-1}\ \overline{\mathbf{m}}_n(\theta),$$

where $\widehat{\mathbf{\Omega}}$ is a moment weighting matrix.

Following Kuersteiner and Prucha (2020), we employ the efficient GMM estimator using as weighting matrix $\widehat{\mathbf{\Omega}} = (\mathbf{V}_l + 2\mathbf{V}_q)$, where $\mathbf{V}_l = n^{-1}\mathbf{Z}^{\top}\mathbf{Z}$ and $\mathbf{V}_q = n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^{\top}a_{ij}$ are extended with zeros to fit the dimensions. Finally, we obtain the variance-covariance matrix $\widehat{\mathbf{\Sigma}} = (\mathbf{G}^{\top}\widehat{\mathbf{\Omega}}^{-1}\mathbf{G})^{-1}$ using the Jacobian matrix $\mathbf{G} = \frac{\partial\overline{\mathbf{m}}_n(\theta)}{\partial\theta}$ of the vector of moments.

4.3 Social Effects

Estimation and inference of the social marginal effects requires identifying and estimating the parameters β correctly. The matrix of estimated social marginal effects, denoted as $\mathbf{S}(\hat{\beta})$, is defined as $(1-\hat{\beta}_1-\hat{\beta}_2)(\mathbf{I}-\hat{\beta}_1\mathbf{W}_1-\hat{\beta}_2\mathbf{W}_2)^{-1}$ when using the estimates $\hat{\beta}$. Here, the element $\hat{s}ij$ of this matrix represents $\partial y_i/\partial x_{j,k}$. These effects can amplify or diminish the impacts of any changes in the set of covariates, such as the introduction of a new policy. Therefore, any potential treatment effect would be distorted by the spillovers generated by these social connections.

Our focus is on the aggregate social effects for each individual j, which we define as $\hat{s}_j = \sum_{j\neq i}\hat{s}_{ij}$. We interpret this as a measure of the influence that individuals have over the sample. Hence, individuals with high $\hat{s}_j$ are labeled as *influencers*. However, note that this characterization of influence is induced by the peer pressure mechanism governing risky behaviors in adolescents. It is worth mentioning that this definition of individual social effects excludes the own marginal effect s_{ii}. Since the adjacency matrices $\mathbf{W}_1$ and $\mathbf{W}_2$ are row-normalized, including s_{ii} in the aggregate social effect would sum to 1 for each individual.

We can also re-scale these aggregate social effects into averages. Since the individual social effects depend only on the parameters $\hat{\beta}_1$ and $\hat{\beta}_2$, we can use the delta method to obtain standard errors for the average social effects $\hat{s}_j/n$. The details of the delta method application are provided in Appendix C. To apply the delta method, we rewrite the matrix $\mathbf{S}(\hat{\beta})$ in terms of the infinite sum of its elements. After bounding the expression for $\mathbf{S}(\hat{\beta})$, we apply the delta method to the vector of average social effects $\bar{\mathbf{s}} = n^{-1}[\mathbf{S}(\hat{\beta}) - \operatorname{diag}\mathbf{S}(\hat{\beta})]^\top \cdot \iota$, where ι is a conformable vector of ones.

5 Results and Discussion

5.1 Friends and Classmates Effects

In Section 4, we employ an estimation strategy called GMM-ML, which uses multilayer networks as instruments based on the distance in the network space. In addition, we compare our results with those obtained using the GMM estimator assuming exogeneity of friends and classmates interactions, which requires the use of the networks $\mathbf{W}_1^2$ and $\mathbf{W}_2^2$ as instruments. We also present OLS estimates for comparison purposes. Appendix A includes robustness checks using other outcome variable definitions.

Table 2 displays our main findings on friends' and classmates' peer effects. The full set of results, including controls, can be found in Table A8. Our results reveal significant and precise peer effects for both friends and classmates. Specifically, a one-day increase in the average smoking behavior of an individual's friends is associated with a 0.34 increase in her own smoking behavior. Furthermore, we find substantial effects for classmates that are consistent with other estimates of this type of social interaction. Even without including addiction effects as in Arduini et al. (2019), we obtain very similar classmates' effects, with estimates around 0.18, using our new set of instruments. Notably, the size of our classmates' effects doubles when using our instruments compared to when endogenous networks are employed, increasing from 0.22 to 0.48.

Comparing cigarette smoking with marijuana consumption, we observe that the friends' effect almost doubles, while the classmates' effect remains roughly the same. However, the estimation results using endogenous networks do not yield any significant findings. Our estimates for friends' effects appear to be higher than those reported in Card and Giuliano (2013), who find effects ranging from 0.32 to 0.46 and 0.10 to 0.25 for intermediate and high levels of activity, respectively. Nevertheless, our estimates fall within the same range

Table 1. Peer Effects on Cigarette Smoking

	OLS		2SLS		GMM	
Peer Effects						
Friends Use	0.396***	(0.021)	-0.009	(1.593)	0.469***	(0.021)
Classmates Use	0.178	(0.151)	0.191	(4.162)	0.351***	(0.003)
Contextual Effects						
Friends Risk	-1.888*	(0.992)	-2.680	(19.68)	-1.179***	(0.258)
Classmates Risk	2.717	(2.708)	11.29	(77.82)	2.177***	(0.818)
Additional Controls	Yes		Yes		Yes	
Instruments			Smoker Parent		Smoker Parent	

Note: This table reports the peer effect estimates (β_m) for (2). The complete set of results with controls is located in table A8. OLS estimation uses clustered standard errors at the school level. GMM estimation uses the endogenous networks $\mathbf{W}_m^2$ to calculate peer effects while the GMM-ML estimation incorporates the networks of distant individuals $\mathcal{W}_{m,\lambda}$. For both GMM estimations, we use the efficient variance-covariance matrix for the standard errors.

as their structural estimates when considering the collective effects of friends' marijuana consumption (intermediate and high).

Table A8 in the Appendix also shows the direct effects of the two outcomes of interest. In line with prior research, we find negative effects of cigarette smoking among female and black students in the Add Health data. Additionally, we find strong evidence that risk-averse and less future-oriented adolescents engage more in both risky behaviors. Regarding parents' characteristics, having a smoker parent or a parent with a college degree increases cigarette smoking and marijuana use. However, the effects of having two parents in the household differ, with adolescents engaging in less cigarette smoking but more marijuana use when living with both parents.

5.2 Identifying Influencers

Using the peer effects estimates obtained previously, we can now calculate social effects for the sample of adolescents. Table 3 presents summary statistics for the individual social effects $\hat{s}_j$. The calculated social effects are expected to fall between 0 and 1 since the matrices of social interactions are row-normalized. On average, we find that the social effects of cigarette and marijuana consumption are approximately 0.26, which represents the aggregate effect an adolescent has on their peers when there is a change in her own characteristics. Thus, the social effects measure the spillover effect due to the presence of friends and classmate

Table 2. Peer Effects on Marijuana Use

	OLS		2SLS		GMM	
Peer Effects						
Friends Use	0.152*	(0.080)	0.227	(2.495)	0.226***	(0.023)
Classmates Use	-0.067	(0.231)	0.005	(2.421)	0.281***	(0.012)
Contextual Effects						
Friends Risk	0.133	(0.144)	-2.041	(11.82)	0.157	(0.168)
Classmates Risk	-0.786	(0.795)	4.241	(12.98)	1.172	(0.976)
Additional Controls	Yes		Yes		Yes	
Instruments			Smoker Parent		Smoker Parent	

Note: This table reports the peer effect estimates (β_m) for (2). The complete set of results with controls is located in table A8. OLS estimation uses clustered standard errors at the school level. GMM estimation uses the endogenous networks $\mathbf{W}_m^2$ to calculate peer effects while the GMM-ML estimation incorporates the networks of distant individuals $\mathcal{W}_{m,\lambda}$. For both GMM estimations, we use the efficient variance-covariance matrix for the standard errors.

Table 3. Adolescents Social Effects on Risky Behaviors

	mean	std	min	25%	50%	75%	max
Cigarette Influence	0.3	0.18	0.0	0.15	0.2	0.45	0.76
Marijuana Influence	0.31	0.11	0.0	0.22	0.24	0.42	0.49

Note: This table reports summary statistics of the calculated aggregate social effects for each adolescent. We obtain individual social effects by summing over the columns of the matrix of social marginal effects $\mathbf{S}(\hat{\beta})$.

connections.

There is significant heterogeneity in the aggregate social effect that individuals possess. For cigarette smoking, the interquartile range of social effects is between 0.15 to 0.43, while for marijuana use, it is between 0.09 to 0.47. Although both behaviors have similar summary statistics for social effects, the top percentile of the distribution reveals a higher influence on marijuana use than on cigarette smoking. This discrepancy may be due to the stronger effect of friends on marijuana use compared to cigarette smoking.

These calculated social effects enable us to explore the characteristics of a specific group of individuals. For example, if policymakers have the ability to distinguish whom to treat, examining the subpopulation of “influencers” (individuals with high social effects) could be of great interest. Figure 1 shows the empirical distribution of individual social effects, distinguished by sex. We can identify two groups of adolescents from the histogram, with the group on the right being the influencers. We also separate the two groups by running a k-means community detection algorithm. For both cigarette smoking and marijuana use, there appears to be a slightly higher proportion of females among influencers.

Figure 1. Histogram of Adolescents Social Effect by Sex

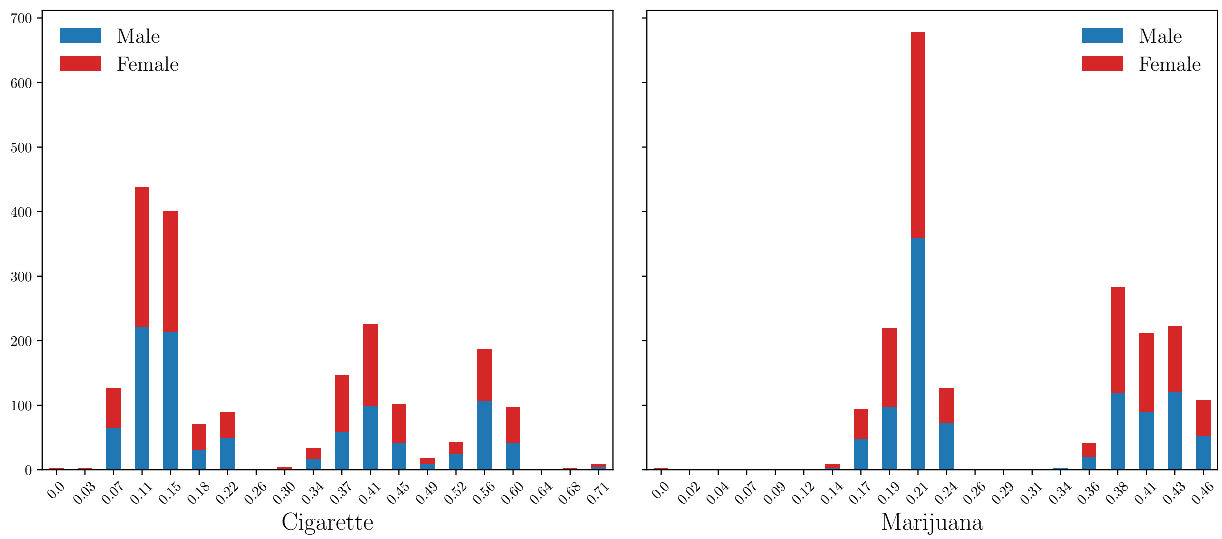
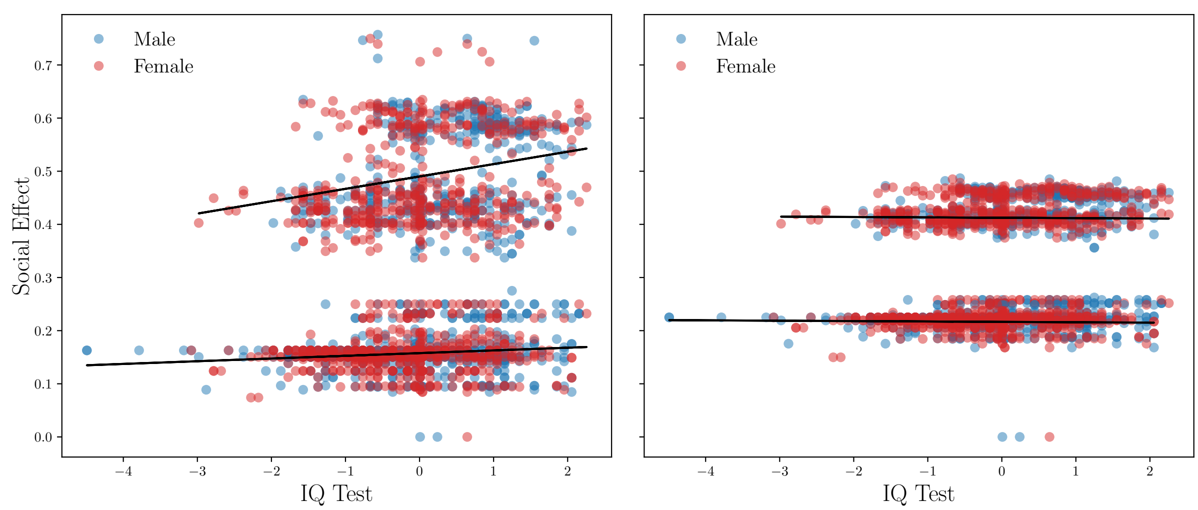


Figure 2. Adolescents Social Effect and PVT Test Scores (normalized)



We can also examine the relationship between PVT test scores and social effects. Figure 2 displays the social effects for each individual across their test scores. The figure suggests that there is no significant difference between high and low-influence high school students. However, we observe different slopes for the subgroups in both outcomes. Additionally, using the formula we provide for standard errors on average social effects, we can establish statistically significant differences among groups of individuals.

Table 4. Difference between High and Low-Influence Adolescents

	Difference
Risk Aversion	-0.03
Future Orientation	-0.03
Physical Activity	-0.0
IQ Test	0.29***
Female	0.05**
Age	-0.08
Black	-0.1***
Asian	-0.0
Other Race	-0.04***
College Parent	0.02
Two Parents	0.06***
Smoker Parent	0.03

Note: This table reports differences in means for high-influencing individuals vs low-influencing individuals. The p-value is calculated using a t-test for two independent samples with different variances.

Lastly, we can assess the differences in observed characteristics between the two groups of adolescents. There is a statistically significant difference in PVT test scores, with the influencers scoring higher than the non-influencers. However, the two groups are not significantly different in the amount of physical activity they practice, their future orientation, or their level of risk aversion. With regards to demographics, they are only significantly different by age and the proportion of blacks. No significant differences were found in their parental characteristics. Finally, there is a difference in the proportion of students with prevention programs in their schools.

As a robustness check, we investigate how the estimated social effects change when using only one type of social connection. Figures A1 and A2 compare the estimated social effects when only one of the social networks is included. We find that both types of social interactions contribute to the variation in the estimated individual social effect, but the variation from friendships is higher than that from classmates. This could be due to the fact that the estimands for the effects of friends are larger than for classmates and the highly clustered nature of the classmates network, which creates a lot of homogeneity in the variation of individual social effects. These results demonstrate the robustness of our findings and reinforce the importance of considering multiple types of social connections when estimating peer effects.

6 Conclusion

This paper investigates how to determine the individuals that have high influence over the network under the assumption that there is peer pressure in preferences. We study peer pressure in the context of high school students and risky behaviors, where we consider that these assumptions are plausible. The empirical strategy uses friends and classmates creating instruments across networks to identify peer effects. We find a group of high and-low influence individuals and we show their characteristics by sex and ability test scores.

It would be valuable to compare the results of targeting influential individuals to the effects of a randomized intervention to evaluate the model's ability to describe reality accurately. Additionally, a crucial question for future research is how to design the optimal policy. A policymaker aiming to maximize welfare would need to decide whether the goal is to minimize risky behaviors at the intensive margin, the extensive margin, or both.

References

The numbers at the end of every reference link to the pages citing the reference.

Arduini, Tiziano, Alberto Bisin, Onur Özgür, and Eleonora Patacchini, "Dynamic Social Interactions and Health Risk Behavior," Working Paper, National Bureau of Economic Research sep 2019. 2, 6, 8, 11

Blume, Lawrence E., William A. Brock, Steven N. Durlauf, and Rajshri Jayaraman, "Linear Social Interactions Models," *Journal of Political Economy*, apr 2015, *123* (2), 444–496. 3, 27

Boucher, Vincent and Ismael Mourifié, "My friend far, far away: a random field approach to exponential random graph models," *The Econometrics Journal*, oct 2017, *20* (3), S14–S46. 2

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin, "Identification of peer effects through social networks," *Journal of Econometrics*, may 2009, *150* (1), 41–55. 2

Calvó-Armengol, Antoni, Eleonora Patacchini, and Yves Zenou, "Peer Effects and Social Networks in Education," *Review of Economic Studies*, oct 2009, *76* (4), 1239–1267. 2, 5

Card, David and Laura Giuliano, "Peer Effects and Multiple Equilibria in the Risky Behavior of Friends," *The Review of Economics and Statistics*, oct 2013, *95* (4), 1130–1149. 2, 11

Cawley, John and Christopher J. Ruhm, "The Economics of Risky Health Behaviors," 2011, pp. 95–199. 1

Comola, Margherita and Silvia Prina, "Treatment Effect Accounting for Network Changes," *The Review of Economics and Statistics*, jul 2021, *103* (3), 597–604. 2

Estrada, Juan, "Causal Inference in Multilayer Social Networks," 2021. 2, 8, 9, 27, 29

Goldsmith-Pinkham, Paul and Guido W. Imbens, "Social Networks and the Identification of Peer Effects," *Journal of Business Economic Statistics*, jul 2013, *31* (3), 253–264. 2

Hsieh, Chih-Sheng and Xu Lin, "Social interactions and social preferences in social networks," *Journal of Applied Econometrics*, dec 2020, *36* (2), 165–189. 2

Kojevnikov, Denis, Vadim Marmer, and Kyungchul Song, "Limit theorems for network dependent random variables," *Journal of Econometrics*, jun 2021, *222* (2), 882–908. 29

Kuersteiner, Guido M. and Ingmar R. Prucha, "Dynamic Spatial Panel Models: Networks, Common Shocks, and Sequential Exogeneity," *Econometrica*, 2020, *88* (5), 2109–2146. 2, 8, 9, 10, 29, 30

Lin, Xu, "Identifying Peer Effects in Student Academic Achievement by Spatial Autoregressive Models with Group Unobservables," *Journal of Labor Economics*, oct 2010, *28* (4), 825–860. 2

Liu, Xiaodong, Eleonora Patacchini, and Yves Zenou, "Endogenous peer effects: local aggregate or local average?," *Journal of Economic Behavior Organization*, jul 2014, *103*, 39–59. 2

Manski, Charles F., "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, jul 1993, *60* (3), 531. 2

Manta, Alexandra, Anson T.Y. Ho, Kim P. Huynh, and David T. Jacho-Chávez, "Estimating social effects in a multilayered Linear-in-Means model with network data," *Statistics & Probability Letters*, apr 2022, *183*, 109331. 2

Mele, Angelo, "A Structural Model of Homophily and Clustering in Social Networks," *Journal of Business Economic Statistics*, jun 2021, *40* (3), 1377–1389. 2

Patacchini, E. and Y. Zenou, "Juvenile Delinquency and Conformism," *Journal of Law, Economics, and Organization*, dec 2009, *28* (1), 1–31. 2

Solá, Luis, Miguel Romance, Regino Criado, Julio Flores, Alejandro García del Amo, and Stefano Boccaletti, "Eigenvector centrality of nodes in multiplex networks," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, sep 2013, *23* (3), 033131. 5

Udry, J. Richard, Rose Maria Li, and Janet Hendrickson-Smith, "Health and Behavior Risks of Adolescents with Mixed-Race Identity," *American Journal of Public Health*, nov 2003, *93* (11), 1865–1870. 7

Ushchev, Philip and Yves Zenou, "Social norms in networks," *Journal of Economic Theory*, jan 2020, *185*, 104969. 1, 3

A Appendix Tables and Figures

Appendix Table A1. Descriptive Statistics of Networks

Network	Friends	Classmates
Edges	735	182167
Density	0.0004	0.0913
Average Cluster	1.5	32.2
Shortest Path	12.5	1.0

Appendix Table A2. Descriptive Statistics of the Sample

	Mean	SD
Cigarette Use	4.25	9.44
Friends Cigarette	1.95	6.15
Classmates Cigarette	4.25	3.42
Marijuana Use	1.61	7.22
Friends Marijuana	0.72	4.68
Classmates Marijuana	1.61	1.29
Friends Risk	0.17	0.34
Classmates Risk	0.41	0.1
Risk Aversion	0.41	0.49
Future Orientation	0.49	0.5
Physical Activity	0.54	0.5
IQ Test	0.01	0.97
College Parent	0.65	0.48
Two Parents	0.72	0.45
Smoker Parent	0.64	0.48

Appendix Table A3. Friends and Classmates Effects with Controls

	OLS		2SLS		GMM	
Peer Effects						
Friends Effect	0.396***	(0.021)	-0.009	(1.593)	0.469***	(0.021)
Classmates Effect	0.178	(0.151)	0.191	(4.162)	0.351***	(0.003)
Contextual Effects						
Friends Risk	-1.888*	(0.992)	-2.680	(19.68)	-1.179***	(0.258)
Classmates Risk	2.717	(2.708)	11.29	(77.82)	2.177***	(0.818)
Additional Controls						
Risk Aversion	-1.514***	(0.488)	-1.523	(1.247)	-1.684***	(0.047)
Future Orientation	0.397*	(0.206)	0.284	(0.625)	0.029	(0.045)
Physical Activity	-0.472***	(0.154)	-0.439*	(0.245)	-0.525***	(0.045)
IQ Test	0.023	(0.294)	-0.063	(1.142)	-0.210***	(0.028)
Female	0.156	(1.086)	0.229	(1.830)	0.377***	(0.049)
Age	0.314	(0.231)	0.420	(1.955)	0.207***	(0.021)
Black	-1.828***	(0.326)	-2.077	(1.952)	0.136	(0.089)
Asian	0.284	(0.295)	0.179	(0.754)	0.406***	(0.085)
Other Race	1.000***	(0.348)	0.707	(1.134)	3.162***	(0.119)
College Parent	-0.315	(0.399)	-0.273	(0.430)	-1.477***	(0.051)
Two Parents	-1.062**	(0.418)	-1.342	(2.115)	-1.747***	(0.057)
Smoker Parent	1.827***	(0.380)	1.739	(1.194)	1.324***	(0.049)
Fixed Effects	School		School		School	
Instruments			Smoker Parent		Smoker Parent	

Note: This table reports the estimates for (2). Each estimation also includes contextual effects for PVT test score, physical activity, future orientation, and risk aversion; and covariates for asian, other race, drug/tobacco prevention program, dummies for missing PVT test and prevention program, and a constant. OLS estimation uses clustered standard errors at the school level. GMM estimation uses the endogenous networks $\mathbf{W}_m^2$ to calculate peer effects while the GMM-ML estimation incorporates the networks of distant individuals $\mathcal{W}_{m,\lambda}$. For both GMM estimations we use the efficient variance-covariance matrix for the standard errors.

Appendix Table A4. Friends and Classmates Effects with Controls

	OLS		2SLS		GMM	
Peer Effects						
Friends Effect	0.152*	(0.080)	0.227	(2.495)	0.226***	(0.023)
Classmates Effect	-0.067	(0.231)	0.005	(2.421)	0.281***	(0.012)
Contextual Effects						
Friends Risk	0.133	(0.144)	-2.041	(11.82)	0.157	(0.168)
Classmates Risk	-0.786	(0.795)	4.241	(12.98)	1.172	(0.976)
Additional Controls						
Risk Aversion	-1.059***	(0.262)	-1.068***	(0.368)	-1.132***	(0.047)
Future Orientation	1.000***	(0.227)	0.974**	(0.469)	0.791***	(0.045)
Physical Activity	-0.025	(0.499)	-0.017	(0.467)	0.352***	(0.045)
IQ Test	-0.100	(0.102)	-0.061	(0.083)	0.259***	(0.026)
Female	-1.209***	(0.239)	-1.052	(1.376)	-0.973***	(0.048)
Age	0.012	(0.159)	-0.009	(0.187)	-0.059***	(0.022)
Black	-0.763***	(0.113)	-0.739*	(0.415)	0.253***	(0.086)
Asian	-1.913***	(0.213)	-1.801	(2.163)	-1.903***	(0.087)
Other Race	-0.301	(0.246)	-0.347	(1.387)	1.001***	(0.122)
College Parent	0.372	(0.268)	0.410	(0.816)	0.761***	(0.052)
Two Parents	-0.563**	(0.265)	-0.458	(1.113)	-0.730***	(0.054)
Smoker Parent	0.595***	(0.184)	0.473	(0.818)	0.409***	(0.049)
Fixed Effects	School		School		School	
Instruments			Smoker Parent		Smoker Parent	

Note: This table reports the estimates for (2). Each estimation also includes contextual effects for PVT test score, physical activity, future orientation, and risk aversion; and covariates for asian, other race, drug/tobacco prevention program, dummies for missing PVT test and prevention program, and a constant. OLS estimation uses clustered standard errors at the school level. GMM estimation uses the endogenous networks $\mathbf{W}_m^2$ to calculate peer effects while the GMM-ML estimation incorporates the networks of distant individuals $\mathcal{W}_{m,\lambda}$. For both GMM estimations we use the efficient variance-covariance matrix for the standard errors.

Appendix Table A5. Friends and Classmates Effects with Controls

	OLS		2SLS		GMM	
Peer Effects						
Friends Effect	0.420***	(0.019)	0.171	(0.279)	0.480***	(0.021)
Classmates Effect	0.132	(0.161)	0.547***	(0.205)	0.316***	(0.008)
Contextual Effects						
Friends Risk	-0.777	(0.957)	-1.279	(4.686)	-0.510	(0.689)
Friends Future	-1.935**	(0.965)	-1.819	(4.632)	-0.511	(0.363)
Friends Physical	-0.991	(0.653)	-0.542	(2.829)	4.050***	(0.290)
Classmates Risk	2.162	(2.851)	1.642	(3.106)	-1.773***	(0.445)
Classmates Future	3.283*	(1.889)	0.598	(2.166)	1.639***	(0.391)
Classmates Physical	-2.385	(2.024)	-0.696	(1.787)	-1.801***	(0.385)
Additional Controls						
Risk Aversion	-1.493***	(0.500)	-1.598***	(0.480)	-1.647***	(0.050)
Future Orientation	0.425**	(0.199)	0.307	(0.280)	0.067	(0.047)
Physical Activity	-0.471***	(0.171)	-0.408*	(0.216)	-0.409***	(0.048)
IQ Test	0.037	(0.311)	-0.015	(0.367)	-0.157***	(0.027)
Female	0.184	(1.097)	0.189	(1.061)	0.418***	(0.051)
Age	0.181	(0.287)	0.119	(0.199)	0.065***	(0.023)
Black	-2.062***	(0.290)	-2.258***	(0.396)	-0.440***	(0.097)
Asian	0.387	(0.287)	0.323	(0.447)	0.528***	(0.085)
Other Race	0.814**	(0.334)	0.762**	(0.322)	2.625***	(0.120)
College Parent	-0.266	(0.380)	-0.240	(0.364)	-1.431***	(0.051)
Two Parents	-1.025**	(0.409)	-1.216**	(0.495)	-1.748***	(0.056)
Smoker Parent	1.844***	(0.366)	1.805***	(0.417)	1.414***	(0.050)
Fixed Effects	School		School		School	
Instruments			Smoker Parent Risk Aversion Future Orientation Physical Activity		Smoker Parent Risk Aversion Future Orientation Physical Activity	

Note: This table reports the estimates for (2). Each estimation also includes contextual effects for PVT test score, physical activity, future orientation, and risk aversion; and covariates for asian, other race, drug/tobacco prevention program, dummies for missing PVT test and prevention program, and a constant. OLS estimation uses clustered standard errors at the school level. GMM estimation uses the endogenous networks $\mathbf{W}_m^2$ to calculate peer effects while the GMM-ML estimation incorporates the networks of distant individuals $\mathcal{W}_{m,\lambda}$. For both GMM estimations we use the efficient variance-covariance matrix for the standard errors.

Appendix Table A6. Friends and Classmates Effects with Controls

	OLS		2SLS		GMM	
Peer Effects						
Friends Effect	0.163*	(0.084)	0.639***	(0.147)	0.231***	(0.022)
Classmates Effect	-0.092	(0.255)	0.248	(0.273)	0.252***	(0.013)
Contextual Effects						
Friends Risk	0.543**	(0.251)	-0.500	(4.080)	0.390	(0.879)
Friends Future	-0.928**	(0.383)	-1.816	(1.925)	0.189	(0.564)
Friends Physical	-0.100	(0.385)	-0.136	(2.935)	3.275***	(0.278)
Classmates Risk	-0.999	(0.881)	0.501	(1.706)	-0.929**	(0.422)
Classmates Future	1.655	(1.134)	1.101	(0.818)	1.496***	(0.484)
Classmates Physical	-1.620	(0.987)	-2.403	(1.494)	-2.450***	(0.381)
Additional Controls						
Risk Aversion	-1.059***	(0.262)	-1.098***	(0.243)	-1.139***	(0.052)
Future Orientation	1.014***	(0.224)	1.069***	(0.227)	0.857***	(0.047)
Physical Activity	-0.046	(0.499)	-0.036	(0.521)	0.333***	(0.050)
IQ Test	-0.105	(0.099)	-0.106	(0.126)	0.265***	(0.026)
Female	-1.203***	(0.251)	-0.961***	(0.280)	-0.976***	(0.055)
Age	-0.055	(0.199)	-0.131	(0.162)	-0.125***	(0.024)
Black	-0.854***	(0.116)	-0.900**	(0.415)	0.068	(0.105)
Asian	-1.870***	(0.215)	-1.444***	(0.083)	-1.817***	(0.086)
Other Race	-0.353	(0.217)	0.005	(0.136)	0.968***	(0.124)
College Parent	0.396	(0.284)	0.569**	(0.240)	0.767***	(0.052)
Two Parents	-0.557**	(0.269)	-0.314	(0.251)	-0.740***	(0.055)
Smoker Parent	0.603***	(0.166)	0.457**	(0.184)	0.427***	(0.053)
Fixed Effects	School		School		School	
Instruments			Smoker Parent Risk Aversion Future Orientation Physical Activity		Smoker Parent Risk Aversion Future Orientation Physical Activity	

Note: This table reports the estimates for (2). Each estimation also includes contextual effects for PVT test score, physical activity, future orientation, and risk aversion; and covariates for asian, other race, drug/tobacco prevention program, dummies for missing PVT test and prevention program, and a constant. OLS estimation uses clustered standard errors at the school level. GMM estimation uses the endogenous networks $\mathbf{W}_m^2$ to calculate peer effects while the GMM-ML estimation incorporates the networks of distant individuals $\mathcal{W}_{m,\lambda}$. For both GMM estimations we use the efficient variance-covariance matrix for the standard errors.

Appendix Table A7. Friends and Classmates Effects with Controls

	No FE		School FE		Cohort FE	
Peer Effects						
Friends Effect	0.472***	(0.014)	0.469***	(0.021)	0.474***	(0.020)
Classmates Effect	0.419***	(0.002)	0.331***	(0.004)	0.427***	(0.007)
Contextual Effects						
Friends Risk	-1.145***	(0.187)	-1.142***	(0.262)	-1.032***	(0.241)
Classmates Risk	2.418***	(0.311)	2.215**	(0.870)	2.521***	(0.478)
Additional Controls						
Risk Aversion	-1.616***	(0.046)	-1.711***	(0.047)	-1.599***	(0.046)
Future Orientation	0.068	(0.045)	-0.002	(0.045)	0.005	(0.045)
Physical Activity	-0.460***	(0.044)	-0.566***	(0.045)	-0.442***	(0.045)
IQ Test	-0.201***	(0.025)	-0.190***	(0.028)	-0.342***	(0.027)
Female	0.222***	(0.049)	0.380***	(0.049)	-0.138***	(0.050)
Age	0.144***	(0.007)	0.258***	(0.022)	-0.872***	(0.042)
Black	-0.580***	(0.077)	0.108	(0.089)	-0.766***	(0.077)
Asian	0.041	(0.078)	0.322***	(0.085)	0.027	(0.080)
Other Race	2.670***	(0.115)	3.107***	(0.119)	2.278***	(0.116)
College Parent	-1.541***	(0.051)	-1.406***	(0.051)	-1.541***	(0.051)
Two Parents	-1.904***	(0.058)	-1.700***	(0.058)	-1.979***	(0.057)
Smoker Parent	1.470***	(0.048)	1.309***	(0.049)	1.537***	(0.049)
Fixed Effects	School		School		School	
Instruments	Smoker Parent		Smoker Parent		Smoker Parent	

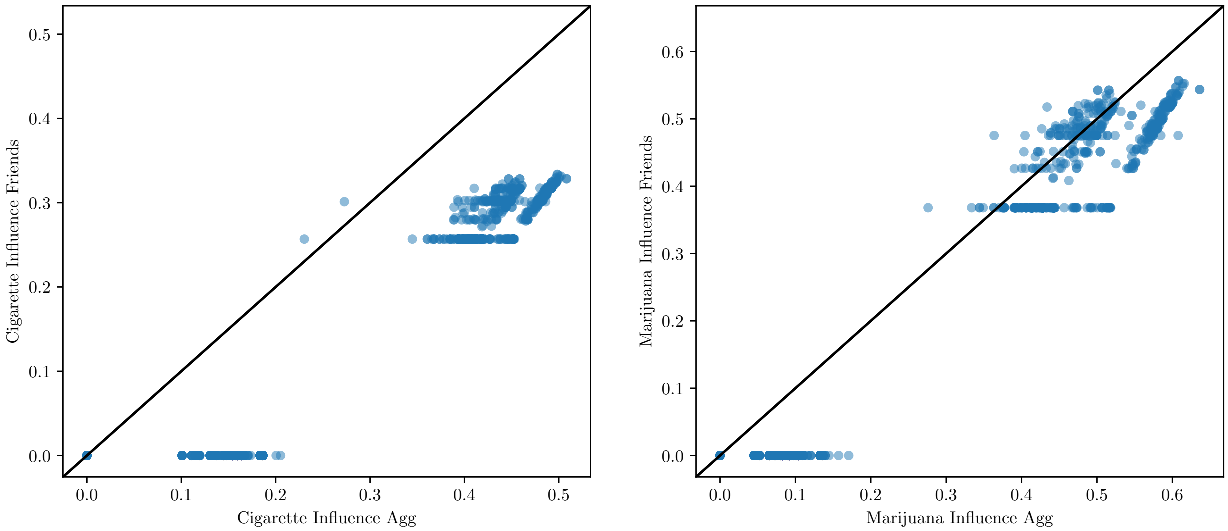
Note: This table reports the estimates for (2). Each estimation also includes contextual effects for PVT test score, physical activity, future orientation, and risk aversion; and covariates for asian, other race, drug/tobacco prevention program, dummies for missing PVT test and prevention program, and a constant. OLS estimation uses clustered standard errors at the school level. GMM estimation uses the endogenous networks $\mathbf{W}_m^2$ to calculate peer effects while the GMM-ML estimation incorporates the networks of distant individuals $\mathcal{W}_{m,\lambda}$. For both GMM estimations we use the efficient variance-covariance matrix for the standard errors.

Appendix Table A8. Friends and Classmates Effects with Controls

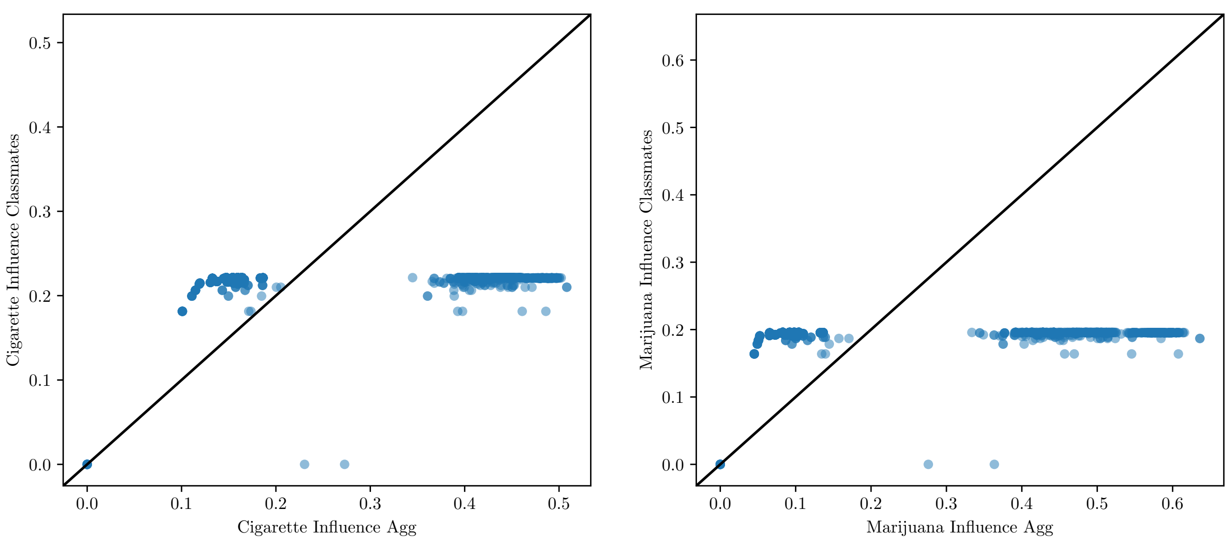
	No FE		School FE		Cohort FE	
Peer Effects						
Friends Effect	0.229***	(0.029)	0.227***	(0.023)	0.229***	(0.027)
Classmates Effect	0.335***	(0.012)	0.280***	(0.012)	0.321***	(0.022)
Contextual Effects						
Friends Risk	0.202	(0.174)	0.167	(0.168)	0.154	(0.182)
Classmates Risk	0.448	(0.291)	1.139	(0.983)	-0.754**	(0.341)
Additional Controls						
Risk Aversion	-1.090***	(0.046)	-1.132***	(0.048)	-1.169***	(0.046)
Future Orientation	0.766***	(0.045)	0.792***	(0.046)	0.805***	(0.046)
Physical Activity	0.441***	(0.044)	0.352***	(0.045)	0.420***	(0.045)
IQ Test	0.167***	(0.025)	0.260***	(0.026)	0.197***	(0.026)
Female	-0.942***	(0.048)	-0.973***	(0.048)	-0.967***	(0.050)
Age	0.075***	(0.007)	-0.059***	(0.022)	0.283***	(0.040)
Black	0.426***	(0.068)	0.251***	(0.086)	0.682***	(0.072)
Asian	-1.548***	(0.079)	-1.902***	(0.087)	-1.425***	(0.081)
Other Race	1.588***	(0.115)	0.997***	(0.122)	1.928***	(0.118)
College Parent	0.700***	(0.053)	0.763***	(0.052)	0.732***	(0.053)
Two Parents	-0.716***	(0.055)	-0.730***	(0.054)	-0.796***	(0.056)
Smoker Parent	0.427***	(0.049)	0.412***	(0.049)	0.362***	(0.049)
Fixed Effects	School		School		School	
Instruments	Smoker Parent		Smoker Parent		Smoker Parent	

Note: This table reports the estimates for (2). Each estimation also includes contextual effects for PVT test score, physical activity, future orientation, and risk aversion; and covariates for asian, other race, drug/tobacco prevention program, dummies for missing PVT test and prevention program, and a constant. OLS estimation uses clustered standard errors at the school level. GMM estimation uses the endogenous networks $\mathbf{W}_m^2$ to calculate peer effects while the GMM-ML estimation incorporates the networks of distant individuals $\mathcal{W}_{m,\lambda}$. For both GMM estimations we use the efficient variance-covariance matrix for the standard errors.

Appendix Figure A1. Social Effects of Friends vs Friends & Classmates Networks



Appendix Figure A2. Social Effects of Classmates vs Friends & Classmates Networks



B Social Interactions Model

In this section, we present a model of social interactions where preferences feature peer pressure, also called social conformity, with heterogeneous effects depending on the type of the social connection. Following Blume et al. (2015), we assume that individuals exhibit a quadratic utility function highlighting contextual effects, peer effects, and the cost of taking action. The setup of the game has two stages. First, there is a network formation process for $\{G_1, G_2\}$ networks, which are friendships and classmates. In the second stage, agents choose the optimal action of risky behavior y_i. We solve the game by backward induction. First, we solve for the optimal action, and then for the network formation process given the optimal actions.

The set $V = \{1, \ldots, N\}$ denotes the players of the game. They have individual characteristics (x_i, z_i). The characteristics x_i, such as physical activity and grades, are publicly observed. Instead, the characteristics z_i are private information about each individual that only they know about themselves. An example of this private information could be the family history on substance abuse. Therefore, we define an individual type as $(x, z_i) \in \mathbb{R}^{N+1}$. Therefore individual i chooses action $y_i \in \mathbb{R}$ to maximize

$$U_i\left(y_i,\ y_{-i}\right) = \left(\gamma^* x_i + z_i + \frac{1}{2}\sum_{m=1}^{2}\sum_{j=1}^{N} w_{ij,m}\ x_j\ \delta_m^*\right) y_i - \frac{1}{2}y_i^2 - \frac{1}{2}\sum_{m=1}^{2}\beta_m^*\left(y_i - \frac{1}{2}\sum_{j=1}^{N} w_{ij,m}\ y_j\right)^2,$$

and the best reply function[5] is

$$\mathbf{y} = \left[\left(1 + \sum_{m=1}^{M}\beta_m^*\right)\mathbf{I} - \sum_{m=1}^{M}\beta_m^*\mathbf{W}_m\right]^{-1}\left[\sum_{m=1}^{M}\delta_m^*\mathbf{W}_m\mathbf{x} + \gamma^*\mathbf{x} + \mu(\mathbf{x},\mathbf{z}) + \mathbf{z}\right],$$

where $\mathbf{S}(\beta^*) = \left[\left(1 + \sum_{m=1}^{M}\beta_m^*\right)\mathbf{I} - \sum_{m=1}^{M}\beta_m^*\mathbf{W}_m\right]^{-1}$ is the matrix of marginal effects and $\mu_i(\mathbf{x},\mathbf{z})$ depends only on $\mathbf{x}$ and z_i.

In the first stage, individual i choose to connect with j following a sequential multilayer network formation process. To simplify the process, we will assume that individuals are myopic and only consider her connections in layer $m-1$ when forming a relationship in m. In our case, the classrooms network G_2 is formed exogenously first. Then, adolescents choose whom to be friends with on the friendship network G_1, based on the connections from

[5]Details of the proof can be found in Blume et al. (2015) and Estrada (2021) but for the case when preferences display strategic complementarities.

classmates.

Individuals i and j choose to connect on the friends layer ($m = 1$) to maximize

$$U_{i,1}(\mathbf{W}_2) = \sum_{j=1}^{N} \alpha_x |x_i - x_j| + \alpha_z |z_i - \mathbb{E}[z_j | x_i, x_j, \rho]| + \alpha_1 \mathbf{1}\{i \in C_k\}\mathbf{1}\{j \in C_l\} \sum_j w_{ij,2} + v_{ij,1} \ ,$$

where α_x represent a homophily parameter over the public characteristics x, α_z for the private characteristics z, and α_1 describe the tendency of individuals to bridge structural holes over the disjoint clusters C_k and C_l. The more classmates individual j has, the more utility individual i earns if becoming friends. This simplified two-stage game of social interactions provides the microfoundation for peer pressure on risky behaviors.

C Identification and Estimation Details

C.1 Identification

The following assumptions are based on Estrada (2021) adapted to the identification results in Kuersteiner and Prucha (2020).

Assumption 1 (Invertibility). *Define the matrix* $\mathbf{S}(\beta^*) = \left[\left(1 + \sum_{m=1}^{M} \beta_m^*\right)\mathbf{I} - \sum_{m=1}^{M} \beta_m^* \mathbf{W}_m\right]^{-1}$ *and assume that* $\mathbf{S}(\beta^*)$ *exist for* $\beta^* \in \Theta_{\beta^*}$*, where* $\Theta_{\beta^*} = \Theta_{\beta_1^*} \times \cdots \times \Theta_{\beta_M^*}$ *is a compact set.*

Assumption 1 is used in equation 1 to put a restriction on the parameter space of β^*. This assumption helps for the identification of the structural parameters θ.

Assumption 2 (Moment Restriction). $\mathbb{E}\left[\mathbf{x}_i e_i\right] = 0$ *for all* $i = 1, \ldots, n$ *and* $s \leq t$*.*

Assumption 2 put a restriction between errors and covariates only for the same individual, in contrast to the classical assumption of $\mathbb{E}\left[\mathbf{x}\ e\right] = 0$. Between individuals i and j we will impose a weak neighborhood dependence assumption instead.

With the multilayer measure of distance $d^{\mathcal{M}}(i,j)$, we define the collection of pairs $\mathcal{P}(a,b,s) = \{(A,B) : A, B \in \mathcal{N}_n, |A| = a, |B| = b, d_{n,t}(A,B) \geq d\}$, where $d_{n,t}(A,B) = \min\{d^{\mathcal{M}}(i,j) : i \in A, j \in B\}$

To establish Kojevnikov et al. (2021) definition of weak dependence, define the $(k \times 1)$ vector $\mathbf{r}_i = [\mathbf{x}_i^\top, e_i] \in \mathbb{R}^{k+1}$, and $\mathbf{r}_A = (\mathbf{r}_i : i \in A)$. Let $\mathcal{L}_d, a$ denote the set of bounded real Lipschitz functions mapping $\mathbb{R}^{d \times a} \to \mathbb{R}$.

Definition 1 ($\psi-$dependence). *A triangular array* $\{\mathbf{r}_i\}_{i=1}^n$ *is* ψ*-dependent if there exists (1) a sequence* $\{\phi_{n,s}\}_{s,n\in\mathbb{N}}$ *with* $\phi_{n,0} = 1$ *such that* $\sup_n \phi_{n,s} \to 0$ *as* $s \to \infty$ *and* $\exists S \leq n$ *such that if* $s > S$ *then* $\phi_{n,s} = 0$*; (2) collection of functionals* $\{\psi_{a,b}\}_{a,b\in\mathbb{N}}$ *with* $\psi_{a,b} : \mathcal{L}_{v,a} \times \mathcal{L}_{v,b} \to [0,\infty)$ *such that*

$$|\operatorname{Cov}\left(f(\mathbf{r}_{A,t}), g(\mathbf{r}_{B,t})\right)| \leq \psi_{a,b}(f,g)\ \phi_{n,s}$$

for all $A, B \in \mathcal{P}_{n,t}(a,b,s)$*, and* $f \in \mathcal{L}_{v,a}, g \in \mathcal{L}_{v,b}$*.*

Definition 1 uses the sequence ϕ_n as the dependence coefficients of $\{\mathbf{r}_i\}$. It also states that when two set of nodes are apart from at least a distance S then they are independent. With this definition of $\psi-$dependence we state our weak dependence assumption.

Assumption 3 (Weak Dependence). *For all networks* $\mathcal{M}$ *that occur with positive probability in* $\mathcal{F}$*, the conditional distribution* $\mathcal{F}(\mathcal{M}, \mathbf{X}, \mathbf{e})$ *is such that*

(i) $\{\mathbf{r}_i\}$ *is* $\psi-$*dependent with dependent coefficients* ϕ_n

(ii) For some constant $C > 0$, $\psi_{a,b}(f,g) \leq C \times ab\left(\|f\|_\infty + \text{Lip}(f)\right)\left(\|g\|_\infty + \text{Lip}(g)\right)$

Assumption 4 states necessary conditions for identification. Define $\eta_{m,i}$ and $\eta_{m,i,\lambda}$ as an indicator of whether individual i is non-isolated in the network $\mathbf{W}_m$ and $\mathcal{W}_{m,\lambda}$ respectively. In addition, let $\mathbf{D}$ be the matrix with the variables of the rhs of equation 2.

Assumption 4 (Relevance)**.** *Suppose*

(i) the event $\eta_{m,i} = 0$ and $\eta_{m,i,\lambda} = 0$ for all m, i, λ, happens with probability zero.

(ii) $\mathbf{Q}_{ZD} = \text{plim}\,\frac{1}{n}\mathbf{Z}^\top\mathbf{D} < \infty$ and $\mathbf{Q}_{ZX} = \text{plim}\,\frac{1}{n}\mathbf{Z}^\top\mathbf{X} < \infty$

Before showing the theorem for identification, define $\mathbf{S} = n^{-1}\left[\mathbf{y}^\top\mathbf{M}_z^\top\mathbf{A}_r\mathbf{M}_z\mathbf{y},\ \mathbf{W}\mathbf{y}^\top\mathbf{M}_z^\top\mathbf{A}_r\mathbf{M}_z\mathbf{W}\mathbf{y}\right]$ and $\mathbf{M}_z = \mathbf{I} - \mathbf{D}\left(\mathbf{D}^\top\mathbf{P}_z\mathbf{D}\right)^{-1}\mathbf{D}^\top\mathbf{P}_z$ with $\mathbf{P}_z = \mathbf{Z}\left(\mathbf{Z}^\top\mathbf{Z}\right)^{-1}\mathbf{Z}^\top$.

Theorem 1 (Identification)**.** *Let Assumptions 1, 2, 3, and 4 hold for some K_c and K_d such that $K_d \geq K_c + 1$. The parameters $\theta^0 = [\beta^0, \delta^0, \gamma^0]^\top$ are identifiable if*

(i) $\mathbf{Q}_{ZD}$ has full column rank then $\text{plim}\,\frac{1}{\sqrt{n}}\mathbf{m}_l(\theta) = 0$ (linear moment conditions) has a unique solution at $\theta = \theta^0$.

(ii) Only $\mathbf{Q}_{ZX}$ and $\mathbf{S}$ has full column rank then $\text{plim}\,\frac{1}{\sqrt{n}}\mathbf{m}(\theta) = 0$ (linear and quadratic moment conditions) has a unique solution at $\theta = \theta^0$.

The proof for theorem 1 follows from Kuersteiner and Prucha (2020) under Lemma EX1.

C.2 Average Social Effects

In this section, we detail the process to obtain standard errors for the individual social effects. Recall that the matrix of social effects is defined as $\mathbf{S}(\beta^*) = \left[\left(1 + \sum_{m=1}^M \beta_m^*\right)\mathbf{I} - \sum_{m=1}^M \beta_m^*\mathbf{W}_m\right]^{-1}$. However, we want to rewrite the matrix of social effects using the estimated coefficients as $\mathbf{S}(\hat{\beta}) = (1 - \hat{\beta}_1 - \hat{\beta}_2)(\mathbf{I} - \hat{\beta}_1\mathbf{W}_1 - \hat{\beta}_2\mathbf{W}_2)^{-1}$. We aim to obtain standard errors for the vector of average social effects $\bar{s}_j = n^{-1}\sum_{j\neq i}\hat{s}_{ij}$. To apply the delta method we need $\partial\bar{s}_j/\partial\beta$, which involves to take the derivate of the inverse of a sum of matrices. To overcome that, we express the matrix of social effects in terms of the infinite sum of the product of the different adjacency matrices, given by

$$\mathbf{S}(\hat{\beta}) = (1 - \hat{\beta}_1 - \hat{\beta}_2)\sum_{r=0}^{\infty}\left(\hat{\beta}_1\mathbf{W}_1 + \hat{\beta}_2\mathbf{W}_2\right)^r.$$

Therefore, we can calculate the variance-covariance matrix for the average social effects as

$$\mathbf{V}(\bar{\mathbf{s}}) = \frac{\partial\bar{s}}{\partial\beta}^\top\widehat{\boldsymbol{\Sigma}}_\beta\frac{\partial\bar{s}}{\partial\beta}$$

where

$$\frac{\partial}{\partial \beta_m}\mathbf{S}(\hat{\beta}) = -\sum_{r=0}^{\infty}\left(\hat{\beta}_1\mathbf{W}_1 + \hat{\beta}_2\mathbf{W}_2\right)^r + (1-\beta_1-\beta_2)\left[\sum_{r=1}^{\infty} r\mathbf{W}_m\left(\hat{\beta}_1\mathbf{W}_1 + \hat{\beta}_2\mathbf{W}_2\right)^{r-1}\right]$$