Dynamic Social Interactions and Health Risk Behavior*

Tiziano Arduini †
University of Bologna
tiziano.arduini@unibo.it

Alberto Bisin‡
New York University, NBER
alberto.bisin@nyu.edu

Onur Özgür§
Melbourne Business School
onur.ozgur@mbs.edu

Eleonora Patacchini¶
Cornell University
ep454@cornell.edu

September 14, 2019

*We are grateful to Steven Durlauf, Matt Jackson, David Levine for their extensive comments and suggestions at different stages. We also thank Larry Blume, Pierre-André Chiappori, Aureo De Paula, Marcel Fafchamps, Bryan Graham, Bob Lucas, Jen Overbeck, Tom Sargent, José Scheinkman, Giorgio Topa, and seminar participants at Yale, UPenn, NYU, Chicago, Cornell, Queens, University of Toronto, University of Queensland, Melbourne Business School, Université Laval, Australian National University, University of Western Australia, Workshop on Social Identity and Social Interactions in Economics in Quebec City, and the Australasian Economic Theory Workshop in Canberra. Özgür thanks the Economics Department at Sciences Po in Paris, where a big part of this research was done, for its hospitality. Özgür is also grateful for financial support to Melbourne Business School through MBS Competitive Grant schemes in 2013-2017, to "La Chaire du Canada en Économie des Politiques Sociales et des Ressources Humaines" at Université Laval, and to SSHRC, FQRSC, and IFM2.

†Department of Economics, University of Bologna, Piazza Scaravilli 2, Bologna, 40126 Italy.

‡Department of Economics, New York University, 19 West Fourth Street, New York, NY, 10012, USA; CIREQ, IZA, NBER.

§Melbourne Business School, 200 Leicester Street, Carlton, VIC 3053, Australia; CIREQ, CIRANO, CIRPÉE.

¶Cornell University, Ithaca, NY, USA; EIEF, CEPR and IZA.

Abstract

We study risky behavior of adolescents. Concentrating on smoking and alcohol use, we structurally estimate a dynamic social interaction model in the context of students' school networks included in the National Longitudinal Study of Adolescent Health (Add Health). The model allows for forward-looking behavior of agents, addiction effects, and social interactions in the form of preferences for conformity in the social network.

We find strong evidence for forward looking dynamics and addiction effects. We also find that social interactions in the estimated dynamic model are quantitatively large. A misspecified static model would fit data substantially worse, while producing a much smaller estimate of the social interaction effect. With the estimated dynamic model, a temporary shock to students' preferences in the 10th grade has effects on their behavior in grades 10, 11, 12, with estimated social multipliers 1.53, 1.03, and 0.76, respectively. The multiplier effect of a permanent shock is much larger, up to 3.7 in grade 12. Moreover, (semi-) elasticities of a permanent change in the availability of alcohol or cigarettes at home on child risky behavior implied by the dynamic equilibrium are 25%, 63%, and 79%, in grades 10, 11, 12, respectively.

Journal of Economic Literature Classification Numbers: C18, C33, C62, C63, C73, I12.
Keywords: Dynamic social interactions, conformity, addiction, identification, model validation, dynamic social multiplier, Heckman selection.

1 Introduction

Smoking and alcohol use are widespread among adolescents. According to the 2018 National Youth Tobacco Survey, more than 1 in 4 high school students and about 1 in 14 middle school students had used a tobacco product in the past 30 days.[1] Smoking and alcohol are a serious policy concern in that they involve severe risks in terms of health outcomes. The World Health Organization's Global Health Risks Report 2009 establishes that tobacco and alcohol account for, respectively, 10.7 and 6.7% of the global burden of disease and injury in high income countries, as measured in disability-adjusted life years (DALYs); see WHO (2009), Table 2. Furthermore, smoking and especially alcohol are responsible for a large amount of socio-economic costs, in terms of, e.g., poor academic performance (Renna, 2007), earnings and unemployment (Kenkel and Ribar, 1994; Mullahy and Sindelar, 1996; Terza, 2002), and criminal victimization (French and Maclean, 2006).

The empirical literature on risky health behavior in economics and in the social sciences emphasizes several fundamental aspects of smoking and alcohol use. First of all, they respond to dynamic incentives, such as, e.g., future price changes and anticipated future consumption, and have an addictive component (Becker, Grossman, and Murphy, 1991; Chaloupka, 1991; Waters and Sloan, 1995). Furthermore, they are social behaviors, in the sense that they depend on the behavior of relevant peers (Argys and Rees, 2008; Lundborg, 2006; Duncan et al., 2005; Balsa and Diaz, 2018).[2]

In this paper, we study the smoking and alcohol consumption decisions of adolescents. In accordance with the empirical literature, we account for the dynamic forward-looking aspect of the decision problem, allowing adolescents to consider the addictive characteristics of tobacco and alcohol in evaluating the consequences of their behavior. Furthermore, we embed the dynamic choice of adolescents regarding alcohol and smoking use in a school environment characterized by rich social interactions. The joint consideration of dynamic choice and social interactions highlights interesting novel dimensions of the choice problem, allowing students e.g., to anticipate a change in their social network after high-school, which in turn may affect the importance of peer effects over schooling age.

More specifically, we formulate and structurally estimate a dynamic social interactions model.

[1]The National Survey on Drug Use and Health reports that, in 2017, about 19% of underage people (ages 12-20) were current alcohol users, and about 11.9% of the underage were binge drinkers.

[2]See Cawley and Ruhm (2011) and Kenkel and Sindelar (2011) for extensive surveys.

Agents' preferences over choices at any time depend on their own previous choices, to capture habits and addiction. Agents interact in their social reference group, the social network, and display preference externalities: each individual's preferences depend on the current choices of the agents in his/her network, to capture preferences for conformity to the social reference group. This dynamic interaction structure induces each individual's choice to depend on previous choices and current preference shocks of all other individuals in his/her social network.[3] We bring the model to data in the context of students' school networks included in the Add Health. The data collected by this survey includes information about each student's health-risk behavior as well as his/her social network, repeatedly, in distinct school-years. We use the panel dimension of the data to structurally estimate our dynamic social interaction model. We estimate the system of linear policy rules describing the equilibrium. In turn, the equilibrium characterization of the dynamic game allows us to back out the structural preference parameters from our estimates of the policy rules.

There are well-known inferential problems in the study of social interactions.[4] In our context, three main issues arise due to: (i) the endogeneity of previous choices as explanatory variables for current choice, in the absence of any restrictions on the intertemporal correlation structure of errors, (ii) the existence of common shocks or common unobserved factors affecting all individuals' choices in a network, independently of social interactions, and (iii) the endogeneity of the network. All these issues translate into correlation between the regressors and the errors. For all three issues, we offer solutions that allow us to construct a consistent estimator in our environment by using the moment restrictions imposed by the dynamic equilibrium.

Our empirical analysis confirms the main thrust of our model regarding smoking and alcohol use in the adolescent population. The preference parameters driving the addiction effect and the social interaction effect are estimated both to be significantly different from 0. Furthermore, a significant forward looking component characterizes students' decision making: namely, the discount rate also is positive and significant. Indeed, we measure a relevant bias associated with estimating (i) a mis-specified *myopic* model (which allows for addiction but not for forward looking choice); as well as (ii) a mis-specified *static* model (with no addiction nor forward looking behavior). Notably, the static model produces a much smaller estimate of the social interaction

[3]Formally, the model is reduced to a dynamic game which, under our assumptions, we show has a unique Subgame Perfect Equilibrium. We characterize equilibrium behavior as a system of linear non-stationary Markovian policy rules, for each individual and in each time period.

[4]See e.g., Blume et al. (2015), Brock and Durlauf (2001b), and Manski (2008).

effect.[5] As for social interactions, we estimate sizable quantitative effects, as measured by the implied *social multiplier* and by the induced correlation between the students' choices across the network.[6] More precisely, the social multiplier, in our context, has a fundamental dynamic component: a temporary (one-period) shock to agents' preferences in the 10th grade has effects on their behaviors in grades 10, 11, 12, with multipliers 1.53, 1.03, and 0.76, respectively. The multiplier of a permanent shock in grade 10 is 1.55 in the same grade, 2.6 in grade 11, and 3.68 in grade 12. Finally, the same-period multiplier effect of a temporary shock in the different grades encodes the importance of the number of periods to the end of school in students' choice. In this respect, our estimates imply that the same-period social multiplier in grade 12 (the last year of high-school) is about 1.7, whereas it is 1.56 and 1.53, respectively, in grades 11 and 10. This is evidence that students anticipate a change in their social network after high-school and this affects the importance of peer effects over schooling age: as the time to the end of high-school increases, the students' policy functions weigh more heavily future shocks and the current shock has a smaller effect.

We also implement a validation exercise of our empirical strategy and our structural estimates. We use the structural estimates to make out-of-sample predictions of students' equilibrium behavior in networks that are not included in the estimation sample. More precisely, we use the structural parameters off of our estimation network sample consisting of grades 10, 11, and 12 to predict equilibrium behavior for students in the network sample consisting of grades 7, 8, and 9. We compare predictions and actual choice data to validate the model and demonstrate that the model performs very well in the validation exercise.

The remainder of the paper is organized as follows. Section 1.1 places the paper in context vis-à-vis the related literature. Section 2 presents the structural model of dynamic health risk choices with network interactions. Section 3 describes the Add Health data set we use to estimate the model. Section 4 describes the empirical implementation of the model. Section 5 presents our main empirical findings, model validation outcomes, as well as our results regarding the social multiplier and cross-sectional behavioral correlation effects, obtained using actual student networks. Section 6 investigates the sensitivity of the results to an alternative definition of peers.

[5] The empirical analysis demonstrates also the importance of controlling for endogeneity, hence allowing for selection into the social network. Neglecting the endogeneity of the social network leads to a large significant downward bias in the estimate of addiction effects and an upward bias in the estimate of conformity effects.

[6] The social multiplier measures the amplification on individual actions of the effects of an exogenous shock due to peer effects. It is the standard metric adopted for social interactions. Similar results are obtained also using as a measure of social interactions the correlation between students' choices in each period and the average level of those of their peers at different social distances (smallest number of links between them in the network).

Finally, Section 7 concludes.

1.1 Related Literature

As we noted in the Introduction, risky health behaviors have been extensively studied in economics and more generally in the social sciences. We refer to Cawley and Ruhm (2011) and Kenkel and Sindelar (2011) for extensive and detailed surveys of the literature. A fundamental aspect of both the theoretical and the empirical literatures in economics involves i) distinguishing *rational addiction* models, as introduced by Becker and Murphy (1988) and developed by Orphanides and Zervos (1995), from *behavioral* models, as in Gruber and Kosegi (2001), Bernheim and Rangel (2004) and others; ii) dealing with the inferential problems plaguing the empirical study of social interactions, as noticed by Manski (1993) and addressed by Blume et al. (2015), Brock and Durlauf (2001b) and many others.

With respect to rational and behavioral models of addiction, it should be noted that in both classes of models agents respond to dynamic incentives, such as, e.g., future price changes and anticipated future consumption. But the implied responses are different along several dimensions. Agents' choices in behavioral models are driven by preferences for immediate gratification, impulsivity, and cue-triggered addiction which have no role in models of rational addiction. The distinction between rational and behavioral addiction manifest itself most clearly, therefore, in high frequency decisions, over days. It is much less relevant when studying, as in our case, low frequency decisions, over years. For this reason we postulate rational agents in our analysis.

With respect to inference in social interactions models, as we noted in the Introduction, we try and address the main issues in the literature: (i) the endogeneity of previous choices as explanatory variables for current choice, in the absence of any restrictions on the intertemporal correlation structure of errors, (ii) the existence of common shocks or common unobserved factors affecting all individuals' choices in a network, independently of social interactions, and (iii) the endogeneity of the network. First of all, regarding (i) we instrument previous choices, using the characteristics of connected individuals in the social network as well as own lagged characteristics. In the absence of dynamics, the quest for valid instruments is conducted necessarily at the cross-sectional level and exclusion restrictions are translated into necessary conditions on the structure of the adjacency matrix.[7] In our dynamic environment, we are not restricted to the cross-section.

[7] The characteristics of friends and friends of friends are valid instruments under appropriate restrictions on the structure of the adjacency matrix; see e.g. Bramoullé et al. (2009), Calvo-Armengol et al. (2009).

In particular, we have access, for each period, to the strictly exogenous lagged variables from the information set in the previous periods. These variables are informative for the lagged choice variables by virtue of the intertemporal linkages formed by the moment restrictions of dynamic equilibrium of our social interactions model. To sum up, exploiting the equilibrium restrictions that jointly employ interactions in "space" as well as rational expectations interactions in "time" provides us with much richer possibilities for identification. As for (ii), we tackle the issue by including network fixed effects. Because in the Add Health data networks are small (composed on average of 25 students, with 10 as the median), this strategy reasonably accounts for the presence of unobserved factors common to groups of friends. Finally, regarding (iii), to control for the endogeneity due to selection into the friendship network, we add a Heckman-correction term to the structural equations we estimate as also recommended by Blume et al. (2015). More precisely, we estimate an extended version of our model, by adopting an approach developed by Qu and Lee (2015), in which we explicitly account for a possible correlation between unobserved factors jointly affecting both network formation and equilibrium outcomes.[8]

From a theoretical point of view the main novelty in the analysis of this paper consists in the study of the theoretical properties of equilibrium in an economy displaying both dynamic forward-looking agents and social interactions. In this respect, a related model is introduced in Reiff (2018), to characterize the theoretical properties of addiction in a dynamic forward-looking model with social interactions. Social interactions, however, are modeled in a reduced form, by having agents' preferences depend on the average action in the economy, without a specification of the structure of interactions on the network. Differently than in our model, therefore, in Reiff (2018) agents need not anticipate the effects of their actions on those of their peers in their decision problems. Various theoretical properties of models of social interactions in linear dynamic economies are also studied in Özgür, Bisin, and Bramoullé (2019). In the current paper, however, the analysis of social interactions is extended to allow for a general network topology. This is important in particular because it changes identification conditions. More specifically,

[8] A growing literature on social interactions has resorted to modeling the formation of social networks, to provide a more satisfactory solution to iii); see e.g., Apicella, Marlowe, Fowler, and Christakis (2012), Badev (2018), de Paula, Richards-Shubik, and Tamer (2018), Hsieh and Lee (2015), Sheng (2018), Mele (2017a,b). Empirical work along these lines also exploits Add Health data. Embedding network formation in a fully specified dynamic forward-looking choice model is theoretically daunting. Mele (2017a,b), for instance, estimates a network formation model to fit the observed networks' statistical properties, such as, e.g. homophily. But the paper does not study equilibrium choices in the network. Badev (2018), while estimating a network formation model in the context of smoking choice, restricts the analysis to a static choice model (though the network is allowed to change following an evolutionary process). For this reason, to be able to allow for dynamic choice, in this paper we take the simpler but admittedly reduced form approach of estimating a Heckman-correction term.

an incomplete network structure provides a source of non-linearity (intransitive triads) that can be exploited for identification purposes (Bramoullé et al., 2009; Calvo-Armengol et al., 2009) in addition to lagged values of exogenous variables as suggested by the moment restrictions of the dynamic equilibrium.

In terms of the empirical analysis, the main contribution of this paper still consists in estimating structurally a model which allows jointly for both dynamic forward-looking agents and social interactions. Indeed, most studies of risky health behaviors have examined either peer group effets or addiction and dynamic effects; see the literature surveyed by Cawley and Ruhm (2011) and Kenkel and Sindelar (2011), and, more recently, e.g., Nakajima (2007), Card and Giuliano (2013), Eisenberg et al. (2014), Lee et al. (2014) and Hsieh and Van Kippersluis (2018). On this dimension, the closest paper to ours is Dahl, Løoken, and Mogstad (2014), on the influence of peers in the take up of social programs (specifically, paid paternity leave in Norway). Using information transmission as the channel for social interactions, Dahl, Løoken, and Mogstad (2014) estimates "snowball effects", that is, peer effects which have a dynamic component as in our dynamic social multiplier. The analysis however does not allow for forward-looking behavior in the dynamic choice of agents, and the dynamics of peer effects is due to the exogenous spreading of interactions over the network.

2 Dynamic Interactions on Networks

In this section, we introduce the theoretical structure we shall adopt in the paper to study dynamic interactions on networks. Agents make choices over time. Their preferences over choices at any time t depend on their own previous choices at $t-1$. In the context of health risk behavior we study in this paper, this dependence represents the costs associated to behavioural changes due, e.g., to habits and addictions. Agents interact in their social reference group, the social network, and display preference externalities: each agent's preferences at any time t depend on the current choices of agents in her network. In the context of health risk behavior, this effect represents agents preferences for conformity with the social reference group. This dynamic interaction structure induces each agent's *optimal* choice to depend on all other agents' previous choices and current preference shocks.

2.1 The model

The economy is populated by a finite number of agents $i = 1, \dots, N$ for $t = 1, \dots, T$ periods. Each agent i chooses an action y_{it} at time t after having observed a preference shock $\theta_{it} \in \Theta$.[9] Let $\mathbf{y}_t$ and θ_t denote the corresponding N-dimensional vectors stacking all agents. Let $\theta := (\theta_t) := (\theta_{it})_{i=1,\dots,N,\, t\geq 1}$ be the stochastic process of agents' preference shocks.

The economy's *social network* is represented by an $N \times N$ matrix $\mathbf{G} = [g_{ij}]$, where g_{ij} indicates the friendship relationship between i and j. We consider a *directed network*, in which each agent interacts directly with her friends, and friendship of i with j does not imply friendship of j with i. Following the convention in the social networks literature, i) $g_{ij} > 0$ if i nominates j as one of her friends, otherwise $g_{ij} = 0$; ii) $\sum_j g_{ij} = 1$; iii) $g_{ii} = 0$.[10]

The preferences of an agent i at time t are represented by the utility function

$$
\begin{aligned}
u_i(y_{it-1}, \mathbf{y}_t, \theta_t, \mathbf{G}) \quad := \quad & -\alpha_1 (y_{it-1} - y_{it})^2 - \alpha_2 (\theta_{it} - y_{it})^2 \\
& -\alpha_3 \sum_{j=1}^{N} g_{ij} (y_{jt} - y_{it})^2,
\end{aligned}
\tag{1}
$$

where $\alpha_1, \alpha_2, \alpha_3 \geq 0$ are parameters. The utility function u_i represents the trade-offs that each agent i faces in her choice at time t. Each agent i obtains utility from matching her individual choice y_{it} with her previous choice y_{it-1}, her preference shock θ_{it}, and with the current choices of her peers $\{y_{jt}\}_{j: g_{ij} \neq 0}$. We refer to α_1 as the *addiction effect*, to α_2 as the *own effect*, and to α_3 as the *peer effect*. While $(\alpha_1, \alpha_2, \alpha_3)$ are restricted to be homogeneous across agents, preference heterogeneity is captured in the formulation of the stochastic processes θ_{it}.[11]

Agents maximize expected present discounted utility, with discount rate $\delta < 1$. Before her choice at time t, each agent observes i) the history of previous choices, $\mathbf{y}^{t-1} = (\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{t-1})$, and ii) the history of preference shocks, $\theta^t = (\theta_1, \dots, \theta_t)$ (including the period-t realization).

[9] See Appendix A for the formal introduction of the model, where all the technical assumptions are well-specified.

[10] By assuming a directed network, the matrix $\mathbf{G}$ is asymmetric. None of our theoretical results, however, hinge on this assumption. That is, they hold also in the case of a symmetric netwok structure.

[11] While we model preferences for conformity directly as a preference externality, we intend this as a reduced form of models of behavior in groups that induce indirect preferences for conformity, as e.g., Jones (1984) , Cole et al. (1992), Bernheim (1994), and Peski (2007).

2.2 Equilibrium

We consider Subgame Perfect Nash equilibria of this economy. At a Subgame Perfect Nash equilibrium, agents make optimal choices simultaneously at each time t. The equilibrium is represented by a family of maps $\{y_i^*\}_{i=1}^N$ such that for all $i = 1, \ldots, N$ and for all $(\mathbf{y}^{t-1}, \theta^t)$,

$$y_{it}^*\left(\mathbf{y}^{t-1},\, \theta^t\right) \in argmax_{y_{it}\in Y}\ E\left[\sum_{t=1}^{T} \delta^{t-1}\, u_i(y_{it-1}, \mathbf{y}_t, \theta_t, \mathbf{G})\right] \qquad (2)$$

for $(\mathbf{y}^0, \theta^1)$ given.

The economy displays a *unique Subgame Perfect Equilibrium.*[12] The first order condition for agent i's problem can be written as

$$\mathbf{y}_t = \alpha_1\, \mathbf{B}_t \mathbf{y}_{t-1} + \alpha_2\ \mathbf{B}_t \left(\mathbf{D}_t + \theta_t\right), \quad t = 1, ..., T, \qquad (3)$$

where $(\mathbf{y}^0, \theta^1)$ is given and the $N \times N$ matrix $\mathbf{B}_t$ and the $N \times 1$ matrix $\mathbf{D}_t$ can be computed recursively: $\mathbf{B}_t$, $t < T$ depends only on the future equilibrium coefficient matrices $(\mathbf{B}_\tau)_{\tau > t}$; while $\mathbf{D}_t$ represents the discounted sum of the effects of expected future θ_τ's, $\tau > t$.[13]

3 Data

Our data source is the Add Health, a dataset on adolescents' behavior in the United States. The dataset collects self-reported demographic and behavioral characteristics from students in grades 7-12 from a nationally representative sample of roughly 130 private and public schools in years 1994-95.[14] Every student attending the sampled schools on the interview day was asked to complete a questionnaire (*in-school questionnaire*) containing questions on respondents' demographic

[12] In Appendix A, we formally state and prove the equilibrium existence and uniqueness result, as well as the details of recursive algorithm to compute equilibria. Uniqueness requires $\alpha_1 + \alpha_2 > 0$ to anchor agents' preferences on their own private types or past choices. Clearly, without such an anchor, actions are driven only by social interactions, own past behavior and types have no effect on the outcomes, and a large multiplicity of equilibria would arise.

[13] See Appendices B and C for closed form characterizations of $\mathbf{B}_t$ and $\mathbf{D}_t$, as well as a recursive algorithm to compute them.

[14]The Add Health is a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Information on how to obtain the Add Health 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 analysis.

and basic family background characteristics. A subset of students randomly selected from the rosters of the sampled schools was then asked to complete at home a longer questionnaire containing more sensitive individual and household information (*in-home questionnaire*), including detailed questions about risky behaviors. In 16 randomly sampled schools, all students are interviewed at home (the so-called *saturated sample*). Our analysis focusses on this data. Specifically, we restrict our sample to include all students in the schools where information on risky behaviors is collected for the entire school.[15]

A unique feature of the Add Health data is that students are asked to identify their best friends from a school roster, so that the geometry of the friendship networks within each school is observed.[16] The nominations are asked in two waves of the survey (wave I and wave II) one year apart. The demographic and behavioral characteristics of the nominated friends of each of the students in the sample are also observed. As a result, each student's information is collected twice, in two waves in consecutive grades. Hence the sample has a panel dimension. In our analysis, we use Wave II to measure the students' risky behavior a time t and the network topology, and wave I to get information on the students' risky behavior at time $t-1$. In Section 6, we show the robustness of our results when adopting as network definitions the network topology at wave I.[17]

We focus our analysis on smoking and alcohol use, which are the most common risky behaviors among teenagers. We focus on students who declare having tried smoking and alcohol to avoid the difficult and arbitrary classification of adolescents as "consumers" or "non-consumers." To measure smoking, we use the answers to the question "During the past 30 days, on the days you smoked, how many cigarettes did you smoke each day?". To measure alcohol consumption, we use the most similar question: "Think of all the times you have had a drink during the past 12 months. How many drinks did you usually have each time?". We construct an index of risky behavior for each individual by first standardizing each component (cigarette and alcohol consumption) in [0,1] and then averaging. An higher value of the index is interpreted as a riskier health choice. Figure 1 shows the distribution of the teenagers in our sample by cigarette consumption, alcohol consumption, and the composite index. The distributions have a mode (of about 40% of students)

[15]This is done to avoid the complex inferential issues due to the missing observations in the friendship data for the schools where a random sampling scheme is adopted.

[16]The number of nominations is limited to five males and five females, but the limit is not binding not even by gender. Less than 1 percent of the students in our sample report a list of ten best friends, less than 3 percent report a list of 5 boys, and roughly 4 percent a list of 5 girls.

[17]In the saturated sample, about 30 percent of the wave I nominations are not confirmed in wave II and about 30 percent of new nominations appear in Wave II.

corresponding to minimal consumption of cigarettes and alcohol.[18]

Our set of control variables includes the variables indicated by the literature (see e.g., Cawley and Ruhm, 2011; Lee et al., 2014) as determinants of teenagers' risky behavior, such as age, gender, parental education, race, and indicators of the social structure of families as well as variables measuring the susceptibility of a teenager to engage in risky behavior (whether alcohol or cigarettes are easily available at home and whether the child receives a weekly allowance from the parents).[19][20][21] Finally, we include height, since it has been indicated as a predictor of participation in social activities, such as team sports (see Persico, Postlewaite, and Silverman, 2004).[22] The uniqueness of our data where both respondents and friends are interviewed allows us to control for peers' characteristics, thus disentangling the effects of endogenous from exogenous effects. More precisely peers' characteristics are defined as the average value of the above controls over the nominated friends.

The sample counts 2336 students, about 543 of which are in middle schools (grades 7, 8, and 9) and 1793 in high-schools. We use the (larger) sample of high school students to structurally estimate our model and the sample of middle school students for its validation. In the sample of high school students without missing values in the variables (1759 out of 1793 students), there are 552 students who do not indicate any best friend in the nomination roster. In Table G1, we report summary statistics for the entire sample (Panel (a))), for the sample without observations with missing values (Panel (b)) and for the sample with no isolates (Panel (c)). It appears that the composition of the sample is roughly unaffected, thus revealing that the average

[18] About 40 percent of students smoke one cigarette or less and/or have one drink or less.

[19] In the wave I in-home questionnaire students are asked, for each parent, to select how far each of their biological parents went in their education, with possible answers: "never went to school", "not graduate from high school","high school graduate","graduated from college or a university", "professional training beyond a four-year college". If the information is available for both residential parent, we select the father level of education. We construct a variable "Parents College degree", which is coded as 1 if the parent is "graduated from college or a university" or "professional training beyond a four-year college". The base category is "never went to school".

[20] From the wave I in-home questionnaire, we construct a variable "Two-parents family" using the respondent' answers about household composition. In particular, this variable is coded as 1 if students report to have two parents (both biological or not) that are currently living in their household, and 0 otherwise.

[21] Students are asked to answer yes or no to the questions "Are cigarettes easily available to you in your home?" and "Is alcohol easily available to you in your home?". We construct a variable "Alcohol/tobacco at home", which is coded as 1 if the respondent answers yes in at least one of the questions above, and 0 otherwise. We construct a variable "Pocket money", from student responses to the question: "How much is your allowance each week? If you don't receive your allowance weekly, how much would it be each week?". The three questions together measure the accessibility to alcohol and tobacco, either directly (i.e. stealing a cigarette from the mother's purse) or indirectly, by buying them and are also related with the price elasticities of demand for cigarettes (see e.g., Gallet and List, 2003; Cook and More, 2000). We measure these variables in both wave I and Wave II, from the in-home questionnaire.

[22] The respondents' height in feet and inches is available in both wave II and wave I *in-home* questionnaire.

student who do not indicate any best friend and are not nominated by anyone is not dissimilar from the average student who names at least one best friend. Because a smaller sample size substantially decreases computing time in the estimation of our model, we focus our analysis on students who are connected in a social network and show the robustness of our main results when including isolated individuals in Section 6. Our final sample of high school students counts 1207 individuals (Table G1, Panel (c)). Consistently with the epidemiological literature (see e.g., Malone et al., 2012) which find persistence of alcohol and tobacco consumption after initiation, our risky behavior index increases from wave I to wave II given the young age of students (from 14 to 20 with an average of 17 in wave II. Girls make up about 50 percent of the sample. Around 60 percent of the sample is White, 11 percent is Black or African American, 18 percent is Hispanic or Latino, and the reminder is Asian or with unclear racial background. The average height in the sample is 67 feet and the average allowance around 8 dollar per week. Finally, 26 percent of our adolescents have cigarettes or alcohol easily available at home. Regarding student family background, about 73 percent of the adolescents in our sample have two parents living in the household and roughly 25 percent have parents who are college graduates or above. These percentages are in line with data from other national representative surveys. We report in Table G2 in Appendix G information from the 1994 Current Population Survey (CPS) that is re-weighted to match the age distribution of the Add Heath sample. As shown, the Add Health population is broadly similar to the U.S. population as calculated from the CPS.[23]

4 Empirical Methodology

We use the panel dimension of our data to structurally estimate our dynamic social interaction model. The structural system of equations we estimate is the first order condition system describing the equilibrium, Equation 3,

$$\mathbf{y}_t = \alpha_1 \mathbf{B}_t \mathbf{y}_{t-1} + \alpha_2 \ \mathbf{B}_t \left(\mathbf{D}_t + \theta_t\right), \ \ t = 1, ..., T.$$

which expresses the vector of outcomes $\mathbf{y}_t$ as a function of the outcomes in the previous period, $\mathbf{y}_{t-1}$, and contemporaneous variables (including expectations about the future).

In our empirical exercise, the empirical counterpart of this equation system at time $t = T$ is constructed by considering students at grade 12 in wave II and the same individuals in grade 11

[23]The IPUMS-CPS database is freely available since 1962. See Flood et al. (2018) for further information.

at wave I. In other words, we use outcomes of 12th grade students as $\mathbf{y}_T$ and outcomes of the same students in the 11th grade as $\mathbf{y}_{T-1}$. The empirical counterpart of the system at $t = T - 1$ is then constructed by considering individuals in grade 11 in wave II and the same individuals in grade 10 in wave I; and so on until $t = T - 4$ with students in 8th grade (wave II) and the 7th (wave I). In Table 1, we summarize the structure of our sample.

We estimate Equation 3 at $t = T$, $t = T - 1$ and $t = T - 2$ with data referring to high school students (grades 10th to 12th). We then exploit the structural equation at $t = T-3$ and $t = T-4$ with data referring to middle school students (grades 7th to 9th) to validate the model.

Equation 3 depends on the structure of the stochastic process of preference shocks θ_T, which captures the heterogeneity across agents in the model. In our empirical exercise, we implement heterogeneous preference shocks as follows. Let the index $k = 1, ..., K$ account for the k-th distinct component of individual i's observable characteristics (e.g. age, gender, parent's education) and let $x_{it}^{(k)}$ denote its value for agent i at time t. The preference shock θ_{it} is allowed to depend on individual i's characteristics and on those of the member of her network. It is decomposed as follows:

$$\theta_{it} \quad := \quad \underbrace{\sum_{k=1}^{K} \beta_k \, x_{it}^{(k)} + \sum_{k=1}^{K} \sum_{j=1}^{N} \phi_k \, g_{ij} \, x_{jt}^{(k)}}_{\text{Observed exogenous heterogeneity } (a_{it})} + \underbrace{u_{it},}_{\text{Unobserved component}} \tag{4}$$

where β_k and ϕ_k are parameters. In matrix form:

$$\theta_t = \mathbf{X}_t \beta + \mathbf{G} \mathbf{X}_t \, \phi + \mathbf{u_t}, \tag{5}$$

where β and ϕ are $K \times 1$ vectors of parameters, $\mathbf{X}_t$ is an $N \times K$ matrix, and $t = 1, \ldots, T$.

4.1 Identification

In this section, we derive conditions under which the dynamic model with social interaction we have introduced is identified when the number of individuals N, the horizon of the economy T, and the social structure $\mathbf{G}$, are fixed and known to the econometrician.

The structural equation system defining the equilibrium of the dynamic social interaction economy is obtained by substituting the equation for the preference shocks, Equation 5, into the first order condition system, Equation 3:

$$\mathbf{y}_t = \alpha_1 \mathbf{B}_t \mathbf{y}_{t-1} + \alpha_2 \mathbf{B}_t \left(\mathbf{D}_t^x + \mathbf{X}_t \beta + \mathbf{G} \mathbf{X}_t \phi \right) + \varepsilon_t, \quad t = 1, 2, \ldots, T \tag{6}$$

where the conditional expectations in $\mathbf{D}_t$ are conveniently split into an observable and an unobservable component, respectively denoted $\mathbf{D}_t^x$ and $\mathbf{D}_t^u$ and $\varepsilon_t = \alpha_2 \mathbf{B}_t (\mathbf{D}_t^u + \mathbf{u}_t)$.

The parameters of the economy are the utility parameters $(\alpha_1, \alpha_2, \alpha_3)$, the discount factor δ, the own and social effects parameters $\beta = (\beta_1, \ldots, \beta_K)'$ and $\phi = (\phi_1, \ldots, \phi_K)'$. Utility functions are unique up to positive affine transformations and hence $(\alpha_1, \alpha_2, \alpha_3)$ are normalized so that $\sum_i \alpha_i = 1$. We assume the following standard conditions:

Exogeneity: $E[\mathbf{u}_s \mid \mathbf{X}_t] = 0$, for any $s, t = 1, \ldots, T$;

Full rank: the moment matrix generated by the elements of $\{\mathbf{X}_T, \mathbf{X}_{T-1}, \mathbf{X}_{T-2}\}$ has full rank.

Exogeneity requires that observable covariates and unobservables are orthogonal, contemporaneously and intertemporally. Full rank requires lack of multicollinearity and enough intertemporal variation in the observable covariates. We also assume that

Regularity: there exists $k \in \{1, \ldots, K\}$ such that $\alpha_1 \alpha_2 (\beta_k + \phi_k) \neq 0$.

This assumption requires that the lagged variables from the information set in period $t-1$, $\{\mathbf{X}_{t-1}, \mathbf{X}_{t-2}, \ldots, \mathbf{X}_0, \mathbf{y}_0\}$, are potentially informative for the lagged choice variable $\mathbf{y}_{t-1}$, which is the endogeneous variable, in the structural equation for period t. Importantly, all these assumptions are consistent with substantial correlation over time and across the network, both in observables and unobservables.

Under these assumptions, the structural equations of our dynamic model with social interaction, Equation 6, for $T \geq 2$, are identified.[24] The proof of this identification result proceeds in two steps: i) we prove that the coefficients of the structural equation, Equation 6, can be consistently estimated, and ii) we show that the map from the structural parameters $(\alpha_1, \alpha_2, \alpha_3, \beta, \phi, \delta)$ to the coefficients of Equation 6 is injective. We provide here the main arguments and a general discussion of how we can implement step i) of the identification result. Technical arguments needed for proving i) and a description of the recursive algorithm for step ii) are detailed in Appendix D.

[24] More precisely,

$$F_p(\mathbf{y}, \mathbf{X}) = F_{p'}(\mathbf{y}, \mathbf{X}) \;\; \Rightarrow \;\; p = p';$$

where $p = (\alpha_1, \alpha_2, \alpha_3, \beta, \phi, \delta)$ and $F_p(\mathbf{y}, \mathbf{X})$ is the joint probability distribution of observables $(\mathbf{y}, \mathbf{X})$ induced by the parameters p .

Consider the system in Equation 6 at $t = T, T-1$. Because $\mathbf{D}_t$ contains expectations about future shocks, $\mathbf{D}_T = 0$. The system in 6 can be reduced to

$$\mathbf{y}_T = \alpha_1 \mathbf{B}_T \mathbf{y}_{T-1} + \alpha_2 \mathbf{B}_T (\mathbf{X}_T \beta + \mathbf{G}\mathbf{X}_T \phi) + \varepsilon_T, \tag{7}$$

$$\mathbf{y}_{T-1} = \alpha_1 \mathbf{B}_{T-1} \mathbf{y}_{T-2} + \alpha_2 \mathbf{B}_{T-1} (\mathbf{X}_T \beta + \mathbf{G}\mathbf{X}_T \phi) + \alpha_2 \mathbf{B}_{T-1} \mathbf{D}^x_{T-1} + \varepsilon_{T-1}. \tag{8}$$

The endogeneity of $\mathbf{y}_{T-1}$ in Equation 7 and of $\mathbf{y}_{T-2}$ in Equation 8 requires to find suitable instrumental variables, $\mathbf{q}_{T-1}$ and $\mathbf{q}_{T-2}$. Consider selecting $\mathbf{q}_t = [\mathbf{X}_t, \mathbf{G}\mathbf{X}_t]$, $t = T-1, T$. Predicted values of $\mathbf{y}_{t-1}$ are formed by projecting them on the space spanned by the set of instrumential variables $\mathbf{q}_{t-1}$, $t = T-1, T$. These are valid instruments by construction since: (i) Regularity implies $E[\mathbf{q}_{t-1}\mathbf{y}_{t-1}] \neq 0$, $t = T-1, T$; (ii) Full rank implies explanatory variables are not collinear; (iii) Exogeneity guarantees that exclusion restrictions are satisfied, i.e., $E[\mathbf{q}_{t-1}\varepsilon_t] = 0$, $t = T-1, T$. Finally, Equation 7 is independent of δ and hence the condition $T \geq 2$ is necessary to identify it.

The set of instruments $\mathbf{q}_t$, $t = T-1, T$ is larger than necessary allowing for flexibility and power in empirical implementation. More specifically, our dynamic model with social interactions can be identified using the dynamic equilibrium restriction imposed by the model with only past variables as instruments, $\mathbf{q}_{t-1} = [\mathbf{X}_{t-1}]$. The model can also be identified with only the topology of the social network, $\mathbf{q}_{t-1} = [\mathbf{G}\mathbf{X}_{t-1}]$, though this requires some extra regularity conditions on the network, e.g., the linear independence of the powers of $\mathbf{G}$, as in Bramoullé et al. (2009).

4.2 Estimation

Because of the dynamic recursive structure of the theoretical framework, we can jointly estimate the reduced-form equilibrium equations 6, for $t = T, T-1, T-2$ (as in Table 1). In particular, we implement a nonlinear optimal Generalized Method of Moments (GMM) estimator (Hansen (1982) and Cameron and Trivedi (2005)).[25] The definition of the instrument matrix is based on our identification result and it is detailed in Appendix E. We consider two modelling strategies. First, we estimate the model assuming exogeneity of the network $\mathbf{G}$. Because of the dynamic recursive structure of the theoretical framework, we can jointly estimate the reduced-form equilibrium equations 6, for $t = T, T-1, T-2$ (as in Table 1). Second, we tackle a possible endogeneity of the network by implementing a two-step procedure à la Heckman, as recommended by Blume

[25] See Appendix E for details on the GMM procedure.

et al. (2015).[26] Specifically, we estimate an extended version of our model in which we explicitly account for a possible correlation between unobserved factors jointly affecting both network formation and equilibrium outcomes. To this end, we adapt an approach developed by Qu and Lee (2015) for the estimation of a spatial autoregressive model with geographical data.

In our context, the approach consists in modeling friendship ties between students using a standard dyadic model of link formation (see, e.g., Fafchamps and Gubert (2007)). Let g_{ij} denote the probability that two students i and j are linked as friends. We postulate a linear probability model in terms of the distance between the agents in terms of their characteristics, $[x_i^k, x_j^k]$, for $k = 1, ..., K$:

$$g_{ij} = \lambda_0 + \sum\nolimits_k \lambda_k |x_i^{(k)} - x_j^{(k)}| + \eta_{ij} \tag{9}$$

where

$$E(\eta_{ij}\eta_{ik}) = \begin{cases} \sigma_\eta^2 & \forall j = k \\ = 0 & \forall j \neq k \end{cases}.$$

The *selection effect* is the correlation between unobservable characteristics determining link formation η_{ij} and the unobservable characteristics u_{it} in the preference shock θ_{it} in equation 4. It is assumed to be homogeneous across agents:

$$E(u_{it}\eta_{ij}) = \sigma_{u\eta}, \ \forall j \neq i.$$

Under these assumptions and linearity of the conditional expectation of u_{it} given η_{ij}, $E(u_{it}|\eta_{i2}, \ldots, \eta_{in}) = \psi\xi_t$, where $\psi = \sigma_{u\eta}\sigma_\eta^2$ and $\xi_t = \sum_{j \neq i} \eta_{ij}$ Equation 6 can then be rewritten as:[27]

$$\mathbf{y}_t = \alpha_1 \mathbf{B}_t \mathbf{y}_{t-1} + \alpha_2 \mathbf{B}_t \left(\mathbf{D}_t^x + \mathbf{X}_t\beta + \mathbf{G}\mathbf{X}_t \phi\right) + \alpha_2 \mathbf{B}_t \left(\psi\xi_t\right) + \mathbf{e}_t, \tag{10}$$

where $\mathbf{e}_t = \alpha_2 \mathbf{B}_t \epsilon_t$ with $\epsilon_t = (\mathbf{u}_t - \psi\xi_t)$, $\xi_\mathbf{t} = (\xi_1, ...\xi_n)$ and the term $\psi\xi_t$ captures the selectivity bias.

We implement a two-stage estimation procedure of equation 10:

Stage 1 Estimate $\widehat{\xi_t}$ from an OLS regression of 9;

[26] Alternative avenues to address this issue would embed a network formation game into our dynamic recursive behavioral model. We reserve such challenging avenues for future research.

[27] Observe that normality is not required. Linearity of the conditional expectation of u_{it} given η_{ij} suffices. For example, exponential and uniform distributed errors also satisfy the requirement. See also the discussion of Assumption 2 in Qu and Lee (2015).

Stage 2 Estimate by nonlinear optimal GMM the reduced-form equilibrium equations 10 for $t = T, T-1, T-2$, after replacing ξ_t with its estimated counterpart in Stage 1, $\widehat{\xi}_t$.[28]

Inference is however complicated in this estimation procedure because the selectivity term is a generated regressor from a previous estimation and no closed form solution is available for the nonlinear optimal GMM standard errors estimates. We use bootstrapped standard errors with 1000 replications. Because of the inherent structural dependency of network data, the design of the resampling scheme needs special consideration. The residuals in the vector $\hat{\mathbf{e}}_t$ are not i.i.d., and thus one cannot sample with replacement from this vector. We thus use a residual bootstrap procedure, resampling on the structural errors ϵ_t which are assumed to be i.i.d.[29] The extended model 9-10 is identified even if the x_i^k variables used in the link formation and in the outcome equation completely overlap, as in our case (see Goldsmith-Pinkham and Imbens (2013), and Hsieh and Lee (2015)).[30]

5 Empirical network effects

Estimating our structural model of dynamic interaction on networks allows us to recover individual preferences from health risk behavior data. The main preference parameters in the model are the *discount rate* δ, the *addiction effect* parameter α_1, the *own effect* parameter α_2, and the *peer effects* parameter α_3, which we normalize without loss of generality so that $\sum_{i=1}^{3} \alpha_i = 1$. In the specifications in which the network is endogenous, we also estimate the selectivity parameter ψ, a measure of the importance of unobservables leading individuals both to risky behavior and to form friendship ties.

Table 2 presents the estimates of the structural parameters: in Column 1 the network $\mathbf{G}$ is assumed to be exogenous, while in Column 2 network formation is assumed to satisfy the linear dyadic structure in equation 9.

[28] In this stage we use the same instrument matrix of the exogenous network case detailed in Appendix E.

[29] This procedure is commonly used in spatial econometrics; see e.g., Anselin (1990). In practice, having in hand the residual vector $\hat{\mathbf{e}}_t$ one can derive the estimates of the structural errors from $\hat{\epsilon}_t = \widehat{\mathbf{B}}_t^{-1}\hat{\mathbf{e}}_t$. These estimates are then resampled school by school.

[30] Observe that the way identification is achieved, however, is not by functional form as in the traditional Heckman selection model (that is based on the use of a 0-1 dependent variable in the first stage). The identification strategy here exploits non-linearities specific to the network structure of our model. In our approach, the dyad-specific regressors used in the first stage (the network formation stage) are expressed in absolute values of differences of individual characteristics, $|x_i^{(k)} - x_j^{(k)}|$. These differences do not appear in the outcome equation. Finding pure "excluded variables" is notoriously difficult in environments like ours.

The estimate for δ is significant and relatively stable, independently of whether network formation is endogenized. The point estimate is low but not abnormally so with respect to the results in the experimental literature; see e.g., Frederick, Lowenstein, and O'Donoghue (2002). Most importantly, the peer effect α_3 and the addiction effect α_1 are also significant; they are similar in size, but they are 2-3 times larger (with respect to the own effect α_2) when controlling for the endogeneity of network formation. The selectivity parameter ψ is small and not significantly different from 0. The validation exercise we implement next in Section 5.1 gives however statistical support to our approach to control for the endogeneity of the network. Interestingly, the negative sign of the point estimate for ψ is consistent with the presence of unobservables which lead individuals to form friendship ties while reducing their utility for risky behavior.[31]

5.1 Out-of-sample Validation

In this section, we implement a model validation exercise of the structural estimates of our dynamic recursive model, comparing out-of-sample predicted equilibrium behavior to actual behavioral data across different networks not employed in the estimation sample. We believe that by demonstrating that our dynamic model's predictions perform well when applied to pertinent new data, this validation exercise would give the reader more confidence that a mechanism of social interactions (the underlying structure) is uncovered from the data rather than imposed on the data.

To fullfil this objective, we follow an approach inspired by Todd and Wolpin (2006).[32] More precisely, as we explain in Section 4.2, we first estimate the structural parameters using the (larger) sample of students in grades 10, 11, and 12 under two specifications: (i) under the exogeneity of the network $\mathbf{G}$, using the reduced-form equilibrium equations 6 jointly, for $t = T, T-1, T-2$, by linking the structural equations and the sample as in Table 1; and (ii) under the possible endogeneity of the network $\mathbf{G}$, using the reduced-form equilibrium equations 10 jointly, for $t = T, T-1, T-2$, once again by linking the structural equations and the sample as in Table 1 and by explicitly accounting for a possible correlation between unobserved factors jointly affecting both network formation and equilibrium outcomes. Those parameter estimates under both specifications are reported in Table 2.

[31] This could be the case if, for example, individuals practice a sport together with their peers.

[32] In the context of a randomized social experiment in Mexico, Todd and Wolpin (2006) estimate a dynamic model without using post-program data and then compare the model's predictions about program impacts to the experimental impact estimates.

Next, we predict equilibrium risky behavior index for students in the hold-out sample consisting of grades 7, 8, and 9. Specifically, (i) under the exogeneity of the network **G**, we use the reduced-form equilibrium equations 6 jointly, for $t = T-3, T-4$, by linking the structural equations and the sample as in Table 1, and (ii) under the possible endogeneity of the network **G**, we use the reduced-form equilibrium equations 10 jointly, for $t = T-3, T-4$, once again by linking the structural equations and the sample as in Table 1. For both specifications, the predicted student behaviors are generated recursively, using the estimated parameters in Table 2, the baseline controls (and network **G**) for grades 7, 8 and 9, and students' initial risky behavior index values for $t = T-5$, observed in the data.

Finally, we compare the predictions of our dynamic model for grades 8 and 9 to the actual behavioral choices of those adolescents observed in the data. Further, comparing the out-of-sample predictions obtained through the exogenous and endogenous network estimates allows us to statistically discriminate between the two modeling strategies.

Figure 2 presents the distributions of actual and predicted risky behavior index values. The distribution of the behaviors generated by the dynamic model, using parameter estimates in Table 2, are close to the distribution of the actual risky behavior index values, particularly so for the estimates controlling for the endogeneity of the network. The Pseudo$-R^2$ scores, computed as $Corr^2(y, \hat{y})$, where y is the observed and $\hat{y}$ the predicted behavior, are equal to 0.0917 and 0.1952 for the exogenous and the endogenous network estimates, respectively.

Table 3 reports the comparison of mean predicted and mean actual risky behavior index values for the overall sample and for several subgroups defined by gender, race, parental education, and others. Mean predicted values are close to the actual values in the data, for the whole sample and all the chosen sub-groups; again particularly so when predictions are obtained from the estimates allowing for endogeneity of the network. Indeed, for the specification allowing for endogeneity of the network, the p-values for t-tests for differences between predicted and actual means are uniformly higher across subgroups and we never reject the zero null.

We believe that the results presented in Figure 2 and in Table 3 deliver enough confidence in our dynamic model and produce a clear ranking of the goodness of fit of predictions for the two competing estimates in favor of controlling for network endogeneity. We therefore proceed to study further the implications of our structural estimates, with the endogenous network case as the baseline.

5.2 Social Multiplier

The social multiplier is a measure of the amplification of the effects of an exogenous shock on individual actions which is due to peer effects. It is a fundamental metric used to represent the implications of peer effects. In this section, we report on the social multiplier implied by our baseline structural estimates, allowing for endogeneity in network formation. Importantly, the social multiplier, in our context, has a fundamental dynamic component: a shock to agents' preferences in the 10th grade has effects on their behavior in all grades 10, 11, and 12.

Consider an exogenous shock to $\theta_{i,t}$, say by $\Delta\theta_{i,t} = \pi > 0$, for all agents i. This shock could represent, e.g., the outcome of a policy geared towards affecting students' behavior directly or indirectly through information or preferences. Consider first the case in which the shock is temporary at t, that is, preferences are now shocked once at any given grade t. From the structural equation 3 at t, the total change on the risky behavior index at t, accounting for peer interaction effects in the network at equilibrium, is $\Delta\mathbf{y_t} = \alpha_2\,\pi\mathbf{B_t}\,\mathbf{1}$, where $\mathbf{1}$ is the $N\times 1$ matrix of ones. The direct effect, abstracting from peer effects, that is, for $\alpha_3 = 0$, would be $\Delta\mathbf{y_t}\,|_{direct} = \alpha_2\,\pi\mathbf{1}$. Therefore, we define the multiplier of an exogenous preference shock at t on behavior at t as

$$m_{t,t} = \frac{\Delta\mathbf{y_t}}{\Delta\mathbf{y_t}\,|_{direct}} \quad = \quad \mathbf{B_t}\,\mathbf{1}.$$

Iterating the structural equation 3 for $\tau = t,\ldots,T$, we trace out the effects of an exogenous preference shock at t on behavior at all $\tau > t$:

$$\Delta\mathbf{y}_\tau \quad = \quad \pi\,\alpha_1^{\tau-t}\,\alpha_2\,(\mathbf{B}_t\times\cdots\times\mathbf{B}_\tau)\,\mathbf{1};$$

and the (*dynamic forward-looking*) social multiplier of an exogenous preference shock at t on behavior at τ is

$$m_{t,\tau} \quad = \quad \alpha_1^{\tau-t}\,(\mathbf{B}_t\times\cdots\times\mathbf{B}_\tau)\,\mathbf{1} \tag{11}$$

for any period $\tau = t,\ldots,T$.[33]

Similarly, for a permanent shock from time t to the end-time τ, we can compute the multiplier

[33] Please see Appendix F for detailed derivation of the dynamic social multiplier.

effect as:

$$m_{t,\tau} \;=\; \sum_{s=t}^{\tau} \alpha_1^{\tau-s} \left(\mathbf{B}_\tau \times \cdots \times \mathbf{B_s}\right) \left(\mathbf{1} + \tilde{D}_s\right) \tag{12}$$

The sample means of multiplier values ($\bar{m}_{t,\tau}$) for the baseline estimated parameter values, with endogenous network, are reported in Table 4. In Figure 3 the instantaneous multiplier $\bar{m}_{t,t}$ is reported for $t = 10, 11, 12$ and for both a temporary and a permanent shock.

Several interesting properties of the calibrated multipliers are worth noticing. First of all, the multiplier effects of a temporary preference shock at t decline over time: $m_{t,\tau}$ decreases with τ, for all t. For instance, a temporary shock at $t = 10$ has multiplier 1.53, 1.03, and 0.76, respectively at $\tau = 10, 11, 12$. The multiplier of permanent shocks instead increase over time. A permament shock in grade 10 has multiplier 1.55 in the same grade, 2.6 in grade 11, and 3.68 in grade 12. Furthermore, permanent shocks have larger multiplier effects than temporary shocks, both instantaneously and over time.[34] Perhaps most interestingly, the same-period multiplier effect decreases with the number of periods to the end of high-school T: $\bar{m}_{10,10} < \bar{m}_{11,11} < \bar{m}_{12,12}$. This is the case for both temporary and permanent shocks. In the case of temporary shocks, our estimates imply that the same-period social multiplier in grade 12 (the last year of high-school) is about 1.7, whereas it is 1.56 and 1.53, respectively, in grades 11 and 10. The same-period multiplier effects in the different grades encode the importance of the number of periods to the end of school in students' choice. Our estimates, therefore, are evidence that students anticipate a change in their social network after high-school and that this affects the importance of peer effects over schooling age: as the time to the end of high-school $T - t$ increases, the students' policy functions weigh more heavily future shocks and the current shock has a smaller effect.

Multiplier effects operate also through an *expectations channel*: forward-looking agents change their contemporaneous behavior in response to an anticipated shock in the future. Namely, we can compute the multiplier effect of an exogenous preference shock at a future date t on behavior at $\tau < t$:

$$m_{t,\tau} \;=\; \mathbf{B}_t \tilde{D}_t \tag{13}$$

where $\tilde{D}_t$ capture the sum of the expected effects on period τ marginal utility of a unit future

[34] Note that, obviously, a permanent shock at $T = 12$ is equivalent to a temporary shock and hence the multipliers are the same.

shock that is anticipated to change the random component of preferences, θ_t.[35]

In Table 5, we report the expected multiplier effects, in grade 10, 11 and 12 , induced by an anticipated shock to the preferences of all agents in grade 12; that is, we report the multipliers $m_{t,\tau}$ for $t = 12$ and $\tau = 10, 11, 12$. In anticipation of an increased preference for risky behavior in grade 12 agents increase risky behavior in grade 11 ($\bar{m}_{12,11} = 0.0540$) and decrease it slightly in grade 10 ($\bar{m}_{12,10} = -0.0055$). These effects are subtle but can be intuitively be explained as follows. In grade 10, agents anticipate that they will increase risky behavior at 12 and anticipate that so will do all their peers in the network ($\bar{m}_{12,12} = 1.7132$). This entails an adjustment cost in terms of utility, because the addiction effect penalizes behavioral changes over time. As a consequence, with strictly concave preferences, the agents will have an incentive to smooth these adjustment costs over time, increasing risky behavior starting from grade 10 and then in grade 11 and 12 as well. Interestingly, however, a compensating effect could dominate in grade 10 inducing the agents to decrease risky behavior first in grade 10 and then to increase in grades 11 and 12. This occurs in particular at the estimated parameter values, and more generally when peer effects are particularly strong, that is, α_3 is relatively large. Indeed, peer effects induce agents to engage in more risky behavior than their own preferences would induce them to in grade 12; the more so the stronger the peer effects. That is, generally $y_{i,12} > \theta_{i,12}$; and this also implies an adjustment cost in terms of utility for the agents. The negative multiplier in grade 10, therefore, is the result of the trade-off between smoothing the adjustment costs via the addiction channel and via the own preference channel in the presence of large peer effects.[36]

The salience of the social multiplier dynamics implied by our estimates can be illustrated by comparing the different (semi-) elasticities of a change in the availability of alcohol or cigarettes at home on child risky behavior implied by the direct effect ($\Delta E(y_{12})/E(y_{12}) \mid_{direct}$) versus the equilibrium effect ($\Delta E(y_{12})/E(y_{12})$).[37] Thus, 12th graders adolescents that have alcohol or

[35] $\tilde{D}_t$'s can be computed recursively. The explicit formula for $\tilde{D}_t = \pi^{-1}\Delta\mathbf{D}_t$ is given using

$$\begin{aligned}\Delta\mathbf{D}_t \quad := \quad & \sum_{s=t+1}^{t} \delta^{s-t}\Bigg(-\alpha_1 \operatorname{diag}\left(\Lambda_{t,s-1} - \Lambda_{t,s}\right)\left(\Delta\Gamma_{t,s-1} - \Delta\Gamma_{t,s}\right) \\ & + \operatorname{diag}\left(\Lambda_{t,s}\right)\left(\Delta\bar{\theta}_s - \Delta\Gamma_{t,s}\right) \\ & - \alpha_3 \sum_{k=1}^{N} \operatorname{diag}\left(G_{\bullet k}\, \iota_N'\right) \operatorname{diag}\left(\iota_N\, \Lambda_{k\bullet,t,s} - \Lambda_{t,s}\right)\left(\Delta\Gamma_{k,t,s}\mathbf{1} - \Delta\Gamma_{t,s}\right)\Bigg) \end{aligned} \qquad (14)$$

Please see Appendix F for the explicit derivation of the multiplier formulas recursively.

[36] Simulations of the anticipation effect at different values of the parameters, which we do not report, validate this intuition regarding the occurrance of a negative multiplier in grade 10.

[37] For all the computations we report the sample means of the semi-elasticities. Semi-elasticity is the percentage

tobacco easily available at home (directly) increase their risky behavior index by roughly 44% (with respect to adolescents that do not have alcohol or tobacco easily available at home). This effect almost doubles when we consider network peer effects in equilibrium (roughly 78%). Please refer to Table 7 for analogous results for the simulated (semi-) elasticities of a temporary as well as a permanent change in the availability of alcohol or cigarettes at home in any given grade on child risky behavior implied by the direct effect versus the equilibrium effect on behaviors in grades 10, 11, and 12.

5.3 Static and myopic bias

The main thrust of our theoretical analysis consists in modeling health risk behavior as the outcome of dynamic choice by forward looking agents. Empirically, the dynamic choice component of this approach is validated by the fact that the addiction effect α_1 is estimated to be significantly different than 0. The forward looking (as opposed to myopic) component of this approach is instead validated by the fact that the discount rate δ is estimated to be significantly different than 0.

To better gauge at the relevance of dynamic forward looking behavior, in this section we measure the bias associated to estimating (i) a mis-specified *myopic* model, obtained by restricting the structural equation 3 by imposing $\delta = 0$; (ii) a mis-specified *static* model, obtained restricting the structural equation 3 by imposing $\delta = \alpha_1 = 0$. We then compare the resulting parameter estimates and goodness of fit of these models with those of our *dynamic* baseline model with no restrictions. All models are estimated controlling for the endogeneity of the network. Results are reported in Table 8 and Figure 6. Table 8 compares the parameter estimates across specifications. The striking feature is that while the static model produces an estimate of the peer effect α_3 which is small, about a half of the own effect α_2, in the dynamic model, the peer effect estimate is on the contrary more than twice as large as the own effect. Furthermore, the sign of the selectivity effect parameter is positive in the static model estimates and negative in the dynamic model estimates. On the other hand, the estimates of the preference parameters obtained from the myopic model are in line with those associated with the dynamic baseline. Nonetheless, the performance of the two models is different, as shown in Figure 6.[38] Figure 6 compares the fit of the three models

change in a function relative to an absolute change in its parameter. Algebraically, the semi-elasticity of a function f at point x is $\frac{f'(x)}{f(x)} = \frac{d \ln f(x)}{dx}$.

[38]Observe that the point estimate of the selectivity parameter is of the opposite sign and especially of much larger magnitude in the myopic model estimates with respect to the dynamic baseline. In fact, the bias associated

with the data. It plots the actual data and the predictions that are obtained when using different models. It appears that the predictions of the dynamic baseline model are closer to the actual data than the predictions of both the static and the myopic models. Indeed, the pseudo$-R^2$ is equal to 0.3190, 0.2488 and 0.0571 for the dynamic, the myopic and the static models, respectively.

In addition to this evidence, observe that the restrictions imposed by both the static and the myopic models induce a dramatic bias in the estimation of the social multipliers. Firstly, as can be seen immediately from the construction of the multiplier measures in (11), (13) and (12), a misspecified static model cannot generate any intertemporal multiplier effects since the link across two consecutive periods is broken by setting $\alpha_1 = 0$. Second and perhaps more importantly, if a shock in question has not realized yet but will in the future, forward-looking agents anticipate that shock and change their contemporaneous behavior accordingly. However, myopic agents do not care about their future behavioral paths and consequently do not change their behavior accordingly. For a misspecified *myopic* model, as $\delta \to 0$, $|\mathbf{B}_T - \mathbf{B}_t| \to 0$ for any period t, and the model becomes one of a sequence of myopic period economies. Consequently, the instantaneous multiplier effect generated under this specification is constant across periods, yielding a bias, which is increasing in the time-to-end $T - t$, relative to our benchmark dynamic specification.

5.4 Spatial Correlation

We conclude our investigation by looking at the performance of our model in terms of spatial correlation, which is a traditional metric used to represent the implications of peer effects in terms of the spillovers they induce. We demonstrate that the correlations between risky behavior of agents at different social distances predicted by our dynamic model (with endogenous network formation) track very closely the correlations in the actual data.

In order to do that, we introduce a linear order along which space, and hence distance between agents, is defined. Let the social distance between any pair of students be characterized by the smallest number of links establishing a connection between them. Formally, we construct a minimum-path algorithm that maps any given network $\mathbf{G}$ into a series of geodesic adjacency matrices $\mathbf{G}_k$, $k = 1, 2, \ldots, d(\mathbf{G})$, where $\mathbf{G}_k$ is the matrix of agents who can reach each other through a shortest path of length at least k and $d(\mathbf{G})$ is the diameter of the entire friendship network $\mathbf{G}$.[39] We then pick the the largest network component in our sample which consists of

to the myopic model with respect to our dynamic baseline model appears especially in the estimated effects of the conditioning characteristics, which we report in Table G4 in Appendix G.

[39]The geodesic adjacency matrices are weighted by the number of connections each student has.

286 students and whose diameter is 24.[40] Figure 4 presents the actual series of geodesic student adjaceny matrices $\mathbf{G}_k$ of increasing social distance with $k = 1, 5, 10$, and 20, obtained using our algorithm, starting with the largest student network component in our sample. As is apparent in the Figure, and is consistent with theory, $\mathbf{G}_k$ first gets denser as k increases and later becomes sparse, as k gets closer to the diameter of the entire network component.

Using this definition of social distance between students, we follow an approach similar to that of Todd and Wolpin (2006) to predict the individual equilibrium paths of risky behavior of students by using the reduced-form equilibrium equations 10 jointly, for $t = T, T-1, T-2$, once again by linking the structural equations and the sample as in Table 1, and with parameters calibrated using the baseline estimates from Table 2, Column 2. Then, we compute the correlation between students' choices in each period and the average level of those of their peers at different social distances, as represented by adjacency matrix $\mathbf{G}_k$, $k = 1, 2, \ldots, d(\mathbf{G})$.[41] Figure 5 reports the results with 95 percent confidence interval, showing that spatial correlation declines quickly and is effectively zero at and after social distance $k = 3$. The trends in the estimated and actual correlations are remarkably similar.

6 Robustness

In our analysis, we control for a possible endogeneity of the social network by implementing a Heckman correction procedure. In this section, we investigate the sensitivity of the results to an alternative definition of peers, which does not require dealing with network endogeneity. We experiment by defining as peers all students of the same gender within the same grade at a school. This definition is grounded on the sociological literature documenting that adolescents are more likely to have same gender friends (see, e.g. McPherson, 2001). We inspect the validity of this hypothesis in our data in Figure 8. Figure 8 depicts friendship linkages in the larger network in our data (286 nodes with diameter 24) by using different colors for nodes indicating students of different gender. The picture reveals that indeed social interactions are assortative by gender.

[40] A component of a network is a nonempty subnetwork in which every two nodes belonging to it are connected to each other by paths, and which is connected to no additional nodes outside of the subnetwork. The diameter of a network is the shortest distance between the two most distant nodes in the network.

[41] Our method of internal validation works by predicting students' risky behavior indices at different time periods. The setup is the same as the one used in Section 5.1. In particular, predicted behaviors are generated recursively for each grade starting from grade 10 ($t = T-2$) and going up to grade 12 ($t = T$). We use the actual and the predicted dependent variables to construct the correlation between each individual choice and the average level of individual's peers' choices at different social distances represented by k.

Formally,let $\tilde{\mathbf{G}} = [\tilde{g}_{ij}]$ be the "new" interaction network,

$$\tilde{g}_{ij} = \begin{cases} \frac{1}{|N_{\tilde{\mathbf{G}}}(i)|}, & \text{if i and j have the same gender} \\ 0, & \text{otherwise.} \end{cases} \tag{15}$$

where $|N_{\tilde{\mathbf{G}}}(i)|$ denotes the number of friends (of the same gender) that agent i has. Hence the social term $\tilde{\mathbf{G}}\mathbf{X}_t$, in the preference shock decomposition in equation (5), is an average, for each agent i, of the covariates of all other agents of the same gender. Note also that as in our benchmark dynamic model, student i is removed (i.e. $\tilde{g}_{ii} = 0$) when the average risky behavior index of same-gender students is computed from the school-grade distribution. While this is a much coarser network definition, it has the advantage of being exogenous. Infact, when parents choose the school for their children, they are unlikely to be aware of how the intensity of risky behavior by gender vary by grade within a particular school. The idea is to treat the composition of students by grade and gender within a school as quasi-random and to use this quasi-random variation to identify peer effects.[42] We show that the the variation in the average risky behavior of same grade and gender school mates is unrelated to the variation in a number of predetermined student characteristics in Table G5 in Appendix G. We consider the same controls as the baseline specification listed in Table G1. We run separate regressions with each of these variables as alternate dependent variables, and add school fixed effects to control for differences in average student characteristics across schools as well as for other aspects of school quality. As shown in the table, none of the estimated correlations appear to be significantly different from zero, supporting the notion that our model specification identifies an exogenous source of variation. We estimate the model using the same strategy of Subsection 4.2. Specifically, we use the nonlinear optimal GMM estimator detailed in Appendix E, and replacing network fixed effects with school fixed effects.[43]

Results are reported in Table 9. The first column reports results of the estimated coefficients in the original sample using both middle and high school students. The second column report the estimated coefficients of the model using our regression sample (high school students). The

[42] This approach has been first proposed by Hoxby (2000b) to estimate the impact of class size, and subsequently widely used in studying peer effects in education (e.g. Angrist and Lang (2004); Friesen and Krauth (2007); Hanushek, Kain, and Rivkin (2002); Lavy and Schlosser (2011); Lavy, Paserman, and Schlosser (2012); Olivetti, Patacchini, and Zenou (2018)). Also, Patacchini and Zenou (2016) use a similar approach to investigate the impact of peer religiosity in the intergenerational transmission of religion.

[43] The matrix of instruments for the GMM estimation is defined as the one that we used in Section 4.2 with $\tilde{\mathbf{G}}$ instead of $\mathbf{G}$. Details are shown in Appendix E.

estimates of the preference parameters of the model confirm the salience of our theoretical model: the presence of the dynamic choice component is validated by the fact that the estimate of the addiction effect α_1 remains positive and significantly different than 0; the forward looking component is validated by the fact that the discount rate δ also remains positive and significantly different than 0; and the importance of the peer effect component is reflected in the positive and significant estimate α_3. Perhaps unsurprisingly, the magnitudes of the effects are different due to the different definition of peers. Notably, the estimated peer effect is much smaller, perhaps a consequence of a less precisely specified network.

Finally, we further perform two robustness checks. We estimate our dynamic model when considering the friendship nominations at wave I (rather than wave II) and when including also students who do not nominate any best friend and are not nominated by anyone. Table G6 in Appendix G reports the estimates of the structural parameters for these two cases: the network with wave I friendship nominations (Column 1), and the sample including isolated individuals (Column 2). For each robustness exercise the network is assumed to be endogenous. Results remains qualitatively unchanged when compared with the baseline estimates reported in Table 2.[44]

7 Concluding Remarks

Dynamic social interactions provide a rationale for several important phenomena at the intersection of economics and sociology. The theoretical and empirical study of economies with long-lived social interactions has been hindered by both mathematical and conceptual problems.

In this paper, we show how some of these obstacles can be overcome while studying the risky behavior of adolescents. We formulate and structurally estimate a dynamic social interaction model in the context of students' school networks included in Add Health. The equilibrium characterization of the dynamic game allows us to offer solutions to the well-known inferential problems in the study of social interactions. We construct a consistent estimator in our environment, by using the moment restrictions imposed by the dynamic equilibrium to back out the structural preference parameters. Our empirical analysis confirms the main thrust of our exercise regarding smoking and alcohol use in the adolescent population. We find strong evidence for forward looking dynamics, addiction effects, and social interaction effects. Social interactions in

[44]When we include isolated individuals in the sample we also add in the specification a dummy, which indicates if an individual is isolated or not.

the estimated dynamic model are indeed quantitatively large.

These results have important policy implications. The importance of social interactions for policy analysis relies on the fact that when social interactions are quantitatively important, well-targeted policy interventions at a smaller scale might have much larger effects at the aggregate through the social multiplier channel for those interactions. In this respect, we show that a misspecified static model would a much smaller estimate of the social interaction effect than the dynamic model. Furthermore, our empirical analysis implies that the effects of policy interventions to affect adolescents' risky behavior display over time, indeed increase over time when policy interventions are permanent. Finally, it also implies that the design of policy interventions should depend on the students' network structure and should consider the dynamics of the network itself, as students anticipate its natural breaks.

References

ALIPRANTIS, D., AND BORDER, K. C. (2006), *Infinite Dimensional Analysis*, third edition. Springer-Verlag, Berlin.

ALTONJI, J.G., ELDER, T.E. AND C.R. TABER (2005), "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools," *Journal of Political Economy* 113, 151-184.

ANSELIN, L. (1990), "Some Robust Approach to Testing and Estimation in Spatial Econometrics," *Regional Science and Urban Economics*, **20**, 141-163.

ANGRIST, J.D. AND K. LANG (2004),"Does school integration generate peer effects? Evidence from Boston's Metco program," *American Economic Review* 94, 1613-1634.

APICELLA, C. L., MARLOWE, F. W., FOWLER, J. H., AND CHRISTAKIS, N. A. (1994), "Social Networks and Cooperation in Hunter-gatherers," *Nature*, **481** (7382), 497-501.

ARGYS, L.M. AND REES, D.I. (2008), "Searching for Peer Group Effects: A Test of the Contagion Hypothesis," *Review of Economics and Statistics*, **90** (3), 442–58.

ATTANASIO, O., A. BARR, J. C. CARDENAS, G. GENICOT, AND C. MEGHIR (2012), "Risk Pooling, Risk Preferences, and Social Networks," *American Economic Journal: Applied Economics*, **4** (2), 134-167.

AXTELL, R., S. DURLAUF, J.M. EPSTEIN, R. HAMMOND, B. KLEMENS, J. PARKER, Z. SONG, T. VALENTE, AND H.P. YOUNG (2006): "Social Influences and Smoking Behavior: Final Report to the American Legacy Foundation," mimeo, Center on Social and Economic Dynamics, The Brookings Institution.

BABB S., MALARCHER A., SCHAUER G., ASMAN K., JAMAL A. (2017): 'Quitting Smoking Among Adults United States, 20002015," MMWR Morb Mortal Wkly Rep 2017;65:14571464. DOI: http://dx.doi.org/10.15585/mmwr.mm6552a1External

BADEV, A. (2015), "Discrete Games in Endogenous Networks: Theory and Policy," Unpublished manuscript, Federal Reserve Board, Washington D.C.

BALSA, A., DAZ, C. (2018): "Social interactions in health behaviors and conditions.," *Oxford Research Encyclopedia of Economics and Finance*. Oxford University Press.

Battaglini, M. and Patacchini, E. (2016), "Influencing Connected Legislators," CEPR Discussion paper DP11571.

Becker, G.S. and Murphy, K. M. (1988): "A Theory of Rational Addiction," *Journal of Political Economy*, **96**, (4), 675–700.

Becker, G.S., Grossman, M., and Murphy, K. M. (1994): "An Empirical Analysis of Cigarette Addiction," *American Economic Review*, **84**, 3, 396–418.

Benhabib, J., Bisin, A., and Jackson, M. O. (eds.) (2011): *Handbook of Social Economics*, Vol. 1A, The Netherlands: North-Holland, 339–416.

Bernheim, B. D. (1994), "A Theory of Conformity," *Journal of Political Economy*, **102**, 841-877.

Bernheim, B. D. and Rangel, A. (1994), "Addiction and Cue-Triggered Decion Processes," *American Economic Review*, **94**, (5), 1558–90.

Blume, L. E., Brock, W. A., Durlauf, S. N., and Jayaraman, R. (2015), "Linear Social Interactions Models," *Journal of Political Economy*, **123**, 2, 446-496.

Bramoullé, Y., Djebbari, H., and Fortin, B. (2009), "Identification of Peer Effects through Social Networks," *Journal of Econometrics*, **150**, 41-55.

Bramoullé, V. Boucher, Y., Djebbari, H., and Fortin, B. (2014), "Do Peers Affect Student Achievement? Evidence from Canada Using Group Size Variation," *Journal of Applied Econometrics*, **29**, 91-109.

Broadhead, R.S., Heckathorn, D. D., Weakliem, D. L., Anthony, D. L., Madray, H., Mills, R. J. (1998), "Harnessing Peer Networks as an Instrument for AIDS Prevention: Results from a Peer-driven Intervention," *Public Health Reports*, **113** (Suppl. 1): 42-57.

Brock, W. and Durlauf, S. (2001b), "Interactions-Based Models," in J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics, Vol. V*, North-Holland, Amsterdam.

Calvo-Armengol, T., E. Patacchini and Y. Zenou (2009), "Peer Effects and Social Networks in Education", *Review of Economic Studies*, **76**, 4, 1239-1267.

Cameron, A. C. and Trivedi, P. K. (2005), *Microeconometrics: Methods and Applications*, Cambridge University Press, Cambridge.

Card, D. and Giuliano, L. (2013), "Peer Effects and Multiple Equilibria in the Risky Behavior of Friends," *Review of Economics and Statistics*, **95**, 4, 1130-1149.

Carpenter, C. and Dobkin, C. (2011), "The Minimum Legal Drinking Age and Public Health," *Journal of Economic Perspectives*, **25**, 2, 133-156.

Case, A. (1991), " Spatial Patterns in Household Demand," *Econometrica*, **59**, 4, 953-965.

Cawley, J. and Ruhm, C. J. (2011), "The Economics of Risky Health Behaviors," in Mark V. Pauly, Thomas G. Mcguire and Pedro P. Barros (Eds.), *Handbook of Health Economics, Vol. 2*, North-Holland, Amsterdam.

Centers for Disease Control and Prevention (CDC) (2016), *Tobacco Use Among Middle and High School Students â€' United States, 2011-2015.*, Morbidity and Mortality Weekly Report, April 15, 2016, **65** (14), 361-367, http://www.cdc.gov/mmwr/volumes/65/wr/mm6514a1.htm

Centers for Disease Control and Prevention (CDC) (2016), *Global Adult Tobacco Survey (GATS) Atlas* (2015), Atlanta, GA, http://www.cdc.gov/tobacco/global/gtss/tobacco_atlas

Centers for Disease Control and Prevention (CDC) (2014), *The Health Consequences of Smoking - 50 Years of Progress: A Report of the Surgeon General*, Atlanta: U.S., http://www.cdc.gov/tobacco/data_statistics/sgr/50th-anniversary/index.htm

Chaloupka, F.J. (1991), "Rational Addictive Behavior and Cigarette Smoking," *Journal of Political Economy*, **99** (4), 722–42.

Chandrasekhar, A. G. and Lewis, R. (2016), "Econometrics of Sampled Networks," mimeo Stanford University.

Cole, H. L., Mailath, G. J. and Postlewaite, A. (1992), "Social Norms, Savings Behaviour, and Growth," *Journal of Political Economy*, **100**, 1092-1125.

Cook, Philip J., and Michael J. Moore. (2000), "Alcohol," In A.J. Culyer and J.P Newhouse (eds.) Handbook of Health Economics, Volume 1. (New York: Elsevier), pp. 1629-1674.

Cutler, D. M. and Glaeser, E. L. (2007), "Social Interactions and Smoking," mimeo Harvard University and NBER working paper No. 13477.

Dahl, G., Løken, K., and Mogstad, M. (2014), "Peer Effects in Program Participation," *American Economic Review*, **104** (7), 2049-2074.

Dao, J. M., McQuown, S. C., Loughlin, S. E., Belluzzi, J. D., and Leslie, F. M. (2014), "Nicotine alters limbic function in adolescent rat by a 5-HT1A receptor mechanism," *Neuropsychopharmacology*, **36**(7), 1319.

Duncan, G.J., Boisjoly, J., Kremer, M., Levy, D.M., and Eccles, J. (2005), "Peer Effects in Drug Use and Sex among College Students", *Journal of Abnormal Child Psychology*, **33** (3), 375–85.

Eisenberg, D., Golberstein, E., and Whitlock, J. L. (2014), "Peer effects on risky behaviors: New evidence from college roommate assignments," *Journal of health economics*, **33**, 126-138.

Fafchamps, M. and Gubert, F. (2007), "The Formation of Risk Sharing Networks," *Journal of Development Economics*, **83** (2), 326-350.

Flood S., King M., Rodgers R., Ruggles S., and Warren J. R. (2018), *Integrated Public Use Microdata Series, Current Population Survey: Version 6.0* [dataset], Minneapolis, MN: IPUMS, 2018. https://doi.org/10.18128/D030.V6.0

Frederick, S., G. Lowestein, and T. O'Donoghue (2002), "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, **40** (2), 351-401.

French, M.T. and Maclean, J.C. (2006), "Underage Alcohol Use, Delinquency, and Criminal Activity," *Health Economics*, **15**, 1261–81.

Friesen, J. and B. Krauth (2007), "Sorting and inequality in Canadian schools," *Journal of Public Economics* 91, 2185-2212.

Gallet, C. and List, J. A. (2003), "Cigarette Demand: A Meta Analysis of Elasticities, " *Health Economics*, **12**, 821-835.

Gaviria, A., and Raphael, S. (2001), "School-based Peer Effects and Juvenile Behavior, " *Review of Economics and Statistics*, **83**(2), 257-268.

Glaeser, E. and J. Scheinkman (2003), "Non-Market Interactions," in *Advances in Economics and Econometrics: Theory and Applications, Eight World Congress, Vol. I*, (M. Dewatripont, L.P. Hansen, and S. Turnovsky, eds.), Cambridge University Press, 339-369.

Glaeser, E. and J. Scheinkman (2001), "Measuring Social Interactions," in S. Durlauf and P. Young (Eds.), *Social Dynamics*, Brookings Institution Press and MIT Press, Cambridge, MA.

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

Graham, B. S. (2017), "An Econometric Model of Link Formation With Degree Heterogeneity," *Econometrica*, **85** (4), 1033-1063.

Graham, B. S. (2008), "Identifying Social Interactions through Excess Variance Contrasts," *Econometrica*, **76**, 643-660.

Graham, B. S., and Hahn, J. (2005), "Identification and Estimation of the Linear-in-Means Model of Social Interactions," *Economics Letters* , **88**,1,1-6.

Gruber, J. and Kosegi, B. (2001), "Is Addiction Rational? Theory and Evidence," *Quarterly Journal of Economics*, **116**, (4), 1261–303.

Hansen, L. P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 1029-1054.

Hanushek, E.A., Kain, J.F. and S.G. Rivkin (2002), "Inferring program effects for special populations: Does special education raise achievement for students with disabilities?" *Review of Economics and Statistics*, 84, 584-599.

Horn, R. A. and Johnson, C. R. (2013), *Matrix Analysis*, second edition, Cambridge: Cambridge University Press.

Hoxby, C. M. (2000a), "The Effects Of Class Size On Student Achievement: New Evidence From Population Variation," *The Quarterly Journal of Economics*, **115**, 4, 1239-1285.

Hoxby, C. M. (2000b), "Peer Effects in the Classroom: Learning from Gender and Race Variation," NBER Working Paper 7867.

Hsieh, C-S. and Lee, L. F. (2015), "A Social Interactions Model with Endogenous Friendship Formation and Selectivity," *Journal of Applied Econometrics*, **31** (2), 301-319.

Hsieh, C. S., and Van Kippersluis, H. (2018), 'Smoking initiation: Peers and personality," *Quantitative Economics*, **9**(2), 825-863.

Jackson, M. O. (2014), "Networks in the Understanding of Economic Behaviors," *Journal of Economic Perspectives*, **28**, 4, 3-22.

Jones, S. R. G. (1984), *The Economics of Conformism*, Oxford: Basil Blackwell.

Johnson, R. A. and Wichern, D. W. (2002), *Applied Multivariate Statistical Analysis*, 6th edition, Prentice Hall. Upper Saddle River, NJ.

Kandel, D. B. (1978), "Homophily, Selection, and Socialization in Adolescent Friendships," *American Journal of Sociology*, **84**, 427-436.

Katz, L., Kling, A. and Liebman, J. (2001), "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment," *Quarterly Journal of Economics*, **CXVI**, 607-654.

Kelejian, H. H. and Prucha, I. R. (1999), "A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model," *International Economic Review*, **40**, 509-533.

Kendel, D.S. and Ribar, D. (1994), "Alcohol Consumption and Young Adults' Socioeconomic Status," *Brookings Papers on Economic Activity - Micro*, 119–61.

Kendel, D.S. and Sindelar, J. (2011), "Economics of Health Behavior and Addictions: Contemporary Issues and Policy Implications,"in S. Glied and P.C. Smith (eds.), *The Oxford Handbook of Health Economics*, Oxford, Oxford University Press.

Kling, J., Ludwig, J. and Katz, L. (2005), "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment," *Quarterly Journal of Economics*, **120**(1), 87-130.

Koopmans, T. (1953), "Identification Problems in Economic Model Construction," In Studies in Econometric Method, edited by W. C. Hood and T. Koopmans, 27- 48. New York: Wiley.

Koopmans, T. C. and Reiersøl, O. (1950), "The Identification of Structural Characteristics," *Annals of Mathematical Statistics*, **21**.

Koopmans, T. C.. Rubin, H., and Leipnik, R. B. (1950), "Measuring the Equation Systems of Dynamic Economics," Chapter II of *Statistical Inference in Dynamic Economic Models* (T. C. Koopmans, ed.),New York: John Wiley and Sons, 1950 (Cowles Commission Monograph No.10)

Krauth, B. V. (2006), "Simulation-based estimation of peer effects," *Journal of Econometrics*, **133**, 243-271.

Kremer, M. and Levy, D. (2008), "Peer Effects and Alcohol Use among College Students," *The Journal of Economic Perspectives*, **22**,3, 189-206.

Lavy, V., Paserman, D. and A. Schlosser (2012), "Inside the black box of ability peer effects: Evidence from variation in low achievers in the classroom," *Economic Journal* 122, 208-237.

Lavy, V. and A. Schlosser (2011), "Mechanisms and impacts of gender peer effects at school,"*American Economic Journal: Applied Economics* 3, 1-33.

Lee, L. F. (2003), "Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances", *Econometric Reviews*, **22**, 307-335.

Lee, L. F., Li, J., and Lin, X. (2014), "Binary choice models with social network under heterogeneous rational expectations", *Review of Economics and Statistics*, **96** (3), 402-417.

Liu, X. and L.F. Lee (2010), "GMM Estimation of Social Interaction Models with Centrality,", *Journal of Econometrics*, **159**, 99-115.

Liu, X., Patacchini, E., and Rainone, E. (2017), "Peer effects in Bed Time Decisions among Adolescents: a Social Network Model with Sampled Data,", forthcoming in *The Econometric Journal*.

Ludwig, J., Hirschfeld, P. and Duncan, G. (2001), " Urban Poverty and Juvenile Crime: Evidence from a Randomized Housing-Mobility Experiment," *Quarterly Journal of Economics*, **CXVI** (2), 665-679.

Lundborg (2006), " Having the Wrong Friends? Peer Effects in Adolescent Substance Use," *Journal of Health Economics*, **25** (2), 214–33.

McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001), "Birds of a feather: Homophily in social network," *Annual review of sociology*, **27**(1), 415-444.

Malone, P. S., Northrup, T. F., Masyn, K. E., Lamis, D. A., and Lamont, A. E. (2012), "Initiation and persistence of alcohol use in United States Black, Hispanic, and White male and female youth," *Addictive behaviors*, **37**(3), 299-305.

Manski, C. (1993), "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, **60**, 531-42.

Manski, C. (2008), *Identification for Prediction and Decision*, Harvard University Press, Cambridge.

Matzkin, R. (2007), "Nonparametric Identification," In *Handbook of Econometrics*, vol. **6**, ed. J. Heckman and E. Leamer. Amsterdam: North-Holland.

Mayer, A. and Puller, S. L. (2008), "The Old Boy (and Girl) Network: Social Network Formation On University Campuses," *Journal of Public Economics*, **92**, 329-347.

Mele, A. (2017a), "A Structural Model of Dense Network Formation," *Econometrica*, **85**, 3, 825-850.

Mele, A. (2017b), "Segregation in Social Networks: A Structural Approach," working paper, Johns Hopkins University.

Moffitt, R. A. (2001), "Policy Interventions, Low-Level equilibria, and Social Interactions," in S. Durlauf and P. Young (Eds.), *Social Dynamics*, Brookings Institution Press and MIT Press, Cambridge, MA.

Mokdad, A. H., Marks, J. S., Stroup, D. F., and Gerberding, J. L. (2004), "Actual Causes of Death in the United States, 2000", *JAMA*, **291** (10), March 10, 1238-1245.

Mokdad, A. H., Marks, J. S., Stroup, D. F., and Gerberding, J. L. (2005), "Correction: Actual Causes of Death in the United States, 2000.", *JAMA*, **293** (3).

Moody, J., (2001), "Race, School Integration, and Friendship Segregation in America", *American Journal of Sociology*, **107**, 3, 679–716.

Mullahy, J. and Sindelar, J.L., (1996), "Employment, Unemployment, and Problem Drinking", *Journal of Health Economics*, **15**, 3, 409–34.

Nakajima, R. (2007), "Measuring Peer Effects on Youth Smoking Behavior", *Review of Economic Studies*, **74**, 897-935.

National Institute on Alcohol Abuse and Alcoholism (NIAAA) (2006), "Underage Drinking,", *Alcohol Alert,* **67**, http://pubs.niaaa.nih.gov/publications/AA67/AA67.htm

National Institute on Alcohol Abuse and Alcoholism (NIAAA) (2015), "2015 National Survey on Drug Use and Health,", available at https://www.samhsa.gov/data/sites/default/files/NSDUH-TrendBreak-2015.pdf

Neyman, J. and Scott, E. L. (1948), "Consistent Estimates Based on Partially Consistent Observations", *Econometrica,* **16**(1), 1-32.

Olivetti, C., Patacchini, E. and Y. Zenou (2018). "Mothers, Peers and Gender-Role Identity", mimeo Cornell University.

Orphanides, A. and Zervos, D. (1948), "Rational Addiction with Learning and Regret", *Journal of Political Economy*, **103**, (4), 739–58.

Özgür, O., Bisin, A. and Bramoullé, Y. (2019), " Dynamic Linear Economies with Social Interactions," mimeo Melbourne Business School, and NYU.

Patacchini, E., Rainone, E., and Zenou, Y. (2011), "Dynamic Aspects of Teenage Friendships and Educational Attainment," CEPR working paper DP8223.

Patacchini, E., and Y. Zenou (2016), "Social networks and parental behavior in the intergenerational transmission of religion," *Quantitative Economics* 7(3), 969-995.

de Paula, À., Richards-Shubik, S., and Tamer, E. (2018), "Identifying Preferences in Networks with Bounded Degree," forthcoming in *Econometrica*, CeMMAP Working Paper 54/16.

Persico, N., Postlewaite, A., and Silverman, D. (2004), "The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height," *Journal of Political Economy*, **112** (5), 1019-1053.

Peski, M. (2007), "Complementarities, Group Formation, and Preferences for Similarity," mimeo.

Qu, X. and Lee, L. (2015), "Estimating a spatial autoregressive model with an endogenous spatial weight matrix," *Journal of Econometrics*, **184** (2), 209-232.

Reiff, J. (2018), "A Model of Addiction and Social Interactions," forthcoming in *Economic Inquiry*.

Renna, F. (2001), "The Economic Cost of Teen Drinking: Late Graduation and Lowered Earnings," *Health Economics*, **16**, 407–19.

Robert, C. and G. Casella (2004), *Monte Carlo Statistical Methods*, 2nd ed, Springer-Verlag, New York.

Sacerdote, B. (2001), "Peer Effects with Random Assignments: Results from Dartmouth Roommates," *Quarterly Journal of Economics*, **116**, 681-704.

Saluter, Arlene F., (1994), *Marital Status and Living Arrangements: March 1994*, U.S. Bureau of the Census, Current Population Reports, Series P20-484, U.S. Government Printing Office, Washington, DC, 1996.

Sheng, S. (2018), "A Structural Econometric Analysis of Network Formation Games Through Subnetworks," forthcoming in *Econometrica*, mimeo UCLA.

Stocker, C. M. and Youngblade, L. (1999), "Marital Conflict and Parental Hostility: Links with Children's Sibling and Peer Relationships," *Journal of Family Psychology*, **13**(4), 598-609.

Terza, J.V. (2002), "Alcohol Abuse and Employment: A Second Look," *Journal of Applied Econometrics*, **17**, 393–404.

Todd, P. E. and Wolpin, K. I. (2006), "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility," *American Economic Review*, **96** (5), 1384-1417.

U.S. Department of Agriculture and U.S. Department of Health and Human Services (2015), "2020 Dietary Guidelines for Americans, 2015", 8th ed., Appendix 9, http://health.gov/dietaryguidelines/2015/guidelines/appendix-9/

Wasserman, S. and Faust, K. (1994), *Social Network Analysis: Methods and Applications*, Cambridge University Press.

Waters, T.M. and Sloan, F.A. (1995), "Why Do People Drink? Test of the Rational Addiction Model," *Applied Economics*, **27** (5), 727–36 .

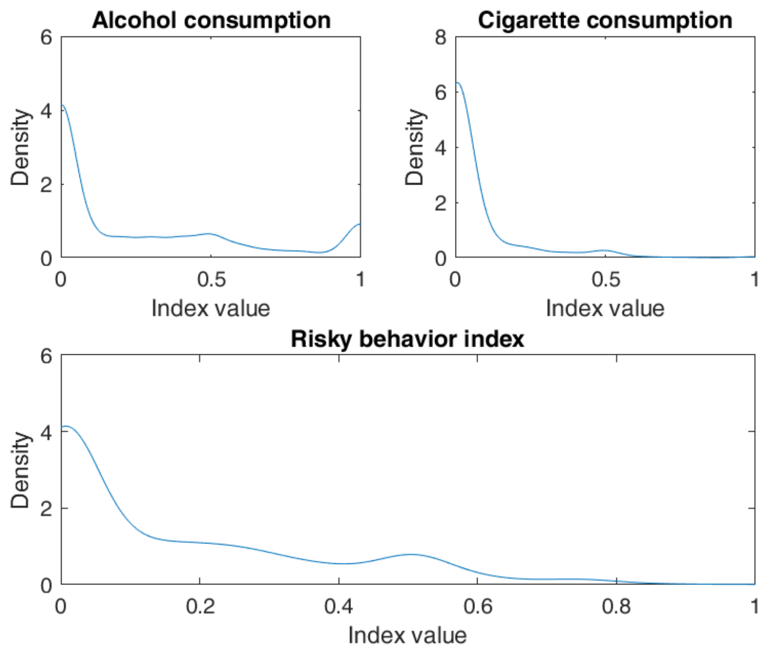
World Health Organization (2009), *WHO Global Health Risks, 2009*, https://www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf

World Health Organization (2012), *WHO Global Report: Mortality Attributable to Tobacco, 2012*, http://apps.who.int/iris/bitstream/10665/44815/1/9789241564434_eng.pdf

World Health Organization (2014), *WHO Global Status Report on Alcohol and Health, 2014*, http://www.who.int/substance_abuse/publications/global_alcohol_report/msb_gsr_2014_1.pdf?ua=1

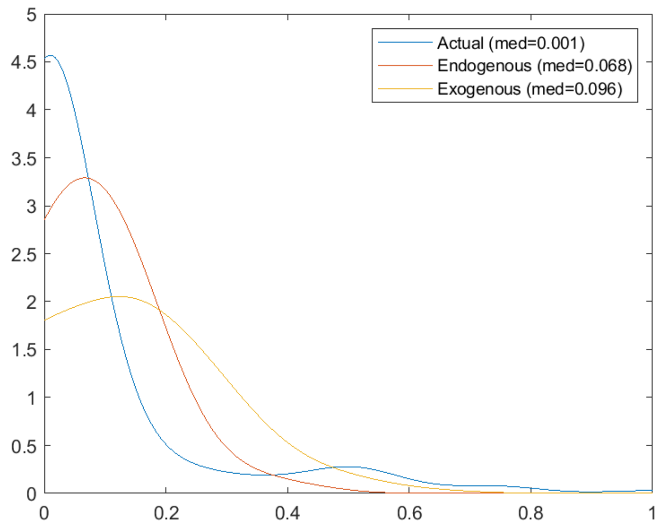
Figures and Tables

Figure 1: **Distribution of students by alcohol and cigarette consumption**



This figure graphs the empirical densities of alcohol consumption, cigarette consumption and the risky behavior index. Densities are smoothed using a kernel density estimator.

Figure 2: **Distributions of actual and predicted values of the risky behavior index**



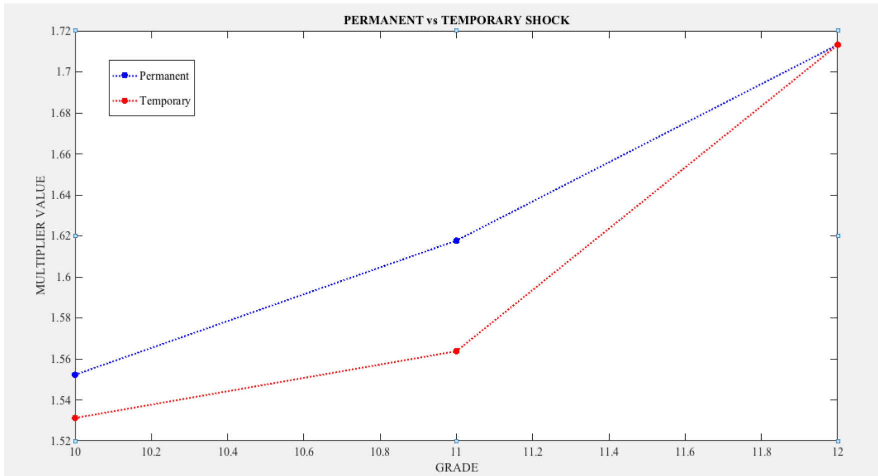
This figure graphs the distributions of actual and predicted values of the risky behavior index. we predict equilibrium risky behavior index for students in the hold-out sample consisting of grades 7, 8, and 9. Specifically, (i) under the exogeneity of the network $\mathbf{G}$, we use the reduced-form equilibrium equations 6 jointly, for $t = T-3, T-4$, by linking the structural equations and the sample as in Table 1, and (ii) under the possible endogeneity of the network $\mathbf{G}$, we use the reduced-form equilibrium equations 10 jointly, for $t = T-3, T-4$, once again by linking the structural equations and the sample as in Table 1. For both specifications, the predicted student behaviors are generated recursively, using the estimated parameters in Table 2, the baseline controls (and network $\mathbf{G}$) for grades 7, 8 and 9, and students' initial risky behavior index values for $t = T-5$, observed in the data.

Table 1: **Linking the structural equations and the sample**

	Y_{T-5}	Y_{T-4}	Y_{T-3}	Y_{T-2}	Y_{T-1}	Y_T
Eq'n 3 at $t = T$					g11 WI	g12 WII
Eq'n 3 at $t = T-1$				g10 WI	g11 WII	
Eq'n 3 at $t = T-2$			g9 WI	g10 WII		
Eq'n 3 at $t = T-3$		g8 WI	g9 WII			
Eq'n 3 at $t = T-4$	g7 WI	g8 WII				

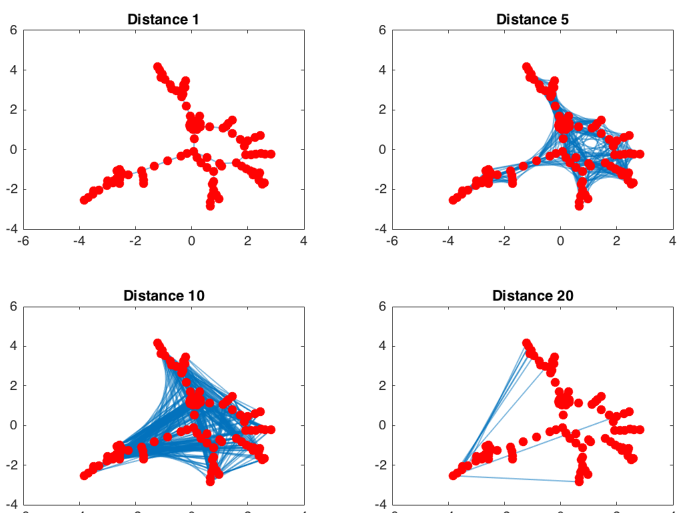
This table summarizes the structure of our sample.

Figure 3: **Temporary and permanent instantaneous multiplier effects**



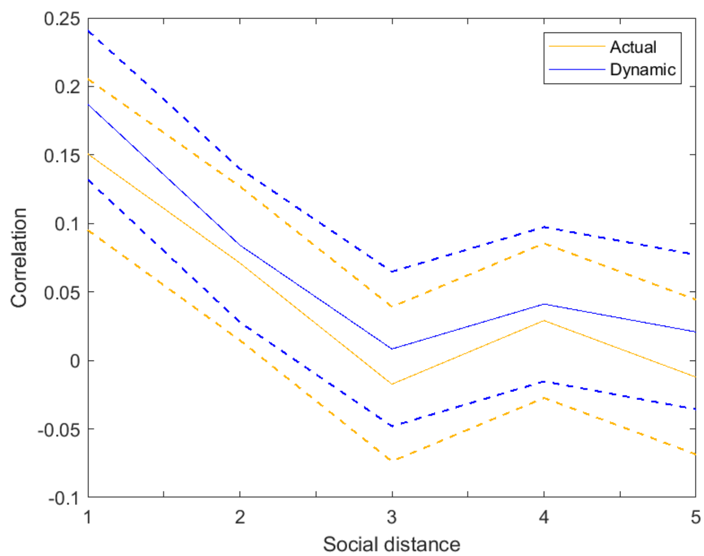
This figure illustrates temporary and permanent instantaneous multiplier effects. The sample means of dynamic multiplier values are computed across grades using actual parameter estimates from Table 2 column 2. Here, we plot the permanent effect relative to the temporary effect.

Figure 4: **Adjacency matrices by social distance**



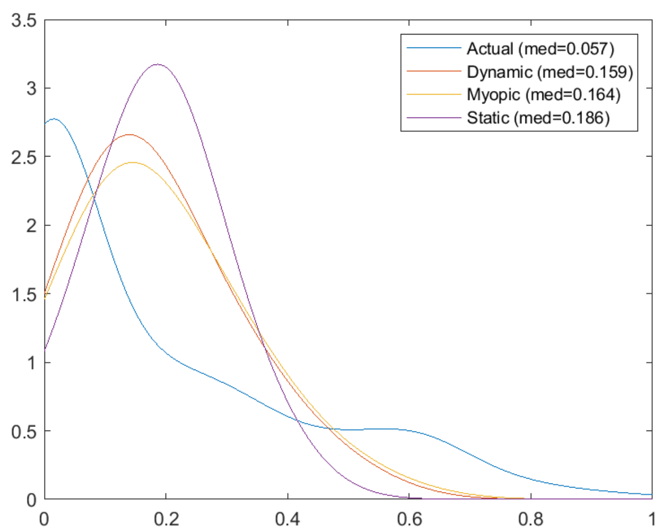
This figure graphs the actual series of geodesic student adjaceny matrices $\mathbf{G}_k$ of increasing social distance with $k = 1, 5, 10,$ and 20, obtained using our algorithm, starting with the largest student network component in our sample.

Figure 5: **Behavioral corrrelation and social distance**



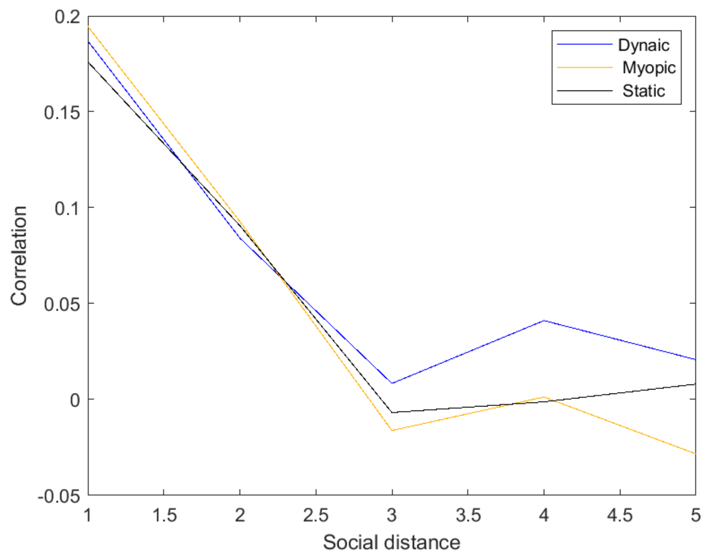
This figure illustrates the correlation between students' choices in each period and the average level of those of their peers at different social distances, as represented by adjacency matrix $\mathbf{G}_k$, $k = 1, 2, \ldots, d(\mathbf{G})$. Our method of internal validation works by predicting students' risky behavior indices at different time periods. The setup is the same as the one used in Section 5.1. We use the reduced-form equilibrium equations 10 jointly, for $t = T-3, T-4$, by linking the structural equations and the sample as in Table 1. In particular, predicted behaviors are generated recursively for each grade starting from grade 10 ($t = T-2$) and going up to grade 12 ($t = T$). Predicted values for dynamic model are obtained by equation 10 with parameters calibrated using the values of Table 2 Column 2. We use the actual and the predicted dependent variables to construct the correlation between each individual choice and the average level of individual's peers' choices at different social distances represented by k. Dotted lines represents 95 percent confidence interval for the estimated correlations.

Figure 6: **Distributions of actual and predicted values of the risky behavior index**



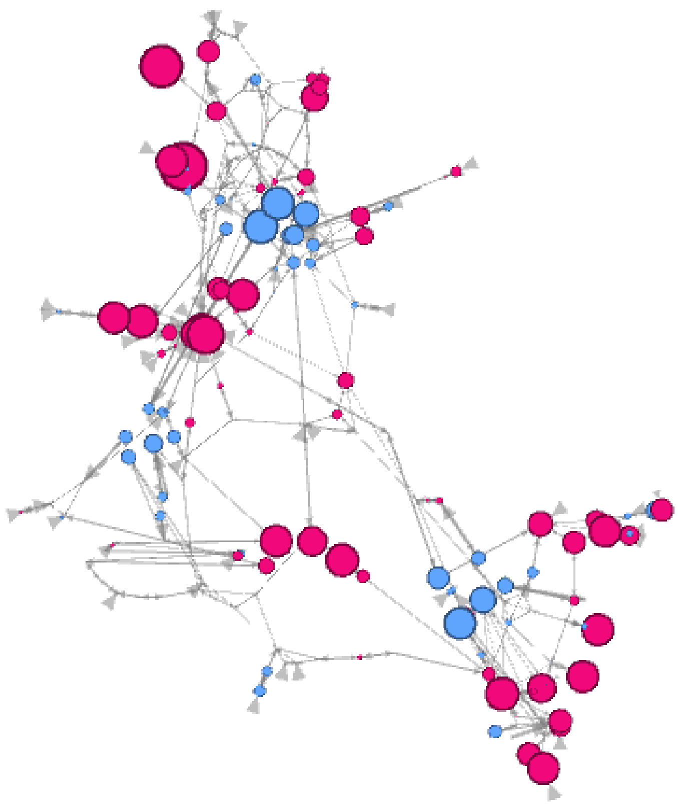
This figure graphs distributions of actual and predicted values of the risky behavior index. In particular, predicted behaviors are generated recursively for each grade starting from grade 10 ($t = T - 2$) and going up to grade 12 ($t = T$). Predicted values for the dynamic baseline model are obtained by the structural model 10 with parameters calibrated using the values of Table 2 Column 2. Predicted values for the myopic and static model are obtained by the structural model 10 with parameters calibrated using the values of Table 8 Column 3 and 4.

Figure 7: **Behavioral correlation and social distance**



This figure illustrates the correlation between students' choices in each period and the average level of those of their peers at different social distances, as represented by adjacency matrix $\mathbf{G}_k$, $k = 1, 2, \ldots, d(\mathbf{G})$. Our method of internal validation works by predicting students' risky behavior indices at different time periods. The setup is the same as the one used in Section 5.1. We use the reduced-form equilibrium equations 10 jointly, for $t = T-3, T-4$, by linking the structural equations and the sample as in Table 1. In particular, predicted behaviors are generated recursively for each grade starting from grade 10 ($t = T-2$) and going up to grade 12 ($t = T$). Predicted values for dynamic model are obtained by equation 10 with parameters calibrated using the values of Table 2 Column 2. Predicted values for myopic and static model are obtained by equation 10 with parameters calibrated using values of Table 8 columns 1 and 2.

Figure 8: **Gender assortativeness**



This figure depicts friendship linkages in the larger network in our data (286 nodes with diameter 24) by using different colors for nodes indicating students of different gender. The picture reveals that indeed social interactions are assortative by gender. Nodes represented by a light blue (resp., purple) dot correspond to female students (resp., male students).Sizes of nodes are proportional to individuals' risky index levels.

Table 2: **Dynamic recursive model**

Dep. Var. Risky Behavior Index	Exogenous network (1)	Endogenous network (2)
Addiction effect (α_1)	0.3692***	0.4299***
	(0.0095)	(0.1552)
Own effect (α_2)	0.3499***	0.1538***
	(0.0170)	(0.0210)
Peer effect (α_3)	0.2809***	0.4163**
	(0.0172)	(0.1665)
Discount factor (δ)	0.5038***	0.4925**
	(0.0640)	(0.2017)
Selectivity (ψ)		-0.0620
		(0.5942)
Student characteristics	Yes	Yes
Peers' characteristics	Yes	Yes
Networks fixed effects	Yes	Yes
N. Obs.	1,207	1,207

This table reports GMM estimates of the structural models 6 and 10. In Column 1 the network **G** is assumed to be exogenous, while in Column 2 network formation is assumed to satisfy the linear dyadic structure in equation 9. Students' characteristics are listed in Table G1. The peers' characteristics are calculated as friends' averages of the included variables. Heteroskedasticity-robust numerical standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: **Model validation**

	Data	Model Exogenous network	*p-value*	Model Endogenous network	*p-value*
Overall	0.0720	0.0612	0.3717	0.0719	0.9932
	(0.1649)	(0.1673)		(0.1040)	
If female = 1	0.1027	0.0550	0.0533	0.0684	0.1064
	(0.1856)	(0.1753)		(0.1178)	
If Parent college degree = 1	0.0952	0.0573	0.0385	0.0757	0.2259
	(0.1839)	(0.1631)		(0.1154)	
If two-parents = 1	0.0924	0.0588	0.0721	0.0736	0.261
	(0.1806)	(0.1550)		(0.1142)	
If Black or African American=1	0.1366	0.0665	0.1384	0.1153	0.6291
	(0.2322)	(0.1505)		(0.1170)	
If pocket money = 1	0.0932	0.0648	0.0932	0.0799	0.3656
	(0.1784)	(0.1635)		(0.1129)	
If Alcohol/tobacco at home = 1	0.0847	0.0141	0.2674	0.0362	0.3642
	(0.2038)	(0.1705)		(0.0929)	
If height > median	0.1116	0.0690	0.0758	0.0791	0.1358
	(0.1963)	(0.1539)		(0.1128)	
N. Obs.	381	381	381	381	

This table reports the comparison of mean predicted and mean actual risky behavior index values for the overall sample and for several subgroups defined by gender, race, parental education, and others. Means and standard deviations (in parentheses) of observed and predicted risky behavior index are reported. The reported p-values refer to zero null t-tests for differences in means between observed and simulated risky behavior.

Table 4: **The multiplier**

	$\bar{m}_{10,10}$	$\bar{m}_{10,11}$	$\bar{m}_{10,12}$	$\bar{m}_{11,11}$	$\bar{m}_{11,12}$	$\bar{m}_{12,12}$
Temporary shock	1.5312	1.0293	0.7581	1.5638	1.1517	1.7132
Permanent shock	1.5523	2.6070	3.6767	1.6177	2.9042	1.7132

This table reports the sample means of multiplier values ($\bar{m}_{t,\tau}$) for the baseline estimated parameter values, with endogenous network in Table 2, Column 2.

Table 5: **The expectation multiplier**

$\bar{m}_{12,10}$	$\bar{m}_{12,11}$	$\bar{m}_{12,12}$
-0.006	0.054	1.7132

This table reports the sample means of multiplier values ($\bar{m}_{t,\tau}$) in grade 10, 11 and 12 induced by an anticipated shock to the preferences of all agents in grade 12, calibrated to the baseline estimated parameters with endogenous network in Table 2, Column 2.

Table 6: **The expectation multiplier - under different parameter values**

$a_2 = 0.1$	$\alpha_3 = 0.8(\alpha_1 = 1 - \alpha_2 - \alpha_3)$	$\alpha_3 = 0.7$	$\alpha_3 = 0.6$	$\alpha_3 = 0.5$	$\alpha_3 = 0.4$
$\bar{m}_{12,10}$	-0.089	-0.089	-0.067	-0.048	-0.0346
$\bar{m}_{12,11}$	-0.633	-0.275	-0.090	0.004	0.052
$\bar{m}_{12,12}$	5	3.3333	2.5	2	1.6667
$a_2 = 0.2$	$\alpha_3 = 0.6(\alpha_1 = 1 - \alpha_2 - \alpha_3)$	$\alpha_3 = 0.5$	$\alpha_3 = 0.4$	$\alpha_3 = 0.3$	$\alpha_3 = 0.2$
$\bar{m}_{12,10}$	-0.023	-0.014	-0.006	-0.001	-0.001
$\bar{m}_{12,11}$	0.059	0.109	0.135	0.146	0.148
$\bar{m}_{12,12}$	2.5	2	1.6667	1.4286	1.25
$a_2 = 0.4$	$\alpha_3 = 0.5(\alpha_1 = 1 - \alpha_2 - \alpha_3)$	$\alpha_3 = 0.4$	$\alpha_3 = 0.3$	$\alpha_3 = 0.2$	$\alpha_3 = 0.1$
$\bar{m}_{12,10}$	0.024	0.035	0.043	0.048	0.050
$\bar{m}_{12,11}$	0.201	0.227	0.238	0.240	0.238
$\bar{m}_{12,12}$	1.6667	1.4286	1.25	1.1111	1
$a_3 = 0$	$\alpha_1 = 0.1(\alpha_2 = 1 - \alpha_1 - \alpha_3)$	$\alpha_1 = 0.2$	$\alpha_1 = 0.3$	$\alpha_1 = 0.4$	$\alpha_1 = 0.5$
$\bar{m}_{12,10}$	0.007	0.019	0.031	0.038	0.038
$\bar{m}_{12,11}$	0.082	0.137	0.172	0.192	0.198
$\bar{m}_{12,12}$	1	1	1	1	1
$a_3 = 0$	$\alpha_1 = 0.6(\alpha_2 = 1 - \alpha_1 - \alpha_3)$	$\alpha_1 = 0.7$	$\alpha_1 = 0.8$	$\alpha_1 = 0.9$	
$\bar{m}_{12,10}$	0.0299	0.013	-0.011	-0.031	
$\bar{m}_{12,11}$	0.192	0.172	0.137	0.082	
$\bar{m}_{12,12}$	1	1	1	1	

This table reports the simulated sample means of expectation multiplier values ($\bar{m}_{t,\tau}$) with actual school networks $\mathbf{G}$ and with $\delta = 0.99$.

Table 7: **The (semi-) elasticities of "availability of alcohol or cigarettes at home"**

	10 on 10	10 on 11	10 on 12	11 on 11	11 on 12	12 on 12
Temporary shock						
Direct Effect	0.0492	0.0212	0.0091	0.2014	0.0866	0.4393
Equilibrium Effect	0.0996	0.0736	0.0556	0.3398	0.2580	0.7895
Permanent shock						
Direct Effect	0.0492	0.0704	0.0795	0.0004	0.2880	0.4393
Equilibrium Effect	0.0996	0.1757	0.2502	0.3399	0.6301	0.7895

This table reports the sample means of the (semi-) elasticities of a change of "availability of alcohol or cigarettes at home" on child risky behavior implied by the direct effect versus the equilibrium effect for the baseline estimated parameter values, with endogenous network in Table 2, Column 2.

Table 8: **Myopic and static vs. dynamic baseline estimates**

Dep. Var. Risky Behavior Index	Endogenous network		
	Myopic	*Static*	*Dynamic baseline*
	$\delta = 0$	$\alpha_1 = 0$, $\delta = 0$	
	(1)	(2)	(3)
Addiction effect (α_1)	0.4233***	-	0.4299***
	(0.1225)		(0.1552)
Own effect (α_2)	0.1468***	0.6453***	0.1538***
	(0.0134)	(0.0999)	(0.0210)
Peer effect (α_3)	0.4299***	0.3547***	0.4163**
	(0.1347)	(0.0999)	(0.1665)
Discount factor (δ)	-	-	0.4925**
			(0.2017)
Selectivity (ψ)	0.2615	0.0888	-0.0620
	(1.1107)	(0.1680)	(0.5942)
Student characteristics	Yes	Yes	Yes
Peers' characteristics	Yes	Yes	Yes
Networks fixed effects	Yes	Yes	Yes
N. Obs.	1,207	1,207	1,207

This table reports GMM estimates of the structural model 10. In Column 1 we restrict the model by setting $\delta = 0$, while in Column 2 we restrict the model by setting $\alpha_1 = 0$ and $\delta = 0$. Column 3 reports baseline estimates presented in Table 2 Column 2. Students' characteristics are listed in Table G1. The peers' characteristics are calculated as friends' averages of the included variables. Heteroskedasticity-robust numerical standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: **Alternative definition of peers**

Dep. Var. Risky Behavior Index	*All Sample* (1)	*Regression Sample* (2)
Addiction effect (α_1)	0.4139***	0.2937***
	(0.0002)	(0.0174)
Own effect (α_2)	0.3260***	0.3603***
	(0.0003)	(0.0265)
Peer effect (α_3)	0.2601***	0.3460***
	(0.0004)	(0.0381)
Discount factor (δ)	0.4210***	0.5206***
	(0.0009)	(0.0833)
Student characteristics	Yes	Yes
Peers' characteristics	Yes	Yes
School fixed effects	Yes	Yes
N. Obs.	1,759	1,207

This table reports GMM estimates of the structural models using the alternative definition of peers defined in 15 (Column 1). In Column 2 we condition the estimation to the sample used in the baseline estimation (Table 2). Students' characteristics are listed in Table G1. The peers' characteristics are calculated as friends' averages of the included variables. Heteroskedasticity-robust numerical standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Technical Appendix

A Formal Model

A dynamic linear economy with social interactions is populated by a finite number of agents $i = 1, \ldots, N$. Agents live for the whole duration of the economy $t = 1, \ldots, T$. Each agent i chooses an action y_{it} at time t from a closed and convex set $Y \subset \mathbb{R}$ after having observed a preference shock $\theta_{it} \in \Theta \subset \mathbb{R}$, a closed and convex set of possible types (we denote with $\mathbf{y}_t \in \mathbf{Y}$ and $\theta_t \in \mathbf{\Theta}$ the corresponding N-dimensional vectors stacking all agents).[45] Let $\theta := (\theta_t) := (\theta_{it})_{i=1,\ldots,N,\, t\geq 1}$ be the stochastic process of agents' types, which is assumed, with no loss of generality, to be defined, on the canonical probability space $(\mathbf{\Theta}, \mathcal{F}, \mathbb{P})$, where $\mathbf{\Theta} := \left\{(\theta_1, \theta_2, \cdots) : \theta_t \in \Theta^N\,,\, t = 1, 2, \cdots, T\right\}$ is the space of sample paths. The sequence $(\mathcal{F}_1, \mathcal{F}_2, \cdots, \mathcal{F}_T)$ of Borel sub-σ-fields of $\mathcal{F}$ is a filtration in $(\mathbf{\Theta}, \mathcal{F})$, that is $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}$. Finally, the process $\theta = (\theta_1, \theta_2, \cdots, \theta_{\mathbf{T}})$ is adapted to the filtration $(\mathcal{F}_t : t \geq 1)$, that is, for each t, θ_t is measurable with respect to $\mathcal{F}_t$. Finally, $P : \mathcal{F} \to [0, 1]$ is a probability measure where $P((\theta_1, \ldots, \theta_t \in A) := P(\{\theta \in \Theta : (\theta_1, \ldots, \theta_t) \in A\})$, all $A \in \mathcal{F}_t$.

The *social network* is represented by an $N \times N$ matrix $\mathbf{G} = [g_{ij}]$, where g_{ij} indicates the friendship relationship between i and j. Following the convention in the social networks literature, $\mathbf{G}$ has a main diagonal of zeros. We consider *row-normalized* $\mathbf{G}$'s, i.e., if i nominates j as one of his friends, then $g_{ij} > 0$, otherwise $g_{ij} = 0$, and $\sum_j g_{ij} = 1$. In other words, we consider a *directed network*, in which each agent interacts directly with his friends, and friendship of i with j does not imply friendship of j with i.

The instantaneous preferences of an agent $i \in N$ are represented by the utility function

$$\begin{aligned} u_i(y_{it-1}, \mathbf{y}_t, \theta_t, \mathbf{G}) \quad := \quad & -\alpha_1(y_{it-1} - y_{it})^2 - \alpha_2(\theta_{it} - y_{it})^2 \\ & -\alpha_3 \sum_{j=1}^{N} g_{ij}(y_{jt} - y_{it})^2 \end{aligned} \tag{A.1}$$

and $\alpha_1, \alpha_2, \alpha_3 \geq 0$ are parameters. We require that either α_1 or α_2 be strictly positive.

The precise timing of events is as follows: Before each agent's time t choice, the history of previous choices, $\mathbf{y}^{t-1} = (\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_{t-1})$, and the history of preference shocks, $\theta^t = (\theta_1, \ldots, \theta_t)$ (including the period-t realization), are observed by all agents. After time t choices are made, $\mathbf{y}_t = (y_{it})_{i=1}^N$ becomes common knowledge and the economy moves to time $t + 1$.

Each agent i chooses strategy $y_i = (y_{it})$, where for each t, $y_{it} : \mathbf{Y}^t \times \mathbf{\Theta}^t \to Y$, to maximize

$$E\left[\sum_{t=1}^{T} \delta^{t-1}\, u_i(y_{it-1}, \mathbf{y}_t, \theta_t, \mathbf{G}) \,\middle|\, (\mathbf{y}^0, \theta^1)\right] \tag{A.2}$$

given $\{y_j\}_{j\neq i}$, the strategies of other agents, and any finite initial history $(\mathbf{y}^0, \theta^1) \in \mathbf{Y} \times \mathbf{\Theta}$.

[45] All of our results are easily extended to the case in which choice and type variables are multidimensional.

Definition A. 1 *A Subgame Perfect Equilibrium of a dynamic linear conformity economy is a family of maps* $\{y_i^*\}_{i=1}^N$ *such that for all* $i = 1, \dots, N$ *and for all* $(\mathbf{y}^{t-1}, \theta^t) \in \mathbf{Y}^t \times \mathbf{\Theta}^t$

$$y_{it}^* \left(\mathbf{y}^{t-1},\, \theta^t\right) \in argmax_{y_{it} \in Y}\; E\left[\sum_{t=1}^{T} \delta^{t-1}\, u_i(y_{it-1}, \left(y_{it}, \{y_{jt}^*\}_{j \neq i}\right), \theta_t, \mathbf{G}) \,\middle|\, (\mathbf{y}^0, \theta^1)\right] \tag{A.3}$$

B Proof of Proposition 1 (Existence and Uniqueness):

Proposition 1 (Equilibrium Existence and Uniqueness) *Consider a* dynamic linear economy with social interactions and preferences for conformity, *with* $\alpha_1 + \alpha_2 > 0$. *There exists a unique subgame perfect equilibrium. Individuals' equilibrium choices at time* T *are uniquely determined by*

$$\mathbf{y}_T = \underbrace{\left[\Delta_T \mathbf{I} - \alpha_3\, \boldsymbol{G}\right]^{-1}}_{\mathbf{B}_T} \times \left(\alpha_1\, \mathbf{y}_{T-1} + \alpha_2\, \theta_T\right) \tag{B.1}$$

where $\mathbf{B}_T := [b_{ijT}]$ *is an* $N \times N$ *matrix of equilibrium coefficients. For any* $t = 1, \dots, T-1$*, individuals' optimal choices in equilibrium are uniquely given by*

$$\mathbf{y}_t = \mathbf{B}_t \left(\alpha_1\, \mathbf{y}_{t-1} + \alpha_2\, \theta_t + \alpha_2\, \mathbf{D}_t\right). \tag{B.2}$$

Each $\mathbf{B}_t$, $t < T$, *depends only on the future equilibrium coefficient matrices* $(\mathbf{B}_\tau)_{\tau > t}$ *and is computed recursively as the unique fixed point of contraction maps induced by the first-order conditions of problem (A.3).*

Proof: - Step 1: Existence and uniqueness at $t = T$. Let any history of previous choices, $\mathbf{y}^{T-1} = (\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{T-1})$ and of preference shocks, $\theta^T = (\theta_1, \dots, \theta_T)$, and other agents' choices $(y_{jT})_{j \neq i}$ be given. Agent i solves

$$\max_{y_{iT} \in Y} \left\{ -\alpha_1 (y_{iT-1} - y_{iT})^2 - \alpha_2 (\theta_{iT} - y_{iT})^2 - \alpha_3 \sum_{j=1}^{N} g_{ij} (y_{jT} - y_{iT})^2 \right\} \tag{B.3}$$

The first order condition

$$2\left[\alpha_1 \left(y_{iT-1} - y_{iT}\right) + \alpha_2 \left(\theta_{iT} - y_{iT}\right) + \alpha_3 \sum_{j=1}^{N} g_{ij} (y_{jT} - y_{iT})\right] = 0$$

implies that

$$y_{iT} = \Delta_T^{-1} \left(\alpha_1\, y_{iT-1} + \alpha_2\, \theta_{iT} + \alpha_3 \sum_{j=1}^{N} g_{ij}\, y_{jT}\right) \tag{B.4}$$

where $\Delta_T := \alpha_1 + \alpha_2 + \alpha_3 > 0$. This choice is feasible (in Y) since it is a convex combination of elements of

Y, a convex set by assumption. The objective function (B.3) is strictly concave in y_{iT}, thus the right-hand side of (B.4) is the unique optimizer.

Let $\mathbf{B}$ be the class of bounded $\mathcal{F}_T$-measurable functions $y : (\mathbf{Y} \times \mathbf{\Theta})^T \to Y$. The right hand side of (B.4) can be seen as an operator, call it FOC_{iT}, that maps any given collection $f = \{f_j\}$ of bounded and $\mathcal{F}_T$-measurable functions in $\mathbf{B}$ to the $\mathcal{F}_T$-measurable function $FOC_{iT}(f)$, defined as

$$FOC_{iT}(f)(\mathbf{y}^{T-1}, \theta^T) := \Delta_T^{-1} \left(\alpha_1\, y_{iT-1} + \theta_{iT} + \alpha_3 \sum_{j=1}^{N} g_{ij}\, f_j(\mathbf{y}^{T-1}, \theta^T) \right) \tag{B.5}$$

FOC_{iT} is a self-map for any i. Thus, the map $FOC_T := (FOC_{iT})_i : \mathbf{B}^n \to \mathbf{B}^n$. Endow both $\mathbf{B}$ and $\mathbf{B}^n$ with the sup norm which makes $(\mathbf{B}^n, ||\cdot||_\infty)$ a Banach space. Showing the existence of an equilibrium in the continuation given history $(\mathbf{y}^{T-1}, \theta^T)$ is equivalent to finding the fixed point of the operator $FOC_T := (FOC_{iT})_i : \mathbf{B}^n \to \mathbf{B}^n$. To that end, we show next that the map FOC_T is a contraction map. Pick $f, \hat{f} \in \mathbf{B}^n$. We have for all $\left(\mathbf{y}^{T-1}, \theta^T\right)$

$$\begin{aligned}
&\left| FOC_{iT}\left(f\right)\left(\mathbf{y}^{T-1}, \theta^T\right) - FOC_{iT}\left(\hat{f}\right)\left(\mathbf{y}^{T-1}, \theta^T\right) \right| \\
&\quad = \Delta_T^{-1} \Big| \alpha_1\, y_{iT-1} + \theta_{iT} + \alpha_3 \sum_{j=1}^{N} g_{ij}\, f_j(\mathbf{y}^{T-1}, \theta^T) \\
&\quad - \alpha_1\, y_{iT-1} - \theta_{iT} - \alpha_3 \sum_{j=1}^{N} g_{ij}\, \hat{f}_j(\mathbf{y}^{T-1}, \theta^T) \Big| \\
&\quad = \left(\frac{\alpha_3}{\Delta_T}\right) \left| \sum_{j=1}^{N} g_{ij} \left(f_j(\mathbf{y}^{T-1}, \theta^T) - \hat{f}_j(\mathbf{y}^{T-1}, \theta^T) \right) \right|
\end{aligned}$$

The coefficient $\left(\frac{\alpha_3}{\Delta_T}\right) < 1$ since either α_1 or α_2 is nonzero by assumption. But then the expression in the last line

$$\begin{aligned}
&\left(\frac{\alpha_3}{\Delta_T}\right) \left| \sum_{j=1}^{N} g_{ij} \left(f_j(\mathbf{y}^{T-1}, \theta^T) - \hat{f}_j(\mathbf{y}^{T-1}, \theta^T) \right) \right| \\
&\leq \left(\frac{\alpha_3}{\Delta_T}\right) \sum_{j=1}^{N} g_{ij} \left| f_j(\mathbf{y}^{T-1}, \theta^T) - \hat{f}_j(\mathbf{y}^{T-1}, \theta^T) \right| \\
&\leq \left(\frac{\alpha_3}{\Delta_T}\right) \sum_{j=1}^{N} g_{ij} \left\| f_j - \hat{f}_j \right\|_\infty \\
&\leq \left(\frac{\alpha_3}{\Delta_T}\right) \left\| f - \hat{f} \right\|_\infty
\end{aligned}$$

Hence FOC_T is a contraction mapping on $(\mathbf{B}^n, ||\cdot||_\infty)$. Thus, by Banach Fixed Point Theorem (see e.g., Aliprantis and Border (2006), p.95), FOC_T has a unique fixed point f^* in $\mathbf{B}^n$.

Consider now $\mathbf{B}_c$ the subset of $\mathbf{B}$ that includes families of bounded measurable linear maps as in

$$\mathbf{B}_c := \left\{ \begin{array}{l} f : \mathbf{y}^{t-1}, \theta^t \to Y \text{ s.t.} \\ f(\mathbf{y}^{t-1}, \theta^t) = \sum_{j=1}^N c_j\, y_{j,t-1} + \sum_{j=1}^N d_j\, \theta_{j,t} + \sum_{\tau=t+1}^T \sum_{j=1}^N e_{j,\tau-t}\, E\left[\theta_{j,\tau}|\theta^t\right] \\ \text{with } c_j, d_j, e_j \geq 0 \text{ and } \sum_{j=1}^N \left(c_j + d_j + \sum_{\tau=t+1}^T e_{j,\tau-t}\right) \leq 1 \end{array} \right\} \tag{B.6}$$

where each element is a linear combination of one-period before history, current and expected future preference shocks. Thanks to the linearity and inequality constraints, $\mathbf{B}_c^n$ is a closed subset of $\mathbf{B}^n$ and that FOC_T in (B.4) maps $\mathbf{B}_c^n$ into itself. Since FOC_T is a contraction mapping, its unique fixed point then lies necessarily in $\mathbf{B}_c^n$. Moreover, the existence of the unique fixed point for FOC_T in (B.4) written in matrix form

$$\Delta_T \mathbf{y}_T = \alpha_1\, \mathbf{y}_{T-1} + \alpha_2\, \theta_T + \alpha_3\, \mathbf{G}\mathbf{y}_T \tag{B.7}$$

is equivalent to the invertibility of this matrix equation.[46] Hence, the equilibrium choices vector takes the form

$$\mathbf{y}_T = \underbrace{\left[\Delta_T \mathbf{I} - \alpha_3\, \mathbf{G}\right]^{-1}}_{\mathbf{B}_T} \times (\alpha_1\, \mathbf{y}_{T-1} + \alpha_2\, \theta_T) \tag{B.8}$$

This proves that the statement of the Proposition is true for the last period (1-period economies). Next, we demonstrate that this result holds for any finite-horizon, T-period economy. Hence, the rest of the proof will use an induction argument. In any period $t = 1, \ldots, T-1$, future equilibrium policy matrices $\mathbf{B}_{t+1}, \ldots, \mathbf{B}_T$ are known. The first-order condition for agent i's problem takes the form

$$\begin{aligned} 0 \quad = \quad & \alpha_1\,(y_{i,t-1} - y_{i,t}) + \alpha_2\,(\theta_{i,t} - y_{i,t}) + \alpha_3 \sum_{j=1}^N g_{ij}\,(y_{j,t} - y_{i,t}) \\ & + E\left[\sum_{\tau=t+1}^T \delta^{\tau-t}\left(-\alpha_1\,(y_{i,\tau-1} - y_{i,\tau})\left(\frac{\partial y_{i,\tau-1}}{\partial y_{i,t}} - \frac{\partial y_{i,\tau}}{\partial y_{i,t}}\right) + \alpha_2\,(\theta_{i,\tau} - y_{i,\tau})\,\frac{\partial y_{i,\tau}}{\partial y_{i,t}} \right.\right. \\ & \left.\left. - \alpha_3 \sum_{j=1}^N g_{ij}\,(y_{j,\tau} - y_{i,\tau})\left(\frac{\partial y_{j,\tau}}{\partial y_{i,t}} - \frac{\partial y_{i,\tau}}{\partial y_{i,t}}\right)\right)\right] \end{aligned} \tag{B.9}$$

[46] Another way to see this is that since $\frac{\alpha_3}{\Delta_T} < 1$, $\Delta_T \mathbf{I} - \alpha_3\, \mathbf{G}$ is invertible. See Case (1991), footnote 5.

By iterating through future policy functions, we can write y_τ, for $\tau = t+1, \ldots, T$, as

$$\begin{aligned}
y_\tau &= \mathbf{B}_\tau \left[\alpha_1\, y_{\tau-1} + \alpha_2\, \theta_\tau + \alpha_2\, \mathbf{D}_\tau\right] \\
&= \alpha_1^2\, (\mathbf{B}_\tau \times \mathbf{B}_{\tau-1})\; y_{\tau-2} + \alpha_1\, \alpha_2\, \mathbf{B}_\tau \times \mathbf{B}_{\tau-1}\, (\theta_{\tau-1} + D_{\tau-1}) + \alpha_2\, \mathbf{B}_\tau\, (\theta_\tau + \mathbf{D}_\tau) \\
&\;\;\vdots \\
&= \alpha_1^{\tau-t}\, (\mathbf{B}_\tau \times \cdots \times \mathbf{B}_{t+1})\; y_t + \sum_{s=t+1}^{\tau} \alpha_1^{\tau-s}\, \alpha_2\, \underbrace{(\mathbf{B}_\tau \times \cdots \times \mathbf{B}_s)}_{\tau-s+1 \text{ terms}}\, (\theta_s + \mathbf{D}_s) \qquad \text{(B.10)}
\end{aligned}$$

Define $\Lambda_{t,\tau}$, for any $\tau = t+1, \ldots, T$, as

$$\Lambda_{t,\tau} := \alpha_1^{\tau-t}\, \mathbf{B}_\tau \times \ldots \times \mathbf{B}_{t+1} \qquad \text{(B.11)}$$

with the convention that $\Lambda_{t,t} := I_N$, the identity matrix. Using this latter, one can obtain the intertemporal partial derivatives as

$$\frac{\partial y_{j,\tau}}{\partial y_{i,t}} \;=\; \alpha_1^{\tau-t}\, B_{j\bullet,\tau} \times \cdots \times B_{\bullet i,t+1} \;=\; \Lambda_{ji,t,\tau} \qquad \text{(B.12)}$$

where $B_{j\bullet,\tau}$ denotes the j'th row of the $N \times N$ matrix $\mathbf{B}_\tau$ and $B_{\bullet i,t+1}$ denotes the i'th column of the $N \times N$ matrix $\mathbf{B}_{t+1}$, and $\Lambda_{ji,t,\tau}$ denotes the entry at the j'th row and i'th column of the $N \times N$ matrix $\Lambda_{t,\tau}$. Similarly, define $\Gamma_{t,\tau}$, for any $\tau = t+1, \ldots, T$, as

$$\Gamma_{t,\tau} := \sum_{s=t+1}^{\tau} \alpha_1^{\tau-s}\, (\mathbf{B}_\tau \times \cdots \times \mathbf{B}_s)\, \left(\bar{\theta}_s + \mathbf{D}_s\right) \qquad \text{(B.13)}$$

with the convention that $\Gamma_{t,t} := 0_N$, the $N \times 1$ matrix of zeros, and where for notational simplicity, $\bar{\theta}_s$ is the expected value of θ_s, conditional on period-t information. The first-order condition is linear hence we know that the total coefficient of $y_{j,t}$ is going to be given by the cross partial derivative of the objective function with respect to $y_{j,t}$ and $y_{i,t}$, i.e.,

$$\begin{aligned}
\Delta_{ii,t} &:= \alpha_1 + \alpha_2 + \alpha_3 + \sum_{\tau=t+1}^{T} \delta^{\tau-t} \Bigg(\alpha_1 \left(\frac{\partial}{\partial y_{i,t}} (y_{i,\tau-1} - y_{i,\tau}) \right)^2 + \alpha_2 \left(\frac{\partial}{\partial y_{i,t}}\, y_{i,\tau} \right)^2 \\
&\qquad + \alpha_3 \sum_{k=1}^{N} g_{ik} \left(\frac{\partial y_{k,\tau}}{\partial y_{i,t}} - \frac{\partial y_{i,\tau}}{\partial y_{i,t}} \right)^2 \Bigg) \\
&= \alpha_1 + \alpha_2 + \alpha_3 \qquad \text{(B.14)} \\
&+ \sum_{\tau=t+1}^{T} \delta^{\tau-t} \left(\alpha_1\, (\Lambda_{ii,t,\tau-1} - \Lambda_{ii,t,\tau})^2 + \alpha_2\, (\Lambda_{ii,t,\tau})^2 + \alpha_3 \sum_{k=1}^{N} g_{ik}\, (\Lambda_{ki,t,\tau} - \Lambda_{ii,t,\tau})^2 \right)
\end{aligned}$$

Similarly, for any $j \neq i$,

$$
\begin{aligned}
\Delta_{ij,t} &:= \alpha_3\, g_{ij} - \sum_{\tau=t+1}^{T} \delta^{\tau-t} \Bigg[\alpha_1 \left(\frac{\partial}{\partial y_{j,t}} (y_{i,\tau-1} - y_{i,\tau}) \frac{\partial}{\partial y_{i,t}} (y_{i,\tau-1} - y_{i,\tau}) \right) \\
&+ \alpha_2 \left(\frac{\partial}{\partial y_{j,t}} y_{i,\tau} \frac{\partial}{\partial y_{i,t}} y_{i,\tau} \right) \\
&\quad + \alpha_3 \sum_{k=1}^{N} g_{ik} \left(\frac{\partial y_{k,\tau}}{\partial y_{j,t}} - \frac{\partial y_{i,\tau}}{\partial y_{j,t}} \right) \left(\frac{\partial y_{k,\tau}}{\partial y_{i,t}} - \frac{\partial y_{i,\tau}}{\partial y_{i,t}} \right) \Bigg] \\
&= \alpha_3\, g_{ij} \qquad\qquad \text{(B.15)} \\
&- \sum_{\tau=t+1}^{T} \delta^{\tau-t} \Bigg[\alpha_1 \left(\Lambda_{ij,t,\tau-1} - \Lambda_{ij,t,\tau} \right) \left(\Lambda_{ii,t,\tau-1} - \Lambda_{ii,t,\tau} \right) + \alpha_2\, \Lambda_{ij,t,\tau}\, \Lambda_{ii,t,\tau} \\
&+ \alpha_3 \sum_{k=1}^{N} g_{ik} \left(\Lambda_{kj,t,\tau} - \Lambda_{ij,t,\tau} \right) \left(\Lambda_{ki,t,\tau} - \Lambda_{ii,t,\tau} \right) \Bigg]
\end{aligned}
$$

Let $\text{diag}\,(A)$ be the $N \times N$ diagonal matrix whose non-zero entries are the diagonal elements of the matrix A. So, in matrix form the matrix Δ_t is defined in two-steps as

$$
\begin{aligned}
\tilde{\Delta}_t &:= -\sum_{\tau=t+1}^{T} \delta^{\tau-t} \Bigg[\alpha_1\, \text{diag}\left(\Lambda_{t,\tau-1} - \Lambda_{t,\tau} \right) \left(\Lambda_{t,\tau-1} - \Lambda_{t,\tau} \right) + \alpha_2\, \text{diag}\left(\Lambda_{t,\tau} \right)\, \Lambda_{t,\tau} \\
&+ \alpha_3 \sum_{k=1}^{N} \text{diag}\left(G_{\bullet k}\, \iota_N' \right) \text{diag}\left(\iota_N\, \Lambda_{k\bullet,t,\tau} - \Lambda_{t,\tau} \right) \left(\iota_N\, \Lambda_{k\bullet,t,\tau} - \Lambda_{t,\tau} \right) \Bigg] \qquad \text{(B.16)}
\end{aligned}
$$

where ι_N is an $N \times 1$ column-vector of ones and ι_N' is an $1 \times N$ row-vector of ones; $G_{\bullet k}$ is the k'th column of the $N \times N$ matrix $\mathbf{G}$; $\Lambda_{k\bullet,t,\tau}$ is the k'th row of the $N \times N$ matrix $\Lambda_{t,\tau}$. Now,

$$
\Delta_t \quad := \quad \alpha_3\, \mathbf{G} + (\alpha_1 + \alpha_2 + \alpha_3)\, I_N + \tilde{\Delta}_t - 2\, \text{diag}\left(\tilde{\Delta}_t \right) \qquad \text{(B.17)}
$$

Finally, let D_t capture the sum of the effects on the current period (period t) marginal utility of future θ_τ's. D_t's can be computed recursively beginning with $t = T$, setting $\mathbf{D}_T = \mathbf{0}$, $N \times 1$ vector of zeros (no future period). Then, for $t < T$, let D_t be defined as

$$
\begin{aligned}
\alpha_2\, D_{i,t} &:= \alpha_2 \sum_{\tau=t+1}^{T} \delta^{\tau-t} \Bigg(-\alpha_1 \left(\Gamma_{i,t,\tau-1} - \Gamma_{i,t,\tau} \right) \left(\Lambda_{ii,t,\tau-1} - \Lambda_{ii,t,\tau} \right) \\
&+ \left(\bar{\theta}_{i,\tau} - \Gamma_{i,t,\tau} \right) \Lambda_{ii,t,\tau} \\
&- \alpha_3 \sum_{k=1}^{N} g_{ik} \left(\Gamma_{k,t,\tau} - \Gamma_{i,t,\tau} \right) \left(\Lambda_{ki,t,\tau} - \Lambda_{ii,t,\tau} \right) \Bigg)
\end{aligned}
$$

Hence, in matrix form

$$
\begin{aligned}
\mathbf{D}_t \quad := \quad & \sum_{\tau=t+1}^{T} \delta^{\tau-t} \Bigg(-\ \alpha_1 \,\mathrm{diag}\left(\Lambda_{t,\tau-1} - \Lambda_{t,\tau}\right) \left(\Gamma_{t,\tau-1} - \Gamma_{t,\tau}\right) \\
& + \mathrm{diag}\left(\Lambda_{t,\tau}\right)\ \left(\bar{\theta}_\tau - \Gamma_{t,\tau}\right) \\
& - \alpha_3 \sum_{k=1}^{N} \mathrm{diag}\left(G_{\bullet k}\, \iota_N'\right) \mathrm{diag}\left(\iota_N\, \Lambda_{k\bullet,t,\tau} - \Lambda_{t,\tau}\right) \left(\Gamma_{k,t,\tau}\mathbf{1} - \Gamma_{t,\tau}\right) \Bigg)
\end{aligned}
\tag{B.18}
$$

where $\mathbf{1}$ is an $N \times 1$ column vector of ones. For $t = T-1$, this translates into

Now define

$$\bar{\Delta}_t := \mathrm{diag}(\Delta_t)$$

and

$$\bar{\bar{\Delta}}_t := \Delta_t - \bar{\Delta}_t$$

using which we can rewrite the system of first-order conditions in matrix form as

$$\bar{\Delta}_t \mathbf{y}_t = \alpha_1\, \mathbf{y}_{t-1} + \alpha_2\, \theta_t + \alpha_2\, \mathbf{D}_t + \bar{\bar{\Delta}}_t \mathbf{y}_t \tag{B.19}$$

As we did in the beginning of the proof for the final period, the right hand side of (B.19) can be seen as an operator, call it FOC_{it}, that maps any given collection $f = \{f_j\}$ of bounded and $\mathcal{F}_t$-measurable functions in $\mathbf{B}$ to the $\mathcal{F}_t$-measurable function $FOC_{it}(f)$. Hence, showing the existence of a linear equilibrium policy for the first period of a $T-t+1$-period economy is equivalent to finding the fixed point of the operator FOC_{it}. Using straightforward modifications of the arguments in the proof for the last period, FOC_{it} is a contraction mapping and that it maps the closed subset $\mathbf{B}_c^n$ of $\mathbf{B}^n$ into itself; hence its unique fixed point then lies necessarily in $\mathbf{B}_c^n$. Thus, the equilibrium choice vector is linear in period $t-1$ choices, period-t shocks, and future expected shocks. Moreover, the existence of the unique fixed point for FOC_t in (B.19) is equivalent to the invertibility of this matrix equation. Hence, the equilibrium choices vector takes the form

$$\left(\bar{\Delta}_t - \bar{\bar{\Delta}}_t\right) \mathbf{y}_t = \alpha_1\, \mathbf{y}_{t-1} + \alpha_2\, \theta_t + \alpha_2\, \mathbf{D}_t \tag{B.20}$$

and the optimal policy then is given by

$$\mathbf{y}_t = \underbrace{\left(\bar{\Delta}_t - \bar{\bar{\Delta}}_t\right)^{-1}}_{\mathbf{B}_t} \left(\alpha_1\, \mathbf{y}_{t-1} + \alpha_2\, \theta_t + \alpha_2\, \mathbf{D}_t\right) \tag{B.21}$$

where $\mathbf{B}_t := [b_{ij,t}]$ is an $N \times N$ matrix of equilibrium coefficients for period t. Therefore, in any period

$t = 1, \ldots, T-1$, the system of first-order conditions in matrix form can be written as

$$\left(\bar{\Delta}_t - \bar{\bar{\Delta}}_t\right) \mathbf{y}_t = \left[\alpha_1 \, \mathbf{y}_{t-1} + \alpha_2 \, \theta_t + \alpha_2 \, \mathbf{D}_t\right] \tag{B.22}$$

and the optimal policy then is given by

$$\mathbf{y}_t = \underbrace{\left(\bar{\Delta}_t - \bar{\bar{\Delta}}_t\right)^{-1}}_{\mathbf{B}_t} \left(\alpha_1 \, \mathbf{y}_{t-1} + \alpha_2 \, \theta_t + \alpha_2 \, \mathbf{D}_t\right); \tag{B.23}$$

which concludes the proof of the Proposition. In the next section of this appendix we provide a recursive algorithm to compute $\bar{\Delta}_t, \bar{\bar{\Delta}}_t$ (and hence $\mathbf{B}_t = \left(\bar{\Delta}_t - \bar{\bar{\Delta}}_t\right)$, and D_t.

■

C Recursive Algorithm

Below is the recursive algorithm that follows the steps of the recursive characterization argument of the last section. We use this algorithm to compute the equilibrium policy weights when we simulate our model.

1. Compute $\mathbf{B}_T$ from the last period ($T = 12$), assuming that $\mathbf{D}_T = \mathbf{0}$ is the $N \times 1$ vector of zeros.
2. Define $\bar{\theta}_t := \mathbf{X}_t\beta + \mathbf{G}\mathbf{X}_t\phi + \eta \, \iota_N$ as the $N \times 1$ vector of non-stochastic part of period-t shocks, for all $t = 8, \ldots, 12$.
3. Let $t = 11$.
4. Compute $\Lambda_{t,t+1}, \ldots, \Lambda_{t,T}$ using equation (F.1).
5. Compute $\Gamma_{t,t+1}, \ldots, \Gamma_{t,T}$ using equation (B.13).
6. Compute Δ_t using (B.17).
7. Compute D_t using (B.18).
8. Compute $\bar{\Delta}_t := \text{diag}(\Delta_t)$ and $\bar{\bar{\Delta}}_t := \Delta_t - \text{diag}(\Delta_t)$.
9. Compute $\mathbf{B}_t$ from (B.21).
10. Let $t = t - 1$. If $t \neq 8$ then go to Step 3. Otherwise Stop.

D Proof of Identification

Proposition 2 *Suppose that $T \geq 2$, and Full Rank, Exogeneity, and Regularity assumptions of Section 4.1 are satisfied. Then, our dynamic linear economy with social interactions is identified.*

Proof: : Based on the equilibrium characterization in Proposition 1 and using the decomposition in (4) and (5), the following reduced-form equations hold

$$\mathbf{y}_T = [\Delta_T \mathbf{I} - \alpha_3 \mathbf{G}]^{-1} \times \left[\alpha_1 \mathbf{y}_{T-1} + \alpha_2 \left(\sum_{k=1}^{K} (\beta_k \mathbf{I} + \phi_k \mathbf{G}) \mathbf{x}_T^{(k)} + \mathbf{u}_T\right)\right]$$

$$\mathbf{y}_{T-1} = \left[\bar{\Delta}_{T-1} - \bar{\bar{\Delta}}_{T-1}\right]^{-1} \left(\alpha_1 \mathbf{y}_{T-2} + \alpha_2 \left(\sum_{k=1}^{K} (\beta_k \mathbf{I} + \phi_k \mathbf{G}) \mathbf{x}_{T-1}^{(k)} + \mathbf{u}_{T-1}\right) + \mathbf{D}_{T-1}\right)$$

We split the term $\mathbf{D}_{T-1}$ that includes the conditional expectations given period-$T-1$ information into observable and unobservable (by the econometrician) parts, namely, $E\left[\sum_{k=1}^{K} (\beta_k \mathbf{I} + \phi_k \mathbf{G}) \mathbf{x}_T^{(k)} \,|\, \mathbf{X}^{T-1}\right]$ and $E\left[\mathbf{u}_T \,|\, \mathbf{u}^{T-1}\right]$. Agents observe both $\mathbf{X}$ and $\mathbf{u}$ and these two are not correlated by the Exogeneity Assumption. Furthermore, $E\left[\sum_{k=1}^{K} (\beta_k \mathbf{I} + \phi_k \mathbf{G}) \mathbf{x}_T^{(k)} \,|\, \mathbf{X}^{T-1}\right]$ is a function of $\mathbf{X}^{T-1}$ which is known by the econometrician. Hence, using the definition of D_{T-1} in equation (B.18) in Appendix B and letting $\bar{\theta}_T^x := E\left[\sum_{k=1}^{K} (\beta_k \mathbf{I} + \phi_k \mathbf{G}) \mathbf{x}_T^{(k)} \,|\, \mathbf{X}^{T-1}\right]$ and $\bar{\theta}_T^u := E\left[\mathbf{u}_T \,|\, \mathbf{u}^{T-1}\right]$,

$$\begin{aligned}
\mathbf{D}_{T-1} &:= \mathbf{D}_{T-1}^x + \mathbf{D}_{T-1}^u \\
&= \delta\Bigg(-\alpha_1 \operatorname{diag}(\mathbf{I}_N - \alpha_1 \mathbf{B}_T)\left(-\mathbf{B}_T \bar{\theta}_T^x\right) + \operatorname{diag}(\alpha_1 \mathbf{B}_T)\left(\bar{\theta}_T^x - \mathbf{B}_T \bar{\theta}_T^x\right) \\
&\quad -\alpha_3 \sum_{l=1}^{N} \operatorname{diag}(G_{\bullet l}\, \iota_N')\, \alpha_1 \operatorname{diag}(\iota_N \mathbf{B}_{l\bullet,T} - \mathbf{B}_T)\left((\mathbf{B}_T \bar{\theta}_T^x)_{l,\bullet} \mathbf{1} - \mathbf{B}_T \bar{\theta}_T^x\right)\Bigg) \\
&+ \delta\Bigg(-\alpha_1 \operatorname{diag}(\mathbf{I}_N - \alpha_1 \mathbf{B}_T)\left(-\mathbf{B}_T \bar{\theta}_T^u\right) + \operatorname{diag}(\alpha_1 \mathbf{B}_T)\left(\bar{\theta}_T^u - \mathbf{B}_T \bar{\theta}_T^u\right) \\
&\quad -\alpha_3 \sum_{l=1}^{N} \operatorname{diag}(G_{\bullet l}\, \iota_N')\, \alpha_1 \operatorname{diag}(\iota_N \mathbf{B}_{l\bullet,T} - \mathbf{B}_T)\left((\mathbf{B}_T \bar{\theta}_T^u)_{l,\bullet} \mathbf{1} - \mathbf{B}_T \bar{\theta}_T^u\right)\Bigg)
\end{aligned}$$

Substituting these back into the reduced-form equation for $T-1$ above, we get the following system of linear simultaneous econometric equations with N endogenous variables on the right hand side of each equation,

$$\mathbf{y}_T = \alpha_1 \mathbf{B}_T \mathbf{y}_{T-1} + \alpha_2 \mathbf{B}_T \left(\sum_{k=1}^{K} (\beta_k \mathbf{I} + \phi_k \mathbf{G}) \mathbf{x}_T^{(k)}\right) + \varepsilon_T \tag{D.1}$$

$$\varepsilon_T = \mathbf{B}_T \alpha_2 \mathbf{u_T}.$$

$$\mathbf{y}_{T-1} = \alpha_1 \mathbf{B}_{T-1} \mathbf{y}_{T-2} + \alpha_2 \mathbf{B}_{T-1} \left(\sum_{k=1}^{K} (\beta_k \mathbf{I} + \phi_k \mathbf{G}) \mathbf{x}_{T-1}^{(k)}\right) + \alpha_2 \mathbf{B}_{T-1} D_{T-1}^x + \varepsilon_{T-1} \tag{D.2}$$

$$\varepsilon_{T-1} = \mathbf{B}_{T-1} \alpha_2 \mathbf{u_{T-1}} + \mathbf{B}_{T-1} \alpha_2 D_{T-1}^u.$$

where the error terms $\varepsilon_{T-1}, \varepsilon_T$ are known linear combinations of own and friends' current unobservables and expectations of future unobservables.

The endogeneity of $\mathbf{y}_{T-1}$ in equation (D.1) and of $\mathbf{y}_{T-2}$ in equation (D.2) require us to find suitable instrumental variables $\mathbf{z}_{T-1}$ and $\mathbf{z}_{T-2}$. Thanks to the Regularity Assumption, one of the characteristic dimensions does affect the individual choice y_{it}, either through direct own effects $\beta_k x_{it}^{(k)}$, or through social effect $\phi_k \mathbf{G}_i \mathbf{x}_t^{(k)}$. Assume for simplicity of presentation that it is the former. The argument for the latter is identical. Hence, we can define $z_{it} := x_{it}^{(k)}$ for $t = T-1, T-2$. This way, we have N instruments for each period. Predicted values of $\mathbf{y}_t$, $t = T-1, T-2$, are formed by projecting them on to the space spanned by the set of instrumential variables $\mathbf{z}_t$, $t = T-1, T-2$. These are valid instruments by construction since:

1. They are uncorrelated with the errors, hence satisfy exclusion restrictions: Thanks to the Exogeneity Assumption and using iterated expectations, for $t = T$

$$\begin{aligned} E[\varepsilon_T \mid \mathbf{z}_{T-1}] &= \alpha_2 \mathbf{B}_T \, E[\, \mathbf{u}_T \mid \mathbf{x}_{T-1}^{(k)}] \\ &= \alpha_2 \mathbf{B}_T \, E[\, E[\, \mathbf{u}_T \mid \mathbf{X}_1, \ldots, \mathbf{X}_{T-1}] \mid \mathbf{x}_{T-1}^{(k)}] = E[0 \mid \mathbf{x}_{T-1}^{(k)}] = 0 \end{aligned}$$

   and for $t = T-1$, similar arguments lead to

$$E[\varepsilon_{T-1} \mid \mathbf{z}_{T-2}] = \alpha_2 \mathbf{B}_{T-1} \, E[\, E[\, \mathbf{u}_{T-1} + D_{T-1}^u \mid \mathbf{X}_1, \ldots, \mathbf{X}_{T-1}] \mid \mathbf{x}_{T-1}^{(k)}]$$

   Note that ε_t is a linear function of two sets of variables: u_t and $E\left[u_\tau | u^t\right]$, with $\tau \geq t+1$. By the Exogeneity Assumption, we have $E\left[u_t | x_{t-1}\right] = 0$ and, since $E\left[u_\tau | u^t\right]$ is a function of u^t, $E\left[E\left[u_\tau | u^t\right] | x_{t-1}\right] = 0$.

2. They are informative about the explanatory variable, i.e. $E[\mathbf{z}_{T-1}\mathbf{y}_{T-1}] \neq 0$, thanks to $\alpha_2 \beta_k \neq 0$ by the Regularity Assumption.

Moreover, they are not collinear with $\mathbf{X}_t$, $t = T-1, T$, thanks to the Full Rank Assumption. With $T \geq 2$, we can consistently estimate $\mathbf{B}_T$ and $\mathbf{B}_{T-1}$ using the constructed instrumental variables.

So far, we have demonstrated that we can estimate the reduced form equilibrium coefficients consistently under the stated assumptions. In the second part, we show that the map from the utility parameters into the reduced form coefficients is injective. Consider now two sets of structural parameters $\gamma = (\alpha_1, \alpha_2, \alpha_3, \beta, \phi)$ and $\gamma' = (\alpha_1', \alpha_2', \alpha_3', \beta', \phi')$ leading to the same reduced form in equation (7).[47] Coefficient estimates would imply

$$\alpha_1 \, \mathbf{B}_T(\gamma) \, \mathbf{y}_{T-1} = \alpha_1' \, \mathbf{B}_T(\gamma') \, \mathbf{y}_{T-1} \quad \Longrightarrow \quad \alpha_1 \;\, \mathbf{B}_T(\gamma) = \alpha_1' \, \mathbf{B}_T(\gamma')$$

due to observational equivalence, where $\mathbf{B}_T(\gamma) := \left[\Delta_T \, \mathbf{I} - \alpha_3 \, \mathbf{G}\right]^{-1}$, and $\mathbf{B}_T(\gamma') := \left[\Delta_T' \, \mathbf{I} - \alpha_3' \, \mathbf{G}\right]^{-1}$. Since

[47] If needed, one can simply add a constant term to the structural equations to make the comparison with the previous works easier. This addition would not alter any of the results or the proof argument.

$\alpha_1, \alpha_1' \neq 0$ by the Regularity assumption, and $\mathbf{B}_T(\gamma)$ and $\mathbf{B}_T(\gamma')$ are invertible, we obtain

$$\frac{1}{\alpha_1}\,[\Delta_T\,\mathbf{I} - \alpha_3\,\mathbf{G}] \quad = \quad \frac{1}{\alpha_1'}\,[\Delta_T'\,\mathbf{I} - \alpha_3'\,\mathbf{G}] \tag{D.3}$$

Since $g_{ii} = 0$, the diagonal entries on left and right hand sides of the equation give $\Delta_T/\alpha_1 = \Delta_T'/\alpha_1'$. Moreover, the row sums on both sides off the diagonal yield $\alpha_3/\alpha_1 = \alpha_3'/\alpha_1'$. Under the normalization $\Delta_T = \sum_i \alpha_i = 1$, we obtain

$$\frac{\alpha_1}{\Delta_T} = \alpha_1 \quad = \quad \alpha_1' = \frac{\alpha_1'}{\Delta_T'}$$

which would in turn imply, by substituting back into (D.3), that $\alpha_3 = \alpha_3'$. Therefore, $\alpha_2 = 1 - \alpha_1 - \alpha_2$.

Consistent estimate of the reduced form equation (7) yields further observable equivalence restrictions, namely, for $k = 1, \ldots, K$

$$(\beta_k\mathbf{I} + \phi_k\mathbf{G})\,\mathbf{x}_T^{(k)} \;=\; (\beta_k'\mathbf{I} + \phi_k'\mathbf{G})\,\mathbf{x}_T^{(k)} \quad \Longrightarrow \quad (\beta_k\mathbf{I} + \phi_k\mathbf{G}) \;=\; (\beta_k'\mathbf{I} + \phi_k'\mathbf{G})$$

which is equivalent to

$$(\beta_k - \beta_k')\,\mathbf{I} + (\phi_k - \phi_k')\,\mathbf{G} \quad = \quad 0 \tag{D.4}$$

Since $\mathbf{I}$ and $\mathbf{G}$ are linearly independent (remember that $g_{ii} = 0$), this yields $\beta_k = \beta_k'$ and $\phi_k = \phi_k'$.

Similarly, for $t = T - 1$, we obtain consistent coefficient estimates in equation (8) and can recover $\mathbf{B}_{T-1}$ since we already recovered the true α_1 using (7) above. We can then recover Δ_{T-1} by reversing the operations in Step 8 of the recursive algorithm we used to obtain $\mathbf{B}_{T-1}$ from Δ_{T-1}, that we presented in Section C, namely by

$$\Delta_{T-1} \quad = \quad \mathrm{diag}(\mathbf{B}_{T-1}^{-1}) - \big(\mathbf{B}_{T-1}^{-1} - \mathrm{diag}(\mathbf{B}_{T-1}^{-1})\big)$$

and using the expression in (B.17), we can also recover

$$\Lambda \;:=\; \tilde{\Delta}_{T-1} - 2\,\mathrm{diag}\left(\tilde{\Delta}_{T-1}\right) \quad = \quad \Delta_{T-1} - \alpha_3\,\mathbf{G} - (\alpha_1 + \alpha_2 + \alpha_3)\,\mathbf{I}_N \tag{D.5}$$

since we already recovered everything to the right of the second equality sign. Hence, we can also obtain $\tilde{\Delta}_{T-1}$ by reverse operations, namely by $\tilde{\Delta}_{T-1} = \Lambda - 2\,\mathrm{diag}\,(\Lambda)$. Moreover, we know by substituting period-T equilibrium into (B.16)that, $\tilde{\Delta}_{T-1}$ takes the form

$$\begin{aligned}\tilde{\Delta}_{T-1} &:= -\delta\Bigg[\alpha_1\,\text{diag}\,(\mathbf{I}_N-\alpha_1\mathbf{B}_T)\,(\mathbf{I}_N-\alpha_1\;\mathbf{B}_T)+\alpha_2\,\text{diag}\,(\alpha_1\mathbf{B}_T)\;\alpha_1\mathbf{B}_T\\ &+\;\alpha_3\sum_{l=1}^{N}\text{diag}\,(G_{\bullet l}\,\iota_N')\,\text{diag}\,(\iota_N\,\mathbf{B}_{l\bullet,T}-\mathbf{B}_T)\,(\iota_N\,\mathbf{B}_{l\bullet,T}-\;\mathbf{B}_T)\Bigg]\\ &= -\delta\,M\end{aligned}$$

where M represents everyhting inside the brackets, which we recovered using period-T equilibrium restrictions. Hence, δ is recovered as well. This concludes the proof. ■

E Optimal GMM Estimator - G exogenous

Let $\mathbf{z}_t = [\mathbf{y}_{t-1}, \mathbf{X}_t, \mathbf{GX}_t]$ and $\mathbf{q}_t = [\mathbf{X}_{t-1}, \mathbf{GX}_{t-1}, \mathbf{G}^2\mathbf{X}_t, \mathbf{G}^2\mathbf{X}_{t-1}, \mathbf{X}_t, \mathbf{GX}_t]$ be vectors of explanatory variables and instruments, respectively. [48] Let $\gamma = [\alpha_1, \alpha_2, \alpha_3, \beta', \phi', \delta]'$.

Let the stacked vectors be $\mathbf{Y} = [\mathbf{y}'_T, \dots, \mathbf{y}'_{T-4}]'$, $\mathbf{Z} = [\mathbf{z}'_T, \dots, \mathbf{z}'_{T-4}]'$, $\mathbf{Q} = [\mathbf{q}'_T, \dots, \mathbf{q}'_{T-4}]'$ and $F(\mathbf{Z}, \gamma) = \begin{bmatrix} \alpha_1\,\mathbf{B}_5\,\mathbf{y}_4 + \alpha_2\,\mathbf{B}_5\,(\beta\,\mathbf{X}_5 + \phi\,\mathbf{GX}_5\; + \mathbf{D}_5 + \mathbf{u_5}) \\ \vdots \\ \alpha_1\,\mathbf{B}_1\,\mathbf{y}_0 + \alpha_2\,\mathbf{B}_1\,(\beta\,\mathbf{X}_1 + \phi\,\mathbf{GX}_1\; + \mathbf{D}_1 + \mathbf{u_1}) \end{bmatrix}$. Finally, let $\mathbf{h}(\mathbf{Z}, \mathbf{Q}, \gamma) = \mathbf{Q}'[\mathbf{Y} - F(\mathbf{Z}, \gamma)]$. The moment conditions are then,

$$E(\mathbf{h}(\mathbf{Z}, \mathbf{Q}, \gamma)) = E(\mathbf{Q}'[\mathbf{Y} - F(\mathbf{Z}, \gamma)]) = 0; \tag{E.1}$$

The optimal GMM estimator of γ satisfies

$$\widehat{\gamma} \in \arg\min\; \mathbf{h}(\mathbf{Z}, \mathbf{Q}, \gamma)'\,\mathbf{W}^{-1}\,\mathbf{h}(\mathbf{Z}, \mathbf{Q}, \gamma), \tag{E.2}$$

where the weight matrix $\mathbf{W}$ is the variance of the moment condition, $\mathbf{W} = \mathbf{S} = E[\mathbf{h}(\mathbf{Z}, \mathbf{Q}, \gamma)\mathbf{h}(\mathbf{Z}, \mathbf{Q}, \gamma)']$.

In practice, we

Step 1 Obtain a GMM estimator using (the suboptimal choice of the weight matrix) $\mathbf{W} = \mathbf{I}_P$; obtain also the consistent estimate $\widehat{\mathbf{S}} = \frac{1}{N}\sum_{i=1}^{N}\widehat{u}_i^2\mathbf{q}_i\mathbf{q}_i'$; where $\widehat{u}_i = y_i - F(\mathbf{z}_i, \widehat{\gamma})$, $\mathbf{q}_i$ is the i-th row of $\mathbf{Q}$ and $\widehat{\gamma}_1$ is the (inefficient) GMM estimator of γ_0 in the first step.

Step 2 Obtain a GMM estimator, again, but using (the optimal weight matrix) $\mathbf{W} = \widehat{\mathbf{S}}^{-1}$.

The optimal GMM estimator $\widehat{\gamma}_{OGMM}$ is consistent and asymptotically normally distributed with mean

[48] Observe that in principle we would not need to add $\mathbf{G^2X_t}$ as excluded instrument. Given that empirically $\mathbf{G}$ and $\mathbf{G^2}$ are linear independent we exploit Bramoullé et al. (2009) conditions including this additional instrument in order to improve efficiency.

γ_0 and with estimated asymptotic variance $\widehat{\mathbf{V}} = N\left(\widehat{\mathbf{D}}'\mathbf{Q}\widetilde{\mathbf{S}}^{-1}\mathbf{Q}'\widehat{\mathbf{D}}\right)^{-1}$, where $\widetilde{S}$ denotes $\widehat{S}$ evaluated at $\widehat{\gamma}_{GMM}$ and a consistent estimate $\widehat{\mathbf{D}}$ of $\mathbf{D}_0$ can be obtained from $\widehat{\mathbf{D}} = \frac{\partial \mathbf{u}}{\partial \gamma}$ evaluated at $\gamma = \widehat{\gamma}_{GMM}$. For the model in Section 6, where we use $\tilde{\mathbf{G}}$ instead of $\mathbf{G}$ (see (15)), the matrix of the instrument is defined as $\tilde{\mathbf{q}}_t = [\mathbf{X}_{t-1}, \tilde{\mathbf{G}}\mathbf{X}_{t-1}, \tilde{\mathbf{G}}^2\mathbf{X}_t, \tilde{\mathbf{G}}^2\mathbf{X}_{t-1}, \mathbf{X}_t, \tilde{\mathbf{G}}\mathbf{X}_t]$.

F Recursive Computation of the Dynamic Multiplier

Remember that $\Lambda_{t,\tau}$, for any $\tau = t+1, \ldots, T$, as

$$\Lambda_{t,\tau} := \alpha_1^{\tau-t}\, \mathbf{B}_\tau \times \ldots \times \mathbf{B}_{t+1} \tag{F.1}$$

with the convention that $\Lambda_{t,t} := I_N$, the identity matrix. Moreover, using equation (B.12) from the Appendix for the definition of $\Gamma_{t,\tau}$, for any $\tau = t+1, \ldots, T$, we obtain

$$\Delta\Gamma_{t,\tau} := \sum_{s=t+1}^{\tau} \alpha_1^{\tau-s}\, (\mathbf{B}_\tau \times \cdots \times \mathbf{B}_s)\left(\Delta\bar{\theta}_s + \Delta\mathbf{D}_s\right) \tag{F.2}$$

with the convention that $\Gamma_{t,t} := 0_N$, the $N \times 1$ matrix of zeros, and where $\bar{\theta}_s$ is the expected value of θ_s, conditional on period-t information. Similarly, equation (B.17) yields

$$\begin{aligned} \Delta\mathbf{D}_t \quad := \quad & \sum_{\tau=t+1}^{T} \delta^{\tau-t}\Bigg(-\, \alpha_1 \,\mathrm{diag}\left(\Lambda_{t,\tau-1} - \Lambda_{t,\tau}\right)\left(\Delta\Gamma_{t,\tau-1} - \Delta\Gamma_{t,\tau}\right) \\ & + \mathrm{diag}\left(\Lambda_{t,\tau}\right)\, \left(\Delta\bar{\theta}_\tau - \Delta\Gamma_{t,\tau}\right) \\ & - \alpha_3 \sum_{k=1}^{N} \mathrm{diag}\left(G_{\bullet k}\, \iota_N'\right) \mathrm{diag}\left(\iota_N\, \Lambda_{k\bullet,t,\tau} - \Lambda_{t,\tau}\right)\left(\Delta\Gamma_{k,t,\tau}\mathbf{1} - \Delta\Gamma_{t,\tau}\right)\Bigg) \end{aligned} \tag{F.3}$$

where $\mathbf{1}$ is an $N \times 1$ column vector of ones. These are the two variables we need for the computation.

For $t = T$, $\Delta\mathbf{D}_T = 0$ by construction. For $t = T-1$, (F.2) simplifies to

$$\Delta\Gamma_{T-1,T} = \mathbf{B}_T\, \Delta\bar{\theta}_T \;=\; \pi\, \mathbf{B}_T\, \mathbf{1} \tag{F.4}$$

Hence, $\Delta\mathbf{D}_{T-1}$, for example, can be obtained as

$$\begin{aligned} \Delta\mathbf{D}_{T-1} \quad = \quad & \pi\, \delta\Bigg(\alpha_1 \,\mathrm{diag}\left(\Lambda_{T-1,T-1} - \Lambda_{T-1,T}\right)\mathbf{B}_T\mathbf{1} \\ & + \mathrm{diag}\left(\Lambda_{T-1,T}\right)\, \left(\mathbf{1} - \mathbf{B}_T\mathbf{1}\right) \\ & - \alpha_3 \sum_{k=1}^{N} \mathrm{diag}\left(G_{\bullet k}\, \iota_N'\right) \mathrm{diag}\left(\iota_N\, \Lambda_{k\bullet,t,\tau} - \Lambda_{T-1,T}\right)\left((\mathbf{B}_T\,\mathbf{1})_{k\bullet}\mathbf{1} - \mathbf{B}_T\,\mathbf{1}\right)\Bigg) \end{aligned}$$

Here is the **recursive algorithm** to compute these variables for the remaining periods $t = 1, \dots, T-1$:

1. Compute $\Delta\Gamma_{t,\tau}$, $\tau = t+1, \dots, T$ using equation (F.2).
2. Compute $\Delta\mathbf{D}_t$ using equation (F.3).
3. Repeat until $t = 0$.

G Data and Additional Tables Checks

Table G1: **Sample Selection**

	Panel (a) Initial sample (N. obs.: 1793)		Panle (b) Sample without missing values (N. obs.: 1759)		Panel (c) Students connected in a social network (N. obs.: 1207)	
Variable	Mean	SD	Mean	SD	Mean	SD
Risky behavior index (wave II)	0.1535	0.2023	0.1545	0.2026	0.1579	0.204
Risky behavior index (wave I)	0.1396	0.1947	0.1399	0.1947	0.1345	0.1877
Female	0.4992	0.5001	0.4997	0.5001	0.4979	0.5002
Black or African American	0.1439	0.3511	0.145	0.3522	0.1052	0.307
White	0.5438	0.4982	0.5469	0.3522	0.6081	0.4884
Hispanic or Latino	0.2166	0.4121	0.2166	0.412	0.1806	0.3849
Parents College degree	0.2627	0.4402	0.2609	0.4393	0.2759	0.4471
Two-parent family	0.7317	0.4432	0.7351	0.4414	0.7664	0.4233
Age (wave II)	16.9844	0.9389	16.9835	0.9397	16.9246	0.9437
Pocket money (wave II)	8.7582	12.2948	8.7822	12.3681	7.9039	11.6601
Alcohol/tobacco at home (wave II)	0.2651	0.3497	0.2649	0.3493	0.2668	0.3475
Height (wave II)	67.1150	3.9558	67.133	3.9549	67.1814	3.941

This table reports means and standard deviations of students' characteristics for the initial sample (Panel (a)), for the sample without missing values in observations (Panel (b)) and the sample with no isolates (Panel (b)).

Table G2: **Sample Representativness**

	Add Health (N. obs.: 1207)		CPS (N. obs.: 14257)	
Variable	Mean	SD	Mean	SD
Female	0.4979	0.5002	0.5021	0.5000
Black or African American	0.1052	0.307	0.1268	0.3327
White	0.6081	0.4884	0.6587	0.4742
Hispanic or Latino	0.1806	0.3849	0.1639	0.3702
Parents College degree	0.2759	0.4471	0.2328	0.4227

This table reports summary statistics for the Add Health data sample used in the paper and the 1994 CPS. Person weights are used in the 1994 Current Population Surveys (CPS). The CPS sample is restricted to those aged 14-20 and re-weighted to match the age distribution of the Add Health sample.

Table G3: **Dynamic recursive model- Controls and contextual effects**

Dep. Var. Risky Behavior Index	Endogenous networks	
	(1)	(2)
Addiction effect (α_1)	0.5084**	0.4299***
	(0.2129)	(0.1552)
Own effect (α_2)	0.1017***	0.1538***
	(0.0227)	(0.0210)
Peer effect (α_3)	0.3899*	0.4163**
	(0.2243)	(0.1665)
Discount factor (δ)	0.4829*	0.4925**
	(0.2614)	(0.2017)
Selectivity (ψ)	0.1488	-0.0620
	(0.7404)	(0.5942)
Female	-0.3266	-0.2897
	(0.9793)	(0.2546)
Black or African American	-0.1789	-0.1920
	(1.0958)	(0.3811)
Age (wave II)	-0.0021	0.0041
	(0.1931)	(0.0447)
Parents College degree	-0.9833	-0.5259
	(1.7672)	(0.7302)
Two-parent family	0.7263	0.3738
	(1.5658)	(0.6258)
Pocket money	0.0065	0.0029
	(0.0137)	(0.0060)
Alcohol/tobacco at home (wave II)	0.4332	0.1698
	(1.0185)	(0.5274)
Height (wave II)	0.0008	-0.0026
	(0.0556)	(0.0122)
Peers' characteristics	No	Yes
Networks fixed effects	Yes	Yes
N. Obs.	1,207	1,207

This table reports GMM estimates of the structural model 10. The network formation is assumed to satisfy the linear dyadic structure in equation 9. The peers' characteristics are calculated as friends' averages of the included variables. Heteroskedasticity-robust numerical standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G4: **Constrained "myopic" and "static" models- Controls**

Dep. Var. Risky Behavior Index	Endogenous network		
	Myopic	*Static*	*Benchmark*
	$\delta = 0$	$\alpha_1 = 0$, $\delta = 0$	
	(1)	(2)	(3)
Addiction effect (α_1)	0.4233***		0.4299***
	(0.1225)		(0.1552)
Own effect (α_2)	0.1468***	0.3547***	0.1538***
	(0.0134)	(0.0999)	(0.0210)
Peer effect (α_3)	0.4299***	0.6453***	0.4163**
	(0.1347)	(0.0999)	(0.1665)
Discount factor (δ)			0.4925**
			(0.2017)
Selectivity (ψ)	0.2615	0.0888	-0.0620
	(1.1107)	(0.1680)	(0.5942)
Female	-0.2602	-0.0953	-0.2897
	(0.5906)	(0.0712)	(0.2546)
Black or African American	-0.0133	-0.0931	-0.1920
	(0.7322)	(0.1761)	(0.3811)
Age (wave II)	0.0016	-0.0019	0.0041
	(0.2155)	(0.0238)	(0.0447)
Parents College degree	-1.0193	-0.3309	-0.5259
	(0.9485)	(0.2497)	(0.7302)
Two-parent family	0.4994	0.1888	0.3738
	(0.8497)	(0.1347)	(0.6258)
Pocket money	0.0056	0.0010	0.0029
	(0.0092)	(0.0019)	(0.0060)
Alcohol/tobacco at home (wave II)	0.1601	0.0517	0.1698
	(0.8999)	(0.1634)	(0.5274)
Height (wave II)	0.0015	0.0002	-0.0026
	(0.0489)	(0.0026)	(0.0122)
Student characteristics	Yes	Yes	Yes
Peers' characteristics	Yes	Yes	Yes
Networks fixed effects	Yes	Yes	Yes
N. Obs.	1,207	1,207	1,207

his table reports GMM estimates of the structural model 10. The network formation is assumed to satisfy the linear dyadic structure in equation 9. In Column 1 we restrict the model by setting $\delta = 0$, while in Column 2 we restrict the model by setting $\alpha_1 = 0$ and $\delta = 0$. Column 3 reports baseline estimates presented in Table 2 Column 2. The peers' characteristics are calculated as friends' averages of the included variables. Heteroskedasticity-robust numerical standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G5: **Balancing tests**

	Parents College degree	Two-parent family	Black or African American	Pocket money (wave II)	Alcohol/tobacco at home (wave II)	Height (wave II)
Same gender average risky behavior	-0.324	-0.1404	-0.3234	-21.1289	0.2696	5.4013
	(0.8734)	(0.8512)	(0.3681)	(21.2361)	(0.7274)	(5.4155)
School fixed effects	yes	yes	yes	yes	yes	yes
N. Obs.	1,207	1,207	1,207	1,207	1,207	1,207

The figures in each row are coefficients from separate regressions of students' characteristics (parent college degree, two parent family, black or African American, pocket money, alcohol/tobacco at home and height) on peers' average risky behavior, a gender indicator and the students' age. Heteroskedasticity-robust numerical standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G6: **Dynamic recursive model - Robustness**

Dep. Var. Risky Behavior Index	WI links (1)	Isolated nodes (2)
Addiction effect (α_1)	0.4040***	0.3897***
	(0.1177)	(0.1106)
Own effect (α_2)	0.1529***	0.2061***
	(0.0207)	(0.0275)
Peer effect (α_3)	0.4431***	0.4042***
	(0.1290)	(0.1164)
Discount factor (δ)	0.5088***	0.4989**
	(0.1595)	(0.2292)
Selectivity (ψ)	0.1059	-0.1545
	-0.2519	(0.3356)
Student characteristics	Yes	Yes
Peers' characteristics	Yes	Yes
Networks fixed effects	Yes	Yes
Dummy for isolated individuals	No	Yes
N. Obs.	1286	1759

This table reports GMM estimates of the structural model 10. The network formation is assumed to satisfy the linear dyadic structure in equation 9. Students' characteristics are listed in Table G1. The peers' characteristics are calculated as friends' averages of the included variables. Heteroskedasticity-robust numerical standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.