# Dynamic Social Interactions and Smoking Behavior*

Tiziano Arduini †
*University of Rome Tor Vergata*
tiziano.arduini@uniroma2.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

February 3, 2024

*We are grateful to Steven Durlauf, Matt Jackson, David Levine for their extensive comments and suggestions at different stages. We thank Ana Balsa, Larry Blume, John Cawley, Pierre-André Chiappori, Marcel Fafchamps, Bob Lucas, Jen Overbeck, Julian Reif, Tom Sargent, José Scheinkman, Giorgio Topa, Chao Yang, Roman Andres Zarate, 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, Catholic University of Uruguay, 2021 Workshop on Peer Effects, Virtual Seminar on the Economics of Risky Health Behaviors, 2021 Annual Conference of the International Association for Applied Econometrics, Workshop on Social Identity and Social Interactions in Economics in Quebec City, and the Australasian Economic Theory Workshop in Canberra. We also thank Angela Cools for helping us with the replication package.

Patacchini acknowledges financial support from NSF grant 2049207. Özgür thanks the Economics Department at Sciences Po in Paris, where a substantial 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 2013-2023, 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 and Finance, Tor Vergata University of Rome, Via Columbia 2, Rome, 00133 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.

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

**Abstract**

We study the risky behavior of adolescents. Concentrating on smoking, 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. The estimated dynamic model allows us to decompose the effect on smoking of a permanent shock to students' preferences in the 10th grade - e.g., a shock to tobacco availability at home or family income - in its own direct component and in the component due to social interactions. Indeed we find large relative social effects in grade 10th, declining in grade 11th and 12th.

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

# 1 Introduction

Smoking behavior is widespread among adolescents. According to the National Youth Tobacco Survey in 2021, approximately 2.55 million (9.3%) students in the U.S. reported to have used a tobacco product in the past 30 days: 2.06 million (13.4%) high school students and 470,000 (4.0%) middle school students (CDC, 2021). Smoking is a serious policy concern because it involves severe risks in terms of health outcomes. According to WHO, the tobacco epidemic is one of the biggest public health threats the world has ever faced killing more than 8 million people a year around the world (WHO, 2021). Furthermore, smoking is responsible for a large amount of socio-economic costs, in terms of, e.g., poor academic performance (König, 2016), earnings and unemployment (Ekpu and Brown, 2015; Weng et al., 2015), and criminal victimisation (Lewis et al., 2016). The empirical literature on risky health behavior in economics and social sciences emphasizes several fundamental aspects of smoking behavior. First of all, it responds to dynamic incentives, such as, e.g., future price changes and anticipated future consumption, and has an addictive component (Becker, Grossman, and Murphy, 1991; Chaloupka, 1991). Furthermore, smoking is a social behavior, in the sense that it depends on the behavior of relevant peers (Argys and Rees, 2008; Lundborg, 2006; Duncan et al., 2005; Balsa and Diaz, 2018).[1]

In this paper, we study the smoking 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 consumption in evaluating the consequences of their behavior. Furthermore, we embed the dynamic choice of adolescents regarding smoking 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 may affect the importance of peer effects over schooling age. More specifically, we formulate and structurally estimate a dynamic social interactions model. Agents' preferences over choices at any time depend on their own previous decisions 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 their 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

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

shocks of all peers.[2] 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 different school years. We use the panel dimension of the data to estimate our dynamic social interaction model structurally. 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.[3] 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 a 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 behavior in the adolescent population. The preference parameters driving the addiction effect and the social interaction effect are estimated to be significantly different from zero. Furthermore, a significant forward-looking component characterizes students' decision-making: the discount rate is also 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 preference parameter driving the social interaction effect.

In the dynamic model we estimate sizeable quantitative social interaction effects of potential shocks to the observable covariates we use in our estimation - e.g., *Tobacco at home* or *Family income*. Specifically, these effects are measured by i) the relative strength of the social component of the effect of the shock with respect to the own direct effect;[4] and by ii) the induced correlation

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

[3] See e.g., Blume, Durlauf and Jayaraman (2015), Brock and Durlauf (2001b), and Manski (2008).

[4] This decomposition between the social and the own effect is related to the concept of *social multiplier* (Becker and Murphy (2001), Glaeser and Scheinkman (2003)), though this concept is plagued by subtle identification issues, as shown by Boucher and Fortin (2016); see Section 5.2 for a detailed discussion of this issues

between the students' choices across the network. Importantly, in our context, social interaction effects have 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, and 12 (the effects of permanent shocks are naturally larger). Finally, our theoretical findings point to the importance of the same-period effects of a temporary shock in the different grades encoding the importance of the number of periods to the end of school in students' choices. Students anticipate a change in their social network after high school, which 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 on future shocks and less heavily the current shock.[5]

Finally, 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 sample consisting of grades 10, 11, and 12 to predict equilibrium behavior for students in the sample consisting of grade 9. We compare predictions and actual choice data to validate the model and demonstrate that the model performs very well in the validation exercise.

## 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, Durlauf and Jayaraman (2015), Brock and Durlauf (2001b) and many others.

For 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'

[5] Interestingly, at the estimated parameter values, this effect is softened by a counteracting effect: forward-looking rational agents smoothen consumption paths to the final period to minimize the convex costs of behavioral change due to a very high addiction parameter estimate.

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. Therefore, the distinction between rational and behavioral addiction manifests itself most clearly in high-frequency decisions over days. It is much less relevant when studying, as in our case, low-frequency choices over the 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), 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.[6] In our dynamic environment, we are not restricted to the cross-section. In particular, as we discuss in more detail in Section 4.2, we have access, for each period, to a set of external instrumental 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. Finally, for (ii) and (iii), we tackle them by exploiting quasi-random variation across cohorts within a school, given that the definition of peers in this paper is all students of the same gender in the same grade and school. The idea is to estimate a model with school and grade fixed effects by presuming that neither students nor parents can anticipate perfectly the composition of schoolmates' characteristics across grades when choosing schools or residential neighborhoods. This is a well-known empirical strategy used with AddHealth data because multiple cohorts are observed within schools. Furthermore, in our model, each equation (i.e. each grade) contains a grade-specific intercept together with school fixed effects. As a result, our model can account for time-varying common shocks at the school level (Angrist, 2014). We explain the implementation of this strategy in-depth in Section 4.2.

[6] 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, Djebbari and Fortin (2009), Calvo-Armengol, Patacchini and Zenou (2009).

To provide a more satisfactory (less reduced-form solution to the endogeneity of the network issue iii), a growing literature on social interactions has resorted to modeling the formation of social networks, ; see e.g. Battaglini, Patacchini and Rainone (2021), Badev (2021), Boucher (2020), Christakis et al. (2020), Hsieh and Lee (2015), Hsieh , König and Liu (2022), König (2016), and Mele (2017a,b). Embedding network formation in a fully specified dynamic forward-looking choice model is however 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 in the network. Badev (2021), instead, while estimating a network formation model where the network is allowed to change following an evolutionary process, has to restrict the analysis by severely limiting the rationality of his agents' choices.

Another important inference issue for social interaction models has been raised by Boucher and Fortin (2016), with regards to the identification of pure conformity models with respect to models displaying also social complementarities. This issue is important because when society displays pure conformity (and there is no addiction effect), the preference parameter driving the social interaction effect has no bearing on the total effect of an exogenous shock on agents' behavior; that is, in Boucher and Fortin (2016)'s definition, the *social multiplier* is zero. In this paper we (do not attempt at identifying social complementarities and) postulate pure conformity, but we exploit a decomposition of the own and the social effect of exogenous shocks to provide a measure of social interaction effects, which is related to the *social multiplier* with social complementarities, as in Becker and Murphy (2001) and Glaeser and Scheinkman (2003); we provide more details on this issue in Section 5.2 and in an online Appendix.

From a theoretical point of view, this paper's main novelty consists in studying 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 Reif (2019), 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 from our model, therefore, in Reif (2019) 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 (2020). However, in the current paper, the analysis of social interactions is extended to allow for general network topology. This is important in particular because it changes identification conditions. More specifically, an incomplete network structure provides a source of non-linearity (intransitive triads) that can be exploited for identification purposes (Bramoullé, Djebbari and Fortin, 2009; Calvo-Armengol, Patacchini and Zenou, 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 that allows jointly for both dynamic forward-looking agents and social interactions. Indeed, most studies of risky health behaviors have examined either peer group effects 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, Golberstein and Whitlock (2014), Lee, Li and Lin (2014) and Hsieh and Van Kippersluis (2018). In 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 that have a a dynamic component. Its analysis however does not allow for forward-looking behavior in the dynamic choice of agents, and the dynamics of peer effects are due to the exogenous spreading of interactions over the network.

## 2 Dynamic Interactions on Networks

This section introduces 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 with behavioral 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 $N$ agents $i = 1, \ldots, N$ who live for $t = 1, \ldots, T$ periods. Each agent $i$ chooses an action $y_{it}$ at time $t$ after having observed a preference shock $\theta_{it} \in \Theta$.[7] Let $\mathbf{y}_t$ and $\theta_t$ denote the corresponding $N$-dimensional vectors stacking all agents' choices and preference shocks respectively. Let $\theta := (\theta_t) := (\theta_{it})_{i=1,\ldots,N,\, t\geq 1}$ be the stochastic process of agents' preference shocks.

The general *social network* embedded in our model 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) $g_{ij}$ are row-normalised, i.e., $\sum_j g_{ij} = 1$; and iii) $g_{ii} = 0$. [8]

In our empirical application, we use the following special case of this general network structure as the friendship network:

$$g_{ij} = \begin{cases} \frac{1}{|N_G(i)|}, & \text{if } i \text{ and } j \text{ have the same gender and are in the same grade-school} \\ 0, & \text{otherwise,} \end{cases}$$

where $|N_G(i)|$ denotes the number of students in the reference group of $i$.

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 $\theta_{it}$, and with the current choices of her peers $\{y_{jt}\}_{j:g_{ij}\neq 0}$. We refer to $\alpha_1$ as the *addiction effect*, to $\alpha_2$ as the *own effect*, and to $\alpha_3$ as the *peer effect*. While $(\alpha_1, \alpha_2, \alpha_3)$ are restricted to be homogeneous across agents, preference

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

[8] The general theoretical results this model yields are all obtained under this general social network definition and are presented in Appendices A, B, C, and D). 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 network structure.

heterogeneity is captured in the formulation of the stochastic processes $\theta_{it}$.[9]

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 choice profiles, $\mathbf{y}^{t-1} = (\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_{t-1})$, and ii) the history of preference shocks, $\theta^t = (\theta_1, \ldots, \theta_t)$ (including the period-$t$ realization).

### 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 $(\mathbf{y}^{t-1}, \theta^t)$,

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

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

The economy displays a *unique Subgame Perfect Equilibrium.*[10] The equilibrium choice profile for period $t$ can be written as

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

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 $\theta_\tau$'s, $\tau > t$.[11]

[9] 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, Mailath and Postlewaite (1992), Bernheim (1994), and Peski (2007). Furthermore, as already noticed, we postulate pure conformity; that is, we do not add a component of preferences of the form $\sum_{j=1}^N 2\gamma_{ij}\, y_{it}\, y_{jt}$, which would capture social complementarities, but with no distinguishable aggregate equilibrium behavior (Boucher and Fortin (2016) show this in a static model but their result extends in our set-up).

[10] In Appendices B and C respectively, 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.

[11] 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.

## 3 Data

Our data source is the Add Health, a dataset on adolescents health' 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 the years 1994-95.[12]

Every student attending the sampled schools on the interview day was asked to complete a questionnaire (*in-school questionnaire*) containing questions on respondents' demographic 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 cigarette smoking behaviors. In 16 randomly selected schools, all students are interviewed at home (the so-called *saturated sample*). Our analysis focuses on this data. Specifically, we restrict our sample to include all students in the schools where information on cigarette smoking behaviors is collected for the entire school.[13] Those subjects were interviewed again one year apart in 1995–1996 (Wave II) so that each student's information was collected twice, in two waves in consecutive grades. Hence the sample has a panel dimension.

In our analysis, we define a student's peers as all other students of the same gender in the same grade at the same school. This definition is grounded in the sociological literature documenting that adolescents are more likely to have same-gender friends (see, e.g. McPherson, 2001). Figure 1 depicts friendship linkages in the larger network in our data (286 nodes with a diameter of 24) by using different colors for nodes indicating students of a different gender. The picture reveals that, indeed, social interactions observed in the data are highly assortative by gender.[14]

To estimate our dynamic social interaction model (which we discuss in detail in Section 4),

[12] Add Health is funded by grant P01 HD31921 (Harris) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill. Add Health data are restricted by contractual agreement. Instructions on how to obtain them and all codes necessary for replication are available from the authors. IRB approval: IRB0148126.

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

[14] In the Add Health questionnaire, the students are asked to identify their best friends from a school roster. This paper does not use this information since unobserved characteristics may drive both behavior and friend choices. Variations in the average peer characteristics of the same grade and gender schoolmates are instead reasonably exogenous, as we will show in Section 4.2.

we use the reduced-form linear system describing the equilibrium (Equation 3), which expresses the vector of outcomes (cigarette smoking behavior) of individuals as a linear function of the outcomes in the previous grade, and contemporaneous variables. In order to do that, we need to measure cigarette smoking behavior. To measure it, we follow an approach close to that in Badev (2021) and use the answers to the question "During the past 30 days, on how many days did you smoke cigarettes?". Figure 2 shows the distribution of students by the number of days where they smoke cigarettes. Badev (2021) constructs a smoking indicator that is equal to one if students answer one or more days (about 37% of the sample), and zero otherwise. In this paper, we exploit the richness of the information in the Add Health questionnaire to classify further the large mass of observations at 0 in Figure 2 (i.e., students who declare that they did not smoke in the last 30 days). Specifically, we distinguish students who have never smoked a cigarette from the behavior of students who happened not to smoke in the past 30 days but who have smoked before. The former group is identified using the answers to the following question: "Have you ever tried cigarette smoking, even just 1 or 2 puffs?". To capture this distinction, we construct a smoking behavior indicator which takes one of three possible values: 0 for students who report to have never tried, 1 for the ones who report to have tried and not to have smoked in the last 30 days, and 2 for those who report to have smoked one or more days. We refer to these categories as *never tried smoking*, *occasionally smoking* and *regular smoking*. Figure 3 shows the proportion of students in each category. Students who have never tried are around 49 % and those who smoke regularly are roughly 37% of the sample. We also run a number of robustness checks by changing the definition of the smoking behavior indicator.[15] The results of these robustness checks are reported in Section 6.

Our set of control variables includes the variables indicated by the literature (see e.g., Cawley and Ruhm, 2011; Lee, Li and Lin, 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 smoking behavior (whether

[15]More specifically, we implement the following treatments: the binary smoking indicator in Badev (2021); an indicator with a richer range where we add *daily smoking* as a category (this category is referred to students who report smoking every day and make up around 35% of the *regular smoking* category); and finally a yet richer indicator where the potential values encode the answers to the following question: "During the past 30 days, on the days you smoked, how many cigarettes did you smoke each day?": 0 for students who report to smoke zero cigarettes, 1 for ones who report to smoke from 1 to 5, 2 for students who report to smoke from 6 to 10, and so on up to values greater than 21.

cigarettes are easily available at home and family income).[16][17][18] In addition, we control for Peabody Picture Vocabulary Test (PVT) score from Wave I. Verbal IQ and cognitive skills are found to be predictors of health behaviors (see, e.g., Cawley and Ruhm, 2011). 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).[19] A unique feature of our data set is that both survey respondents, as well as all other schoolmates, are interviewed, which 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 for all students of the same gender in the same grade at a given school.

The sample counts 1727 students excluding missing observations, about 415 of which are in middle schools (grades $7, 8$, and 9) and 1312 in high schools. We use the (larger) sample of high school students to structurally estimate our model and data for other grades not employed in the estimation sample for its validation. We report summary statistics for the entire sample (Table 10, Panel (a))) and for the sample without students who report to skip school without reason for more than two days (see our IV strategy in Section 4.2.1 and Table 10, Panel (b)).[20] It appears that the composition of the sample is roughly unaffected, thus revealing that the average student characteristics in the selected sample are not dissimilar to the average ones for the non-selected. Our final sample of high school students counts 1,043 individuals. Consistently with the epidemiological literature (see e.g., Jackson et al., 2002) which finds the persistence of smoking behavior, our cigarette smoking indicator increases from wave I to wave II given the young age of

[16]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 parents; we select the father level of education. We construct a variable "Parents College degree", which is coded as one 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".

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

[18]We pull information on family income using the declared total income from any sources before taxes. If family income is missing, we imputed the family income as the mean of the sample conditional on gender, race, parental education, and a dummy for missing family income. Students are asked to answer yes or no to the question "Are cigarettes easily available to you in your home?". We construct a variable "Tobacco at home", which is coded as 1 if the respondent answered yes to the question above, and 0 otherwise. The questions together measure the accessibility to tobacco, either directly (i.e., stealing a cigarette from the mother's purse) or indirectly, by buying them and are also related to the price elasticities of demand for cigarettes (see e.g., Gallet and List, 2003; Cook and More, 2000). We measure these variables in both waves I and Wave II, from the in-home questionnaire.

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

[20]We dropped 375 students, which showed inconsistent time answers in the category *never tried smoking* between waves I and II.

students (from 0.61 to 0.88 on average). Girls make up about 52 percent of the sample. Around 63 percent of the sample is White, 14 percent is Black or African American, 16 percent is Hispanic, and the remainder is Asian, American Indian, or unclear racial background. The average height in the sample is 5.61 feet, and the average age is around 17 years old. Finally, around 28 percent of our adolescents have cigarettes easily available at home. Regarding student family background, about 76 percent of the adolescents in our sample have two parents living in the household, and roughly 27 percent have parents who are college graduates or above. These percentages are in line with data from other nationally representative surveys. We report in Table 11 in Appendix H information from the 1994 Current Population Survey (CPS) that is re-weighted to match the age distribution of the Add Health sample. As shown, the Add Health population is broadly similar to the U.S. population as calculated from the CPS.[21]

# 4 Empirical Methodology

We use the panel nature of our data to structurally estimate our dynamic social interaction model. The structural system of equations we estimate

$$\mathbf{y}_{sgt} = \alpha_1 \mathbf{B}_{sgt}\, \mathbf{y}_{sgt-1} \,+\, \alpha_2\, \mathbf{B}_{sgt}\Big[\mathbf{D}_{sgt} + \theta_{sgt}\Big] \,+\, \varepsilon_{sgt} \tag{4}$$

is the reduced-form system describing the equilibrium (Equation 3), implied by the intertemporal first-order conditions of agents' dynamic optimisation problems. Equation 4 expresses the vector of outcomes $\mathbf{y}_{sgt}$ (cigarette smoking behavior) of individuals in school $s$, with gender $g$, and in grade $t$ as a function of the outcomes in the previous grade, $\mathbf{y}_{sgt-1}$, and contemporaneous variables (including expectations about the future, $\mathbf{D}_{sgt}$).[22]

Equation 3 depends on the structure of the stochastic process of preference shocks $\theta_T$, 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 $x^{(k)}_{isgt}$ denote its value for agent $i$ at school $s$, with gender $g$ and in grade $t$. The preference shock $\theta_{sgt}$ is allowed to depend on individual $i$'s characteristics and on those of the member of

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

[22] The interpretation of this regression should naturally account for the discrete nature of the variable $\mathbf{y}_{sgt}$.

her reference group. It is decomposed as follows:

$$\theta_{sgt} \quad := \quad \underbrace{\gamma_t\,\mathbf{1}_{m_{sg}} + \kappa_{st}\,\mathbf{1}_{m_{sg}} + \mathbf{X}_{sgt}\beta + \mathbf{G}_{sg}\mathbf{X}_{sgt}\,\phi}_{\text{Observed exogenous heterogeneity } (a_{it})} + \underbrace{\mathbf{u}_{sgt},}_{\text{Unobserved component}} \tag{5}$$

where $\beta$ and $\phi$ are $k \times 1$ vectors of parameters, $\gamma_t$ is an intercept, and $\kappa_{st}$ is a school effect (specific for each grade) with $\mathbf{1}_{m_{sg}}$ is a conformable vector of ones. More compactly the stochastic process for preference shocks can be written as

$$\theta_t = \mathbf{1}\gamma_t + \iota\kappa_t + \mathbf{X}_t\beta + \mathbf{G}\mathbf{X}_t\,\phi + \mathbf{u}_t, \tag{6}$$

where $\mathbf{X}_t = (\mathbf{X}_{1m}, \ldots, \mathbf{X}_{\bar{r}m}, \mathbf{X}_{1f}, \ldots, \mathbf{X}_{\bar{r}f})$ is an $N_t \times K$ matrix, $\iota = diag(\mathbf{1}_{m_1}, \ldots, \mathbf{1}_{m_{\bar{r}}})$, $\mathbf{1} = (\mathbf{1}_{m_1}, \ldots, \mathbf{1}_{m_{\bar{r}}})$, $\mathbf{G}_g = diag(\mathbf{G}_{1g}, \ldots, \mathbf{G}_{\bar{r}g})$ and $\mathbf{G} = diag(\mathbf{G}_m, \mathbf{G}_f)$ with $g = \{m, f\}$, $\kappa_t = (\kappa_{1t}, \ldots, \kappa_{\bar{r}t})$, where $\bar{r}$ denotes the number of schools, and $N_t$ is the number of individuals at grade $t$. The elements of the social interaction matrix $\mathbf{G} = [g_{ij}]$ are row-normalized and the interactions have the following group structure

$$g_{ij} = \begin{cases} \frac{1}{|N_G(i)|}, & \text{if } i \text{ and } j \text{ have the same gender and are in the same grade-school} \\ 0, & \text{otherwise,} \end{cases} \tag{7}$$

where $|N_G(i)|$ denotes the number of students in the reference group of $i$.

Finally, the structural equation system we estimate, which defines the equilibrium of the dynamic social interaction economy, is obtained by substituting the equation for the preference shocks, Equation 6, into the general reduced-form system, Equation 4:

$$\mathbf{y}_t = \mathbf{B}_t\Big[\alpha_1\,\mathbf{y}_{t-1} + \alpha_2\,\left(\mathbf{D}_t^x + \mathbf{1}\gamma_t + \iota\kappa_t + \mathbf{X}_t\beta + \mathbf{G}\mathbf{X}_t\,\phi\right)\Big] + \varepsilon_t, \quad t = 1, 2, \ldots, T \tag{8}$$

where $\mathbf{y}_t = (\mathbf{y}_{1m}, \ldots, \mathbf{y}_{\bar{r}m}, \mathbf{y}_{1f}, \ldots, \mathbf{y}_{\bar{r}f})$ is $N_t \times 1$ vector of outcomes and $\mathbf{D}_t$ represents the discounted sum of the effects of expected future $\theta_\tau$, $\tau > t$. 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 $\varepsilon_t = \alpha_2\mathbf{B}_t\,(\mathbf{D}_t^u + \mathbf{u}_t)$. $\mathbf{B}_t$ is a nonlinear function of $\alpha_1$, $\alpha_2$, $\alpha_3$, $\delta$ and $\mathbf{G}$. It is computed recursively starting from the last period where $\mathbf{B}_T = (\Delta_T\mathbf{I} - \alpha_\mathbf{3}\mathbf{G})^{-\mathbf{1}}$ with $\Delta_T = \sum_{i=1}^3 \alpha_i$. Also $\mathbf{D}_t^x$ is computed recursively starting from the last period where it is assumed to be zero given that there is no future. It is in general a nonlinear function of $\mathbf{B}_t$, $\delta$ and $E\left[(\beta\mathbf{I} + \phi\mathbf{G})\,\mathbf{X}_T\,|\mathbf{X}^{T-1}\right]$, which is a function of $\mathbf{X}^{T-1} = (\mathbf{X}_1, \ldots, \mathbf{X}_{T-1})$ observed by the econometrician. The detailed definitions of $\mathbf{B}_t$ and $\mathbf{D}_t$ and the description of the recursive algorithm can be found in Appendices B and C. In the absence of future periods, the model in the last period is similar to a reduced form of a static social interaction model

$$\mathbf{y}_T = (\Delta_T \mathbf{I} - \alpha_\mathbf{3}\mathbf{G})^{-\mathbf{1}} \Big[ \alpha_\mathbf{1}\, \mathbf{y_{t-1}} + \alpha_\mathbf{2}\, (\mathbf{1}\gamma_\mathbf{T} + \iota\kappa_\mathbf{T} + \mathbf{X_T}\beta + \mathbf{G}\mathbf{X_T}\,\phi) \Big] \; + \varepsilon_\mathbf{T}. \tag{9}$$

In our empirical exercise, the empirical counterpart of this equation system at time $t = T$ is constructed by considering students in grade 12 ($t = 12$) in wave II and the same individuals in grade 11 ($t = 11$) 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 $t = 11$ in wave II and the same individuals in grade $t = 10$ in wave I; and so on until $t = T - 3$ with students in 9th grade (wave II) and the 8th (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 10, 11, and 12). We then exploit the structural equation at $t = T - 3$ with data referring to middle school students (grades 8, and 9) to validate the model.

## 4.1 Identification

In this section, we derive conditions under which the dynamic model with social interactions we have introduced is identified when the number of individuals $N$ is large enough so that sample averages converge to population expectations, and the horizon of the economy $T$ and the social structure $\mathbf{G}$, are fixed and known to the econometrician.

The parameters of the economy are the utility parameters $(\alpha_1, \alpha_2, \alpha_3)$, the discount factor $\delta$, 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 $\sum_i \alpha_i = 1$.

The main challenge we need to meet is the endogeneity of the lagged smoking choices. Our solution to this challenge is by finding suitable instrumental variables that are correlated with the endogenous lagged smoking choices, and at the same time would satisfy exclusion restrictions by being independent of the unobservable component conditionally on other covariates.

**Exogeneity:** $E[\mathbf{u}_t \mid (\mathbf{X}_t, \mathbf{s}_{t-1})_{t=1}^T] = 0$, for any $t = 1, \ldots, T$.

**Full rank:** The second moment matrix generated by the elements of $\{\mathbf{X}_t, \mathbf{s}_t\}_{t=1}^T$ has full rank.

**Regularity:** $E(\mathbf{y}_{t-1}|\mathbf{s_{t-1}}) \neq \mathbf{0}$, for any $t = 1, \ldots, T$.

Exogeneity requires the existence of variables in last period's information set that are orthogonal to the unobservables contemporaneously and intertemporally. Full rank requires lack of multicollinearity and enough intertemporal variation in the instrumental variables. Finally, Regularity requires that the instrumental variables 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 8, for $T \geq 2$, are identified.[23]

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.[24] In our dynamic environment, we are not restricted to the cross-section. In particular, as we discuss in more detail in Section 4.2, we have access, for each period, to a set of 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.

The proof of this identification result proceeds in two steps: i) we prove that the coefficients of the structural equation, Equation 8, 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 8 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.

[23] 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 $p$ .

[24] 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, Djebbari and Fortin (2009), Calvo-Armengol, Patacchini and Zenou (2009).

Consider the system in Equation 8 at $t = T, T-1$. Because $\mathbf{D}_t$ contains expectations about future shocks, $\mathbf{D}_T = 0$. The system in 8 can be reduced to

$$\mathbf{y}_T = \alpha_1 \mathbf{B}_T \mathbf{y}_{T-1} + \alpha_2 \mathbf{B}_T (\mathbf{1}\gamma_T + \iota\kappa_T + \mathbf{X}_T\beta + \mathbf{G}\mathbf{X}_T \phi) + \varepsilon_T, \tag{10}$$

$$\begin{aligned} \mathbf{y}_{T-1} &= \alpha_1 \mathbf{B}_{T-1} \mathbf{y}_{T-2} + \alpha_2 \mathbf{B}_{T-1} (\mathbf{1}\gamma_{T-1} + \iota\kappa_{T-1} + \mathbf{X}_{T-1}\beta + \mathbf{G}\mathbf{X}_{T-1} \phi) \\ &\quad + \alpha_2 \mathbf{B}_{T-1} \mathbf{D}^x_{T-1} + \varepsilon_{T-1}. \end{aligned} \tag{11}$$

The endogeneity of $\mathbf{y}_{T-1}$ in Equation 10 and of $\mathbf{y}_{T-2}$ in Equation 11 requires to find suitable instrumental variables. Consider selecting $\mathbf{q}_t = [\mathbf{X}_t, \mathbf{G}\mathbf{X}_t, \mathbf{s}_{t-1}]$, $t = T-1, T$. Predicted values of $\mathbf{y}_{t-1}$ are formed by projecting them onto the space spanned by the set of instrumental 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 10 is independent of $\delta$ and hence the condition $T \geq 2$ is necessary to identify it.

### 4.2 Estimation

We jointly estimate the reduced-form equilibrium equations (8), for $t = T, T-1, T-2$, which are related to each other through the the dynamic recursive structure of the equilibrium.[25] In particular, we implement a nonlinear IV estimator (NLIV or nonlinear 2SLS, Amemiya, 1974) where our target parameters are given by $\lambda = [\alpha_1, \alpha_3, \delta, \beta', \phi']'$. Let $\mathbf{z}_t = [\mathbf{y}_{t-1}, \mathbf{X}_t, \mathbf{G}\mathbf{X}_t]$ and $\mathbf{q}_t = [\mathbf{X}_t, \mathbf{G}\mathbf{X}_t, \mathbf{s}_{t-1}]$ be vectors of explanatory variables and instruments, respectively.

More compactly, we stack all variables in matrix format as $\mathbf{Y} = [\mathbf{y}'_T, \ldots, \mathbf{y}'_{T-2}]'$, $\mathbf{Z} = [\mathbf{z}'_T, \ldots, \mathbf{z}'_{T-2}]'$, $\mathbf{Q} = [\mathbf{q}'_T, \ldots, \mathbf{q}'_{T-2}]'$, with $\mathbf{q}_g = diag(\mathbf{q}_{1g}, \ldots, \mathbf{q}_{\bar{r}g})$ and $\mathbf{q}_t = diag(\mathbf{q}_m, \mathbf{q}_f)$ with $g = \{m, f\}$. Finally, let

$$F(\mathbf{Z}, \lambda) = \begin{bmatrix} \alpha_1 \mathbf{B}_T \mathbf{y}_{T-1} + \alpha_2 \mathbf{B}_T (\mathbf{1}\gamma_T + \iota\kappa_T + \mathbf{X}_T\beta + \mathbf{G}\mathbf{X}_T \phi) + \varepsilon_T \\ \vdots \\ \alpha_1 \mathbf{B}_{T-3} \mathbf{y}_{T-3} + \alpha_2 \mathbf{B}_{T-3} (\mathbf{1}\gamma_{T-3} + \iota\kappa_{T-3} + \mathbf{X}_{T-3}\beta + \mathbf{G}\mathbf{X}_{T-3} \phi) + \alpha_2 \mathbf{B}_{T-3} \mathbf{D}^x_{T-3} \end{bmatrix}.$$

The moment conditions are then implied by the exogeneity assumption

$$E(\mathbf{Q}'[\mathbf{Y} - F(\mathbf{Z}, \lambda)]) = E(\mathbf{Q}'\varepsilon) = 0.$$

[25] See Table 1 to see how we link the structural equations and the sample

The nonlinear IV objective function is the following

$$Q_N(\lambda) = \frac{1}{N}\varepsilon'\mathbf{Q}(\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{Q}'\varepsilon. \tag{12}$$

In practice, we implement it in three steps.

**Step 1** Estimate model 9 (i.e., using only the last period) using as starting values estimates from OLS estimation of a reduced form model where the smoking behavior is regressed on its lag, individual's and peers' characteristics controlling for grade dummies and school fixed effects. Results are reported in Table 2. To constraint $\alpha_1, \alpha_3$, we reparameterize them using the function

$$\alpha_i = \frac{1}{\pi}\left(\arctan(\psi_i) + \left(\frac{\pi}{2}\right)\right) \in [0, 1] \tag{13}$$

and then we minimize w.r.t. $\psi_i$. The starting values of $\alpha_1$ and $\alpha_3$ (hence also that of $\alpha_2$) are set uniformly to 0.3333 for each.

**Step 2** Estimate model 8 using as starting values the estimates from **Step 1**.[26] We recover $\hat{\alpha}_2 = 1 - \hat{\alpha}_1 - \hat{\alpha}_3$. Also $\delta$ is reparameterized using the mapping in equation 13.

**Step 3** For inference, we use the standard asymptotic variance (clustered at the group level) for the NLIV estimator where the Jacobian is estimated numerically because no closed form expression is available. In particular, let $\widehat{\mathbf{S_t}} = \sum_{i=1}^{N_t}\sum_{s=1}\sum_{g=1}\widehat{\varepsilon}^2_{isg}\mathbf{q}_{isg}\mathbf{q}'_{isg}$ with $\hat{\varepsilon}_{isgt} = y_{isgt} - F(\mathbf{z}_{isgt}, \widehat{\lambda})$ , $\mathbf{AA} = \widehat{\mathbf{D}}'\mathbf{Q}(\mathbf{QQ})^{-\mathbf{1}}\mathbf{Q}'\widehat{\mathbf{D}}$, and $\mathbf{B} = \widehat{\mathbf{D}}'\mathbf{Q}(\mathbf{QQ})^{-1}\hat{\mathbf{S}}(\mathbf{QQ})^{-1}\mathbf{Q}'\widehat{\mathbf{D}}$ where $\widehat{\mathbf{D}} = \frac{\partial\varepsilon}{\partial\lambda}$ evaluated at $\lambda = \widehat{\lambda}$ and $\widehat{\mathbf{S}} = N^{-1}(\widehat{\mathbf{S}}_1, \ldots, \widehat{\mathbf{S}}_T)'$. The estimated asymptotic covariance matrix is $\widehat{\mathbf{V}} = N\left(\mathbf{AA^{-1}BAA^{-1}}\right)$.[27]

The statistical properties of the NLIV estimator are based on standard high-level assumptions in the recent literature on clustered samples (Hansen and Lee, 2019).[28]

For the main parameters, we test the hypotheses that $\alpha_i = 0$ ($\delta = 0$) versus $\alpha_i > 0$ ($\delta > 0$). We performed the t-test as one-sided, given that the parameters cannot take negative values. For $\alpha_1$ and $\alpha_3$, the standard errors are computed using the Jacobian of the function $\psi_i$ w.r.t $\alpha_i$.

[26] We set the starting value of $\delta = 0.9$.

[27] The variance of $\alpha_2$ is estimated using the variance of a linear combination, that is $var(\alpha_2) = var(1-\alpha_1-\alpha_3) = var(\alpha_1) + var(\alpha_3) + 2cov(\alpha_1, \alpha_3)$.

[28] In this framework, the asymptotic approximation holds for a large number of independent groups with relatively small group sizes. In our empirical application, we have almost 60 groups with median size of 6 students.

Observe that for the case of singular hypothesis one-sided t-test inference is valid, even if the null is on the boundary of the parameter space (Kodde and Palm, 1986).

### 4.2.1 Empirical Challenges

There are two main empirical challenges in the estimation of model 8: the endogeneity of the social interaction structure and the endogeneity of lagged choices. We tackle the first issue by defining peers as all schoolmates in the same cohort and exploiting quasi-random variation of peers' characteristics across cohorts within a school.[29] The idea is to estimate a model with school and grade fixed effects by presuming that neither students nor parents can anticipate perfectly the composition of schoolmates' characteristics in a given grade when choosing schools or residential neighborhoods. This is a well-known empirical strategy used with AddHealth data since multiple cohorts are observed within schools.[30] The challenge in our case is that each FOC represents equilibrium choices of agents in a given grade and so to have variation within grade we define peer groups by gender exploiting the differential exposure to peers with a given characteristic within grade and gender.[31] We, indeed, show that the variation in the average peer characteristics of same grade and gender schoolmates are unrelated to the variation in a number of predetermined student characteristics in Table 12 in Appendix G. We run separate regressions with each of students' background characteristics on peers' average characteristics controlling for the corresponding individual characteristic. Each regression includes a gender indicator, grade dummies and school fixed effects to control for differences in average student characteristics across schools as well as for other aspects of school quality. As we present in Table 12, almost 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.

For the second challenge, namely the endogeneity of the lagged choices, we adopt an identification strategy based on external instrumental variables. We consider each student's exposure

[29] In the Addhealth questionnaire, respondents can nominate up to ten best friends. However, the use of self-reported friends' nominations to define social networks yields endogeneity issues in the estimation of network effects that are difficult to solve.

[30] This approach has been first proposed by Hoxby (2000a) 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, 2020). Also, Patacchini and Zenou (2016) and use a similar approach to investigate the impact of peer religiosity in the intergenerational transmission of religion. In health economics and using the same data, this strategy has been implemented by Arduini, Iorio, and Patacchini (2019) to study the relationship between peers' body size through interpersonal comparisons and the development of eating disorders.

[31] In Section 6, we investigate the robustness of our results to an alternate definition of peers.

to mandatory school-level "Tobacco Use Prevention" programs measured using school absence. First, we collect information on whether the school offers a "Tobacco Use Prevention" program using the information on the mandatory programs in terms of health education, health services, and health policies in each state contained in the School Health Policies and Programs Study (SHPPS).[32] For each school we then construct an indicator variable that assumes the value of one if the school had a tobacco use prevention program in place, and zero otherwise.[33] Next, for each student, we measure her/his exposure to the program by interacting the program indicator with an indicator of school attendance. The Add Health questionnaire asks about school attendance using two questions: "how many times have you been absent from school for a full day with an excuse- for example, because you were sick or out of town?"; and "how many times did you skip school for a full day without an excuse?". The school-attendance index is the average number of school days in the survey period, that is 180 days, minus the sum of the days reported by respondents in the answers to the two questions. We refine our identification strategy by leveraging the unique information contained in the AddHealth data on attendance and risky behaviors. As mentioned in above, we have information on the reasons underlying absence from school. We construct two indicators: *excused absence* and *truancy*. The dummy *excused absence* equals one if the respondent skips school with an excuse. The dummy *truancy* is equal to one if the respondent is absent without reason for more than two days (75th percentile).[34] Table 13 shows that while *truancy* is correlated with other risky behaviors potentially associated with smoking, *excused absence* is not.[35] Therefore, we focus on the sample when excluding students who report truancy (269 observations).

Table 2, column (3) reports linear 2SLS estimates of model 9 for high-school students controlling for grade dummies, school fixed effects, and setting $\alpha_3 = 0$ and $\alpha_2 = 1$.[36] The first-stage F-test has a value 16.60, supporting the instrument relevance. In Table 2, we also report the Anderson-Rubin (AR) test (Anderson and Rubin, 1949) for the significance of $\alpha_1$ since this test is robust to a potential weak instrument problem (see also Kleine and Neil, 2023). The value of

[32] The Add Health contains SHPPS data for Wave I. Data collection was conducted via mail from March to June 1994. All 51 state education agencies completed the questionnaires. At the beginning of 1994, the Centers for Disease Control and Prevention (CDC) guidelines required all the states to offer a tobacco use prevention program at school, but in 1994 not all the states (37 out of 51) had time to adhere to the guideline (CDC, 1994b).

[33] For robustness check, we repeated our analysis when excluding schools in states that did not adhere. Results remain qualitatively unchanged. They are available upon request.

[34] The fraction of truants in the sample is roughly 20%

[35] Regression samples are smaller due to missing values in risky behavior variables. The construction of the risky behavior indicators is detailed in Appendix G.

[36] Table 2, column (1) and (2) reports OLS estimates without and with peers' characteristics.

the AR test is equal to 24.14, confirming that $\alpha_1$ is statistically significant at 1% level.

## 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* $\delta$, the *addiction effect* parameter $\alpha_1$, the *own effect* parameter $\alpha_2$, and the *peer effects* parameter $\alpha_3$, which we normalize without loss of generality so that $\sum_{i=1}^{3} \alpha_i = 1$.

Table 3 presents the estimates of the structural parameters: in Column 1 the peers' average characteristics (*contextual effects*) are not included, whereas in column two they are. In both columns, estimates follow the embedded preference shock structure in equation (5) as we explain carefully in Section 4. Standard errors are clustered at the school level. Our main findings below are consistent across columns.

We find evidence of rational forward-looking behavior: the estimate for $\delta$ is positive, large, significant and relatively stable, independently of our multiple treatments (See Table 9 for robustness checks). Interestingly, the estimate for the own effect, $\alpha_2$, is very small compared to the other deep parameters. Note that the estimated model imposes the structure of the stochastic process of preference shocks in Equation 5. Hence, we also estimate the vector of parameters multiplying different components of individuals' observable characteristics, some of which are very significant and large, and report them in Table 14.[37] Moreover, as we report in Section 5.2, the standard metric (social multiplier) used to represent the implications of peer effects is not directly affected by the value of the own effect estimates; and we report compelling evidence pointing to the existence of strong interaction effects in the data. Perhaps, most importantly, the peer effect $\alpha_3$ and the addiction effect $\alpha_1$ are also significant. The combined evidence shows that dynamics generated by addiction effects ($\alpha_1 > 0$) and social interactions ($\alpha_3 > 0$) are important, quantitatively large, and statistically significant. The validation exercise we implement next in Section 5.1 gives statistical support to the prediction power of our model: comparing out-of-sample predicted equilibrium smoking behavior to actual behavior using data for different grades not employed in the estimation sample, we see that the actual proportions in the data and the predicted ones are remarkably close.

[37] A small estimate of the own effect, $\alpha_2$, also implies that the heterogeneity induced by the stochastic component of $\theta_{it}$ has relatively small direct effects. The stochastic component of $\theta_{it}$, however, affects agents' choices also indirectly through the addiction effect of $y_{i,t}$ on $y_{i,t+1}$, and the peer effect of $y_{i,t}$ on all $y_{j,t}$. Both of these effect then amplify via the equilibrium dynamics.

## 5.1 Out-of-sample Validation

In this section, we validate the structural estimates of our dynamic recursive model, comparing out-of-sample predicted equilibrium smoking behavior to actual behavioral data across different grades, 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 fulfill this objective, we follow an approach inspired by Todd and Wolpin (2006).[38] 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 specification: using the reduced-form equilibrium equations 8 jointly, for $t = T, T-1, T-2$, by linking the structural equations and the sample as in Table 1, with and without contextual effects. Those parameter estimates under both specifications are reported in Table 3.

Next, we predict equilibrium smoking behavior for students in the hold-out sample consisting of grade 9. Specifically, we use the reduced-form equilibrium equations (8), for $t = T-3$, 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 3, the baseline controls (and network $\mathbf{G}$) for grades 8 and 9, and students' initial smoking behavior values for $t = T-4$, as reported in the data.

Table 4 reports the percentage of choices correctly predicted. Predicted choices are generated by splitting the interval $[0, 2]$ uniformly of the mean predicted outcomes for the 9th grader's sample and several subgroups defined by student characteristics. We believe the reported percentages (around 70% for the unconditional sample) deliver enough confidence in our dynamic model and its out-of-sample prediction power.

## 5.2 Social Interaction Effects

In economies where individual preferences incorporate conformity effects, including the current model, a change in the value of an exogenous variable yields a direct effect on behavior and an infinite cascade of indirect effects of the same sign. That is because each agent's action changes

[38] 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.

not only because of the change in the exogenous variable, but also because of the change in the behavior of her peers, and that of their peers and so on. In this section we construct a conceptual measure of these social interaction effects and evaluate it at the estimated parameter values of our dynamic structural model.

As we have already noticed, however, this construction is quite subtle. We discuss all details and produce all relevant computations and results in Appendix E. As an illustration, considering social interaction effects in the last period $t = T$, so as to silence any dynamic effects. Consider an exogenous shock to $\theta_{i,T}$, say by $\Delta\theta_{i,T} = 1$, 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. From the structural equation 3, the total change on the smoking behavior values at $T$, $\Delta\mathbf{y_T}$, can be decomposed into a direct effect, $\Delta\mathbf{y_t}\mid_{direct}$ and a social effect, accounting for peer interaction effects in the network at equilibrium, $\Delta\mathbf{y_t}\mid_{social}$. such that $\Delta\mathbf{y_T} = \Delta\mathbf{y_t}\mid_{social} \times \Delta\mathbf{y_t}\mid_{direct}$. It is then straightforward to show that $\Delta\mathbf{y_t}\mid_{direct} = \alpha_2$ and $\Delta\mathbf{y_t}\mid_{social} = \frac{1}{\alpha_1+\alpha_2}$. Therefore, $\Delta\mathbf{y_T} = 1$ when preferences do not account for addiction, $\alpha_1 = 0$. In this case, then, the total effect is independent of preferences for conformity parameter $\alpha_3$ and the *social multiplier* as defined by Boucher and Fortin (2016), $\Delta y_{i,T} - 1$, is equal to zero. But if we add complementarity effects, that is, if we add a term $\sum_{j=1}^{N} 2\gamma\, y_{it}\, y_{jt}$ to preferences, as discussed in footnote 9, the *social multiplier* as defined by Boucher and Fortin (2016) is equal to $\frac{\alpha_2}{\alpha_1+\alpha_2-\gamma} - 1$, and hence it might be positive and always increasing in $\alpha_3$. Therefore, while the preference parameter driving complementarities $\gamma$ - and hence the social multiplier as defined by Boucher and Fortin (2016) - cannot be identified with aggregate behavior (and indeed we therefore postulate $\gamma = 0$), the ratio of the total effect and the social effect, $\frac{\Delta\mathbf{y_t}}{\Delta\mathbf{y_t}\mid_{direct}} = \Delta\mathbf{y_t}\mid_{social}$ can be interpreted as a measure of the *social interaction effects* in the economy we study.[39] This is in fact how we proceed in this paper.

Going back to the dynamic economy, in this section we report on the social interaction effects implied by our structural estimates, taking into account contextual effects. Importantly, these effects, in our context, have a fundamental dynamic component: a shock to agents' preferences in the 10th grade has social effects in all grades 10, 11, and 12. The summary of the estimated parameters for the dynamic, myopic, and static models we use are in Table 5. See Section 5.3 for a discussion of the myopic and static model specifications. Consider again an exogenous shock to

[39] This is the sense in which in fact both Becker and Murphy (2001) and Glaeser and Scheinkman (2003) use the term *social multiplier*.

$\theta_{i,t}$, say by $\Delta\theta_{i,t} = \pi$, for all agents $i$. Let then social interaction effects at $t$ on behavior at $t$ be defined as

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

From the structural equation (3) for $\tau = t, \ldots, T$, we have $m_{t,t} = \mathbf{B}_t \mathbf{1}$. Iterating, we trace the (*dynamic forward-looking*) social effects of an exogenous preference shock at $t$ at $\tau$:

$$m_{t,\tau} \;\; = \;\; \alpha_1^{\tau-t} \left(\mathbf{B}_t \times \cdots \times \mathbf{B}_\tau\right) \mathbf{1} \tag{14}$$

for any period $\tau = t, \ldots, T$.[40] Similarly, for a permanent shock from time $t$ to the end-time $\tau$, we can compute the social effect as:

$$m_{t,\tau} \;\; = \;\; \sum_{s=t}^{\tau} \alpha_1^{\tau-s} \left(\mathbf{B}_\tau \times \cdots \times \mathbf{B_s}\right) \left(\mathbf{1} + \tilde{D}_s\right) \tag{15}$$

Now, consider a shock to one of the components of students' observable covariates, e.g., a shock to *Tobacco at home*, *Family income*, or to other observable covariates we use in our estimation, as reported in Table 10. As we expressions above show, the social effects are not directly affected by the own effect parameter $\alpha_2$. They are affected by two channels: (i) the addiction channel ($\alpha_1$) which transmits shocks across periods thanks to intertemporally optimising individuals, and (ii) the peer effects channel ($\alpha_3$) that transmits them across the social network thanks to individuals conforming to their peers optimally. The sample means of the social effects ($\bar{m}_{t,\tau}$) are reported in Table 6.

Several interesting properties of the calibrated social effects are worth noticing. First of all, the effects of a temporary preference shock at $t$ decline over time: $m_{t,\tau}$ decreases with $\tau$, for all $t$. The effects of permanent shocks instead increase over time. Furthermore, permanent shocks have larger effects than temporary shocks, both instantaneously and over time.[41] Perhaps most interestingly, the same-period 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 permanent shocks, our estimates imply that the same-period social effect in grade 12 (the last year of high-school) is about 1.1220, whereas it is 2.1256 and 3.0231, respectively, in

[40] Please see Appendix F for the detailed derivation of and the recursive algorithm to compute the dynamic social multiplier.

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

grades 11 and 10. The same-period social 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 social effect. As demonstrated by the analogous social effects, the myopic and static models cannot generate this effect, since agents are assumed not to care about future consequences of their actions; an important distinction that we discuss further in Section 5.3.

Social 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 social effect of an exogenous preference shock at a future date $t$ on behavior at $\tau < t$:

$$m_{t,\tau} \quad = \quad \mathbf{B}_t \tilde{D}_t \tag{16}$$

where $\tilde{D}_t$ capture the sum of the expected effects on period $\tau$ marginal utility of a unit future shock that is anticipated to change the random component of preferences, $\theta_t$.[42]

We consider an anticipated shock to the future real disposable income of a student at time $\tau > t$. As much as one needs to be careful about the income and substitution effects, this is a rough proxy for a decrease in the price of cigarette. A forward-looking agent anticipates this change and changes his/her behavior today. Neither the static nor the myopic model can generate any social effects here, since by definition, they do not care about the future.

In Table 7, we report the expected social effects, in grade 10, 11 and 12 , induced by an anticipated shock to the preferences of all agents in grades 11 and 12; that is, we report $m_{t,\tau}$ for $t = 11, 12$ and $\tau = 10, 11, 12$. In anticipation of an increased preference for risky behavior in grades 11 or 12, agents increase risky behavior in grades 10 and 11 (all values are positive). These

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

$$\begin{aligned} \Delta \mathbf{D}_t \quad := \quad & \sum_{s=t+1}^{t} \delta^{s-t} \Bigg( - \alpha_1 \operatorname{diag}\left(\Lambda_{t,s-1} - \Lambda_{t,s}\right) \left(\Delta\Gamma_{t,s-1} - \Delta\Gamma_{t,s}\right) \\ & + \operatorname{diag}\left(\Lambda_{t,s}\right) \left(\Delta\bar{\theta}_s - \Delta\Gamma_{t,s}\right) \\ & - \alpha_3 \sum_{k=1}^{N} \operatorname{diag}\left(G_{\bullet k}\, \iota_N'\right) \operatorname{diag}\left(\iota_N\, \Lambda_{k\bullet,t,s} - \Lambda_{t,s}\right) \left(\Delta\Gamma_{k,t,s}\mathbf{1} - \Delta\Gamma_{t,s}\right) \Bigg) \end{aligned} \tag{17}$$

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

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 ($\bar{m}_{12,12} = 1.1220$, from Table 6). 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 optimally.

Interestingly, however, a dampening effect could occur in grade 10 inducing the agents to shade the increase in their risky behavior more in the earlier periods and then to increase in grades 11 and 12 at faster rates. This occurs in particular at the estimated parameter values, and more generally when peer effects are particularly strong, that is, $\alpha_3$ 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. This risky behavior smoothening is due to an adjustment cost in terms of utility for the agents: There is a trade-off between smoothing the adjustment costs via the addiction channel and via the own preference channel in the presence of large peer effects.

### 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 $\alpha_1$ 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 $\delta$ is estimated to be significantly different than 0.

To better gauge at the relevance of dynamic forward looking behavior, in this section we estimate (i) a *myopic* model, obtained by restricting the structural equation (3) by imposing $\delta = 0$; (ii) a *static* model, obtained restricting the structural equation (3) by imposing $\delta = \alpha_1 = 0$. We then compare the goodness of fit of these models with those of our *dynamic* baseline model with no restrictions. Results are reported in Table 8.

Although the parameters of both the static and myopic models are imprecise estimates, we begin our analysis with a qualitative comment. Since agents are by definition myopic in the myopic model, the expectations channel of future peer interactions is not captured in the parameter estimates. This absence is compensated by a higher own effect estimate ($\hat{\alpha}_2 = 0.0441$), relative to the dynamic benchmark.

The crucial test to appreciate our contribution with respect to the static peer effect model— which is the go-to model used in the literature— is to compare the fit of the model to the data across the two models. We estimate the static model (no addiction, no forward looking behavior) and demonstrate in Table 8 that there is a substantial jump in goodness of fit from the static to the dynamic.

In addition to this evidence, observe that the restrictions imposed by both the static and the myopic models may induce a bias in the estimation of the social multipliers. Firstly, as can be seen immediately from the construction of the multiplier measures in (14), (16) and (15), a misspecified static model cannot generate any intertemporal social 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 the discount factor $\delta \to 0$, the difference between the last period ($t = T$) and period-$t$ equilibrium maps $|\mathbf{B}_T - \mathbf{B}_t| \to 0$ for any period $t$, and the model becomes one of a sequence of myopic period economies. Consequently, the instantaneous social effect generated under this specification is constant across periods (see Tables 6 and 7), yielding a bias, which is *increasing in the time-to-end* $T - t$, relative to our benchmark dynamic specification.

# 6 Robustness

Next, we check if our main results are robust when we use alternate definitions of smoking behavior, and change the peer group definition.

**Dependent Variable definition.** Table 9 reports NLIV estimates of the structural models 8 under these alternative definitions. In the first column, the dependent variable is the binary smoking indicator in Badev (2021) (37% of students smoke one or more days), which bunches the group of students who report *occasionally smoking* together with the students who report *never tried smoking*. The second column uses a refinement of our baseline definition where we add one more category to our baseline dependent variable: *daily smoking*. This category refers to students who report *to smoke every day* (they are around 35% of the occasional smokers). Finally, the third column uses the finest definition of smoking by using the answers to the following question: "*During the past 30 days, on the days you smoked, how many cigarettes did you smoke each day?*" Hence, we construct a cigarette number indicator which takes 6 potential values: 0 for students who report to smoke zero cigarettes, 1 for those who report to smoke from 1 to 5, 2 for students who report to smoke from 6 to 10, and so on up to values greater than 21 (students smoke daily 2.9 cigarettes on average). The main message of Table 9 is that, as in our benchmark dynamic model, the estimate for $\delta$ is positive, large, and stable across outcomes although the effect is not statistically significant. More importantly, the estimates for peer effect $\alpha_3$ and the addiction effect $\alpha_1$ are also significant across columns, independent of the definition of smoking we use, which is evidence that equilibrium dynamics generated by addiction effects ($\alpha_1 > 0$) and social interactions ($\alpha_3 > 0$) are important, quantitatively large, and statistically significant.

**Peer definition**. We use gender to define groups within cohorts because homophily in friendship networks among adolescents is especially strong along the gender and race dimensions (Shrum, Cheek, and Mac, 1988; Currarini, Jackson, and Pin, 2009). We focus on gender because we need non-overlapping groups and the AddHealth questionnaire allows to report multiple races. For robustness, we have repeated our analysis when defining peer groups by race and using the first reported race for multiracial students. Specifically, we define students' peers as all other students of the same race in the same grade at the same school. Results are reported in Table 9. Table 9, column (4) shows that the results remain qualitatively unchanged.

# 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 adolescents' risky behavior. 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, and social interaction effects. Social interactions in the estimated dynamic model are indeed quantitatively large.

The importance of social interactions for policy analysis relies on the fact that when social interactions are quantitatively meaningful, 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 have a much smaller estimate of the social interaction effect than the dynamic model. Furthermore, our empirical analysis implies that the impact of policy interventions on adolescents' risky behavior increases 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 consider the network's dynamics, as students anticipate its natural breaks.

## References

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

AMEMIYA, T. (1974), "The Nonlinear Two-stage Least-squares Estimator," *Journal of Econometrics*, **2** (2), 105-110.

ANDERSON, T. W., AND RUBIN, H. (1949), "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *The Annals of mathematical statistics*, **20** (1), 46-63.

ANGRIST, J.D. AND LANG K. (2004),"Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program," *American Economic Review* 94, 1613-1634.

ANGRIST, J.D. (2014),"The Perils of Peer Effects," *Labour Economics* 30, 98-108.

ARDUINI, T., IORIO, D., AND PATACCHINI E. (2019), "Weight, Reference Points, and the Onset of Eating Disorders," *Journal of Health Economics*, **65**, 170-188.

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.

BADEV, A. (2021), "Nash Equilibria on (Un)stable Networks," *Econometrica*, **89**(3), 1179-1206.

BAKER, M.L., SIGMAN J.N., AND NUGENT M.E. (2001), "Truancy reduction: Keeping Students in School, Washington, D.C." *U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention.*

BALSA, A. AND DÍAZ, C. (2018): "Social Interactions in Health Behaviors and Conditions," *Oxford Research Encyclopedia of Economics and Finance*. Oxford University Press.

BATTAGLINI, M., PATACCHINI, E., AND RAINONE, E. (2021), 'Endogenous Social Interactions with Unobserved Networks," *The Review of Economic Studies.*

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

BECKER, G.S. AND MURPHY, K. M. (2001): *Social Markets: Market Behavior in a Social Environment*. Cambridge: Belknap-Harvard University Press.

BECKER, G.S., GROSSMAN, M., AND MURPHY, K. M. (1994): "An Empirical Analysis of Cigarette Addiction," *American Economic Review*, **84** (3), 396–418.

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 Decision 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.

BOUCHER, V. (2016), "Equilibrium Homophily in Networks," *European Economic Review,*, 123.

BOUCHER, V. AND FORTIN, B. (2016), "Some Challenges in the Empirics of the Effects of Networks," In Yann Bramoullé, Andrea Galeotti, and Brian W. Rogers, *The Oxford Handbook of the Economics of Networks*, Chapter 12, 277-302.

BOUCHER, V., RENDALL, M., USHCHEV, P., AND ZENOU, Y. (2022), "Toward a General Theory of Peer Effects."

BRAMOULLÉ, Y., DJEBBARI, H., AND FORTIN, B. (2009), "Identification of Peer Effects through Social Networks," *Journal of Econometrics*, **150**, 41-55.

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

BULOW, J., J. GEANOKOPLOS, AND P. KEMPLERER (1985), "Multimarket Oligopoly: Strategic Substitutes and Complements," *Journal of Political Economy*, **93**, 488-511.

CALVO-ARMENGOL, T., E. PATACCHINI AND ZENOU, Y. (2009), "Peer Effects and Social Networks in Education", *Review of Economic Studies*, **76** (4), 1239-1267.

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.

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) (1994b), "Guidelines for school health programs to prevent tobacco use and addiction", *Morbidity and Mortality Weekly Report* **43** (RR-2), 1–18. https://www.cdc.gov/mmwr/pdf/rr/rr4302.pdf

CENTERS FOR DISEASE CONTROL AND PREVENTION (CDC) (2021), "Tobacco Product Use and Associated Factors Among U.S. Middle and High School Students – National Youth Tobacco Survey, United States, 2021,", *Morbidity and Mortality Weekly Report, October 1, 2021,* **70** (39), 361-367, 1387–1389; https://www.cdc.gov/mmwr/volumes/71/ss/ss7105a1.htm?s_cid=ss7105a1_w

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

COELLI, T. (1995), "Estimators and Hypothesis Tests for a Stochastic Frontier Function: A Monte Carlo Analysis," *Journal of productivity analysis*, **6** (3), 247-268.

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 MOORE, M. J. (2000), "Alcohol," In A.J. Culyer and J.P Newhouse (eds.) *Handbook of Health Economics*, Vol. 1. (New York: Elsevier), 1629-1674.

COOPER, R. AND JOHN, A. (1988): "Coordinating Coordination Failures in Keynesian Models," *Quarterly Journal of Economics*, **103**, 441-464.

CHRISTAKIS, N.A., J.H. FOWLER, G.W. IMBENS, AND KALYANARAMAN, K. (2020): "An Empirical Model for Strategic Network Formation," *The Econometric Analysis of Network Data*, 123- 148.

CURRARINI, S., JACKSON, M. O., AND PIN, P. (2009): " An Economic Model of Friendship: Homophily, Minorities, and Segregation," *Econometrica*, **77** (4), 1003-1045.

DAHL, G., LØKEN, K., AND MOGSTAD, M. (2014), "Peer Effects in Program Participation," *American Economic Review*, **104** (7), 2049-2074.

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.

Ekpu, V. U., and Brown, A. K. (2015). “The Economic Impact of Smoking and of Reducing Smoking Prevalence: Review of Evidence,” *Tobacco Use Insights*, **8**, 1-35. https://doi.org/10.4137/TUI.S15628

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

Friesen, J. and Krauth, B. (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.

Glaeser, E. and Scheinkman, J. (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.

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

Hansen, B.E. and Lee, S. (2019), “Asymptotic theory for clustered samples,” *Journal of Econometrics*, **210** (2), 268–290.

Hanushek, E.A., Kain, J.F. and Rivkin, S.G. (2002), “Inferring Program Effects for Special Populations: Does Special Education Raise Achievement for Students with Disabilities?” *Review of Economics and Statistics*, 84, 584-599.

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.

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.

HSIEH C., KÖNIG M. AND LIU, X. (2022), “A Structural Model for the Coevolution of Networks and Behavior,” *The Review of Economics and Statistics*, **104**, 355-367.

JACKSON, K. M, SHER, K. J, COOPER, M. L, AND WOOD, P. K. (2002),“Adolescent Alcohol and Tobacco Use: Onset, Persistence and Trajectories of Use Across Two Samples,” *Addiction*, **97** (5), 517-31.

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

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.

KINNUNEN, J.M., LINDFORS, P., RIMPELÄ, A., SALMELA-ARO, K., RATHMANN, K., PERELMAN, J., FEDERICO, B., RICHTER, M., KUNST, A. E., LORANT, V. (2016),“Academic Well-being and Smoking Among 14- to 17-year-old Schoolchildren in Six European Cities,” *Journal of Adolescence*, **50**, 56-64.

KLEINE, M. AND NEIL, T. (2023), “Instrument Strength in IV Estimation and Inference: A Guide to Theory and Practice,” *Journal of Econometrics*, forthcoming.

KODDE, D. AND PALM, F. (1986), “Wald Criteria for Jointly Testing Equality and Inequality Restrictions,” **5** (54), 1243-1248.

KÖNIG, M. D. (2016), “The Formation of Networks with Local Spillovers and Limited Observability,” *Theoretical Economics*, **11** (3), 813-863.

LAVY, V., PASERMAN, D. AND SCHLOSSER, A. (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 SCHLOSSER, A. (2011), “Mechanisms and Impacts of Gender Peer Effects at School,”*American Economic Journal: Applied Economics* **3**, 1-33.

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.

Lewis, A. S., Oberleitner, L. M., Morgan, P. T., Picciotto, M. R., and McKee, S. A., (2016), "Association of Cigarette Smoking With Interpersonal and Self-Directed Violence in a Large Community-Based Sample," *Nicotine and Tobacco Research : Official Journal of the Society for Research on Nicotine and Tobacco*, **18** (6), 1456-1462. https://doi.org/10.1093/ntr/ntv287

Lundborg, P. (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.

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.

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.

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

Olivetti, C., Patacchini, E. and Zenou, Y. (2020). "Mothers, Peers and Gender-Role Identity", *Journal of the European Economic Association*, **18** (1), 266–301

Orphanides, A. and Zervos, D. (1995), "Rational Addiction with Learning and Regret", *Journal of Political Economy*, **103** (4), 739–58.

Özgür, O., Bisin, A. and Bramoullé, Y. (2020), " Dynamic Linear Economies with Social Interactions," mimeo Melbourne Business School, and NYU.

Patacchini, E., and Zenou, Y. (2016), "Social networks and Parental Behavior in the Intergenerational Transmission of Religion," *Quantitative Economics* 7(3), 969-995.

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.

Reif, J. (2019), "A Model of Addiction and Social Interactions," *Economic Inquiry*, **57** (2), 759-773.

Roebuck, M.C., French, M.T. and Dennis, M.L. (2004), "Adolescent marijuana use and school attendance," *Economics of Education Review*, **23**, 133-141.

Shrum, W., Cheek Jr, N. H. and Mac D, S. (1988), "Friendship in School: Gender and Racial Homophily," *Sociology of Education*, 227-239.

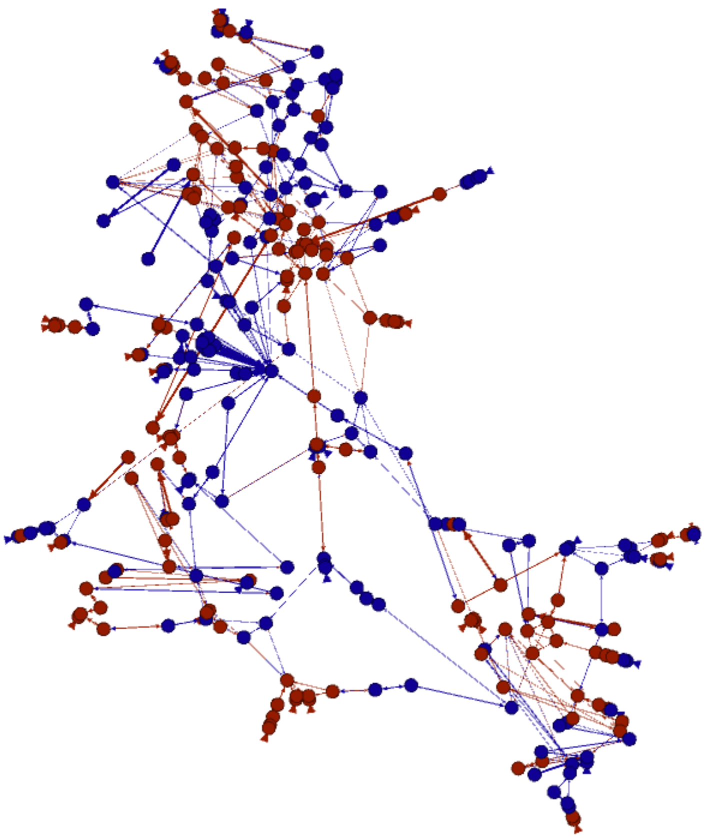
Todd, P. E. and Wolpin, K. I. (2006), "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility," *American Economic Review*, **96** (5), 1384-1417.

Weng, S. F., Ali, S., and Leonardi-Bee, J. (2015), "Smoking and Absence From Work: Systematic Review and Meta-analysis of Occupational Studies," *Addiction*, **108**, 307-319.

World Health Organization (2021), *WHO Fact Sheets Tobacco Web Page*, https://www.who.int/news-room/fact-sheets/detail/tobacco

## Figures and Tables

Figure 1: **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 red (resp., blue) dot correspond to female students (resp., male students).

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.

Table 2: **Reduced form IV model**

| Dep var.: | (1) | (2) | (3) |
|---|---|---|---|
| Smoking behavior (Wave 2) | OLS | OLS | 2SLS |
| Smoking behavior | 0.7147*** | 0.7101*** | 1.0224*** |
| | (0.0703) | (0.0717) | (0.1572) |
| Female | -0.0005 | 0.0212 | 0.1200 |
| | (0.0369) | (0.1709) | (0.2183) |
| Black or African American | -0.1582* | -0.1530* | -0.0652 |
| | (0.0790) | (0.0752) | (0.0566) |
| Asian | -0.0822 | -0.0846* | -0.0579 |
| | (0.0454) | (0.0451) | (0.0367) |
| Hispanic | -0.1491** | -0.1517** | -0.1273** |
| | (0.0531) | (0.0548) | (0.0433) |
| Indian | 0.0453 | 0.0493 | 0.0441 |
| | (0.0284) | (0.0311) | (0.0429) |
| PVT test score | -0.0502*** | -0.0479*** | -0.0344*** |
| | (0.0114) | (0.0085) | (0.0079) |
| Parents College degree | 0.0418 | 0.0449 | 0.0849** |
| | (0.0390) | (0.0343) | (0.0268) |
| Two-parent family | 0.1096*** | 0.1070*** | 0.1051*** |
| | (0.0284) | (0.0257) | (0.0302) |
| Log(family income) | -0.2288** | -0.2208** | -0.3708*** |
| | (0.0910) | (0.0843) | (0.0667) |
| Age (WII) | -0.0206 | -0.0097 | -0.0005 |
| | (0.0190) | (0.0192) | (0.0274) |
| Tobacco at home (WII) | 0.0308 | 0.0297 | -0.0317 |
| | (0.0233) | (0.0242) | (0.0268) |
| Height (WII) | 0.3398* | 0.3264* | 0.4221 |
| | (0.1586) | (0.1706) | (0.2835) |
| Attendance (WII) | -0.0103 | -0.0102 | 0.0016 |
| | (0.0058) | (0.0058) | (0.0117) |
| Peers' characteristics | No | Yes | Yes |
| School fixed effects | Yes | Yes | Yes |
| Grade indicators | Yes | Yes | Yes |
| | | | |
| First-stage F-test | | | 16.60 |
| Anderson-Rubin test | | | 24.14 |
| | | | |
| N. Obs. | 1,043 | 1,043 | 1,043 |

This table reports reduced form Ols and 2SLS estimates of model 9 for high-school students controlling for grade dummies and setting $\alpha_3 = 0$ and $\alpha_2 = 1$. First-stage F-test of excluded instruments statistics and AR test (Anderson and Rubin, 1949) are reported. The peers' characteristics are calculated as friends' averages of the included variables. cluster-robust numerical standard errors in parentheses. Clusters are defined at school level. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: **Dynamic recursive model**

| Dep. Var. Smoking Behavior Index | (1) | (2) |
|---|---|---|
| *Addiction effect* ($\alpha_1$) | 0.8840*** | 0.8912*** |
| | (0.0327) | (0.0273) |
| *Own effect* ($\alpha_2$) | 0.0001 | 0.0001 |
| | (0.0000) | (0.0000) |
| *Peer effect* ($\alpha_3$) | 0.1159*** | 0.1087*** |
| | (0.0327) | (0.0273) |
| *Discount factor* ($\delta$) | 0.8987*** | 0.8948*** |
| | (0.2175) | (0.2175) |
| Student characteristics | Yes | Yes |
| Peers' characteristics | No | Yes |
| Grade indicators | Yes | Yes |
| School fixed effects | Yes | Yes |
| N. Obs. | 1,043 | 1,043 |

This table reports NLIV estimates of the structural models 8. Students' characteristics are listed in Table 10. The peers' characteristics are calculated as friends' averages of the included variables. cluster-robust numerical standard errors in parentheses. Clusters are defined at school level. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: **Model validation**

| | |
|---|---|
| Overall | 0.6927 |
| If female = 1 | 0.6602 |
| If Parent college degree = 1 | 0.72 |
| If two-parents = 1 | 0.7051 |
| If black and african american =1 | 0.75 |
| if Alcohol/tobacco at home = 1 | 0.7143 |
| If height > median | 0.7207 |
| If income > median | 0.6529 |
| If PVT score > median | 0.6972 |
| N. Obs. | 218 |

This table reports the percentage of choices correctly predicted for the 9th grader sample and for several subgroups defined by gender, race, parental education, and others. Predicted choices are generated by splitting the interval $[0, 2]$ uniformly of mean predicted outcomes for the 9th grader's sample and several subgroups defined by student characteristics.

Table 5: **Parameter Estimates**

| | Static | Myopic | Dynamic |
|---|---|---|---|
| *Addiction effect* ($\alpha_1$) | 0 | 0.8480 | 0.8912 |
| *Own effect* ($\alpha_2$) | 0.3408 | 0.0441 | 0.0001 |
| *Peer effect* ($\alpha_3$) | 0.6592 | 0.1079 | 0.1087 |
| *Discount factor* ($\delta$) | 0 | 0 | 0.8948 |

Table 6: **The social effects**

| | $\bar{m}_{10,10}$ | $\bar{m}_{10,11}$ | $\bar{m}_{10,12}$ | $\bar{m}_{11,11}$ | $\bar{m}_{11,12}$ | $\bar{m}_{12,12}$ |
|---|---|---|---|---|---|---|
| Temporary Dynamic | 1.1217 | 1.1215 | 1.1214 | 1.1218 | 1.1217 | 1.1220 |
| Temporary Myopic | 1.1210 | 1.0655 | 1.0129 | 1.1210 | 1.0655 | 1.1141 |
| Temporary Static | 2.9343 | 0 | 0 | 2.9343 | 0 | 2.9343 |
| Permanent Dynamic | 3.0231 | 5.1480 | 6.2694 | 2.1256 | 3.2472 | 1.1220 |
| Permanent Myopic | 1.1210 | 2.1865 | 3.1994 | 1.1210 | 2.1865 | 1.1210 |
| Permanent Static | 2.9343 | 2.9343 | 2.9343 | 2.9343 | 2.9343 | 2.9343 |

This table reports the sample means of the social effects ($\bar{m}_{t,\tau}$) for the three alternative models' estimated parameter values, as summarised in Table 5 .

Table 7: **The expectation multiplier**

| | $\bar{m}_{12,10}$ | $\bar{m}_{12,11}$ | $\bar{m}_{11,10}$ | $\bar{m}_{per10}$ |
|---|---|---|---|---|
| Dynamic | 0.8979 | 1.0035 | 1.9014 | 1.0037 |
| Myopic | 0 | 0 | 0 | 0 |
| Static | 0 | 0 | 0 | 0 |

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 grades 11 and 12, calibrated to the estimated parameters of the three models in Table 3, Column 2.

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

| Dep. Var. Risky Behavior Index | Endogenous network | | |
|---|---|---|---|
| | *Static* | *Myopic* | *Dynamic baseline* |
| | $\alpha_1 = 0$, $\delta = 0$ | $\delta = 0$ | |
| | (1) | (2) | (3) |
| *Addiction effect* ($\alpha_1$) | | 0.8480*** | 0.8912*** |
| | | (0.0419) | (0.0273) |
| *Own effect* ($\alpha_2$) | 0.3408 | 0.0441 | 0.0001*** |
| | (0.4176) | (0.1038) | (0.0000) |
| *Peer effect* ($\alpha_3$) | 0.6592 | 0.1079 | 0.1087*** |
| | (0.5143) | (0.0970) | (0.0273) |
| *Discount factor* ($\delta$) | | | 0.8948*** |
| | | | (0.2175) |
| Mean Squared Error | 2.2907 | 0.3776 | 0.3801 |
| Percentage of correct predictions | 0.4957 | 0.7526 | 0.7632 |
| Student characteristics | Yes | Yes | Yes |
| Peers' characteristics | Yes | Yes | Yes |
| Grade indicators | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes |
| N. Obs. | 1,043 | 1,043 | 1,043 |

This table reports NLIV estimates of the structural model 8. In Column 1 we restrict the model by setting $\alpha_1 = 0$ and $\delta = 0$, while in Column 2 we restrict the model by setting $\delta = 0$. Column 3 reports baseline estimates presented in Table 3 Column 2. Students' characteristics are listed in Table 10. The peers' characteristics are calculated as friends' averages of the included variables. cluster-robust numerical standard errors in parentheses. Clusters are defined at peer groups level. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: **Robustness**

| | *Badev indicator* (1) | *Daily smoking* (2) | *Cigarette numbers* (3) | *Peer Definition* (4) |
|---|---|---|---|---|
| *Addiction effect* ($\alpha_1$) | 0.8050*** | 0.9061*** | 0.8335*** | 0.8786*** |
| | (0.0690) | (0.0584) | (0.0762) | (0.0352) |
| *Own effect* ($\alpha_2$) | 0.0244 | 0.0031 | 0.0462 | 0.0001 |
| | (0.0756) | (0.0504) | (0.1031) | (0.0009) |
| *Peer effect* ($\alpha_3$) | 0.1706** | 0.0909* | 0.1203 | 0.1212*** |
| | (0.0910) | (0.0629) | (0.1133) | (0.0352) |
| *Discount factor* ($\delta$) | 0.8991 | 0.8946 | 0.8986 | 0.8945*** |
| | (0.7229) | (0.8959) | (0.7236) | (0.2186) |
| Student characteristics | Yes | Yes | Yes | Yes |
| Peers' characteristics | Yes | Yes | Yes | Yes |
| Grade indicators | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes |
| N. Obs. | 1,043 | 1,043 | 1,042 | 1,024 |

This table reports NLIV estimates of the structural models 8. In the first column, we recode students classified as *occasionally smoking* into students who *never tried smoking* using the same dependent variable as in Badev (2021). In the second column, we add one category to our baseline dependent variable: *daily smoking*. This category is referred to students who report to smoke every day. In the third column, we also use the answers to the following question: "During the past 30 days, on the days you smoked, how many cigarettes did you smoke each day?". We construct a cigarette number indicator which takes 6 values: 0 for students who report to smoke zero cigarettes, 1 for ones who report to smoke from 1 to 5, 2 for students who report to smoke from 6 to 10, and so on up to values greater than 21. In the fourth column, we define student's peers as all other students of the same racial background in the same grade at the same school. For multiracial students we consider the first race reported. Students' characteristics are listed in Table 10. The peers' characteristics are calculated as friends' averages of the included variables. cluster-robust numerical standard errors in parentheses. Clusters are defined at school level. *** p<0.01, ** p<0.05, * p<0.1.

# Technical Appendix

## A Formal Model under the General Network Topology

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 $\theta_t \in \boldsymbol{\Theta}$ the corresponding $N$-dimensional vectors stacking all agents).[43] 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 $(\boldsymbol{\Theta}, \mathcal{F}, \mathbb{P})$, where $\boldsymbol{\Theta} := \left\{(\theta_1, \theta_2, \cdots) : \theta_t \in \Theta^N \,,\, t = 1, 2, \cdots, T\right\}$ is the space of sample paths. The sequence $(\mathcal{F}_1, \mathcal{F}_2, \cdots, \mathcal{F}_T)$ of Borel sub-$\sigma$-fields of $\mathcal{F}$ is a filtration in $(\boldsymbol{\Theta}, \mathcal{F})$, 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 $t$, $\theta_t$ 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 $\alpha_1$ or $\alpha_2$ 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 $t$, $y_{it} : \mathbf{Y}^t \times \boldsymbol{\Theta}^t \to Y$, 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 $(\mathbf{y}^0, \theta^1) \in \mathbf{Y} \times \boldsymbol{\Theta}$.

[43] 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* $i = 1, \ldots, N$ *and for all* $(\mathbf{y}^{t-1}, \theta^t) \in \mathbf{Y}^t \times \mathbf{\Theta}^t$

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

# B Existence and Uniqueness of Equilibrium

**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*

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

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

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

*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, $\mathbf{y}^{T-1} = (\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_{T-1})$ and of preference shocks, $\theta^T = (\theta_1, \ldots, \theta_T)$, 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

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

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

$$FOC_{iT}(f)(\mathbf{y}^{T-1}, \theta^T) := \Delta_T^{-1} \left( \alpha_1\, y_{iT-1} + \alpha_2 \theta_{iT} + \alpha_3 \sum_{j=1}^{N} g_{ij}\, f_j(\mathbf{y}^{T-1}, \theta^T) \right) \tag{B.5}$$

$FOC_{iT}$ 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 $\left(\mathbf{y}^{T-1}, \theta^T\right)$

$$\begin{aligned}
&\left|FOC_{iT}\left(f\right)\left(\mathbf{y}^{T-1}, \theta^T\right) - FOC_{iT}\left(\hat{f}\right)\left(\mathbf{y}^{T-1}, \theta^T\right)\right| \\
&\quad = \Delta_T^{-1}\Big|\alpha_1\, y_{iT-1} + \alpha_2\, \theta_{iT} + \alpha_3 \sum_{j=1}^{N} g_{ij}\, f_j(\mathbf{y}^{T-1}, \theta^T) \\
&\quad - \alpha_1\, y_{iT-1} - \alpha_2\, \theta_{iT} - \alpha_3 \sum_{j=1}^{N} g_{ij}\, \hat{f}_j(\mathbf{y}^{T-1}, \theta^T)\Big| \\
&\quad = \left(\frac{\alpha_3}{\Delta_T}\right) \left| \sum_{j=1}^{N} g_{ij} \left( f_j(\mathbf{y}^{T-1}, \theta^T) - \hat{f}_j(\mathbf{y}^{T-1}, \theta^T) \right) \right|
\end{aligned}$$

The coefficient $\left(\frac{\alpha_3}{\Delta_T}\right) < 1$ since either $\alpha_1$ or $\alpha_2$ 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.[44] 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

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

[44] 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 $y_\tau$, for $\tau = t+1, \ldots, T$, as

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

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

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

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 $\theta_s$, 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.,

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

Similarly, for any $j \neq i$,

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

Let $\mathrm{diag}\,(A)$ 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 $\Delta_t$ is defined in two-steps as

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

where $\iota_N$ is an $N \times 1$ column-vector of ones and $\iota_N'$ 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,

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

Finally, let $D_t$ capture the sum of the effects on the current period (period $t$) marginal utility of future $\theta_\tau$'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

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

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

Now define

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

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 ($t = T$), 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

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

and the optimal policy then is given by

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

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 $\Delta_t$ using (B.17).
7. Compute $D_t$ using (B.18).
8. Compute $\bar{\Delta}_t :=$ diag$(\Delta_t)$ and $\bar{\bar{\Delta}}_t := \Delta_t -$ 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 Identification of Social Interactions

**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 of the stochastic process of preference shocks as in (6), the following reduced-form equations hold

$$\begin{aligned}
\mathbf{y}_T &= [\Delta_T\,\mathbf{I} - \alpha_3\,\mathbf{G}]^{-1} \times \left[\alpha_1\,\mathbf{y}_{T-1} + \alpha_2\left(\sum_{k=1}^{K}(\beta_k\mathbf{I} + \phi_k\mathbf{G})\,\mathbf{x}_T^{(k)} + \mathbf{u}_T\right)\right] \\
\mathbf{y}_{T-1} &= \left[\bar{\Delta}_{T-1} - \bar{\bar{\Delta}}_{T-1}\right]^{-1}\left(\alpha_1\,\mathbf{y}_{T-2} + \alpha_2\left(\sum_{k=1}^{K}(\beta_k\mathbf{I} + \phi_k\mathbf{G})\,\mathbf{x}_{T-1}^{(k)} + \mathbf{u}_{T-1}\right) + \mathbf{D}_{T-1}\right)
\end{aligned}$$

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 $E\left[\mathbf{u}_T\,|\,\mathbf{u}^{T-1}\right]$. Unlike the econometrician, 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 $\bar{\theta}_T^u := E\left[\mathbf{u}_T\,|\,\mathbf{u}^{T-1}\right]$,

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

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,

$$\begin{aligned}
\mathbf{y}_T &= \alpha_1\,\mathbf{B}_T\,\mathbf{y}_{T-1} + \alpha_2\,\mathbf{B}_T\left(\sum_{k=1}^{K}(\beta_k\mathbf{I} + \phi_k\mathbf{G})\,\mathbf{x}_T^{(k)}\right) + \varepsilon_T &\text{(D.1)} \\
\varepsilon_T &= \mathbf{B}_T\,\alpha_2\,\mathbf{u_T}. \\
\mathbf{y}_{T-1} &= \alpha_1\,\mathbf{B}_{T-1}\,\mathbf{y}_{T-2} + \alpha_2\,\mathbf{B}_{T-1}\left(\sum_{k=1}^{K}(\beta_k\mathbf{I} + \phi_k\mathbf{G})\,\mathbf{x}_{T-1}^{(k)}\right) + \alpha_2\,\mathbf{B}_{T-1}\,D_{T-1}^x + \varepsilon_{T-1} &\text{(D.2)} \\
\varepsilon_{T-1} &= \mathbf{B}_{T-1}\,\alpha_2\,\mathbf{u_{T-1}} + \mathbf{B}_{T-1}\,\alpha_2\,D_{T-1}^u.
\end{aligned}$$

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. Thanks to the Regularity Assumption, our choices of instruments, $\mathbf{s}_{T-1}$ and $\mathbf{s}_{T-2}$, do affect the choice variables $\mathbf{y}_{T-1}$ and $\mathbf{y}_{T-2}$, respectively. This way, we have $N$ instruments for each period. 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 $t = T$

$$E[\varepsilon_T \mid \mathbf{s}_{T-1}] \;=\; \alpha_2 \mathbf{B}_T \, E[\, E[\, \mathbf{u}_T \mid (\mathbf{X}_T, \mathbf{s}_{T-1})_{t=1}^{T}] \mid \mathbf{s}_{T-1}] = E[0 \mid \mathbf{s}_{T-1}] = 0$$

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

$$E[\varepsilon_{T-1} \mid \mathbf{s}_{T-2}] \;=\; \alpha_2 \mathbf{B}_{T-1} \, E[\, E[\, \mathbf{u}_{T-1} + D^u_{T-1} \mid (\mathbf{X}_T, \mathbf{s}_{T-1})_{t=1}^{T}] \mid \mathbf{s}_{T-2}]$$

   Note that $\varepsilon_t$ is a linear function of two sets of variables: $\mathbf{u}_t$ and $E\left[\mathbf{u}_\tau | \mathbf{u}^t\right]$. By the Exogeneity Assumption and since $E\left[\mathbf{u}_\tau | \mathbf{u}^t\right]$ is a function of $\mathbf{u}^t$, $E\left[E\left[\mathbf{u}_\tau | \mathbf{u}^t\right] | \, (\mathbf{X}_T, \mathbf{s}_{T-1})_{t=1}^{T}\right] = 0$.

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

Moreover, they are not collinear with $(\mathbf{X}_t)_{t=1}^{T}$ thanks to the Full Rank Assumption. Therefore, thanks to the hypothesis that $T \geq 2$, 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 (10).[45] 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 $\alpha_1, \alpha_1' \neq 0$ by the Regularity assumption, and $\mathbf{B}_T(\gamma)$ and $\mathbf{B}_T(\gamma')$ are invertible, we obtain

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

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

[45] 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.

$\Delta_T = \sum_i \alpha_i = 1$, 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 (10) 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 (11) and can recover $\mathbf{B}_{T-1}$ since we already recovered the true $\alpha_1$ using (10) above. We can then recover $\Delta_{T-1}$ by reversing the operations in Step 8 of the recursive algorithm we used to obtain $\mathbf{B}_{T-1}$ from $\Delta_{T-1}$, that we presented in Section F, namely by

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

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

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

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

## E Social Multiplier in Economies with Conformity and Complementarities

As we have briefly noticed in the text, the notion of *social multiplier* is quite subtle in the class of economies we study. It is easier to illustrate why by considering the social interaction effects in the last period $t = T$, so as to silence any dynamic effects.

Consider first the pure conformity economy we have studied in the paper. Consider an exogenous shock to $\theta_{i,T}$, say by $\Delta\theta_{i,T} = 1$, for all agents $i$. From individual $i$'s first-order condition in Equation (B.7) which we copy here for ease of readability

$$\mathbf{y}_T = \alpha_1 \, \mathbf{y}_{T-1} + \alpha_2 \, \theta_T + \alpha_3 \, \mathbf{G}\mathbf{y}_T$$

under the identification assumption that $\sum_{i=1}^{3} \alpha_i = 1$, the total equilibrium change on the smoking behavior values at $T$, $\Delta\mathbf{y_T}$ is given by the expression

$$\Delta\mathbf{y}_T \quad = \quad \alpha_2 \, [\mathbf{I} - \alpha_3\mathbf{G}]^{-1} \, \mathbf{1} \tag{E.1}$$

and incorporates changes due to the change in the exogenous variable, as well as changes arising from the change in the behavior of her peers, and that of their peers and so on. To demonstrate this explicitly, we use a power series expansion of the equilibrium change in (E.1) as

$$\begin{aligned}
\Delta\mathbf{y}_T &= \alpha_2 \, [\mathbf{I} - \alpha_3\mathbf{G}]^{-1} \, \mathbf{1} \\
&= \alpha_2 \sum_{k=0}^{\infty} (\alpha_3\mathbf{G})^k \, \mathbf{1} \\
&= \underbrace{\alpha_2 \, \mathbf{1}}_{\textit{direct effect}} + \underbrace{(\alpha_3\mathbf{G}) \, \mathbf{1} + (\alpha_3\mathbf{G})^2 \, \mathbf{1} + \cdots + (\alpha_3\mathbf{G})^k \, \mathbf{1} + \cdots}_{\textit{indirect social interaction effects}} = \underbrace{\frac{\alpha_2}{\alpha_1 + \alpha_2} \, \mathbf{1}}_{\textit{total effect}}
\end{aligned}$$

which in turn reveals that the total equilibrium change, $\Delta\mathbf{y_T} = \frac{\alpha_2}{\alpha_1+\alpha_2} \, \mathbf{1}$, incorporates a direct effect on behavior, $\Delta\mathbf{y_T} \, |_{direct} = \alpha_2 \, \mathbf{1}$, and an infinite cascade of indirect social interaction effects of the same sign. The result of all these indirect effects is what is called the *social multiplier* and can be defined as

$$\Delta\mathbf{y_T} \, |_{social} \quad = \quad \frac{\Delta\mathbf{y_T}}{\Delta\mathbf{y_T} \, |_{direct}} \; = \; \frac{\frac{\alpha_2}{\alpha_1+\alpha_2}}{\alpha_2} \; = = \; \frac{1}{\alpha_1 + \alpha_2} \; > 1$$

It follows than that $\Delta\mathbf{y_T} = 1$ when preferences do not account for addiction; that is, when $\alpha_1 = 0$. Importantly, when $\alpha_1 = 0$, the total effect is independent of preferences for conformity parameter $\alpha_3$ and the *social multiplier* as defined by Boucher and Fortin (2016), $\Delta y_{i,T} - 1$, is equal to zero. With addiction effects, that is, when $\alpha_1 > 0$, the multiplier is negative, but it is increasing in the preferences for conformity parameter $\alpha_3$. In fact, it is proportional to the ratio of the total effect and the social effect, $\frac{\Delta \mathbf{y_t}}{\Delta \mathbf{y_t}|_{direct}} = \frac{1}{1-\alpha_3}$.

Consider now adding a complementarity factor to preferences, as discussed in Footnote 9 in the text; that is, consider adding a term $\sum_{j=1}^{N} 2\gamma\, y_{it}\, y_{jt}$ to preferences. In this case, the *social multiplier* as defined by Boucher and Fortin (2016) is equal to $\frac{\alpha_2}{\alpha_1+\alpha_2-\gamma} - 1$, and hence it might be positive and it is always increasing in $\alpha_3$. Therefore, while the preference parameter driving complementarities $\gamma$ - and hence the social multiplier as defined by Boucher and Fortin (2016) - cannot be identified with aggregate behavior (and indeed we therefore postulate $\gamma = 0$ in the text), the ratio of the total effect and the social effect, $\frac{\Delta \mathbf{y_T}}{\Delta \mathbf{y_T}|_{direct}} = \Delta \mathbf{y_T}\,|_{social}$ can be interpreted as a measure of social interaction effects in the economy we study.[46]

## E.1 General Model of Social Interactions

In this section, we report the results we obtain by allowing both for *conformity* as well as *positive complementarity*. The contemporaneous utility specification takes the form:

$$u(y_{it-1}, y_{it}, \{y_{jt}\}_{j\in N(i)}, \theta_{it}) \quad := \quad \overbrace{-\alpha_1(y_{it-1} - y_{it})^2 - \alpha_2(\theta_{it} - y_{it})^2}^{Private} \\ \underbrace{-\alpha_3 \sum_{j=1}^{N} g_{ij}(y_{jt} - y_{it})^2}_{Conformity} + \underbrace{2\sum_{j=1}^{N} \gamma_{ij}\, y_{it}\, y_{jt}}_{Complementarity} \tag{E.2}$$

where $\alpha_i \geq 0$, $i = 1, 2, 3$, and $\gamma_{ij} \geq 0$, for $i, j = 1, \ldots, N$.

As we have briefly noticed above, the notion of *social multiplier* is quite subtle in the class of economies we study. It is easier to illustrate why by considering the social interaction effects in the last period $t = T$, so as to silence any dynamic effects. Following the exact same recursive induction arguments we used in the Equilibrium Existence proof in Appendix B, these structural ideas extend to earlier periods of a dynamic economy.

The marginal utility takes the form

$$MU(y_{iT-1}, \{y_{jT}\}_{j=1}^{N}, \theta_{iT}) \quad := \quad 2\Bigg[\alpha_1\, y_{iT-1} + \alpha_2\, \theta_{iT} + \alpha_3 \sum_{j=1}^{N} g_{ij} y_{jT} \\ + \sum_{j=1}^{N} \gamma_{ij} y_{jT} - (\alpha_1 + \alpha_2 + \alpha_3)\, y_{iT}\Bigg]$$

which means this specification has *strategic complementarities* built into it, since the marginal utility of an

[46] This is the sense in which in fact both Becker and Murphy (2001) and Glaeser and Scheinkman (2003) use the term *social multiplier*.

agent for an action increases as other agents increase their actions, i.e., $\frac{MU}{\partial y_{jT}} > 0$ for any $j$ where $g_{ij} \neq 0$ or $\gamma_{ij} \geq 0$.[47]. Next, we derive the optimal choice of agent $i$ using her FOC and write her *best response function* as

$$BR^i\left(y_{iT-1}, \theta_{iT}, \mathbf{y}_{-iT},\right) = \left(\frac{\alpha_1}{\Delta_1}\right) y_{i,T-1} + \left(\frac{\alpha_2}{\Delta_1}\right) \theta_{i,T} + \left(\frac{\alpha_3}{\Delta_1}\right) \sum_{j=1}^{N} g_{ij} y_{j,T} + \sum_{j=1}^{N} \left(\frac{\gamma_{ij}}{\Delta_1}\right) y_{j,T}$$

where $\Delta_1 := \alpha_1 + \alpha_2 + \alpha_3$ and $\mathbf{y}_{-iT} := (y_{jT})_{j \neq i}$. We stack all these best responses for $i = 1, \ldots, N$ into an $(N \times 1)$ vector $BR$ as

$$\begin{aligned} BR(\mathbf{y}_{T-1}, \theta_T, \mathbf{y}_T) &= \left(BR^1\left(y_{1T-1}, \theta_{1T}, \mathbf{y}_{-1T}\right), \ldots, BR^N\left(y_{NT-1}, \theta_{NT}, \mathbf{y}_{-NT}\right)\right) \\ &= \left(\frac{\alpha_1}{\Delta_1}\right) \mathbf{y}_{T-1} + \left(\frac{\alpha_2}{\Delta_1}\right) \theta_T + \left(\frac{\alpha_3}{\Delta_1}\right) \mathbf{G}\mathbf{y}_T + \left(\frac{1}{\Delta_1}\right) \Gamma\, \mathbf{y}_T \end{aligned}$$

Define now the function $F$ as

$$F(\mathbf{y}_{T-1}, \theta_T, \mathbf{y}_T) := \mathbf{y}_T - BR(\mathbf{y}_{T-1}, \theta_T, \mathbf{y}_T) \tag{E.3}$$

An equilibrium of the last period exists when $F(\mathbf{y}_{T-1}, \theta_T, \mathbf{y}_T) = 0$ for all $(\mathbf{y}_{T-1}, \theta_T)$. More precisely, the equilibrium is where

$$\begin{aligned} \mathbf{y}_T - \left(\frac{\alpha_1}{\Delta_1}\right) \mathbf{y}_{T-1} - \left(\frac{\alpha_2}{\Delta_1}\right) \theta_T - \left(\frac{\alpha_3}{\Delta_1}\right) \mathbf{G}\mathbf{y}_T - \left(\frac{1}{\Delta_1}\right) \Gamma\, \mathbf{y}_T &= 0 \\ \Longrightarrow \left[\mathbf{I} - \left(\frac{\alpha_3}{\Delta_1}\right) \mathbf{G} - \left(\frac{1}{\Delta_1}\right) \Gamma\right] \mathbf{y}_T &= \left(\frac{\alpha_1}{\Delta_1}\right) \mathbf{y}_{T-1} + \left(\frac{\alpha_2}{\Delta_1}\right) \theta_T \end{aligned} \tag{E.4}$$

A big chunk of what we will demonstrate depends on the structure of the matrix

$$\left[\mathbf{I} - \left(\frac{\alpha_3}{\Delta_1}\right) \mathbf{G} - \left(\frac{1}{\Delta_1}\right) \Gamma\right]. \tag{E.5}$$

First, we present equilibrium behavior when preferences are of the *pure conformity* form. Namely, we set $\Gamma = \mathbf{0}$. After presenting the results for pure conformity, in the rest, we present results when we allow both for *conformity* as well as *positive complementarity*.

Under the assumption $\alpha_1 + \alpha_2 > 0$, our economy has a unique equilibrium. Hence, there exists a unique solution to equation (E.3). The matrix of partial derivatives w.r.t $\mathbf{y}_T$ obtained from that equation

[47]See Bulow, Geanakoplos, and Klemperer (1985) who have coined the term; see also Cooper and John (1988)

has the following structure

$$F_3(\mathbf{y}_{T-1}, \theta_T, \mathbf{y}_T) \;=\; \begin{bmatrix} 1 & -\left(\frac{\alpha_3}{\Delta_1}\right) g_{12} & \dots & -\left(\frac{\alpha_3}{\Delta_1}\right) g_{1N} \\ -\left(\frac{\alpha_3}{\Delta_1}\right) g_{21} & 1 & \dots & -\left(\frac{\alpha_3}{\Delta_1}\right) g_{2N} \\ \vdots & \vdots & \ddots & 0 \\ -\left(\frac{\alpha_3}{\Delta_1}\right) g_{N1} & -\left(\frac{\alpha_3}{\Delta_1}\right) g_{N2} & \dots & 1 \end{bmatrix}$$

Consider w.l.o.g. an exogenous change say by $\Delta\theta_{i,T} = 1$, for all $i = 1, \dots, N$. In equilibrium, the change in the equilibrium choices w.r.t. the exogenous change is given by

$$\frac{\partial \mathbf{y}_T}{\partial \theta_T} \;=\; (F_3)^{-1} (\mathbf{y}_{T-1}, \theta_T, \mathbf{y}_T) \left(\frac{\partial BR^1}{\partial \theta_T}, \dots, \frac{\partial BR^N}{\partial \theta_T}\right)'$$

Since $F_3$ has a dominant diagonal that is equal to one, we may use the Neumann expansion to write:

$$(F_3)^{-1} \;=\; \mathbf{I} + (\mathbf{I} - F_3) + (\mathbf{I} - F_3)^2 + \dots$$

Note that all diagonal elements of $(\mathbf{I} - F_3)$ are zero and the off-diagonal elements are $-F_{3,ij} = \left(\frac{\alpha_3}{\Delta_1}\right) g_{ij} > 0$. Each of the terms in this infinite series is a matrix with non-negative entries, and can be represented as

$$\frac{\partial \mathbf{y}_T}{\partial \theta_T} \;=\; (\mathbf{I} + \mathbf{H}) \left(\frac{\partial BR^1}{\partial \theta_T}, \dots, \frac{\partial BR^N}{\partial \theta_T}\right)'$$

where $\mathbf{H}$ is a matrix with non-negative elements. The non-negativity of the matrix $\mathbf{H}$, means that there is a *social multiplier*. In economies where individual preferences incorporate strategic complementarity, including the current model, a change in the value of an exogenous variable yields a direct effect on behavior and an infinite cascade of indirect effects of the same sign. That is because each agent's action changes not only because of the change in the exogenous variable, but also because of the change in the behavior of her peers, and that of their peers and so on. The result of all these indirect effects is what is called the *social multiplier*.

To quantify this multiplier effect, we measure the ratio of the total (equilibrium) effect $(\mathbf{I} + \mathbf{H}) \frac{\partial BR}{\partial \theta_T} = \left[\mathbf{I} - \left(\frac{\alpha_3}{\Delta_1}\right) \mathbf{G}\right]^{-1} \left(\frac{\alpha_2}{\Delta_1}\right)$ to the direct effect $\frac{\partial BR}{\partial \theta_T} = \left(\frac{\alpha_2}{\Delta_1}\right) \mathbf{I}$ (as in Becker and Murphy (2001) and Glaeser and Scheinkman (2003)).

**Definition 1** *The **social multiplier** of an exogenous preference shock $\Delta\theta_T := \theta'_T - \theta_T$, at $t = T$, on behavior at $t = T$ is given by*

$$\mathbf{m} \;=\; \frac{\Delta \mathbf{y_T}}{\Delta \mathbf{y_T} \,|_{direct}}$$

**Proposition 3** *There exists a social multiplier given by*

$$\mathbf{m} \;=\; \left[\mathbf{I} - \left(\frac{\alpha_3}{\Delta_1}\right)\mathbf{G}\right]^{-1}\mathbf{u} \;=\; \left(\frac{1}{\alpha_1+\alpha_2}\right)\mathbf{u}$$

*and it has the following properties:*

*(i) It is larger than 1 if and only if* $\alpha_3 > 0$

*(ii) It is a monotone increasing function of the strength of interactions* $\alpha_3$*.*

*(iii)* $\lim_{\alpha_3\to 0} m_i = 1$ *and* $\lim_{\alpha_3\to 1} m_i = \infty$*, for any* $i = 1, \ldots, N$*.*

Now is the time to present results allowing both for *conformity* as well as *positive complementarity*. Hence, $\Gamma \neq \mathbf{0}$ anymore. To match the same peer group definition we employ in the paper, we assume that the network structure satisfies, similar to Boucher et al. (2022), the following:

**Assumption 1** *The coefficients of complementarity* $\gamma_{ij}$ *are uniform across agents. Namely,*

$$\gamma_{ij} = \begin{cases} \frac{\gamma}{N-1} > 0 & \text{if } i \neq j \\ 0 & \text{otherwise} \end{cases}$$

As we defined before (in equation (E.4) above), an equilibrium of the last period exists when

$$\underbrace{\left[\mathbf{I} - \left(\frac{\alpha_3}{\Delta_1}\right)\mathbf{G} - \left(\frac{1}{\Delta_1}\right)\Gamma\right]}_{\mathbf{A}}\mathbf{y}_T \;=\; \left(\frac{\alpha_1}{\Delta_1}\right)\mathbf{y}_{T-1} + \left(\frac{\alpha_2}{\Delta_1}\right)\theta_T \tag{E.6}$$

**Proposition 4** *Assume that Assumption 1 holds and that* $\frac{\alpha_3}{\Delta_1} + \frac{\gamma}{\Delta_1} < 1$*. Then, the following are true:*

*(i) The matrix* $\mathbf{A}$ *in (E.6) is nonsingular. Hence, the unique equilibrium policy is given by*

$$\mathbf{y}_T \;=\; \left[\mathbf{I} - \left(\frac{\alpha_3}{\Delta_1}\right)\mathbf{G} - \left(\frac{1}{\Delta_1}\right)\Gamma\right]^{-1}\left[\left(\frac{\alpha_1}{\Delta_1}\right)\mathbf{y}_{T-1} + \left(\frac{\alpha_2}{\Delta_1}\right)\theta_T\right] \tag{E.7}$$

*(ii) The mean outcome is increasing in the strength of complementarity* $\gamma$*. Namely, the average choice can be solved as*

$$\overline{y}_T(\gamma) \;=\; \left(\frac{\alpha_1}{\alpha_1+\alpha_2-\gamma}\right)\overline{y}_{T-1} + \left(\frac{\alpha_2}{\alpha_1+\alpha_2-\gamma}\right)\overline{\theta}_T \tag{E.8}$$

*which is independent of* $\alpha_3$ *as long as the relative ratio* $\left(\frac{\alpha_1}{\alpha_2}\right)$ *stays intact.*

*(iii) The effect on the mean action of a change to a common exogenous variable, say* $\Delta\theta_T = \mathbf{u}$ *or*

$\Delta y_{T-1} = \mathbf{u}$ *is increasing in* $\gamma$*. More formally,*

$$\Delta \bar{y}_T = \begin{cases} \frac{\alpha_1}{\alpha_1+\alpha_2-\gamma}, & \text{if } \Delta y_{T-1} = \mathbf{u} \\ \frac{\alpha_2}{\alpha_1+\alpha_2-\gamma} & \text{if } \Delta\theta_T = \mathbf{u} \end{cases}$$

**Proposition 5** *Assume that Assumption 1 holds and that* $\frac{\alpha_3}{\Delta_1} + \frac{\gamma}{\Delta_1} < 1$*. There exists a social multiplier of an exogenous preference shock at* $t = T$ *on behavior at* $t = T$ *given by*

$$\mathbf{m} = \frac{\Delta \mathbf{y_t}}{\Delta \mathbf{y_t}\,|_{direct}} = \left[\mathbf{I} - \left(\frac{\alpha_3}{\Delta_1}\right)\mathbf{G} - \left(\frac{1}{\Delta_1}\right)\Gamma\right]^{-1}\mathbf{u} = \left(\frac{1}{\alpha_1+\alpha_2-\gamma}\right)\mathbf{u}.$$

*and it has the following properties:*

*(i) It is larger than 1 if and only if* $\alpha_3 > 0$ *or* $\gamma > 0$*.*

*(ii) It is a monotone increasing function of the strength of interactions* $\alpha_3$ *as well as the strength of complementarity* $\gamma$*.*

*(iii) Social multiplier with complementarity (*$\gamma > 0$*) is larger than the social multiplier without.*

*(iv) Moreover,* $\lim_{\alpha_3\to 0} m_i = \frac{1}{1-\gamma}$ *and* $\lim_{(\alpha_3+\gamma)\to 1} m_i = \infty$*, for any* $i = 1, \ldots, N$*.*

## E.2 Proofs of All Results

*Proof:* [Proposition 3] We know from the existence proof in Appendix B of the paper that the unique fixed point of the operator induced by the first-order condition written in matrix form $\Delta_1 \mathbf{y}_T = \alpha_1\,\mathbf{y}_{T-1} + \alpha_2\,\theta_T + \alpha_3\,\mathbf{G}\mathbf{y}_T$ exists. Moreover, this existence is equivalent to the invertibility of this matrix equation. Another way to see this is that since $\frac{\alpha_3}{\Delta_T} < 1$, $\Delta_1 \mathbf{I} - \alpha_3\,\mathbf{G}$ is invertible (see Case (1991), footnote 5). Hence, the social multiplier is well defined. Using a series expansion of the inverse $\left[\mathbf{I} - \left(\frac{\alpha_3}{\Delta_1}\right)\mathbf{G}\right]^{-1} = \sum_{k=0}^{\infty}\left(\frac{\alpha_3}{\Delta_1}\right)^k \mathbf{G}^k$, and using the fact that $\mathbf{G}^k\,\mathbf{u} = \mathbf{u}$ for any $k \geq 1$, we can show that

$$\begin{aligned}
\left[\mathbf{I} - \left(\frac{\alpha_3}{\Delta_1}\right)\mathbf{G}\right]^{-1}\mathbf{u} &= \sum_{k=0}^{\infty}\left(\frac{\alpha_3}{\Delta_1}\right)^k \mathbf{G}^k\,\mathbf{u} \\
&= \sum_{k=0}^{\infty}\left(\frac{\alpha_3}{\Delta_1}\right)^k \mathbf{u} \\
&= \left(\frac{1}{1-\frac{\alpha_3}{\Delta_1}}\right)\mathbf{u} = \left(\frac{1}{\alpha_1+\alpha_2}\right)\mathbf{u}.
\end{aligned}$$

where the last step uses the identification condition $\Delta_1 = \alpha_1 + \alpha_2 + \alpha_3 = 1$. Result (i) follows from the fact that $\Delta_1 := \alpha_1 + \alpha_2 + \alpha_3 = 1$ hence $\alpha_1 + \alpha_2 < 1$ if and only if $\alpha_3 > 0$. Result (ii) follows from the fact that $\Delta_1 := \alpha_1 + \alpha_2 + \alpha_3 = 1$ hence $\alpha_1 + \alpha_2$ goes down when $\alpha_3$ goes up. Result (iii) similarly follows from the fact that as $\alpha_3 \to 0$, $\alpha_1 + \alpha_2 \to 1$ and as $\alpha_3 \to 1$, $\alpha_1 + \alpha_2 \to 0$.

■

*Proof:* [Proposition 4] The nonsingularity of the matrix $\mathbf{A}$ is obtained using the exact same arguments as in the Proof of Proposition 3. Hence the unique equilibrium policy is given by (E.7) as stated. Observe that we can write

$$\left(\frac{\alpha_3}{\Delta_1}\right)\mathbf{G}+\left(\frac{1}{\Delta_1}\right)\Gamma \;=\; \left(\frac{\alpha_3+\gamma}{\Delta_1}\right)\mathbf{G}.$$

With this observation, the average choice is given by $\frac{1}{N}\mathbf{u}'\,\mathbf{y}_T$ where $\mathbf{u}'$ is a $1\times N$ row vector. Therefore, using a power series expansion of the inverse $\left[\mathbf{I}-\left(\frac{\alpha_3+\gamma}{\Delta_1}\right)\mathbf{G}\right]^{-1} = \sum_{k=0}^{\infty}\left(\frac{\alpha_3+\gamma}{\Delta_1}\right)^k\mathbf{G}^k$, and using the fact that $\mathbf{u}'\,\mathbf{G}^k = \mathbf{u}'$ for any $k\geq 1$ since $\mathbf{G}$ is symmetric, and that $\frac{\alpha_3}{\Delta_1}+\frac{\gamma}{\Delta_1}<1$, we can show that

$$\begin{aligned}
\overline{y}_T(\gamma) &= \frac{1}{N}\mathbf{u}'\,\mathbf{y}_T\\
&= \frac{1}{N}\mathbf{u}'\left(\sum_{k=0}^{\infty}\left(\frac{\alpha_3+\gamma}{\Delta_1}\right)^k\mathbf{G}^k\right)\left[\left(\frac{\alpha_1}{\Delta_1}\right)\mathbf{y}_{T-1}+\left(\frac{\alpha_2}{\Delta_1}\right)\theta_T\right]\\
&= \frac{1}{N}\sum_{k=0}^{\infty}\left(\frac{\alpha_3+\gamma}{\Delta_1}\right)^k\mathbf{u}'\mathbf{G}^k\left[\left(\frac{\alpha_1}{\Delta_1}\right)\mathbf{y}_{T-1}+\left(\frac{\alpha_2}{\Delta_1}\right)\theta_T\right]\\
&= \sum_{k=0}^{\infty}\left(\frac{\alpha_3+\gamma}{\Delta_1}\right)^k\mathbf{u}'\frac{1}{N}\left[\left(\frac{\alpha_1}{\Delta_1}\right)\mathbf{y}_{T-1}+\left(\frac{\alpha_2}{\Delta_1}\right)\theta_T\right]\\
&= \sum_{k=0}^{\infty}\left(\frac{\alpha_3+\gamma}{\Delta_1}\right)^k\left[\left(\frac{\alpha_1}{\Delta_1}\right)\overline{y}_{T-1}+\left(\frac{\alpha_2}{\Delta_1}\right)\overline{\theta}_T\right]\\
&= \left(\frac{1}{1-\frac{\alpha_3+\gamma}{\Delta_1}}\right)\left[\left(\frac{\alpha_1}{\Delta_1}\right)\overline{y}_{T-1}+\left(\frac{\alpha_2}{\Delta_1}\right)\overline{\theta}_T\right]\\
\overline{y}_T(\gamma) &= \left(\frac{1}{\alpha_1+\alpha_2-\gamma}\right)\left[\left(\frac{\alpha_1}{\Delta_1}\right)\overline{y}_{T-1}+\left(\frac{\alpha_2}{\Delta_1}\right)\overline{\theta}_T\right]
\end{aligned}$$

hence result (ii) obtains. Moreover, since the fraction is increasing in $\gamma$, as stated in the Proposition, the mean outcome is increasing in the strength of complementarity $\gamma$, and is independent of $\alpha_3$ as long as the relative ratio $\left(\frac{\alpha_1}{\alpha_2}\right)$ stays intact. Finally, result (iii) obtains because of the monotonicity argument we just made.

■

*Proof:* [Proposition 5] Using a series expansion of the inverse $\left[\mathbf{I}-\left(\frac{\alpha_3+\gamma}{\Delta_1}\right)\mathbf{G}\right]^{-1} = \sum_{k=0}^{\infty}\left(\frac{\alpha_3+\gamma}{\Delta_1}\right)^k\mathbf{G}^k$, the social multiplier of an exogenous preference shock at $t = T$ on behavior at $t = T$ is given by

$$\begin{aligned}\mathbf{m} = \left[\mathbf{I} - \left(\frac{\alpha_3}{\Delta_1}\right)\mathbf{G} - \left(\frac{1}{\Delta_1}\right)\Gamma\right]^{-1}\mathbf{u} &= \sum_{k=0}^{\infty}\left(\frac{\alpha_3+\gamma}{\Delta_1}\right)^k \mathbf{G}^k\,\mathbf{u} \\ &= \sum_{k=0}^{\infty}\left(\frac{\alpha_3+\gamma}{\Delta_1}\right)^k \mathbf{u} \\ &= \left(\frac{1}{1-\frac{\alpha_3+\gamma}{\Delta_1}}\right)\mathbf{u} \\ &= \left(\frac{1}{\alpha_1+\alpha_2-\gamma}\right)\mathbf{u}.\end{aligned}$$

Result (i) is obtained as follows: $\left(\frac{1}{\alpha_1+\alpha_2-\gamma}\right) > 1$ if and only if $(\alpha_1+\alpha_2-\gamma) < 1$ if and only if $\alpha_3 = 1-\alpha_1-\alpha_2 > 0$ or $\gamma > 0$ (or both). Result (ii) follows from the fact that $\Delta_1 = \alpha_1+\alpha_2+\alpha_3 = 1$ hence $\alpha_1+\alpha_2$ goes down when $\alpha_3$ goes up and the fraction goes up when $\gamma$ goes up. Result (iii) is obtained by observing that the fraction is monotone increasing in the value of $\gamma$. Result (iv) follows from the fact that as $\alpha_3 \to 0$, $\alpha_1+\alpha_2 \to 1$ and as $\alpha_3+\gamma \to 1$, $\alpha_1+\alpha_2-\gamma = 1-\alpha_3-\gamma \to 0$.

■

# F Recursive Algorithm to Compute the Dynamic Multiplier

We defined $\Lambda_{t,\tau}$, for any $\tau = t+1,\ldots,T$, in Appendix B, 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. Similarly, from Appendix B, the definition of $\Gamma_{t,\tau}$, for any $\tau = t+1,\ldots,T$, is

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

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 $\theta_s$, conditional on period-$t$ information. Finally, equation (B.17) yields

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

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

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

So, to sum all this up, here is the **recursive algorithm** to compute these variables for the remaining periods $t = 1, \ldots, T-1$:

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

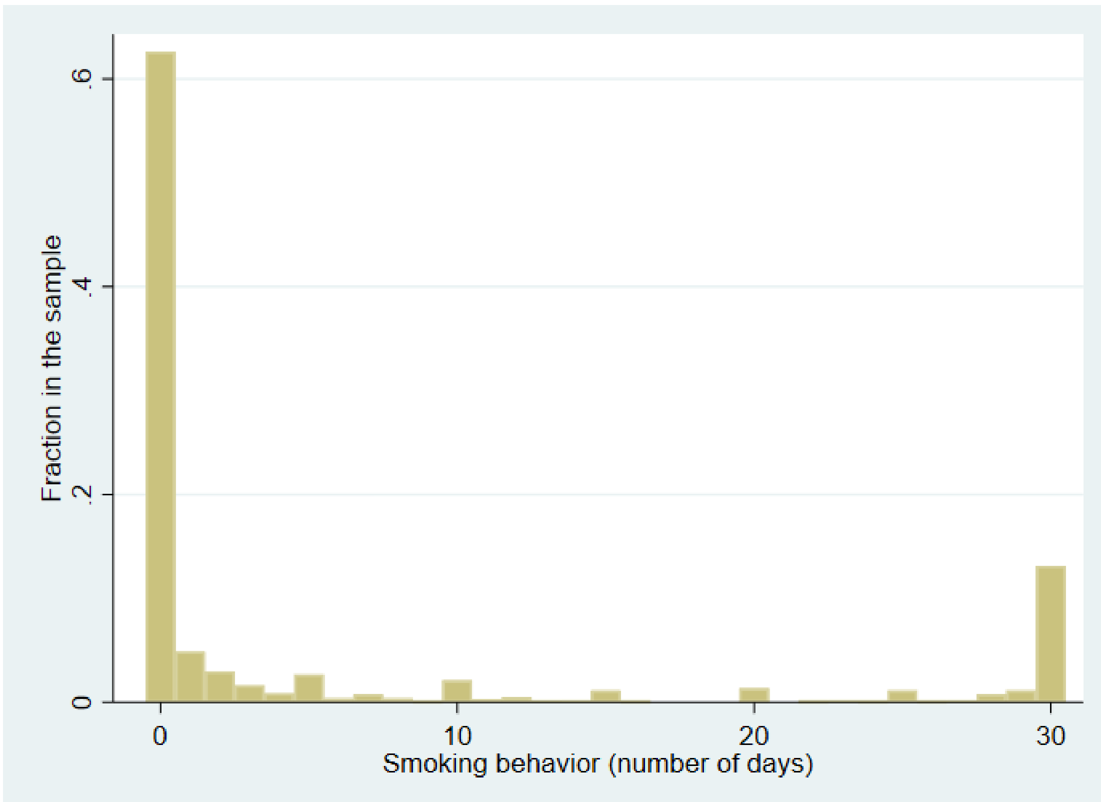
# G Risky Behaviors Indices

Risky behaviors indices are constructed using answers to the questions in the Tobacco, Alcohol, Drugs, and Delinquency scale Sections in the Wave I *in-home questionnaire*. In particular, following the existing literature, we consider variables indicating if students have been attempting alcohol, and indices of how often students participate in a fight, paint graffiti, or steal something.[48] We construct an alcohol consumption indicator using answers to the question: "Have you had a drink of beer, wine, or liquor—not just a sip or a taste of someone else's drink—more than 2 or 3 times in your life?" The variable *Drink ever* is coded as one if students report drinking beer, wine, or liquor and 0 otherwise. We construct indices of how often students participate in a fight, paint graffiti, or steal something using respondent answers to the following questions: "In the past 12 months, how often did you take part in a fight where a group of your friends was against another group? How often did you paint graffiti or signs on someone else's property or in a public place? How often did you steal something worth more than $50?. Specifically, we code the indices *Fight*, *Graffiti*, and *Steal* as 0,1,2 if students report never; 1 or 2 times; 2 or 4 times, respectively.

[48] Baker, Sigman, and Nugent (2001) consider truancy an early warning sign for potential delinquent activity, social isolation, and educational failure. Roebuck, French and Dennis (2004) find that school absenteeism or truancy is a risk factor for substance abuse.

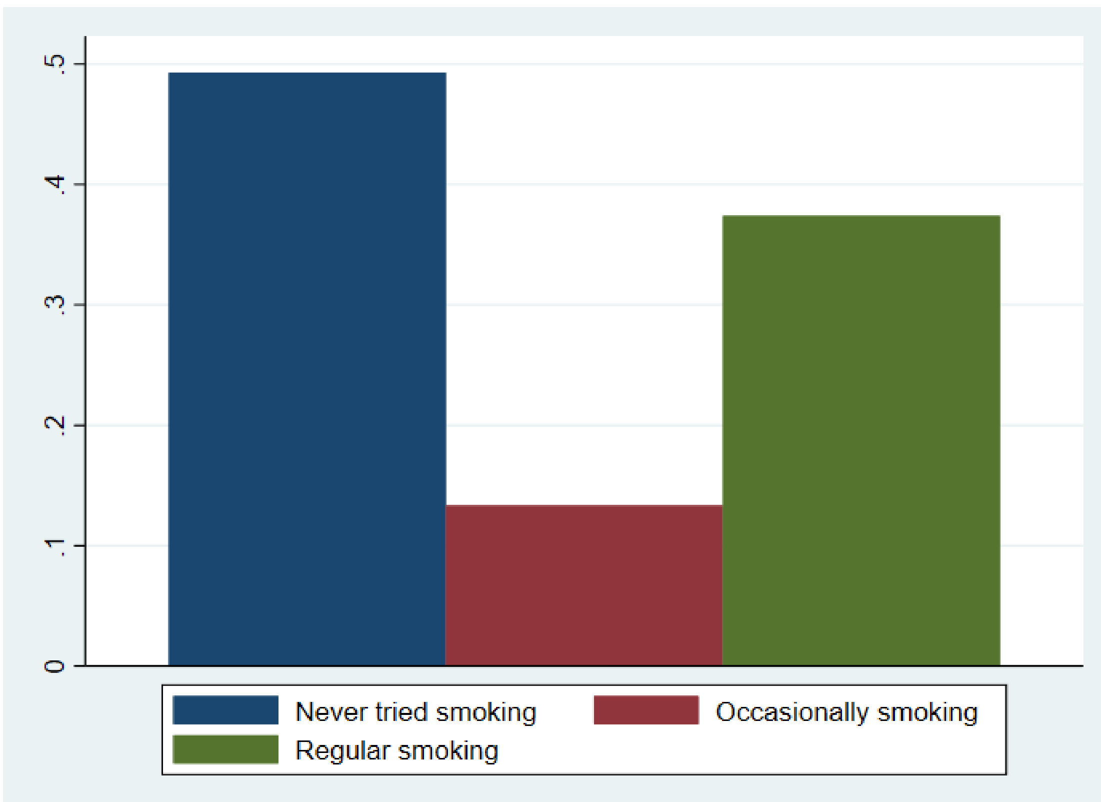
## H Additional Figures and Tables

Figure 2: **Distribution of how many days students smoked**




This figure graphs the empirical density of the answers to the question “During the past 30 days, on how many days did you smoke cigarettes?”. Density is estimated using a histogram density estimator.

Figure 3: **Proportion of students by smoking status**




This figure graphs the proportion of students in our sample who have never tried smoking, who have tried but have reported not to have smoked cigarettes in the last month, and finally who have smoked one or more days in the last month.

Table 10: **Sample selection**

| Years 1994 (WI)-1995 (WII) | Panel (a) Saturated sample (N. obs.: 1312) | | Panel (b) Sample without missing values (N. obs.: 1,043) | |
|---|---|---|---|---|
| Female | 0.5038 | 0.5002 | 0.5158 | 0.5000 |
| Black or African American | 0.1395 | 0.3466 | 0.1371 | 0.3441 |
| White | 0.5930 | 0.4915 | 0.6337 | 0.4820 |
| Asian | 0.1631 | 0.3696 | 0.1419 | 0.3491 |
| Hispanic | 0.1913 | 0.3935 | 0.1592 | 0.3660 |
| Indian | 0.0366 | 0.1878 | 0.0345 | 0.1826 |
| PVT test score | 0.0107 | 0.9196 | 0.0762 | 0.9213 |
| Parents College degree | 0.2614 | 0.4396 | 0.2694 | 0.4439 |
| Two-parent family | 0.7431 | 0.4371 | 0.7622 | 0.4259 |
| Log(family income) | 0.3490 | 0.1738 | 0.3577 | 0.1708 |
| Age (WII) | 16.9558 | 0.9517 | 16.8782 | 0.9576 |
| Tobacco at home (WII) | 0.2980 | 0.4576 | 0.2848 | 0.4515 |
| Height (WII) | 1.7065 | 0.0999 | 1.7067 | 0.1003 |
| Attendance (WII) | 175.6319 | 7.9931 | 176.9808 | 4.6614 |

This table reports means and standard deviations of students' characteristics for the saturated sample (Panel (a)), and for the sample without missing values in observations (Panel (b)).

Table 11: **Sample representativeness**

| | Add Health (N. obs.: 16361) | | CPS (N. obs.: 14257) | |
|---|---|---|---|---|
| Variable | Mean | SD | Mean | SD |
| Female | 0.5158 | 0.5000 | 0.5046 | 0.5000 |
| Black or African American | 0.1371 | 0.3441 | 0.1270 | 0.3330 |
| White | 0.6337 | 0.4820 | 0.6582 | 0.4743 |
| Hispanic or Latino | 0.1592 | 0.3660 | 0.1645 | 0.3707 |
| Parents College degree | 0.2694 | 0.4439 | 0.2292 | 0.4204 |

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 12: **Balancing tests**

| | (1) Two-parent family | (2) Parents College degree | (3) Tobacco at home | (4) Black or African American | (5) Income | (6) PVT test score |
|---|---|---|---|---|---|---|
| Peers characteristics: | | | | | | |
| Parents College degree | -0.1393 | | -0.0374 | 0.0393 | -0.0815 | 0.2700 |
| | (0.1490) | | (0.1657) | (0.1369) | (0.0915) | (0.3332) |
| Black or African American | -0.2924 | 0.8223*** | 0.6204** | | 0.1218 | 0.9507* |
| | (0.3050) | (0.2599) | (0.2423) | | (0.1072) | (0.5275) |
| PVT test score | 0.0210 | 0.1100 | 0.0062 | -0.0032 | -0.0107 | |
| | (0.0775) | (0.0943) | (0.0829) | (0.0475) | (0.0422) | |
| Income | 0.2337 | -0.1226 | -0.7999** | -0.0273 | | -0.4136 |
| | (0.2995) | (0.3689) | (0.3556) | (0.1713) | | (0.6977) |
| Tobacco at home | 0.0460 | 0.1817 | | 0.1292** | -0.0191 | 0.2470 |
| | (0.1576) | (0.1434) | | (0.0605) | (0.0612) | (0.2950) |
| Two-parent family | | -0.2255 | -0.0702 | -0.0593 | -0.0537 | -0.2663 |
| | | (0.1534) | (0.1710) | (0.1019) | (0.0583) | (0.4277) |
| Individual characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| Gender indcator | Yes | Yes | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |

The figures in each row and columns are coefficients from separate regressions of students' background characteristics (parent college degree, two parent family, black or African American, Income, tobacco at home and PVT scores) on peers' average characteristics controlling for the corresponding individual characteristic for a total of 6outcomes $\times$ 5 peers' characteristics regressions. Each regression includes a gender indicator, grade indicators, and school fixed effects. Heteroskedasticity-robust numerical standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 13: **Truancy vs. excused absence: correlated risky behaviors**

| | Risky behavior controls | | | |
|---|---|---|---|---|
| Dep. Var | *Drink ever* | *Fight* | *Graffiti* | *Steal* |
| Skipping school with an excuse | 0.0280 | 0.0190 | 0.0321* | 0.0295 |
| | (0.0208) | (0.0151) | (0.0179) | (0.0204) |
| Obs. | 1,040 | 1,040 | 1,040 | 1,040 |
| Truancy | 0.1860*** | 0.1185*** | 0.1121*** | 0.1853*** |
| | (0.0223) | (0.0166) | (0.0258) | (0.0380) |
| Obs. | 1,304 | 1,305 | 1,305 | 1,306 |
| Individual characteristics | Yes | Yes | Yes | Yes |
| Grade indicators | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes |

This table reports OLS estimates of attendance correlated risky behaviors. The dependent variables are skipping school with an excuse and without (truancy). Definitions of the risky behaviors variables are in Section 6. We drop from the sample in the top panel students with more than 2 days of unexcused absences from school. Students' characteristics are listed in Table 10. heteroschedasticity-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 14: **Dynamic recursive model- Controls**

| Dep. Var. Risky Behavior Index | (1) | (2) |
|---|---|---|
| *Addiction effect* ($\alpha_1$) | 0.8649*** | 0.8853*** |
| | (0.0346) | (0.0305) |
| *Own effect* ($\alpha_2$) | 0.0001*** | 0.0001*** |
| | (0.0000) | (0.0000) |
| *Peer effect* ($\alpha_3$) | 0.1350*** | 0.1147*** |
| | (0.0346) | (0.0305) |
| *Discount factor* ($\delta$) | 0.8980*** | 0.8950*** |
| | (0.2153) | (0.2208) |
| Female | 0.6774*** | -0.7153*** |
| | (0.0000) | (0.0000) |
| Black or African American | -0.0524*** | 0.1027*** |
| | (0.0000) | (0.0000) |
| Asian | -1.6245*** | -0.9000*** |
| | (0.0000) | (0.0000) |
| Hispanic | -1.3635*** | -0.5038*** |
| | (0.0000) | (0.0000) |
| Indian | 0.4200*** | 0.2810*** |
| | (0.0000) | (0.0000) |
| Age (WII) | 1.0812*** | 0.9274*** |
| | (0.0004) | (0.0002) |
| PVT test score | 1.2358*** | -0.7623*** |
| | (0.0001) | (0.0001) |
| Height (WII) | 0.7874*** | 0.3613*** |
| | (0.0000) | (0.0000) |
| Parents College degree | 0.4431*** | 0.5663*** |
| | (0.0000) | (0.0000) |
| Log(family income) | -0.3793*** | -0.4402*** |
| | (0.0000) | (0.0000) |
| Two-parent family | -1.2863*** | -0.6308*** |
| | (0.0000) | (0.0000) |
| Attendance (WII) | 5.7596*** | 4.3002*** |
| | (0.0000) | (0.0001) |
| Tobacco at home (WII) | 3.7696*** | 1.9790*** |
| | (0.0000) | (0.0000) |
| Student characteristics | Yes | Yes |
| Peers' characteristics | No | Yes |
| Grade indicators | Yes | Yes |
| School fixed effects | Yes | Yes |
| N. Obs. | 1,043 | 1,043 |

This table reports NLIV estimates of the structural models 8. Students' characteristics are listed in Table 10. The peers' characteristics are calculated as friends' averages of the included variables. cluster-robust numerical standard errors in parentheses. Clusters are defined at school level. *** p<0.01, ** p<0.05, * p<0.1.

Table 15: **Constrained "myopic" and "static" models- Controls**

| Dep. Var. Risky Behavior Index | *Static* $\alpha_1 = 0$, $\delta = 0$ (1) | *Myopic* $\delta = 0$ (2) | *Dynamic* (3) |
|---|---|---|---|
| *Addiction effect* ($\alpha_1$) | | 0.8566*** | 0.8853 *** |
| | | (0.0421) | (0.0305) |
| *Own effect* ($\alpha_2$) | 0.3520 | 0.0410 | 0.0001 *** |
| | (0.4128) | (0.1011) | (0.0000) |
| *Peer effect* ($\alpha_3$) | 0.6480 | 0.1024 | 0.1147 *** |
| | (0.5115) | (0.0929) | (0.0305) |
| *Discount factor* ($\delta$) | | | 0.8950 *** |
| | | | (0.2208) |
| Female | -0.2965 | -0.4836 | -0.7153*** |
| | (0.4909) | (1.9694) | (0.0000) |
| Black or African American | -0.4972 | -0.7890 | 0.1027*** |
| | (0.6809) | (2.5548) | (0.0000) |
| Asian | 0.0113 | -0.7897 | -0.9000*** |
| | (0.4209) | (2.4980) | (0.0000) |
| Hispanic | -0.1007 | 0.4059 | -0.5038*** |
| | (0.1765) | (2.5280) | (0.0000) |
| Indian | -0.1990 | -0.2124 | 0.2810*** |
| | (0.2621) | (0.8677) | (0.0000) |
| Age (WII) | -0.0605 | -1.4366 | 0.9274*** |
| | (1.0918) | (3.4260) | (0.0002) |
| PVT test score | -0.2233 | 0.5045 | -0.7623*** |
| | (0.2596) | (6.3798) | (0.0001) |
| Height (WII) | 0.2335 | 0.3253 | 0.3613*** |
| | (0.5198) | (1.4983) | (0.0000) |
| Parents College degree | 0.2431 | -0.7001 | 0.5663*** |
| | (0.3431) | (3.2942) | (0.0000) |
| Log(family income) | -0.1079 | 0.3445 | -0.4402*** |
| | (0.1355) | (1.3343) | (0.0000) |
| Two-parent family | 0.4896 | -0.1076 | -0.6308*** |
| | (0.5854) | (0.2496) | (0.0000) |
| Attendance (WII) | 0.1702 | 1.9800 | 4.3002*** |
| | (1.0188) | (4.5968) | (0.0001) |
| Tobacco at home (WII) | -0.1072 | -0.6083 | 1.9790*** |
| | (0.7256) | (3.2156) | (0.0000) |
| Student characteristics | Yes | Yes | Yes |
| Peers' characteristics | Yes | Yes | Yes |
| Grade indicators | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes |
| N. Obs. | 1,043 | 1,043 | 1,043 |

This table reports NLIV estimates of the structural model 8. In Column 1 we restrict the model by setting $\alpha_1 = 0$ and $\delta = 0$, while in Column 2 we restrict the model by setting $\delta = 0$. Column 3 reports baseline estimates presented in Table 3 Column 2. Students' characteristics are listed in Table 10. The peers' characteristics are calculated as friends' averages of the included variables. cluster-robust numerical standard errors in parentheses. Clusters are defined at school level. *** p<0.01, ** p<0.05, * p<0.1.