Friendship Formation and Smoking Initiation Among Teens

by Donna B. Gilleskie and Yun Sean Zhang

May 2009
Preliminary
Do Not Quote

Abstract

In this research, we use a unique data set to examine the effect of peer influence on teen smoking initiation. First, we develop a game theoretic model where friendship networks and smoking decisions are modeled as the equilibrium outcome of a Bayesian Nash game. An important feature of the model is that teens make friendship and smoking decisions simultaneously to maximize utility. Second, we employ an empirical strategy that allows us to estimate the structural equations that arise out of the theoretical model. Identification depends on instrumental variables that exogenously shift peer smoking norms through either friendship probabilities or individual smoking probabilities. We apply the estimator to data from the National Longitudinal Study of Adolescent Health. Estimation results suggest that peer influence is an important determinant of teen smoking. We also find evidence suggesting that friendship sorting based on racial conformity explains why black teens have a lower smoking rate than white teens. Policy simulation results indicate that peer influence, as a social multiplier, amplifies the cigarette tax deterrent effect on smoking. More generally, however, peer influence primarily promotes smoking among teens.

This research is funded by the National Institute on Aging Grant Number R01-AG16600. We appreciate the useful comments of Han Hong, David Guilkey, Sandra Campo, John Akin, Sergio Parreiras, and seminar participants at the Triangle Health Economics Workshop, AddHealth Users Group Workshop, University of Iowa, and University of Michigan. Additional comments are welcome at donna_gilleskie@unc.edu or seanyz@umich.edu.

1 Introduction

Literally, hundreds of studies in the public health, sociology and psychology literatures confirm a strong positive correlation between individual smoking and peer smoking among teens.[1] In a comprehensive review, Bauman and Ennett (1996) conclude that peer smoking is the most important risk factor for teen smoking. These studies strongly suggest that (conforming) peer effects on smoking exist.[2] The smoking behavior of one's peers, however, acts like a double-edged sword. On one hand, it may motivate an individual to smoke when she has smoker friends; on the other hand, it can discourage the same individual from smoking if she has nonsmoker friends. This implies that both the direction and the magnitude of smoking peer effects on individual behavior depend on that individual's peer smoking norm.[3] We further note that a person's peer smoking norm is endogenous because the person chooses her friends. Thus, evaluating the overall impact of peer effects on teen smoking entails understanding teen friendship. So far, most studies fail to model friendship formation among teens and, consequently, the question of whether peer effects promote or contain teen smoking remains unanswered.

The presence of (universal) smoking deterrents further complicate peer effects on smoking. Peer effects lead to a so called "social multiplier effect" (Sheinkman, 2006; Hoxby, 2000; Epple and Romano, 1998). Consider a school that introduces a $25 fine for on-campus smoking. Besides the direct smoking disincentive caused by the $25 fine, a student's expectation of a lower peer smoking norm should serve as an additional smoking disincentive if peer influence exists. Following this "social multiplier effect" argument, peer effects should amplify the deterrent effect of a cigarette tax, for example, on teen smoking. Given this research context, our paper investigates two research questions. First, does the behavior of

[1] A few examples include Kandel (1978), van Roosmalen and McDaniel (1989, 1992), Crane (1991), Graham et al. (1991), Brooks-Gunn et al. (1993), and Bauman and Ennet (1994, 1996).

[2] If peer effects are disconforming, then we should expect that teens with smoker friends are less likely to smoke. In this paper, peer effects consistently refer to conforming peer effects.

[3] Following convention, an agent's peer smoking norm is defined as the average of the agent's friends' smoking actions. Therefore, it is a (nonlinear) function of three endogenous components: the agent's own peer selection action, others' peer selection actions, and others' smoking actions, with the first two actions governing observed friendship outcomes.

one's peers causally affect teen smoking initiation? Second, if peer effects exist, how do they affect teen smoking initiation as the cigarette tax varies?

If peer influence matters, then a rational agent behaves strategically in the sense that the agent makes decisions based on her expectation of others' decisions. Therefore, it is natural to model peer effects using a game theoretic approach. In this paper, we model adolescent students' smoking decisions in an environment with peer effects as a static simultaneous-move pure strategy Bayes game. In the game, a student chooses both a peer selection action and a smoking action.

Decision interdependence creates an identification concern regarding peer effect estimates. Unlike in single-agent utility maximization problems, in a game an agent's actions on all outcomes of interest (say, peer selection and smoking) are functions of not only her own personal exogenous characteristics but also others' personal exogenous characteristics. This dependence implies that an agent's equilibrium peer smoking norm is inherently a function of her own personal (both observed and unobserved) exogenous characteristics. As a consequence, an agent's peer norm is correlated with the agent's unobserved (by the econometrician) heterogeneity leading to an endogeneity (and identification) concern: is a measured correlation between a person's smoking behavior and her peer's smoking behavior spurious or causal?

To address the concern, our empirical strategy uses the recent two-stage method for estimation of discrete games (Bajari et al., 2006; Pesendorfer and Schmidt-Dengler, 2003). The first stage involves instrumenting for agents' endogenous peer smoking norms through reduced-form analyses; the second stage estimates a behavioral model of smoking that allows for smoking peer effects. In the presence of peer effects, even if the endogenous peer smoking norm and additional exogenous characteristics affect an agent's latent utility (or index) of smoking both linearly and additively in the behavioral specification, the corresponding reduced-form representation of endogenous actions is complicated.[4] As such, an agent's peer smoking norm is a complicated function of exogenous characteristics. This suggests that the econometrician wants to instrument endogenous peer norms through flexibly-specified

[4]In a game, the reduced-form representation of the outcome process is, in essence, a Nash equilibrium strategy that maps exogenous inputs into endogenous outcome(s).

reduced-form models; otherwise, the impact of instrumental variables on the peer norm cannot be fully captured. With this caveat in mind, in the first stage of estimation, we instrument endogenous peer smoking norms in three steps sequentially. First, we model individual equilibrium smoking probabilities using a nonparametric bagged (or bootstrapped aggregated) tree classifier that includes variables that may affect one student's propensity to smoke but do not affect someone else's propensity to smoke conditional on the other's smoking behavior (e.g., characteristics of the former's parents). Second, we explain friendship probabilities with a flexibly-specified logit model. Again, these probabilities are explained by some variables that affect friendship formation, but that do not directly affect smoking behavior of the individual conditional on the probability of him making friends with that person (e.g., differences in characteristics of the individual and the potential friend). Last, we recover instrumented peer smoking norms using the instrumented individual smoking probabilities and the instrumented friendship probabilities. Both parental characteristics and school size are among the instrumental variables in the first two steps. We recognize that each of these variables could be correlated with school-level unobserved heterogeneity.[5] Therefore, in estimating the behavioral smoking decision in the second stage, we further control for school-level unobserved heterogeneity through school fixed effects.

The data used in this study are from the in-home wave I survey of the National Longitudinal Study of Adolescent Health (Add Health). Add Health is unique for its detailed measurements of friendship networks among schoolmates. This data advantage allows us to estimate peer influence on smoking based on true peer composition instead of subjectively defined peer groups (such as schools or classes) as used in previous studies (e.g., Norton et al., 1998; Lundborg, 2006).

Our estimation results indicate that peer effects on smoking initiation are significant and homogenous among teens of different grades who have avoided smoking. In all grades (7 to 12), a one percent increase in the peer smoking norm causes a similar increase in the probability of smoking initiation. Interestingly, although the observed smoking rate of

[5]Parents choose schools for their children; therefore, correlation between parental characteristics and school-level unobserved heterogeneity is expected. School size may be correlated with some unobserved school- level heterogeneity also. For example, in small schools, teachers may have more interactions with students; this may help to prevent smoking initiation.

white teens is significantly higher than that of black teens, smoking initiation rates among black teens is found to be seven percentage points higher than that of white teens after controlling for peer influence. Meanwhile, racial conformity is found to be a significant predictor for teen friendship. Collectively, the results suggest that friendship sorting based on racial conformity significantly reverses the racial smoking rate disparity. This finding provides another potential explanation for the racial smoking rate gap puzzle (US Dept of Health and Human Services, 1998). Our work also compares different specifications that attempt to control for peer influence. The results indicate that using the school norms as an explanatory variable to capture peer influence underestimates peer influence by six fold. In addition, even if one can obtain peer norm measures, failing to control for its endogeneity will underestimate peer influence by five to ten fold.

The presence of peer effects complicates policy simulations. In situations with peer effects, a policy perturbation directly affects every agent's actions on both peer selection and smoking, and in turn, affects every agent's peer smoking norm. Consequently, the ceteris paribus style policy simulation is inappropriate because it is conceptually incorrect to hold an agent's peer smoking norm constant while perturbing policy variables. Policy simulation in a smoking game with peer effects entails searching for an equilibrium that satisfies both the smoking equation and the friendship equation. In operation, we iterate an initial smoking probability vector over a behavioral smoking equation and a reduced-form friendship equation until the smoking probability vector converges uniformly across every agent. Policy simulation results indicate two patterns. First, at certain cigarette tax thresholds, a small increase in the cigarette tax causes the smoking rate to drop abruptly suggesting that at those tax thresholds the social multiplier effect is so strong that students make smoking decisions in a herding pattern. Second, although the existence of peer influence significantly amplifies the deterrent effect of cigarette taxes on smoking initiation, peer influence itself promotes teen smoking initiation more severely. Combined, peer influence is a significant promoting factor for teen smoking initiation.

After providing a review of the relevant literature and previewing our contributions in Section 2, we present in Section 3 a theoretical model of optimal friendship formation and decision outcome(s) of interest. In Section 4, we describe the empirical model of smoking initiation that accounts for endogenous peer norms (made up of both friendship selection and peer smoking behavior). The unique data are described in Section 5. Section 6 discusses our empirical findings and concludes.

2 Background and Contributions

Traditionally, econometricians have focused on the effects of price, income, addiction, and various regulations on cigarette consumption. In studying those effects, econometricians typically model the cigarette consumption decision as a single-agent utility maximization problem (Becker and Murphy, 1988; Chaloupka and Warner, 2000; Cook and Moore, 2000). This framework disallows the possibility that an agent's smoking action has a direct impact on another agent's behavior (i.e., peer influence). In a widely cited work on social interaction, Manski (2000) argues that an agent's decision may affect other agents' decisions through three different avenues: preferences, constraints, and expectations. With regard to teen smoking in particular, the potential existence of complementary preferences in the smoking dimension — a teen receives direct utility if conformity between personal smoking and the peer smoking norm exists — cannot be ruled out. Additionally, econometricians have argued that peer effects may increase the magnitude of price elasticity in cigarette consumption (Lewitt et al., 1981; Becker, 1992). While Gilleskie and Strumpf (2005) find that smoking participation among previous smokers is less sensitive to price increases than that of previous nonsmokers, Becker (1992) suggests that both addiction (i.e., previous use) and peer effects would increase the magnitude of the price elasticity. Constraints may be relaxed if teens obtain cigarettes from their smoking friends. Expectations of the negative effects of smoking may be modified by the presence (or absence) of health problems among friends that smoke.

In the peer effects literature, empirical researchers have proposed various strategies to correct for the endogeneity bias associated with estimation of peer effects. Ideally, one could correct for this endogeneity bias by estimating decisions (both the outcome of interest, say smoking, and peer selection) across all agents jointly. Doing so, in general, is infeasible because the high dimensionality of the peer selection decision involves an extremely large number of repeated observations over *all* agents in a self-containing reference group.[6] So far, few studies have modeled peer effects in a game. Bajari et al. (2006) model peer effects on stock recommendation. In the financial market, however, major stock appraising firms do not choose their peers. Hence, peer selection is not a concern in their study. Krauth (2006) models teens' smoking decisions in a pure strategy game of complete information. He controls for peer selection by allowing individual unobserved heterogeneity (in the smoking equation) to be correlated across agents and no instrumental variables are used to instrument endogenous peer norms.[7]

Some studies use exogenous peer arrangements (e.g., random assignment of roommates) or experiments to evaluate peer effects (Kremer and Levy, 2001; Sacerdote, 2001; Katz et al., 2001; Eisenberg, 2004). Whether this approach can successfully purge off correlation between one's endogenous peer norm and her own error term is open to debate. Experiments perturb the distribution of an agent's potential friends' characteristics, and thus "restrict" an agent's peer selection. The degree of this restriction depends on how homogenous agents are within the experimentally-assigned groups and to what extent the experiment can block friendship between agents from different groups. If agents assigned to the same experimental group are quite homogenous and exchanging friendship signals across groups is costly, then peer

[6]To illustrate this point, let us consider a smoking game with only $N = 10$ agents. To simplify the problem, let us further suppose that the peer selection decision is binary in the sense that an agent either sends a friendship signal to another or does not send a friendship signal to another. In such a smoking game, each agent has $2^{N-1} = 2^9$ possible peer selection actions and 2 possible smoking actions, so the size of the entire joint decision space is $[2 \times (2^9)]^{10}$. Therefore, even if the econometrician observes peer selection decisions, these decisions are expected to distribute very sparsely in the decision space, i.e., the curse of dimensionality emerges. As a consequence, estimation of the peer selection decision jointly across agents requires an extremely large number of repeated observations on a reference group.

If a researcher knows *a priori* that the equilibrium will definitely not be established over some subspaces of the entire possible decision space, then the data requirement is less demanding. However, it is hard to obtain this knowledge. In fact, Add Health does not record the peer selection decision; rather, the Add Health data provide realizations of equilibrium friendship probabilities.

[7]Instrumental variables may still be necessary in Krauth (2006). We note that an agent's unobserved heterogeneity (in the smoking equation) affects the agent's equilibrium peer norm because it affects how the agent makes friends with others. For example, suppose a smoker suddenly contracts asthma (unobserved by the econometrician). This unobserved heterogeneity (asthma) not only makes the smoker unwilling to smoke but also, holding all else constant, makes her unwilling to make friends with smokers in order to avoid the utility loss due to smoking disconformity (as well as second hand smoke). In other words, the smoker reselects her peer smoking norm. This example indicates that one's equilibrium peer smoking norm is correlated with her unobserved heterogeneity even after controlling for the correlation between her own unobserved heterogeneity and her peers' as done in Krauth (2006). Hence, using instrumental variables to purge off correlation between the peer norm and unobserved heterogeneity is still necessary.

selection is quite hindered by the experiment; one could argue that peer selection is largely "controlled". However, even in this situation, an agent is still choosing friends from among group members who are similar to her. Hence, strictly speaking, even in experiments, an agent's peer norm is still a function of endogenous peer selection and, in turn, correlated with the agent's own unobserved heterogeneity affecting her peer selection.

Another strategy to purge off the correlation between one's peer norm and her unobserved heterogeneity exploits instrumental variables that directly affect peers' outcomes of interest but not the individual's outcome (Evans, Oates and Schwab, 1992; Gaviria and Raphael, 2001; Hoxby, 2000; Ioannides and Zabel, 2002). This strategy, however, is not completely effective. Although it purges off the correlation between an agent's peers' outcomes and the agent's individual unobserved heterogeneity, the agent's peer selection decision is still correlated with her individual unobserved heterogeneity. In other words, the agent's instrumented peer norm, a function of the smoking behavior of her friends, is still a variable that is conditional on endogenous friendships. Realizing this shortcoming, we instrument an agent's friendships as well as the agent's peers' smoking decisions to address endogeneity of the observed peer norm.

Our research makes several contributions to the literature. First, it provides a theoretical framework that includes both the friendship decision and other decisions (e.g., smoking). Previous theoretical models (Manski, 1995; McMillan, 1995; Bajari, 1996; Brock and Durlauf, 2001; WHO, 2006) begin by considering a reference group of people (say N people) in which an agent's utility associated with an action is affected by *all* the other $N-1$ agents' actions.[8] Unfortunately, these models fail to explain how the reference group comes into being in the first place ignoring that a reference group is endogenously chosen by agents collectively. As a contribution, our model explains not only strategic decisionmaking on actions other than peer selection among members in a reference group but also how agents choose friends strategically to form the relevant reference group.

[8]This implicitly imposes a strong assumption that any one of the N player's utility associated with an action is affected by any one of the rest of the $N-1$ players' actions. So, asymmetric friendships among agents are ruled out in these models.

Additionally, our paper provides insight on the econometric concerns associated with estimation of a Bayes game with peer influence, which cannot be seen easily otherwise. For example, the theoretical model reveals that estimating friendship selection and other actions jointly across all agents, in general, is impossible due to the curse of dimensionality. Also, the derivation of the econometric specification shows exactly which elements are absorbed into the error term. Furthermore, the model suggests why, in estimation of the first stage of the game, the econometrician is willing to trade model interpretability for flexibility.

Policy simulations in the published peer effects literature only allow agents to choose actions given their friendships (e.g. Norton, 1998; Krauth, 2006; Lundborg, 2006; Cooley, 2007). These simulations ignore the fact that a rational agent should update her friendships while choosing other actions in order to maximize her expected utility. Our study contributes to the literature by conducting policy simulations that allow agents to choose friends while also choosing other actions (e.g., smoking).

3 Theoretical Model

The theoretical model described below assumes that agents play a pure strategy simultaneous-move Bayes game. The game endogenizes both the discrete peer selection decision and other discrete decisions.[9] The decision process in the game is the following. At the beginning of a period, each agent receives an action-dependent private shock/information. Then each agent chooses the actions of peer selection and smoking based on public information, the action-dependent private shock, and her expectation of others' private shocks conditional upon her individual private shock. Due to the presence of a stochastic private shock, an agent is not able to tell what actions she should take prior to the realization of the private shock. As such, even with full observability of public information, one can at most predict equilibrium actions in a probability sense. The equilibrium probability distribution of actions on both peer selection and the outcomes of interest, and hence, also equilibrium friendship

[9]Although our main interest is peer effects on smoking, this theoretical model is general in the sense that it allows for peer effects in multiple endogenous dimensions (e.g., smoking, drinking, academic performance) with multiple exogenous interactions (e.g., gender, race, family income).

probabilities among agents, are completely governed by exogenous public information and a common prior.[10]

3.1 Peer Selection and Friendship Production

Consider a self-containing group of players $N_t = \{1, 2, \ldots, n_t\}$ in period $t\,(t = 1, 2, \ldots, T)$. N_t is self-containing in the sense that every player in N_t only makes friends with a subset of N_t. We assume that selected peers are equally important. Players are characterized by a $n_t \times K$ matrix z_t known to all members in N_t at the decision moment in period t. Specifically,

$$z_t = \begin{bmatrix} z'_{1,t} \\ \cdot \\ z'_{i,t} \\ \cdot \\ z'_{n_t,t} \end{bmatrix} = \begin{bmatrix} z_{1,1,t} & \ldots & z_{1,k,t} & \ldots & z_{1,K,t} \\ \cdot & \ldots & \cdot & \ldots & \cdot \\ z_{i,1,t} & \ldots & z_{i,k,t} & \ldots & z_{i,K,t} \\ \cdot & \ldots & \cdot & \ldots & \cdot \\ z_{n_t,1,t} & \ldots & z_{n_t,k,t} & \ldots & z_{n_t,K,t} \end{bmatrix}$$

where $z_{i,t}$ is the column vector that records individual i's exogenous characteristics and predetermined actions.

At the beginning of period t, a generic player i's peer selection action is a $n_t \times 1$ friendship signal vector $(s_{i,t})$, where $s_{i,t} = [s^1_{i,t} ... s^j_{i,t} ... s^{n_t}_{i,t}]'$. We denote all the players in set N_t other than player i by $N_t \backslash \{i\}$. If player i decides to send a friendship signal to player $j \in N_t \backslash \{i\}$ then $s^j_{i,t} = 1$; otherwise, $s^j_{i,t} = 0$. A player does not send a friendship signal to herself; therefore, $s^i_{i,t} = 0 \; \forall i \in N_t$. The cost associated with a particular vector of peer selection actions is reflected in player i's budget constraint in Section 3.3 below. Let S_i denote a generic player i's peer selection action set (i.e., $s_{i,t} \in S_i$). S_i has $2^{n_t - 1}$ distinct elements corresponding to agent i's $2^{n_t - 1}$ different ways of sending friendship signals to $n_t - 1$ other players. Stacking $s'_{1,t}, s'_{2,t}, \ldots, s'_{i,t}, \ldots, s'_{n_t,t}$, we obtain a $n_t \times n_t$ square matrix (s_t) that records peer selection actions across all players in N_t. That is,

$$s_t = \begin{bmatrix} s'_{1,t} \\ \cdot \\ s'_{i,t} \\ \cdot \\ s'_{n_t,t} \end{bmatrix} = \begin{bmatrix} 0 & \ldots & s^j_{1,t} & \ldots & s^{n_t}_{1,t} \\ \cdot & \ldots & \cdot & \ldots & \cdot \\ s^1_{i,t} & \ldots & 0 & \ldots & s^{n_t}_{i,t} \\ \cdot & \ldots & \cdot & \ldots & \cdot \\ s^1_{n_t,t} & \ldots & s^j_{n_t,t} & \ldots & 0 \end{bmatrix}.$$

[10]Though the common prior is known to every player in the game, it is not uniformly included as part of the public information vector notation in order to emphasize its presence in particular equations.

In period t, after the peer selection decision (s_t), the friendship network can be depicted by a $n_t \times n_t$ matrix ζ_t. The network records observed friendship outcomes. A generic element of ζ_t ($\zeta_{i,t}^j$) is either 1 (agent i regards agent j as a peer) or 0 (agent i does not regard agent j as a peer). Since one does not consider herself as a peer, $\zeta_{i,t}^i = 0$, $\forall i \in N_t$. Because friendship can be asymmetric, $\zeta_{i,t}^j$ may not equal $\zeta_{j,t}^i$. More specifically,

$$\zeta_t = \begin{bmatrix} \zeta_{1,t}' \\ \cdot \\ \zeta_{i,t}' \\ \cdot \\ \zeta_{n_t,t}' \end{bmatrix} = \begin{bmatrix} 0 & \dots & \zeta_{1,t}^j & \dots & \zeta_{1,t}^{n_t} \\ \cdot & \dots & \cdot & \dots & \cdot \\ \zeta_{i,t}^1 & \dots & 0 & \dots & \zeta_{i,t}^{n_t} \\ \cdot & \dots & \cdot & \dots & \cdot \\ \zeta_{n_t,t}^{n_t} & \dots & \zeta_{n_t,t}^j & \dots & 0 \end{bmatrix} .$$

A generic element of matrix ζ_t, $\zeta_{i,t}^j$, is the output of the friendship network production function $\zeta(\cdot)$ that depends on the friendship signals of each player i in N_t, as well as their exogenous characteristics and predetermined actions.[11] That is,

$$\zeta_{i,t}^j = \zeta(s_{i,t}^j, s_{j,t}^i; z_{i,t}, z_{j,t}) = \mathbf{1}[\zeta^*(s_{i,t}^j, s_{j,t}^i; z_{i,t}, z_{j,t}) \geq \zeta_c^*(z_{i,t}, z_{j,t})] \tag{1}$$

where $\zeta^*(s_{i,t}^j, s_{j,t}^i; z_{i,t}, z_{j,t})$ is the latent index governing whether agent i regards agent j as a peer and $\zeta_c^*(z_{i,t}, z_{j,t})$ is the friendship cutoff value that varies as characteristics of either agent i or agent j vary. $\mathbf{1}[\cdot]$ is an indicator function that takes on the value 1 if the argument in parentheses is satisfied, and zero otherwise.

Because friendship can be asymmetric, the sequencing of input arguments in functions $\zeta^*(\cdot)$, $\zeta_c^*(\cdot)$, and $\zeta(\cdot)$ matters. In general,

$$\begin{aligned} \zeta_c^*(z_{i,t}, z_{j,t}) &\neq \zeta_c^*(z_{j,t}, z_{i,t}) \\ \zeta^*(s_{i,t}^j, s_{j,t}^i; z_{i,t}, z_{j,t}) &\neq \zeta^*(s_{j,t}^i, s_{i,t}^j; z_{j,t}, z_{i,t}) \\ \zeta(s_{i,t}^j, s_{j,t}^i; z_{i,t}, z_{j,t}) &\neq \zeta(s_{j,t}^i, s_{i,t}^j; z_{j,t}, z_{i,t}) . \end{aligned}$$

We make the following assumptions about the friendship network production function $\zeta(\cdot)$.

- Assumption 1: For all $s_{j,t}^i, z_{i,t}$ and $z_{j,t}$, if $s_{i,t}^j = 0$ then $\zeta_{i,t}^j = 0$.

  If agent i does not send a friendship signal to agent j then agent i does not regard

[11] We assume that friendship between two agents is affected only by their friendship signals to each other and their exogenous characteristics. Then, $z_{\ell,t}$ $\forall \ell \in N_t \backslash \{i, j\}$ and any element in s_t other than $s_{i,t}^j$ and $s_{j,t}^i$ are not arguments of $\zeta(s_{i,t}^j, s_{j,t}^i; z_{i,t}, z_{j,t})$ in equation 1.

individual j as a peer no matter what peer selection decision is made by agent j and what characteristics these two agents have.

- Assumption 2: If $\zeta_{i,t}^{j} = 1$, then $s_{j,t}^{i} = 1$.

  A necessary condition for agent i to regard agent j as a peer is that agent j sends a friendship signal to agent i. This is consistent with the intuition that an agent does not regard another agent as a friend if the latter agent does not spend some time or other resources on the former.

- Assumption 3: For all $s_{i,t}^{j}$, $z_{i,t}$ and $z_{j,t}$, $\zeta^{*}(s_{i,t}^{j}; s_{j,t}^{i} = 1, z_{i,t}, z_{j,t}) \geq \zeta^{*}(s_{i,t}^{j}; s_{j,t}^{i} = 0; z_{i,t}, z_{j,t})$.

  Given agent i's peer selection decision and the exogenous characteristics of agent i and agent j, if agent j sends a friendship signal to agent i, the latent index value of 'agent i regards agent j as a friend' increases.

Assumption 1 and Assumption 2 jointly imply the following:

$$\zeta_{i,t}^{j} = 1 \Rightarrow s_{i,t}^{j} = 1 \text{ and } s_{j,t}^{i} = 1.$$

As long as one agent i regards the other agent j as a peer then these two agents must have exchanged friendship signals. The number of agents regarded by agent i as friends in period t is $\sum\limits_{j \in N_t \setminus \{i\}} \zeta_{i,t}^{j}$.

3.2 Decisions Other Than Peer Selection

Together with the peer selection action, player i also chooses M other actions ($m = 1, 2, ..., M$) summarized in a $M \times 1$ column vector as

$$a_{i,t} = [\ a_{i,t}^{1} \quad a_{i,t}^{2} \quad ... \quad a_{i,t}^{m} \quad ... \quad a_{i,t}^{M}\ ]'.$$

Let A_i denote the decision set of the m^{th} ($m = 1, 2, ..., M$) action for all players, i.e., $a_{i,t}^{m} \in A_i$. A_i has J_m alternatives. For example, if the m^{th} action refers to smoking participation and A_i is the set of (binary) smoking actions, then $A_i = \{0, 1\}$ and $J_m = 2$.

Let column vector $d_{i,t} = [s_{i,t}' \ a_{i,t}']'$ denote an arbitrary decision on both peer selection and the M other choices made by player i at the beginning of period t. Let $D_i \equiv S_i \times \prod_{m=1}^{M} A_i$

be the corresponding decision set where $d_{i,t} \in D_i$. D_i has $N_{D_i} = 2^{n_t-1}\prod_{m=1}^{M} J_m$ elements. D_i is the collection of mutually exclusive and collectively exhaustive actions player i can choose from at decision moment t. In general, since any two different players, say player i and player j, may choose peers from two different peer choice sets, then, $S_i \neq S_j$ and in turn, $D_i \neq D_j$. However, $\forall i \in N_t$, S_i has 2^{n_t-1} elements, so we define $N_D = N_{D_i}\ \forall i \in N_t$.

Let $\overline{a}_{i,t} = [\overline{a}^1_{i,t}\ \overline{a}^2_{i,t} \ldots \overline{a}^m_{i,t} \ldots \overline{a}^M_{i,t}]'$ denote player i's $M \times 1$ peer norm vector in period t, defined as an average of the actions of the peers of player i. We note that player i's peer norm is determined by both her peer selection decision ($s_{i,t}$) and the other $n_t - 1$ players' $M + 1$ decisions (peer selection and actions). The m^{th} element of $\overline{a}_{i,t}$ is defined as

$$\overline{a}^m_{i,t} = \frac{\sum_{j \in N_t \backslash \{i\}} \zeta^j_{i,t}(s^j_{i,t}, s^i_{j,t}; z_{i,t}, z_{j,t}) a^m_{j,t}}{\sum_{j \in N_t \backslash \{i\}} \zeta^j_{i,t}(s^j_{i,t}, s^i_{j,t}; z_{i,t}, z_{j,t})}\ \forall m, m = 1, 2, \ldots, M\ . \tag{2}$$

Having introduced the players' endogenous actions, we next lay out the main elements of this simultaneous game.[12]

3.3 Preferences

Suppose player i knows, at decision moment t, all other $n_t - 1$ players actions $d_{-i,t} = [d_{1,t}, \ldots d_{i-1,t}, d_{i+1,t}, \ldots d_{n_t,t}] \in D_{-i}$ where $D_{-i} \equiv \prod_{j \in N_t \backslash \{i\}} D_j$. She solves the following single-agent constrained utility maximization problem:

$$\begin{aligned} \max_{d_{i,t}} \quad & \sum_{m=1}^{M} u_m(a^m_{i,t}; z_{i,t}) + \mathbf{1}[\sum_{j \in N_t \backslash \{i\}} \zeta^j_{i,t} > 0] \sum_{m=1}^{M} u^p_m(|a^m_{i,t} - \overline{a}^m_{i,t}|; z_{i,t}) \\ & \text{s.t. } C(d_{i,t}; d_{-i,t}, z_{i,t}, p_t) - y_{i,t} \leq 0 \end{aligned} \tag{3}$$

where $u_m(a^m_{i,t}; z_{i,t})$ captures the payoff associated with the m^{th} action ($m = 1, \ldots, M$) other than peer selection. If player i has any friends (i.e., $\mathbf{1}[\sum_{j \in N_t \backslash \{i\}} \zeta^j_{i,t} > 0] = 1$), she also receives utility from the deviation of her m^{th} action from her associated peer norm, $u^p_m(|a^m_{i,t} - \overline{a}^m_{i,t}|; z_{i,t})$. It may be the case that player i maximizes her payoff by choosing to regard no one as a friend. If so, then $\mathbf{1}[\sum_{j \in N_t \backslash \{i\}} \zeta^j_{i,t} > 0] = 0$.

[12]Henceforth, the set $N_t \backslash \{i\}$ and the subscript '$-i$' represent all players in N_t except player i.

Individuals maximize utility subject to a budget constraint that reflects the feasible actions within the decision space, given one's per-period income ($y_{i,t}$) and the price (p_t) of particular actions. Henceforth, income and price variables, such as cigarette prices or prices associated with other actions, are contained in $z_{i,t}$. $C(d_{i,t}; d_{-i,t}, z_t)$ represents the cost of decision $d_{i,t}$ conditional upon other players' decisions $d_{-i,t}$. The presence of $d_{-i,t}$ in player i's budget constraint reflects Manski's (2000) point that social interaction could work not only through preferences (i.e., the utility function) but also through the budget constraint. For example, suppose Tom dislikes Jack. Compared with the case where Jack does not try to make friends with Mary, it probably becomes more financially expensive for Tom to make friends with Mary if Jack tries to make friends with Mary too.[13]

3.4 Information Structure

At the beginning of period t, each of the n_t players receives an (action-specific) private shock (private information). The $n_t \times N_D$ matrix of private shocks, ε_t, is

$$\varepsilon_t = \begin{bmatrix} \varepsilon'_{1,t} \\ \cdot \\ \varepsilon'_{i,t} \\ \cdot \\ \varepsilon'_{n_t,t} \end{bmatrix} = \begin{bmatrix} \varepsilon^1_{1,t} & \dots & \varepsilon^h_{1,t} & \dots & \varepsilon^{N_D}_{1,t} \\ \cdot & \dots & \cdot & \dots & \cdot \\ \varepsilon^1_{i,t} & \dots & \varepsilon^h_{i,t} & \dots & \varepsilon^{N_D}_{i,t} \\ \cdot & \dots & \cdot & \dots & \dots \\ \varepsilon^1_{n_t,t} & \dots & \varepsilon^h_{n_t,t} & \dots & \varepsilon^{N_D}_{n_t,t} \end{bmatrix}$$

where $\varepsilon^h_{i,t}$ is player i's private shock if she chooses the h^{th} decision from her decision choice set D_i. Column vector $\varepsilon_{i,t} = [\ \varepsilon^1_{i,t} \ \ .... \ \ \varepsilon^h_{i,t} \ \ .... \ \ \varepsilon^{N_D}_{i,t}\ ]'$ is player i's private shock vector, which is not revealed to any player $j \neq i$ at the decision moment. In a Bayes game framework, $\varepsilon_{i,t}$ represents "player type". The distribution of private shocks, $\varepsilon_{i,t} \sim F(\cdot)$, is known by every player in the game. Therefore, $f(\varepsilon_{i,t})$ is the common prior. Let $f_{-i|i}(\varepsilon_{-i,t}|\varepsilon_{i,t})$ denote player i's prior over others' private shocks (types) conditional upon hers. Individual priors

[13]The maximization problem above does not reflect a third avenue through which peer effects may be manifested: expectations of future events. That is, individuals may learn about smoking risks, for example, by observing the health outcomes of their peers and updating beliefs about their own probabilities of detrimental health effects. Related, it does not consider the dynamic effects of current smoking decisions on future preferences, constraints, or expectations. In order to focus on peer effects, the model is restricted to myopic behavior.

are consistent with the common prior.[14] Following Bayes' rule, we have

$$f_{-i|i}(\varepsilon_{-i,t}|\varepsilon_{i,t}) = \frac{f(\varepsilon_{i,t}, \varepsilon_{-i,t})}{\int f(\varepsilon_{i,t}, \varepsilon_{-i,t}) d\varepsilon_{-i,t}}, \ \forall i \in N_t.$$

3.5 Allowed Decision Space

Social interactions among agents' budget constraints imply that not only player i's personal decision ($d_{i,t}$) but also her potential peers' decisions ($d_{-i,t}$) affect whether her budget constraint is satisfied. In the game, we assume that there exists at least one nonempty subspace of the largest possible joint decision space, $\prod_{i \in N_t} D_i$, that satisfies the following two conditions simultaneously: (1) the subset can be expressed as a Cartesian product of all n_t players' individual decision space; and (2) each element of the subspace satisfies all n_t budget constraints. We refer to such a subspace as an Allowed Decision Space (ADS) and characterize an ADS in Appendix 1. Hereafter, we abuse notation and use $D = \prod_{i \in N_t} D_i$ to denote an ADS to ease the notational burden.

Consider an ADS, $D = \prod_{i \in N_t} D_i$, and let N_{D_i} denote the number of elements in D_i; then there are $\prod_{i \in N_t} N_{D_i}$ decision elements in D. Solving the maximization problem defined in equation 3, we obtain the following action-specific payoff functions

$$V(d_{i,t}, d_{-i,t}, \varepsilon_{i,t}, z_t, u_m(\cdot), u_m^p(\cdot)) = \sum_{m=1}^{M} V_m(d_{i,t};\ d_{-i,t}, z_t) + \varepsilon_{i,t}^{d_{i,t}} \quad \forall d_t \in D \tag{4}$$

where $\varepsilon_{i,t}^{d_{i,t}}$ is player i's action-dependent private shock. Let I_t denote the public information at decision moment t, such that $I_t = \{z_t, \{V_m(\cdot)\}_{m=1}^{M}, \zeta(\cdot)\} \in \vartheta_{pub}$.

3.6 The Bayes Game and Pure Strategy Bayes Nash Equilibrium

At decision moment t, the following fundamentals ($G_t = \left\langle N_t,\ E,\ I_t,\ f_{-i|i},\ D \right\rangle$) govern the n_t players' decisions:

- Players: $N_t = \{1, 2, \ldots, i, \ldots, n_t\}$
- Private shock set (player type space) : $E = \prod_{i \in N_t} E_i$ where $\varepsilon_{i,t} \in E_i$

[14] We define $f_{-i|i} = (f_{-i|i}(\varepsilon_{-i,t}|\varepsilon_{i,t}))_{i \in N_t}$ where $f_{-i|i}$ is the collection of individual priors consistent with the common prior.

- Public information: $I_t = \{z_t, \{V_m(\cdot)\}_{m=1}^{M}, \zeta(\cdot)\} \in \vartheta_{pub}$

- Priors: $f_{-i|i} = (f_{-i|i}(\varepsilon_{-i,t}|\varepsilon_{i,t}))_{i \in N_t}$ with distribution $F(\cdot)$

- Allowed decision space (ADS): D

A player's strategy profile $(d_{i,t}(\varepsilon_{i,t}, I_t))$ is a correspondence between (both private and public) information and the individual decision; that is, $d_{i,t}(\varepsilon_{i,t}, I_t) : E_i \times \vartheta_{pub} \longmapsto D_i$. At a Nash equilibrium, no player has an incentive to alter her strategy as long as other players do not.

Therefore, a pure strategy Bayes Nash equilibrium (PSBNE) is a collection of n_t mappings from the information set to the decision set, i.e., $d^*_{i,t}(\varepsilon_{i,t}, I_t))_{i \in N_t}$, such that

$$\begin{aligned} d^*_{i,t}(\varepsilon_{i,t}, I_t) \;&\in\; \arg\max_{d_{i,t} \in D_i} \int\limits_{\varepsilon_{-i,t} \in E_{-i}} V(d_{i,t}, d_{-i,t}) f_{-i|i}(\varepsilon_{-i,t}|\varepsilon_{i,t}) d\varepsilon_{-i,t} \\ &\forall i \in N_t \text{ and } \forall \{\varepsilon_{i,t}, I_t\} \in E_i \times \vartheta_{pub} \end{aligned} \tag{5}$$

where $D_i = \bigcup\limits_{\varepsilon_{i,t} \in E_i} d^*_{i,t}(\varepsilon_{i,t}, I_t)$ and $f_{-i|i}(\varepsilon_{-i,t}|\varepsilon_{i,t}) = \frac{f(\varepsilon_{i,t}, \varepsilon_{-i,t})}{\int f(\varepsilon_{i,t}, \varepsilon_{-i,t}) d\varepsilon_{-i,t}}$ are individual priors consistent with a common prior restricted to E_i.

Let us define $d^*(\cdot) : E \times \vartheta_{pub} \longmapsto D$ where $d^*((\varepsilon_{i,t}, I_t)_{i \in N_t}) \equiv (d^*_{i,t}(\varepsilon_{i,t}, I_t))_{i \in N_t}$. By stacking $d^*(\cdot)$ over all players in N_t, equation 5 can be written as a function $\Psi(\cdot) : D^* \mapsto D^*$. More specifically,

$$\begin{aligned} d^*((\varepsilon_{i,t}, I_t)_{i \in N_t}) \;&=\; \Psi(d^*((\varepsilon_{i,t}, I_t)_{i \in N_t});\; N_t,\; I_t,\; f_{-i|i}) \\ &\forall \{\varepsilon_{i,t}, I_t\} \in E_i \times \vartheta_{pub} \,. \end{aligned} \tag{6}$$

According to equation 6, a PSBNE $d^*((\varepsilon_{i,t}, I_t)_{i \in N_t})$ is a fixed point of function Ψ. According to Brouwer's fixed point theorem, such a PSBNE exists.

3.7 Equilibrium Decision Probabilities

Due to the random realization of private information, an exact prediction of whether a player makes a specific decision or not is impossible. However, one can predict the equilibrium decision probabilities based on public information. In this subsection, we characterize the decision probabilities at a PSBNE.

Let the set π_i^* denote the collection of equilibrium probabilities corresponding to equilibrium decision elements in D_i. Let the function $\pi_i^*(\cdot) : D_i^* \longmapsto \pi_i^*$ denote the correspondence between equilibrium decision elements and their probabilities. Then

$$\pi_i^*(d^*) = \int_{\varepsilon_{i,t}} 1[d_{i,t}^*(\varepsilon_{i,t}, I_t) \geq d^*] f_{\varepsilon_{i,t}}(\varepsilon_{i,t}) d\varepsilon_{i,t} \quad \forall i \in N_t, \forall d^* \in D_i$$

where $f_{\varepsilon_{i,t}}(\varepsilon_{i,t})$ is the marginal density function of $\varepsilon_{i,t}$; that is

$$f_{\varepsilon_{i,t}}(\varepsilon_{i,t}) = \int_{\varepsilon_{-i,t} \in E_{-i}} f(\varepsilon_{i,t}, \varepsilon_{-i,t}) d\varepsilon_{-i,t}$$

and $1[\cdot]$ is an indicator function where

$$1[d_{i,t}^*(\varepsilon_{i,t}, I_t) \geq d^*] = \begin{cases} 1 & \text{if } d_{i,t}^*(\varepsilon_{i,t}, I_t) \geq d^* \\ 0 & \text{otherwise.} \end{cases}$$

We note that the following equation holds:

$$\sum_{d^* \in D_i} \pi_i^*(d^*) = 1 \quad \forall i \in N_t.$$

4 Empirical Strategy

In the presence of peer effects, schoolmates making decisions about particular actions (e.g., smoking initiation) are playing a simultaneous game. This simultaneity implies that the ideal approach to understanding what influences their decisions is to estimate the peer selection and smoking actions of all schoolmates jointly. Empirical analysis of the equilibrium decisonmaking means that the econometrician should regard a cross section of peer selection actions and smoking actions made by all schoolmates as a single observation. Herein lies the limitations of that analytic approach and any data set used in estimation. Our analysis sample of the Add Health data contains 124 schools. The sample size would be restricted to 124 observations if we used this approach. Additionally, each school teaches a large number of students. Joint estimation of schoolmates' decisions is computationally infeasible to implement. Another data limitation is that students' peer selection actions are not observed. The Add Health data, while unique in that we can observe the friendship network, provide information on friendship outcomes only. Given these data limitations, we chose to estimate the effects of peer behavior on the dichotomous smoking initiation decision agent by agent (i.e., student i chooses to smoke or not) rather than considering the many combinations of smoking alternatives of all agents in a reference group (i.e., a school in this study) as the polychotomous outcomes of interest. Furthermore, our empirical strategy does not estimate peer selection actions and smoking actions jointly within an agent because we do not observe peer selection actions. Instead we estimate friendship outcome probabilities (i.e., student i considers student j a friend or not).

In what follows, we first derive an empirical specification that is consistent with the theoretical model and is implementable under the data limitations mentioned above. Additionally, we discuss identification concerns in the empirical specification. We also explain how to instrument endogenous peer smoking norms using two separate reduced-form analyses in sequence. Next, we explain how to estimate the structural form empirical specification for smoking initiation using the instrumented peer norm.

4.1 Derivation of Empirical Specification and Identification

In this subsection, we focus the reader on the outcome of interest by applying the theoretical model of peer effects to smoking initiation. We assume players only choose a peer selection action and a smoking action in a one period smoking game.[15] Hence, a generic player's decision is $d_i = [s_i \ a_i]'$ where s_i is the peer selection action and a_i is the smoking action. At a PSBNE, a generic player i expects that other players adopt the equilibrium strategy $d^*_{-i}(\cdot)$. Therefore, we parameterize player i's decision-specific payoff (equation 4) as follows:

$$\begin{aligned} V(d_i = [s_i \ a_i]', d^*_{-i}(\varepsilon_{-i}, I), \varepsilon_i, z, u_{m=1}(\cdot), u^p_{m=1}(\cdot)) & \\ &= V_{m=1}(d_i; d^*_{-i}, z) + \varepsilon_i^{d_i} \\ &= z\beta(a_i; I) + \delta a_i + \gamma \left| a_i - \overline{a}_i(s_i; d^*_{-i}, z) \right| + \varepsilon_i^{d_i} \qquad (7) \end{aligned}$$

where

- $z = \begin{bmatrix} z_i' \\ z_{-i}' \end{bmatrix}$ is a vector of exogenous characteristics of all players in the game;

[15] That is, $M = 1$; there is only one modeled action, in addition to peer selection. Since we model behavior in one period in our application, t subscripts are dropped in this section forward.

- $\beta(a_i; I)$ is the parameter vector corresponding to z and, according to equation 4, $\beta(a_i; I)$ is smoking-action-dependent;

- δ is a parameter measuring the utility of smoking (e.g., stress relief, cool factor) or disutility (e.g., health effects) independent of other players' actions;

- $\overline{a}_i(s_i; d^*_{-i}, z)$ is player i's peer smoking norm when she makes peer selection action s_i. It is a function of s_i reflecting that player i chooses her peers, and, hence, her peer smoking norm;

- $\left|a_i - \overline{a}_i(s_i; d^*_{-i}, z)\right| \geq 0$ is player i's deviation from her peer smoking norm. If she smokes then $\left|a_i - \overline{a}_i(s_i; d^*_{-i}, z)\right| = 1 - \overline{a}_i(s_i; d^*_{-i}, z)$; otherwise, $\left|a_i - \overline{a}_i(s_i; d^*_{-i}, z)\right| = \overline{a}_i(s_i; d^*_{-i}, z)$;

- $\gamma \leq 0$ is the smoking peer effects parameter. It measures the disutility of deviation from the peer smoking norm.

Again, we do not observe students' peer selection actions (s_i) in the Add Health data but we do observe friendship outcomes. To proceed with estimation, we first derive player i's *ex post* expected smoking-action-specific payoff based on equation 7.[16] Recall that $d^*_{-i} \equiv d^*_{-i}(\varepsilon_{-i}, I)$ is a function of the schoolmates' private shocks and public information. We integrate over the player's schoolmates' private shocks (distributed $F(\cdot)$) and condition player i's *ex post* expected smoking-action-specific payoff on her own private shock and the public information. That is,

$$\begin{aligned} EV(\varepsilon_i, I, F(\cdot); a_i) &= \max\nolimits_{s_i} \left[z\beta(a_i; I) + \delta a_i + \gamma \left| a_i - \int\limits_{\varepsilon_{-i}|\varepsilon_i} \overline{a}_i(s_i; d^*_{-i}, z) dF(\varepsilon_{-i}|\varepsilon_i) \right| + \varepsilon_i^{d_i} \right] \\ &= z\beta(a_i; I) + \delta a_i + \gamma \left| a_i - \int\limits_{\varepsilon_{-i}|\varepsilon_i} \overline{a}_i(s_i; d^*_{-i}, z) dF(\varepsilon_{-i}|\varepsilon_i) \right| + \varepsilon_i^{[s_i^*(a_i)\ a_i]'} \end{aligned} \quad (8)$$

where $s_i^*(a_i) = \arg\max_{s_i} \left[z\beta(a_i; I) + \delta a_i + \gamma \left| a_i - \int\limits_{\varepsilon_{-i}|\varepsilon_i} \overline{a}_i(s_i; z, d^*_{-i}) dF(\varepsilon_{-i}|\varepsilon_i) \right| + \varepsilon_i^{[s_i^*(a_i)\ a_i]'} \right]$.
We note that if a_i is player i's smoking action at a PSBNE, then $s_i^*(a_i)$ is her equilibrium peer selection action.

[16] *ex post* means after the realization of the private shock (ε_i); *ex ante* means prior to the realization of the private shock. Note that this expected payoff includes integrating over the distribution of a players' schoolmates private shocks (ε_{-i}).

Substituting the peer smoking norm (equation 2) into equation 8 yields

$$\begin{aligned} EV(\varepsilon_i, I, F(\cdot); a_i) = \quad & z\beta(a_i; I) + \delta a_i \\ & + \gamma \left| a_i - \int_{\varepsilon_{-i}|\varepsilon_i} \frac{\sum\limits_{j \in N_t \backslash \{i\}} \zeta_i^j(s_i^{j*}(a_i), s_j^{i*}; z_i, z_j) a_j^*(\varepsilon_j, I)}{\sum\limits_{j \in N_t \backslash \{i\}} \zeta_i^j(s_i^{j*}(a_i), s_j^{i*}; z_i, z_j)} dF(\varepsilon_{-i}|\varepsilon_i) \right| + \varepsilon_i^{[s_i^*(a_i) \;\; a_i]'}. \end{aligned} \tag{9}$$

Defining the *ex post* expected peer smoking norm as $E[\overline{a}_i^*|\varepsilon_i, I, F(\cdot)]$, where

$$E[\overline{a}_i^*|\varepsilon_i, I, F(\cdot)] = \int_{\varepsilon_{-i}|\varepsilon_i} \frac{\sum\limits_{j \in N_s \backslash \{i\}} \zeta(s_i^*(\varepsilon_i, I), s_j^*(\varepsilon_j, I); I) a_j^*(\varepsilon_j, I)}{\sum\limits_{j \in N_s \backslash \{i\}} \zeta(s_i^*(\varepsilon_i, I), s_{-i}^*(\varepsilon_{-i}, I); I)} dF(\varepsilon_{-i}|\varepsilon_i) \in [0, 1] \,, \tag{10}$$

we can simplify player i's *ex post* expected smoking-action-specific payoff to

$$EV(\varepsilon_i, I, F(\cdot); a_i) = z\beta(a_i; I) + \delta a_i + \gamma \left| a_i - E[\overline{a}_i^*|\varepsilon_i, I, F(\cdot)] \right| + \varepsilon_i^{a_i} \tag{11}$$

where $\left| a_i - E[\overline{a}_i^*|\varepsilon_i, I, F(\cdot)] \right| \geq 0$ is the deviation from the *ex post* expected peer smoking norm.

We note that if the number of players in the game is even modestly large (say 50 players), then estimating parameters related to exogenous terms containing z_{-i} is difficult because there is a large number of different interactions between z_j and $z_{j'}$ ($j \in N\backslash\{i\}$ and $j' \in N\backslash\{i\}$) and between z_i and z_{-i}. As such, we focus on the components of $z\beta(a_i; I)$ that are purely related to z_i (i.e., do not involve interactions of individual observables with peer observables). To do so, we assume that $z\beta(a_i; I)$ can be decomposed as

$$z\beta(a_i; I) = z_i \beta_{self}(a_i; I) + f(z_i, z_{-i}) \beta_{interact}(a_i; I) \tag{12}$$

where $\beta_{self}(a_i; I)$ is the parameter vector corresponding to player i's exogenous characteristics z_i; $f(z_i, z_{-i})$ collects the terms containing interactions of some or all elements in z_{-i} with z_i; and $\beta_{interact}(a_i; I)$ is the parameter vector corresponding to z_{-i}. Let us define as unobservable (or not modeled by the econometrician because of the large number)

$$\upsilon_i^{a_i}(I) \equiv f(z_i, z_{-i}) \beta_{interact}(a_i; I). \tag{13}$$

We note that $\upsilon_i^{a_i}(I)$ is smoking-action-specific because $\beta_{interact}(a_i; I)$ is smoking-action-specific. Substituting equations 12 and 13 into equation 11, we further simplify the *ex* *post* expected smoking-action-specific payoff to

$$EV(\varepsilon_i, I, F(\cdot); a_i) = z_i\beta_{self}(a_i; I) + \delta a_i + \gamma \left| a_i - E[\overline{a}_i^*|\varepsilon_i, I, F(\cdot)]\right| + \varepsilon_i^{a_i} + \upsilon_i^{a_i}(I) \tag{14}$$

Next, we can define a player's *ex post* expected payoff differential ($\Delta EV(\varepsilon_i, I, F(\cdot))$) between smoking and not smoking. It is this differential that governs a player's smoking action. That is,

$$\Delta EV(\varepsilon_i, I, F(\cdot)) \equiv EV(\varepsilon_i, I, F(\cdot); a_i = 1) - EV(\varepsilon_i, I, F(\cdot); a_i = 0) \tag{15}$$

A player's smoking decision is $a_i^* = 1[\Delta EV(\varepsilon_i, I, F(\cdot)) > 0]$.

Substituting equation 14 into equation 15, we have

$$\begin{aligned} \Delta EV(\varepsilon_i, I, F(\cdot)) &= \delta + \gamma + z_i[\beta_{self}(a_i = 1; I) - \beta_{self}(a_i = 0; I)] \\ &\quad -2\gamma E[\overline{a}_i^*|\varepsilon_i, I, F(\cdot)] \\ &\quad +\varepsilon_i^1 - \varepsilon_i^0 + \upsilon_i^1(I) - \upsilon_i^0(I). \end{aligned} \tag{16}$$

The "constant" term $(\delta + \gamma)$ in the equation above is collinear with the intercept (i.e., the β coefficient on the vector of ones contained in z). Thus, the biological smoking disutility parameter (δ) cannot be identified separately from the intercept. Allowing these parameters $(\delta + \gamma)$ to be collapsed into the β coefficient on the vector of ones in z and defining additional notations as

$$\begin{aligned} \Delta\beta_{self} &\equiv \beta_{self}(a_i = 1; I) - \beta_{self}(a_i = 0; I) \\ \Delta\varepsilon_i &\equiv \varepsilon_i^1 - \varepsilon_i^0 \\ \Delta\upsilon_i(I) &\equiv \upsilon_i^1(I) - \upsilon_i^0(I) \\ \lambda &\equiv -2\gamma \ , \end{aligned}$$

we can rewrite the equation 16 as

$$\Delta EV(\varepsilon_i, I, F(\cdot)) = z_i[\Delta\beta_{self}] + \lambda E[\overline{a}_i^*|\varepsilon_i, I, F(\cdot)] + \Delta\varepsilon_i + \Delta\upsilon_i(I). \tag{17}$$

The econometrician does not fully observe all public information (I); that is, $I = [I^o \ I^u]$ where superscript 'o' means observed by the econometrician and superscript 'u' means unobserved by the econometrician. Since z_i is a component of I, then, correspondingly, $z_i = [z^o \ z^u]$. Assume that $z_i \Delta\beta_{self}$ can be decomposed as $z_i \Delta\beta_{self} = z_i^o \Delta\beta_{self}^o + z_i^u \Delta\beta_{self}^u$. Then $\Delta EV(\varepsilon_i, I, F(\cdot))$ can be rewritten as

$$\Delta EV(\varepsilon_i, I, F(\cdot)) = z_i^o \Delta\beta_{self}^o + \lambda E[\overline{a}_i^* | \varepsilon_i, I, F(\cdot)] + \Delta\varepsilon_i + \Delta\upsilon_i(I) + z_i^u \Delta\beta_{self}^u \ . \tag{18}$$

In this game, the common prior $F(\cdot)$, some public information I^u, and private shocks ε_i are unobserved by the econometrician and hence, the private-shock-specific *ex post* expected equilibrium peer norm (i.e., $E[\overline{a}_i^* | \varepsilon_i, I, F(\cdot)]$) cannot be used in equation 18) as a predictor of smoking initiation. In order to proceed with estimation, he can, alternatively, use the *ex ante* equilibrium peer norm to proxy $E[\overline{a}_i^* | \varepsilon_i, I, F(\cdot)]$ because he can recover the *ex ante* equilibrium peer norm from the data. Player i's *ex ante* equilibrium peer norm is

$$E[\overline{a}_i^* | I, F(\cdot)] = \int E[\overline{a}_i^* | \varepsilon_i, I, F(\cdot)] dF(\varepsilon_i). \tag{19}$$

Define

$$\tau(\varepsilon_i, I, F(\cdot)) = \lambda [E[\overline{a}_i^* | \varepsilon_i, I, F(\cdot)] - E[\overline{a}_i^* | I, F(\cdot)]]. \tag{20}$$

Substituting equation 20 into equation 18, we get

$$\begin{aligned} \Delta EV(\varepsilon_i, I, F(\cdot)) \quad &= \quad z_i^o \Delta\beta_{self}^o + \lambda E[\overline{a}_i^* | I, F(\cdot)] \\ &\quad + \Delta\varepsilon_i + \Delta\upsilon_i(I) + z_i^u \Delta\beta_{self}^u + \tau(\varepsilon_i, I, F(\cdot)). \end{aligned} \tag{21}$$

An identification concern regarding the peer effect (λ) arises if correlation exists between $E[\overline{a}_i^* | I, F(\cdot)]$ and either of the four unobserved components: $\Delta\varepsilon_i$, $\Delta\upsilon_i(I)$, $z_i^u \Delta\beta_{self}^u$, and $\tau(\varepsilon_i, I, F(\cdot))$. We note $z_i^u \Delta\beta_{self}^u$ is innocuous because $z_i^o \perp z_i^u$. Meanwhile, we can draw the following three conclusions:

1. Unobserved heterogeneity is expected to be correlated across players for two reasons. First, $E[\Delta\upsilon_i(I), \Delta\upsilon_j(I)] \neq 0$ and $E[\tau(\varepsilon_i, I, F(\cdot)], \tau(\varepsilon_j, I, F(\cdot))) \neq 0$ in general; second, if private shocks are correlated then $E[\Delta\varepsilon_i, \Delta\varepsilon_j) \neq 0$;

2. Correlation between $E[\overline{a}_i^* | I, F(\cdot)]$ and $\Delta\upsilon_i(I)$ is expected because both of them are functions of public information (I);

3. Correlation between $E[\overline{a}_i^*|I, F(\cdot)]$ and $\tau(\varepsilon_i, I, F(\cdot))$ is expected because both of them are functions of public information (I) and the common prior ($F(\cdot)$).

Our interest is to examine the effects of peer behavior on smoking initiation among schoolmates. It is reasonable to assume that school-level unobservables explain some of the correlation in unobserved heterogeneity across students attending the same school. Accordingly, we decompose a student's unobservables into two components: school-level unobservables (μ_s) and individual unobservables ($u_{i,s}$). The final empirical specification is:

$$\Delta EV_{i,s} = \alpha' X_{i,s} + \lambda E[\overline{a}_{i,s}^*|I, F(\cdot)] + \mu_s + u_{i,s} \tag{22}$$

and $a_{i,s}^* = 1[\Delta EV_{i,s} > 0]$. The notation in equation 22 is summarized by

- s indexes schools;
- i indexes students where $i \in N_s = \{1, 2, ..., n_s\}$;
- $X_{i,s}$ is a vector of observed exogenous characteristics of student i in school s ;
- α is the parameter vector corresponding to $X_{i,s}$;
- λ is the peer effect parameter;
- $E[\overline{a}_{i,s}^*|I, F(\cdot)] \in [0, 1]$ is student i's *ex ante* equilibrium peer smoking norm and is a function of public information and the functional form of the common prior;
- μ_s is the school-level unobserved heterogeneity and allows for students' unobservables to be correlated at the school level;
- $u_{i,s}$ is the individual unobserved heterogeneity;
- $a_{i,s}^*$ is the observed equilibrium smoking action;
- $E[\mu_s, \mu_{s'}] = 0$, $E[u_{i,s}, u_{j,s}] = 0$, $E[u_{i,s}, u_{j,s'}] = 0$, and $E[\mu_s, u_{i,s}] = 0$.

4.2 Estimation

In this subsection, we discuss the two stages of estimation of the empirical specification in equation 22. The identification concern is that the peer norm, $E[\overline{a}^*_{i,s}|I, F(\cdot)]$, may be correlated with (school-level and/or individual) unobserved heterogeneity. To address this concern, in the first stage of estimation, we construct the *ex ante* equilibrium peer norm ($\widehat{E}[\overline{a}^*_{i,s}|I, F(\cdot)]$) using instrumental variables that affect the components of an agent's peer smoking norm (i.e., the friendship probabilities and the potential friends' smoking probabilities) and are uncorrelated with the agent's unobserved heterogeneity ($u_{i,s}$). We also use school fixed effects to further control for the potential correlation between the instrumented peer norm and school-level unobserved heterogeneity (μ_s) in the second stage of estimation.

Stage One: Construction of the endogenous peer smoking norm

Before further discussion, let us first examine the functional form of the *ex ante* equilibrium peer smoking norm. Substituting equation 2 into equation 19 yields the reduced-form representation of $E[\overline{a}^*_{i,s}|I, F(\cdot)]$ defined in the footnote below. Apparently, exogenous inputs (I and $F(\cdot)$) of the game affect the *ex ante* equilibrium peer smoking norm nonlinearly and interactively.[17]

[17] To see this through an example, let us think of a 2-agent smoking Bayes game with peer effects. For simplicity, let us assume 1) there is no peer selection decision in this game (two agents are assigned as friends by nature), 2) the latent smoking behavioral outcome process can be specified additively and linearly based on observed public information, and 3) unobserved heterogeneity specified in the behavioral specification of this game follows an *i.i.d.* logistic distribution across agents. Consequently, this game has the following behavioral form

$$\begin{cases} \pi_1 = \frac{\exp(\alpha X_1 + \lambda\pi_2)}{1+\exp(\alpha X_1 + \lambda\pi_2)} \\ \pi_2 = \frac{\exp(\alpha X_2 + \lambda\pi_1)}{1+\exp(\alpha X_2 + \lambda\pi_1)} \end{cases}$$

where π_1 and π_2 are equilibrium smoking probabilities for agent 1 and agent 2, respectively. X_1 and X_2 are public information which have direct impact on the first agent's smoking decision and the second agent's smoking decision, respectively. λ is peer effect behavioral parameter. Solving for $\pi_1$and π_2 in terms of exogenous characteristics, we obtain

$$\begin{cases} \pi_1 = R(X_1, X_2) \\ \pi_2 = R(X_2, X_1) \end{cases}$$

where $R(\cdot)$ is the reduced-form representation. In this example, $R(\cdot)$ cannot be solved in a closed form. We note, X_2 enters the first agent's reduced-form solution and X_1 enters the second agent's reduced-form solution. In addition, X_1 and X_2 enter $R(\cdot)$ nonlinearly and interactively even if they affect the latent outcome in a linear and additive pattern in the behavioral outcome process. We further note that the peer smoking norm is affected by *ex post* equilibrium friendships ($\zeta^*(\varepsilon_i, \varepsilon_j, I)$) and *ex post* equilibrium potential

Because the econometrician does not observe private shocks (ε), he is unable to recover the *ex ante* equilibrium peer smoking norm based on the two *ex post* quantities: the friendship network of student i ($\zeta^*(\varepsilon_i, \varepsilon_j, I)$) and this student's friend's smoking behavior ($a_j^*(\varepsilon_j, I)$). To proceed with estimation, the econometrician can construct an estimate of a student's (say, student i attends school s) *ex ante* equilibrium peer smoking norm based on the student's *ex ante* equilibrium friendship probabilities ($p(\zeta_i^{j*}(I; F(\cdot)) = 1)$, $\forall j \in N_s \backslash \{i\}$) and her schoolmates' *ex ante* smoking probabilities ($p(a_j^*(I; F(\cdot)) = 1)$, $\forall j \in N_s \backslash \{i\}$) because both of them can be recovered from the data. As such, we predict a student's *ex ante* equilibrium peer smoking norm in three sequential steps:[18]

- In a school (say school s), we first determine all students' *ex ante* equilibrium smoking probabilities in order to obtain predictions of the smoking behavior of one's potential friends ($\widehat{p}(a_i^*(I; F(\cdot)) = 1)$, $\forall i \in N_s$).
- Next, we determine all students' *ex ante* equilibrium friendship probabilities so that we can predict the probability of a student being friends with any other student in her school ($\widehat{p}(\zeta_i^{j*}(I; F(\cdot)) = 1)_{i=1}^{n_s}$, $\forall j \in N_s \backslash \{i\}$ and $\forall i \in N_s$).

peers' smoking actions ($a_j^*(\varepsilon_j, I)$).

$$\begin{aligned} E[\overline{a}^*_{i,s} | I, F(\cdot)] &= \int_{\varepsilon_i} \int_{\varepsilon_{-i}|\varepsilon_i} \frac{\sum_{j \in N_s \backslash \{i\}} \zeta(s_i^*(\varepsilon_i, I), s_j^*(\varepsilon_j, I); I) a_j^*(\varepsilon_j, I)}{\sum_{j \in N_s \backslash \{i\}} \zeta(s_i^*(\varepsilon_i, I), s_j^*(\varepsilon_{-i}, I); I)} dF(\varepsilon_{-i}|\varepsilon_i) dF(\varepsilon_i) \\ &= \int_{\varepsilon} \frac{\sum_{j \in N_s \backslash \{i\}} \zeta(s_i^*(\varepsilon_i, I), s_j^*(\varepsilon_j, I); I) a_j^*(\varepsilon_j, I)}{\sum_{j \in N_s \backslash \{i\}} \zeta(s_i^*(\varepsilon_i, I), s_j^*(\varepsilon_{-i}, I); I)} dF(\varepsilon) \\ &= \int_{\varepsilon} \frac{\sum_{j \in N_s \backslash \{i\}} \zeta^*(\varepsilon_i, \varepsilon_j, I) a_j^*(\varepsilon_j, I)}{\sum_{j \in N_s \backslash \{i\}} \zeta^*(\varepsilon_i, \varepsilon_j, I)} dF(\varepsilon) \end{aligned} \tag{23}$$

[18] Instrumenting the *ex ante* equilibrium peer norms in three steps increases the flexibility and, in turn, the goodness of fit. In practice, it is difficult to instrument the *ex ante* equilibrium peer norms in one step because it is difficult to appropriately specify the many interaction terms. It turns out that even when we instrument the *ex ante* equilibrium peer norms in three steps, the interaction terms involved are fairly complicated.

- In the third step, we construct a student's predicted (and instrumented) *ex ante* equilibrium peer smoking norm ($\widehat{E}[\overline{a}^*_{i,s}|I, F(\cdot)]$) from her schoolmates' *ex ante* equilibrium smoking probabilities ($\widehat{p}(a^*_j(I; F(\cdot)) = 1)$, $\forall j \in N_s \backslash \{i\}$) and her *ex ante* friendship probabilities ($\widehat{p}(\zeta_i^{j*}(I; F(\cdot)) = 1)$, $\forall j \in N_s \backslash \{i\}$). That is,

$$\widehat{E}[\overline{a}^*_{i,s}|I, F(\cdot)] = \frac{\sum_{j \in N_s \backslash \{i\}} \left[\widehat{p}(\zeta_i^{j*}(I; F(\cdot)) = 1) \times \widehat{p}(a^*_j(I; F(\cdot)) = 1)\right]}{\sum_{j \in N_s \backslash \{i\}} \widehat{p}(\zeta_i^{j*}(I; F(\cdot)) = 1)} \quad (24)$$
$$\forall j \in N_s \backslash \{i\} \text{ and } \forall i \in N_s$$

We are more specific about each step below.

Step 1: Estimation of potential friends' ex ante equilibrium smoking probabilities. To obtain instrumented *ex ante* equilibrium individual smoking probabilities in the first step, we use a non-parametric bagged tree classifier which is, in essence, an ensemble of 100 fully-grown tree classifiers created from 100 bootstrapped samples (Breiman, 1996). The bagged tree classifier includes the following exogenous characteristics (and their interactions) as explanatory variables: age, grade level, gender, race, family income, parental education, parental smoking behavior, and state cigarette tax.[19] We note that these variables directly affect a student's schoolmates' *ex ante* smoking probabilities ($p(a^*_j(I; F(\cdot)) = 1)$), and in turn, the student's *ex ante* equilibrium peer smoking norm but are assumed to be uncorrelated with the student's individual unobserved heterogeneity (i.e., $u_{i,s}$ in equation 22).[20] Figure 1 provides an example of one branch on one tree using the bootstrapped aggregating method. The large number of interactions of exogenous variables is evident in this example.

Step 2: Estimation of ex ante equilibrium friendship probabilities. To implement the second step, we use a flexibly specified logit model to estimate *ex ante* equilibrium friendship

[19]We found that a logit model does not perform well for several reasons. First, the proportion of smokers (or smoking initiators) is low and a logit estimation model with an unbalanced outcome generally provides poor predictions in the rare states and good predictions in the prevalent states. Second, a quasi-perfect separation problem (e.g., perfect prediction) emerges with the logit estimation model, particularly when interactions of the covariates are included. These covariates explain outcomes so well that variations in binary smoking actions within categories defined by those interaction terms disappear. These concerns further motivate us to use the bagged tree classifier.

[20]The predicted smoking probability for a student may still be correlated with school-level unobserved heterogeneity (μ_s). We return to this concern below.

probabilities ($p(\zeta_i^{j*}(I; F(\cdot)) = 1)$). Three reasons motivate us to adopt a logit model rather than a bagged tree classifier in this second step. First, we have a large number of pairwise observations ($2,987,761$ pairs) where each student is paired with every student in her school; therefore, the quasi-perfect separation problem that plagues the logit model when the sample size is small ($N = 17,844$ in the first step) disappears. Second, conformity theory provides guidance for specifying a logit model, so this step is less "data mining" oriented than the first step (Bernheim, 1994). Third, the large number of pairwise observations makes the computational cost in a bagged tree classifier overly expensive. To be specific, we use a student i's directional deviations from a generic schoolmate j in age, grade level, gender, race, and family income as explanatory variables for student i's friendship with the generic schoolmate j.[21] We also use the instrumented *ex ante* equilibrium smoking probabilities of potential friends from the first step to create directional deviations in smoking that may explain existence of friendships. It is worth mentioning that because the instrumented *ex ante* equilibrium smoking probabilities of the potential friends are functions of exogenous characteristics such as whether or not that students' parents smoke, a student's directional smoking deviations using the instrumented smoking probabilities should be uncorrelated with her own error term in the friendship logit model. Besides these directional deviations, school size is also used as an explanatory variable in the friendship logit model due to the following two considerations. First, as school size increases, the probability that an arbitrary pair of schoolmates run into each other, and become friends, decreases. Second, school size may affect competition (e.g., competition for teachers' attention) among teens, and hence, friendship formation.

Step 3: Construction of ex ante equilibrium peer smoking norm. After modeling *ex ante* equilibrium individual smoking probabilities and *ex ante* equilibrium friendship probabilities, we then construct the instrumented *ex ante* equilibrium peer smoking norm ($\widehat{E}[\overline{a}^*_{i,s}|I, F(\cdot)]$), according to equation 24, as the third step.

[21] To illustrate directional deviations, let us consider a directional deviation in a categorical variable, say, gender, between two agents, Tom and Mary. Tom's directional deviation from Mary in gender is "male-to-female" and Mary's directional deviation from Tom in gender is "female-to-male". Friendship between two agents may not be symmetric; that is, "male-to-female" is not equal to "female-to-male".

Stage two: Estimation of smoking initiation probability

Substituting the instrumented peer norm (equation 24) into the final smoking initiation specification (equation 22), we estimate the following logit probability

$$p(a^*_{i,s} = 1) = \alpha' X_{i,s} + \lambda \widehat{E}[\overline{a}^*_{i,s} | I, F(\cdot)] + \mu_s + u_{i,s} \ . \tag{25}$$

As mentioned above, one's schoolmates' parental characteristics and school size are used to explain her *ex ante* equilibrium peer smoking norm. Schoolmates' parents also choose schools for their children. Consequently, the instrumented *ex ante* equilibrium peer norms could be correlated with school-level unobserved heterogeneity (μ_s). To address this concern, we control for school-level unobserved heterogeneity by including school fixed effects in the second stage of estimation.

5 Description of Data

The theoretical model indicates that an agent's friendships directly affect her peer norm. Therefore, to estimate peer effects, the econometrician should observe an agent's friends in the data.[22] Regarding this data requirement, the National Longitudinal Study of Adolescent Health (Add Health) is a suitable dataset because it provides detailed information on some respondents' friendships within their attending schools. Currently, Add Health has four waves of in-home surveys fielded in 1995, 1996, 2001, and 2004. We use the wave I in-home survey data collected between April 1995 and December 1995 because the subsequent waves lack state cigarette tax information from the survey year and the state identifiers are not released (even in the restricted-use version of the data). The wave I in-home survey sample contains $20,745$ nationally representative 7^{th}-12^{th} graders from 145 schools nationwide.

Among the 145 schools, there are 16 schools in which all students are included in the survey sample and almost every student is asked to nominate five schoolmate friends of each

[22] Observing agents' peer selection actions is unnecessary if the econometrician does not intend to estimate the peer selection action and smoking action jointly. In fact, as explained before, even if the econometrician observes agents' peer selection actions for a couple of periods, estimating peer selection and smoking jointly is typically infeasible due to the high dimensionality of a peer selection action.

If the econometrician cannot observe the friendship network in the data, then the econometrician has to subjectively assign friends to an agent. For example, the econometrician may assign all of a student's classmates as the student's peers. The drawback of doing so is obvious because, in general, a student is not a friend of all of her classmates.

gender.[23] Among the remaining schools, respondents are randomly selected from a gender-grade stratum within each school and some of the respondents are also asked to nominate five schoolmate friends of each gender.[24] Table 1 details derivation of the samples used in estimation. After deleting respondents with missing values, we obtain $17,844$ students (Sample 4 in Table 1). Among the 17,844 students, there are $5,774$ students (Sample 4.1) who were asked to nominate five schoolmate friends of each gender.[25] To model friendship formation, we create $2,987,761$ pairwise observations from the $5,774$ students and their corresponding schoolmates. Consider a created pairwise observation corresponding to a generic student i in the $5,774$ students and one of her schoolmates (say, schoolmate j). The pairwise observation records whether or not i regards j as a friend and student i's directional deviations from schoolmate j of their exogenous characteristics that affect their peer selection actions.

Regarding questions related to smoking behavior, Add Health asked respondents "During the past 30 days, on how many days did you smoke cigarettes?". Based on this question, we dichotomize the smoking decision. We classify respondents who reported smoking cigarettes one or more days during the last 30 days prior to the wave I in-home survey date as smokers; otherwise, they are nonsmokers. The literature suggests that lagged smoking behavior affects an agent's current smoking decision through nicotine tolerance, reinforcement, and withdrawal that alter current period utility (Becker and Murphy, 1988; Bullock et al., 1994; Stolerman and Jarvis, 1995). Without knowledge of a respondent's state of

[23]Due to administrative error, a small fraction (less than five percent) of students in those schools were asked to nominate only one friend of each gender.

[24]Even if a randomly selected student is asked to nominate all her schoolmate friends, the recorded (within-school) friendship network contains missing information. It can be shown that in a very large school, if the sampling rate is r, then the percentage of missing friendship network information is close to $1 - r^2$.

Let the $N \times N$ zero-diagonal matrix, ζ_{pop}, denote the true friendship network among the N individuals in the population of interest. We note, only $N^2 - N$ non-diagonal elements in ζ_{pop} matter. The random sample contains $N \times r$ respondents. Let ζ_{sample} denote the recorded friendship network in the survey sample. Then, $[(N \times r)^2 - (N \times r)]$ non-diagonal elements in ζ_{sample} matter.

Hence, the number of missing (friendship network) elements in the survey is $N^2 - N - [(N \times r)^2 - (N \times r)] = N^2(1 - r^2) - N(1 - r)$ and the corresponding rate of missing information, $R_{missing}$, is then $R_{missing} = \frac{N^2(1-r^2)-N(1-r)}{(N^2-N)} = \frac{N(1-r^2)-(1-r)}{(N-1)}$.

As N approaches infinity, we have $\lim_{N \to +\infty} R_{missing} = \lim_{N \to +\infty} \frac{N(1-r^2)-(1-r)}{(N-1)} = 1 - r^2$.

[25]Ideally, the econometrician would like to have a dataset in which she can track down all of a respondent's friends (not only schoolmate friends but also non-schoolmate friends). However, no such data set exists to our knowledge.

residence, we are unable to find an instrumental variable to control for the endogeneity of lagged smoking behavior (e.g., lagged cigarette prices/taxes). Realizing this data limitation, we restrict the analysis sample in the second stage to respondents who had no regular smoking history before the wave I in-home survey. More specifically, respondents who answered "yes" to "Have you ever smoked cigarettes regularly, that is, at least 1 cigarette every day for 30 days?" in the wave 1 in-home survey are excluded from the analysis of smoking behavior. Hence, this selection focuses our analysis on the smoking initiation decision of 13,924 students (Sample 4.3 in Table 1). The full sample of respondents from each school with complete data (Sample 4 of Table 1) is used in the analysis of friendship formation, potential peers' smoking probabilities, and construction of the peer smoking norm.

Add Health provides a rich set of respondents' characteristics. For our purposes, the following measurements are of particular interest due to their potential impact on smoking. They are age, grade level, gender, race, family income, parental education, parental smoking behavior, and religious importance. Table 2 lists the summary statistics for respondents in Sample 4, 4.1, and 4.3 of Table 1. Almost 12% of the 13,924 students with no smoking history initiated smoking in wave I. Individuals in Sample 4.3 have survived to their grade level without having smoked. To the extent that exposure (reflected by years at risk for smoking) explains smoking decisions, we would expect a higher proportion of younger kids to populate the smaller, selected sample of those with no smoking history. While the distribution of grade levels among those with no smoking history appears similar to the grade level distribution of the full sample, the proportion of 7th graders is 11% higher and of 12th graders is 4% lower. The "non-smoking survivors" are also more likely to be black, to have parents that do not smoke, and to rate religion as very important in their lives.

Figure 2 (created from Sample 4 in Table 1) presents the relationship between individual smoking and peer smoking. As the number of smoker friends among the three closest friends increases from 0 to 3, the individual smoking rate increases from about 10% to about 70%, indicating a strong positive correlation between personal smoking and peer smoking. Figure 3 (created from Sample 4.1 in Table 1) presents the distribution of the number of schoolmate friends among those asked to specify ten friends. 74% of respondents (4,268 out of 5,774) asked to nominate up to ten friends listed at least one schoolmate friend. On average, a respondent had 3.4 such friends.

6 Results

Here we discuss empirical results associated with both the first and second stage of estimation. Construction of the endogenous peer smoking norm requires estimation of the *ex ante* equilibrium smoking probabilities and friendship probabilities. Once the peer norm is calculated, as a function of predicted smoking and friendship probabilities, we discuss estimation of the structural smoking initiation equation (equation 25). We conclude with a discussion of the effect of peer influence on a simulated policy change.

6.1 Peers' Smoking Probabilities

The distribution of instrumented *ex ante* equilibrium smoking probabilities from the bagged tree classifier is shown in Figure 4. The median of the instrumented smoking probabilities among smokers (0.7024) is much larger than that among nonsmokers (0.0268). This is one indication that the bagged tree classifier does a good job differentiating between smokers and nonsmokers. We further examine the performance of the bagged (n=100) tree classifier in comparison with a flexibly specified logit model based on two widely accepted diagnostic tests: the receiver operating characteristic (ROC) curve and the reliability diagram (Jolliffe and Stephenson 2003).[26] The ROC curve provides a graphical representation of the trade-off between the percentage of correctly predicted observed smokers and the percentage of incorrectly predicted non-smokers. The larger the area under the curve the more accurate is the model. As shown in Figure 5, the area under the ROC curve based on smoking probabilities obtained from the bagged tree classifier is much larger than that based on smoking probabilities obtained from the flexibly specified logit model. The bagged tree classifier is much less likely to commit type I or type II prediction errors than the logit model.

[26]The logit model includes the same exogenous explanatory variables as in the bagged tree, along with second order interactions of these variables.

Figure 6 compares the bagged tree classifier and the flexibly specified logit model in terms of reliability.[27] Points generated from the bagged tree classifier (right panel of Figure 6) are clustered around the 45 degree line over the entire range of estimated smoking probabilities. For the logit model (left panel), an apparent problem is that the maximum estimated smoking probability is less than 0.75. This indicates that the flexibly specified logit model underestimates the individual smoking probability of those most likely to smoke. As mentioned previously, the logit model has difficulty predicting less prevalent outcomes in models of unbalanced behavior.[28]

6.2 Friendship Probabilities

Table 3 presents coefficient estimates of the directional friendship logit model. Conformities in age, grade level, gender and race contribute to friendship significantly. Income conformity does not explain friendships except among teens from low income families. Individuals with different smoking propensities are less likely to be friends. This finding is consistent with the theoretical implication that smoking actions and peer selection actions are interdependent. Increases in school size reduce the probability of being friends with someone in your school. This finding is consistent with the intuition that as school size increases a student's probability of making friends with an arbitrary schoolmate falls holding other variables constant. Figure 7 presents the reliability diagram of the estimated directional friendship probabilities. Points representing predicted friendship probabilities and observed friendship percentages are closely clustered along the 45 degree line over the whole range indicating that the logit model performs well.

[27] For a point (x, y) in Figure 6, y is the average observed smoking rate of students falling into an estimated smoking probability bin and x is the center of the estimated smoking probability bin. If the model predicts observed behavior well, we should expect x to be close to y. Therefore, dots should cluster closely to the 45 degree line.

[28] Additionally, the model is estimated on sample 4 (see Table 1) which includes many students with a previous smoking history. To the extent that past behavior explains current behavior, this omitted variable, if included, might produce smoking probabilities closer to one for smokers.

6.3 Structural Smoking Initiation Probabilities

With the instrumented *ex ante* equilibrium smoking probabilities and the instrumented *ex ante* equilibrium directional friendship probabilities in hand, we then construct the instrumented *ex ante* equilibrium peer smoking norms based on equation 24. Figure 8 presents the scatter plot of the instrumented *ex ante* equilibrium individual smoking probabilities and the constructed instrumented *ex ante* equilibrium peer norms; the thick line is the trend curve fitted through a polynomial (up to the 9^{th} order) regression. As we expected, the instrumented *ex ante* equilibrium individual smoking probabilities are positively correlated with the instrumented *ex ante* equilibrium peer smoking norms.

For comparison purposes, we discuss estimation results from six different specifications of the smoking initiation probability (equation 25). Each specification differently controls for the peer smoking norm. Specification 1 (No peer effects) assumes that peer influence does not exist. In all specifications we control for school-level unobserved heterogeneity by using school fixed effects. Specification 2 and 3 model peer influence by using school norms as explanatory variables (Norton, 1998 and Lundborg, 2006). If the econometrician does not have detailed friendship network data but knows the average smoking behavior of the respondents' school membership, then he may run Specification 2 or 3 to capture peer influence. It is worth mentioning that both Specifications 2 and 3 implicitly assume that everyone is everyone else's friend within a school and such a friendship composition is exogenously assigned but not chosen by students. Specification 2 (School norm) uses observed smoking actions to construct school-level peer norms.[29] Specification 3 (IV school norm) uses the instrumented *ex ante* equilibrium smoking probabilities.

Specifications 4, 5 and 6 exploit the detailed peer composition information in Add Health in modeling peer influence. The specifications differ by the effort to correct for bias caused by the endogenous peer norms. In Specification 4 (Peer norm), the peer norms are constructed based on observed friendships and observed smoking actions. In Specification 5 (IV peer norm), the peer norms are constructed based on observed friendships

[29] We are able to construct such school-level peer norms using Add Health because it provides smoking behavior of all students in a school. However, those data are typically unavailable. Rather, researchers have used the reported average school-level smoking rate to represent the peer smoking norm.

and instrumented *ex ante* equilibrium smoking probabilities of those peers. Therefore, both Specifications 4 and 5 fail to control for endogenous friendships. Specification 5 differs from Specification 4 in that it controls for the endogenous smoking action while the latter does not. Some people did not nominate any schoolmate as a friend, therefore, in Specifications 4 and 5, only 4,268 out of the 13,924 observations can be used in analysis because the peer norms are undefined for those students who did not nominate any schoolmate as a friend. Specification 6 (Preferred) is the preferred model because it controls for both the endogenous friendships and the endogenous smoking actions. Because we use predictions of potential peers' smoking behavior and probabilities of friendship to construct the preferred peer norm, we can use the larger sample of 13,924 students (Sample 4.1) to estimate the structural model of smoking initiation.

The marginal effects of all variables in each specification are presented in Tables 4 and 5, while we provide coefficient estimates in Appendix Table A1. A marginal effect is calculated as the average of observation-wise marginal effects on the smoking initiation probability in the analysis sample. Marginal effects are measured in percentages. For example, the overall marginal peer effect is 1.07 (0.09) in Specification 6 of Table 5; this means that a one percentage point increase in the peer norm (or peers' average smoking rate) causes a 1.07 percentage point increase in an individual's smoking initiation rate (with a standard error of 0.09).[30] When interactions were included in the specification, we also report the marginal effect by the interacted variable.[31] In these tables we report the marginal effect and the predicted smoking initiation rates at which marginal effects are evaluated. For example, in Specification 1 of Table 4, the overall marginal tax effect is evaluated at 12.10(0.28); this means the smoking initiation rate is expected to drop to 12.10-3.86=8.24 percent after a 10-cent increase in the cigarette tax.[32]

[30]We calculate bootstrapped standard errors. The presence of interaction terms in each of the six specifications makes it inconvenient to write down the analytical expression of a marginal effect, so we do not use the delta method to calculate standard errors. Alternatively, we randomly sample (n=5,000) a set of model coefficients from their estimated joint distribution (i.e., variance/covariance matrix) and find the standard deviation of the marginal effects from the 5,000 different draws.

[31]When introducing a unit change in a variable that interacts with another variable, we add a unit change to the variable and a corresponding change to all interaction terms that involve the variable.

[32]For continuous variables such as tax and income, the marginal effects are evaluated at the predicted smoking initiation rate of the analysis sample. However, the marginal effect of a non-comparison category of a categorical variable is evaluated at the predicted smoking initiation rate of a hypothetical sample in

An obvious pattern in Tables 4 and 5 is that the grade level effect differs dramatically across different model specifications. Recall that all students in the analysis samples (either N=13,924 or N=4,268) report no smoking history. Therefore, a respondent's grade level is, in fact, perfectly collinear with the respondent's left censored survival time when we have smoking initiation as the event of interest. The starting point of the left censored survival time can be arbitrarily set to a time prior to entering 7th grade depending on research convenience (e.g., one might assume no children experiment with smoking before 3rd grade). A person's left censored survival time is a function of the person's (observed and unobserved) smoking initiation deterrents in the past. Hence, it is reasonable to infer that grade level is positively correlated with the strength of the unobserved (by the econometrician) smoking initiation deterrents in the past. Those unobserved deterrents in the past may be correlated with unobserved smoking initiation deterrents in the present. As such, coefficient estimates on grade levels should be interpreted with the following two cautions. First, they reflect the effects of past unobserved smoking initiation deterrents on smoking initiation in the present. Second, they are biased if serial correlation exists between past unobserved smoking deterrents and present unobserved smoking deterrents. The better we are able to control for the time-invariant deterrents on smoking initiation, the less severe is the bias.

As explained above, grade level is perfectly collinear with the left censored survival time. Hence, we should expect that a marginal change in grade level from 7th grade to a higher grade level causes a larger percentage drop in the smoking initiation rate, holding other covariates constant, because it is reasonable to believe that teens who managed to abstain from smoking a longer period in the past probably are less likely to initiate smoking in the present. In Table 5 we see that estimation results from Specifications 4, 5 and 6 are consistent with such an expectation. In each of these three specifications, as grade level increases (from grade 7), the predicted smoking initiation rate drops significantly (from 26.99 percent in grade 7 to 26.99 - 18.17 = 6.82 percent in grade 12). However, estimation results from Specifications 1, 2 and 3 are inconsistent with such an expected pattern. For example, in

which all observations are assumed to be a member of the comparison (or base) category of the categorical variable. For example, in Specification 6, the marginal effect of being a 12th grader without a smoking history is evaluated at the predicted smoking initiation rate of the hypothetical sample in which every respondent is a 7th grader without a smoking history (26.99 percent).

Specification 1 (No peer effects), changing all teens from those who survive smoking initiation up to at least grade 7 into teens who survive smoking initiation up to at least grade 8 results in a smoking initiation rate increase of 3.30 (1.22) percentage points. On the other hand, our preferred model (Specification 6) suggests that the probability of beginning to smoke would fall by 8.27 (3.22) percentage points (from an average 26.99 percent initiation rate in 7th grade).

6.4 Peer Influence

Turning to the role of peer influence, the preferred model (Specification 6) shows that the peer effect is large and significant. A one percentage point increase in the peer norm causes the smoking initiation rate to increase by 1.07 (0.09) percentage points. All other specifications that use poorly constructed norms as explanatory variables underestimate peer influence. These models underestimate the peer effect 5- to 10-fold with marginal effects ranging from 0.11 to 0.19 percentage points. We find that peer influence using the average smoking behavior of all students in the school (with estimated marginals of 0.18 (0.03) in Specification 2 and 0.16 (0.03) in Specification 3) is much smaller than that using the (expected) average behavior of one's potential friends in Specification 6 (1.07 (0.09)). This larger effect is expected because it is reasonable to believe that a person's chosen friends influence the person more heavily than the person's schoolmates as a whole.

Comparing Specification 4, 5 and 6 in Table 5, we see that peer influence in the former two specifications (0.11 (0.01) and 0.19 (0.04), respectively) is much smaller than that in Specification 6 (1.07 (0.09)). Recall that it is a person's expected peer norm rather than the person's friends' average smoking actions that enters the econometric representation of the outcome process (equation 25). For a person who chooses a finite number of friends at a decision moment, these two quantities, in general, are different. More specifically, the person's expected peer norm is the expectation of her friends' average smoking actions. The smaller the number of friends a person has at equilibrium, the larger the variation in the person's friends average smoking actions. From an econometric perspective, using a person's friends' average smoking actions as an explanatory variable in Specification 4 and 5, in essence, adds measurement error to her expected peer norm.[33] Since an average respondent in Add Health has only 3.4 friends, the variances of the measurement errors among the 4,268 respondents in Specifications 4 and 5 are considerably large. This explains why, when compared with Specification 6, the peer effect estimate in Specification 4 and 5 are smaller.[34]

We can also evaluate the effect of other individual characteristics on smoking initiation. In a model with no peer effects (specification 1), black teens are less likely to initiate smoking (by 4.86 (0.81) percentage points) than white teens (with a smoking initiation rate of 13.76 (0.43) percent). However, in Specification 6 (Preferred) which takes peer influence into account, being black *increases* an individual's smoking initiation probability by 7.01 (1.55) percentage points. This finding suggests that peer influence and friendship sorting based on racial conformity explains why black teens have a lower observed smoking rate than white teens.[35] Having parents who have only a high school education increases a teen's smoking initiation probability by 1.26 (0.62) percentage points (in specification 6 with peer effects) compared to having parents who have a college education. Having parents who have less than a high school degree increases smoking initiation rates by 1.86 (0.85) percentage points. In all specifications, family income does not have a statistically significant effect. Teens of smokers are 2 to 3 percentage points more likely to initiate smoking than teens with no smoking parents. Teens who consider religion to be very important are almost 4 percentage points less likely to begin smoking.

6.5 The Social Multiplier Effect

If peers influence behavior, then policies that affect individual decisions are amplified through the additional peer effect. It should be cautioned, however, that the estimated results from a

[33]The variance of this measurement error is not only affected by the number of friends a person has (as just mentioned) but also by the distribution of equilibrium smoking probabilities among her friends. To see this point, consider a person whose friends' smoking probabilities are all 0s (or 1s), then the measurement error vanishes even if the person has only one friend.

[34]It is correct to note that the analysis sample in Specifications 4 and 5 is different than that used in Specification 6. However, we estimated Specification 6 on the smaller sample in order to rule out differences in sample composition as the reason for the large difference in the estimated peer influence. The marginal effect with the smaller sample is 1.89 (0.22), which is similarly greater than the estimated effects from the other specifications.

[35]In fact, a black teen is 95 percent less likely (than a white teen) to make friends with a white schoolmate and 11 percent more likely than a white teen to make friends with a schoolmate of the same race.

structural smoking equation that includes peer influence cannot be used by itself to provide meaningful policy predictions of the social multiplier effect. It is typically impossible for policy makers to exogenously assign peer norms to teens because teens choose their friends after all. Hence, as different smoking policies alter individual smoking propensities, we would expect friendship probabilities to also change. We discuss how to do policy simulation using both the smoking equation and the friendship formation equation below.

Our estimation results suggest that peer influence amplifies the direct tax deterrent effect on smoking initiation. In Specification 1 (No peer effects), we can see that the overall marginal tax effect is -3.86 (1.81). This implies that if the cigarette tax increases by 10 cents, then the smoking initiation rate for those individuals who had no smoking history will drop by 3.86 (1.81) percentage points. After controlling for peer influences, we see the overall marginal tax effect decreases to -3.68 (2.29). Though the magnitude of the mean marginal effect only drops 0.18 percentage points, the standard error increases quite a lot pulling down the statistical significance from the 3% to 10% level. Recall that the tax effect in Specification 1 is a combination of the direct tax effect and a social multiplier tax effect. Hence, the differential in estimated tax effect implies that the social multiplier effect is playing a role.

To be more precise, consider a policy intervention that changes observed public information (ΔI^o). Since a Nash equilibrium strategy is a function of public information (I), such a policy intervention affects students' actions with regard to both smoking and peer selection and, in turn, their peer smoking norms. In a generic school s with n_s students, the students' *ex ante* equilibrium peer smoking norms with the policy intervention is the n_s by 1 vector $E[\overline{a}^*_{i,s}|I + \Delta I^o, F(\cdot)]$ defined in Equation 23. We note the following three theoretical inferences. First, according to equation 24, the predicted (and instrumented) *ex ante* peer smoking norm vector is completely determined by a $n_s \times (n_s - 1)$ by 1 *ex ante* equilibrium directional friendship probability vector, $\hat{p}(\zeta_i^{j*}(I + \Delta I^o; F(\cdot)) = 1)$, and a n_s by 1 *ex ante* individual equilibrium smoking probability vector, $\hat{p}(a^*_{s,j}(I + \Delta I^o; F(\cdot)) = 1)$. Second, according the equation 25, the *ex ante* individual smoking initiation probability vector is partially determined by the predicted (and instrumented) *ex ante* peer smoking norm.

Third, the *ex ante* equilibrium friendship vector is partially determined by the *ex ante* individual equilibrium smoking probability vector that affects directional deviations in smoking probabilities among schoolmates. Hence, the *ex ante* peer smoking norm vector, the *ex ante* directional friendship probability vector, and the *ex ante* individual smoking probability vector are interdependent at a PSBNE. In other words, these three vectors should be consistent among themselves at a PSBNE. With this caveat in mind, in a policy simulation, we iterate an initial *ex ante* equilibrium individual smoking probability vector over the estimated smoking behavioral equation and the estimated friendship equation. In each iteration the three *ex ante* probability vectors are updated once. An equilibrium emerges when the iterated *ex ante* smoking probability vector converges uniformly across all schoolmates.

Prior to further discussion of the policy simulation, let us first examine how it is affected by data limitations. Due to the lack of knowledge of a respondent's state of residency in Add Health, we do not have instrumental variables (e.g., lagged cigarette price and lagged state cigarette tax) to explain lagged smoking behavior. This data limitation motivates us to only estimate a smoking behavioral equation for students without smoking histories. Regarding friendship formation, however, students without smoking histories are able to choose schoolmates with smoking histories as friends. This implies that in the policy simulation it is appropriate to allow for friendship formation between any two schoolmates regardless of their smoking histories. Thus, in operation we have to use the estimated behavioral smoking equation based on students without smoking histories to update *ex ante* individual smoking probabilities for students with smoking histories. Such a practice is flawed because it ignores the effect of a student's lagged smoking on her current smoking decisions. As a consequence, the simulation results presented below should be interpreted with caution.

We use two schools in our data to illustrate a simulated change in policy with regard to cigarette taxation. The first school (School A) has 55 students, a state cigarette tax of 28 cents, and 5.45 percent of the students smoke. The second school (School B) has 63 students, a tax of 75 cents, and 33.33 percent of the students smoke. In the policy simulation, we perturb cigarette taxes by adding an additional amount of tax (in cents) to the original state cigarette tax. For comparison purposes, we solve for equilibrium smoking initiation rates with and without peer effects at each tax level. In simulating equilibrium smoking initiation rates without peer effects, we set the coefficient estimates corresponding to grade-specific peer effects in the estimated smoking initiation equation (Tables 4 and 5) and the coefficient estimates corresponding to directional deviations in smoking dimension in the friendship equation (Table 3) to zero. Parameters used in simulation are randomly drawn from the estimated distribution of smoking initiation parameters and the estimated distribution of friendship equation parameters.

Figure 9 presents the simulation results in the two schools. Overall, with or without peer effects, as the tax increases the smoking rate decreases. Interestingly, in the presence of peer effects, tax increases in certain ranges may cause "abnormal" increases in the smoking initiation rate (e.g., tax increase equal to forty cents in the first school). We note if the friendship network does not change as taxes increase, then a tax increase should monotonically reduce the smoking initiation rate. These "abnormal" increases in the smoking rate reflect that a variation in the cigarette tax motivates agents to update their smoking decision as well as their friendships. Regarding the social multiplier effect, in both schools, we see that compared with having no peer effects, a given tax increase causes a dramatically larger drop in the smoking initiation rate in the presence of peer effects. This finding suggests that peer effects significantly amplify the tax deterrent effects. Meanwhile, in both schools the presence of peer effects significantly increases the smoking initiation rate over the entire range of the cigarette tax increases. Collectively, simulation results suggest that although peer influence significantly amplifies the cigarette tax deterrent effect, it mainly promotes teen smoking initiation. We also note that in a model with peer effects, the smoking initiation rate drops abruptly at certain tax thresholds (e.g., tax increase equal to seven cents in the first school). This indicates that at particular tax thresholds the social multiplier effect may be particularly large (e.g., herding behavior appears).

References

Bajari, P., H. Hong, J. Krainer, and D. Nekipelov (2006). Estimating Static Models of Strategic Interactions. Working paper.

Bauman, K. and S. Ennett (1994). Peer Influence on Adolescent Drug Use. *American Psychologist*, 49:820-822.

Bauman, K. and S. Ennett (1996). On the Importance of Peer Influence for Adolescent Drug Use: Commonly Neglected Considerations. *Addiction* 91(2):185-198.

Becker, G. (1992). Habits, Addictions, and Traditions. *Kyklos,* vol 45, Fasc 3:327-346.

Becker, G.and Murphy, K.M., (1988). A theory of Rational Addiction. *Journal of Political Economy,* 96:675-701.

Berheim, D. (1994). A theory of Conformity. *Journal of Political Economy*, vol. 102: 841-877

Breiman, L. (1996). Bagging Predictors. *Machine Learning* 26, pp 123-140.

Brock, W. and S. Durlauf (2001). Discrete Choice with Social Interactions. *Review of Economic Studies* 68, pp 235-260.

Bullock, A., K. Barke, and A. Schneider (1994). Nicotine Tolerance in Chromaffin Cell Cultures: Acute and Chronic Exposure to Smoking-Related Nicotine Doses. *Journal of Neurochemistry*62(5): 1863-1869

Case, A. and Katz, L.F., (1991). The Company You Keep: the Effects of Family and Neighborhood on Disadvantaged Youths. National Bureau of Economic Research Working Paper No. 3705.

Chaloupka. F. and K. Warner, (2000). The Economics of Smoking.

Cook, P. and M. Moore (2000). Alcohol. In: Culyer, A.J., Newhouse, J.P. (Eds.), *Handbook of Health Economics*, vol. 1b. Elsevier Science, Netherlands, pp. 1629–1673.

Cooley, J. (2007 March). Desegregation and the Achievement Gap: Do Diverse Peers Help?. *working paper*

Crane, J. (1991). The Epidemic Theory of Ghettos and Neighbourhood Effects on Dropping Out and Teenage Childbearing. *American Journal of Sociology,* 96:1226-1259.

DeCicca, P., D. Kenkel, and A. Mathios (1999). Putting Out the Fires: Will Higher Taxes Reduce Youth Smoking? Mimeo, Cornell University.

Eiserberg, D. (2004). Peer Effects for Adolescent Substance Use: Do They Really Exist? Working Paper.

Epple, D. and R. Romano (1998). Competition between Private and Public Schools, Vouchers, and Peer-Group Effects. *American Economic Review*, 88(1): 33-62

Evans, W., E. Oates, and R. Schwab (1992). Measuring Peer Group Effects: a Study of Teenage Behavior. *Journal of Political Economy*, 100: 966-991.

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

Graham, J., G. Marks, and W. Hansen (1991). Social Influence Processes Affecting Adolescent Substance Use. *Journal of Applied Psychology*, 76:291-298.

Gruber, J. and J. Zinman (2001). "Youth smoking in the U.S.: Evidence and implications" in Jonathan Gruber, ed., Risky Behavior Among Youths: An Economic Analysis, University of Chicago Press.

Gilleskie. D. and K. Strumpf (2005). The Behavioral Dynamics of Youth Smoking. *Journal of Human Resources*. 2005. 40(4): 822-866

Hoxby, C. (2000). Peer Effects in the Classroom: Learning from Gender and Race Variation. National Bureau of Economic Research Working Paper No. 7867.

Ioannides, M. and E. Zabel (2002). Interactions, Neighborhood Selection, and Housing Demand. Working Paper.

Jolliffe I. and Stephenson D (eds) (2003). Forecast Verification: A Practitioner's Guide in Atmospheric Science. Wiley.

Katz, F., R. Kling, and B. Jeffrey (2000). Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment. *Quarterly Journal of Economics,* 116: 607-654.

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

Krauth, B. (2006). Simulation-Based Estimation of Peer Effects. *Journal of Econometrics* 133(1): 243-271.

Kremer, M. and D. Levy (2003). Peer effects and Alcohol Use among College Students. National Bureau of Economic Research Working Paper No. 9876.

Lewitt, E., D. Coate, and M. Grossman (1981). The Effects of Government Regulation on Teenage Smoking. *The Journal of Law and Economics,* 24: 545-569.

Lundborg, P. (2006). Having the Wrong friends? Peer Effects in Adolescent Substance Use. *Journal of Health Economics* 25(2): 214-233.

Lynskey, M. et. al. (2003), Escalation of Drug Use in Early-Onset Cannabis: Users vs Co-twin Controls. *JAMA*, 289, 427-433.

Mackay, J. and Eriksen M. (2002). The Tobacco Atlas. World Health Organziation.

Manski, C. (1995). Identification Problems in the Social Sciences. Cambridge, Mass.: Harvard University Press.

Manski, C. (2000). Economic Analysis of Social Interactions. *Journal of Economic Perspectives* 14, 115-136.

Newcomb, M. and P. Bentler (1989). Substance Use and Abuse among Children and Teenagers. *American Psychologist*, 44: 242-248.

Norton, E., et al. (1998). Controlling For The Endogeneity of Peer Substance Use on Adolescent Alcohol and Tobacco Use. *Health Economics,* 7: 439-453.

Pesendorfer, M. and P.Schmidt-Dengler (2003). Identification and Estimation of Dynamic Games. National Bureau of Economic Research Working Paper No. 9726.

Sacerdote, B. (2001). Peer Effects with Random Assignment: Results for Dartmouth Roommates. *Quarterly Journal of Economics*, 116 (2): 681-704.

Scheinkman, J. (2006). Social Interactions. working paper

Stolerman, I. and M. Jarvis (1995). The scientific case that nicotine is addictive. *Psychopharmacology*, 117: 2–10.

US Dept of Health and Human Services. Tobacco Use Among US Racial/Ethnic Minority Groups–African-Americans, American Indians and Alaska Natives, Asian Americans and Pacific Islanders, and Hispanics: A Report of the Surgeon General. Atlanta, Ga: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office of Smoking and Health; 1998.

Figure 1: Example of One Branch on One Tree using the Bagged Tree Classifier

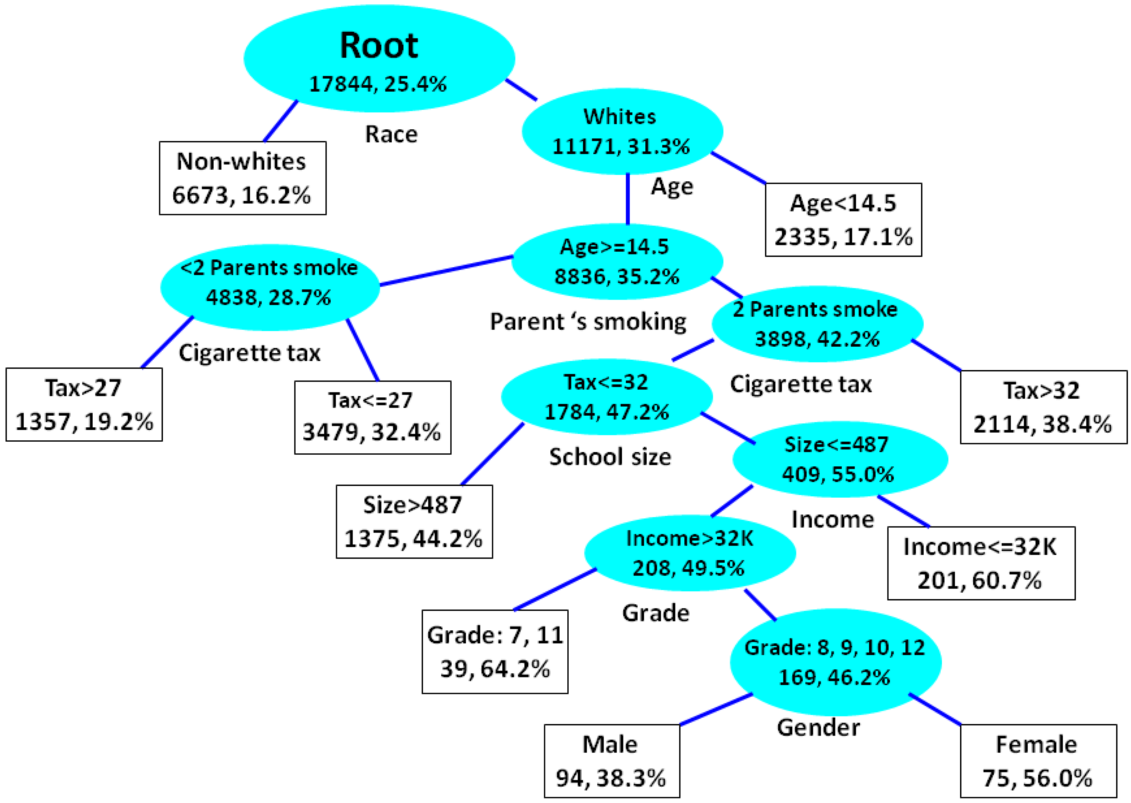

Figure 2: Individual Smoking vs. Peer Smoking

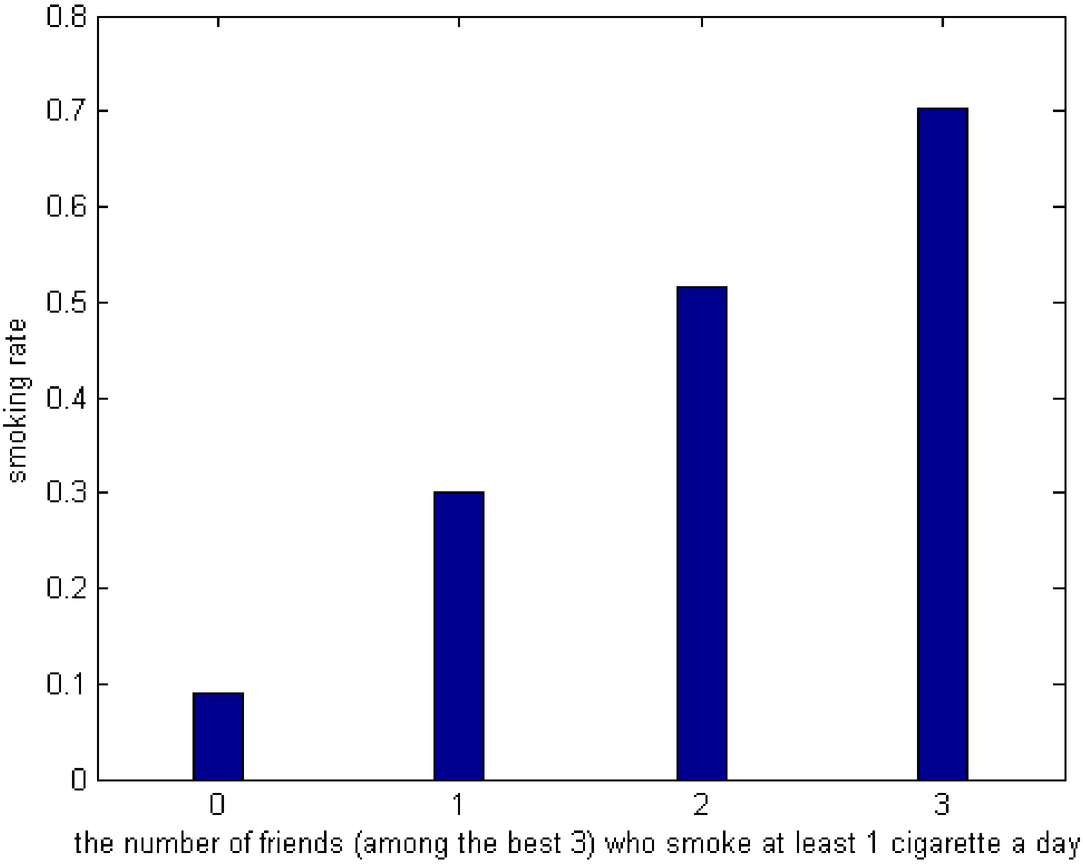

Figure 3: Number of Schoolmate Friends

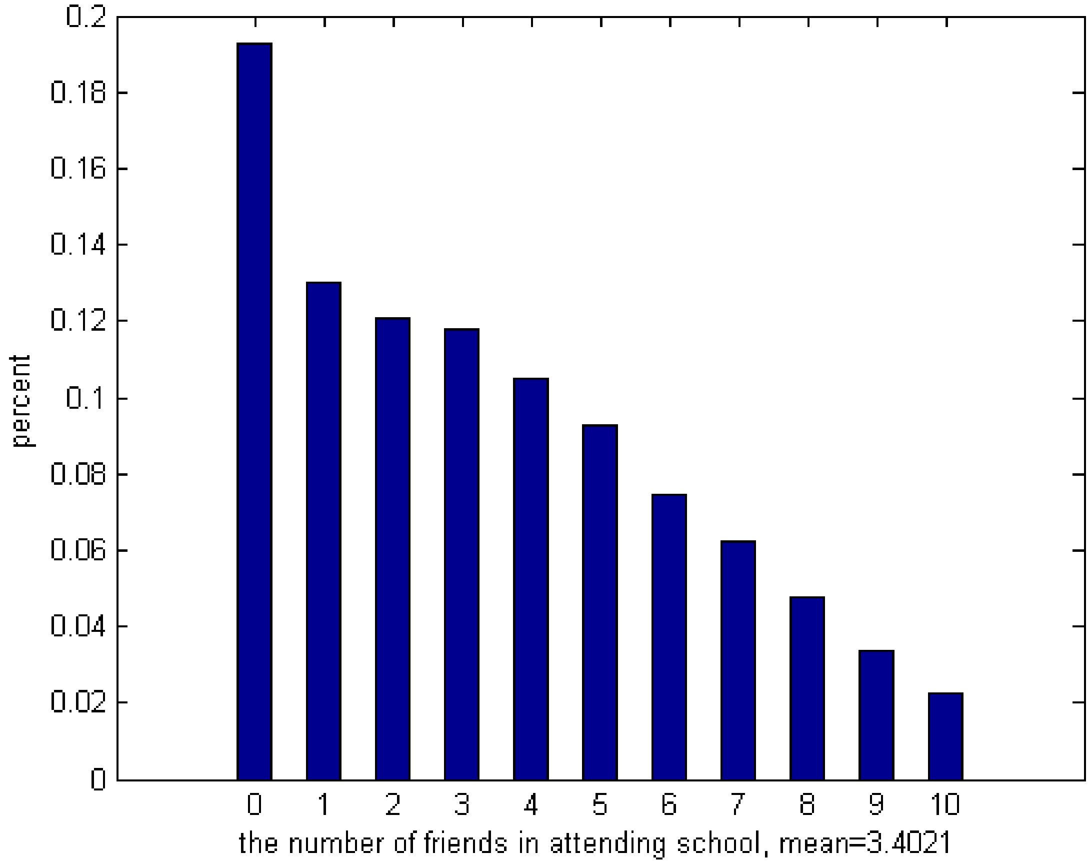

Figure 4: Distribution of Smoking Probabilities using the Bagged Tree Classifier

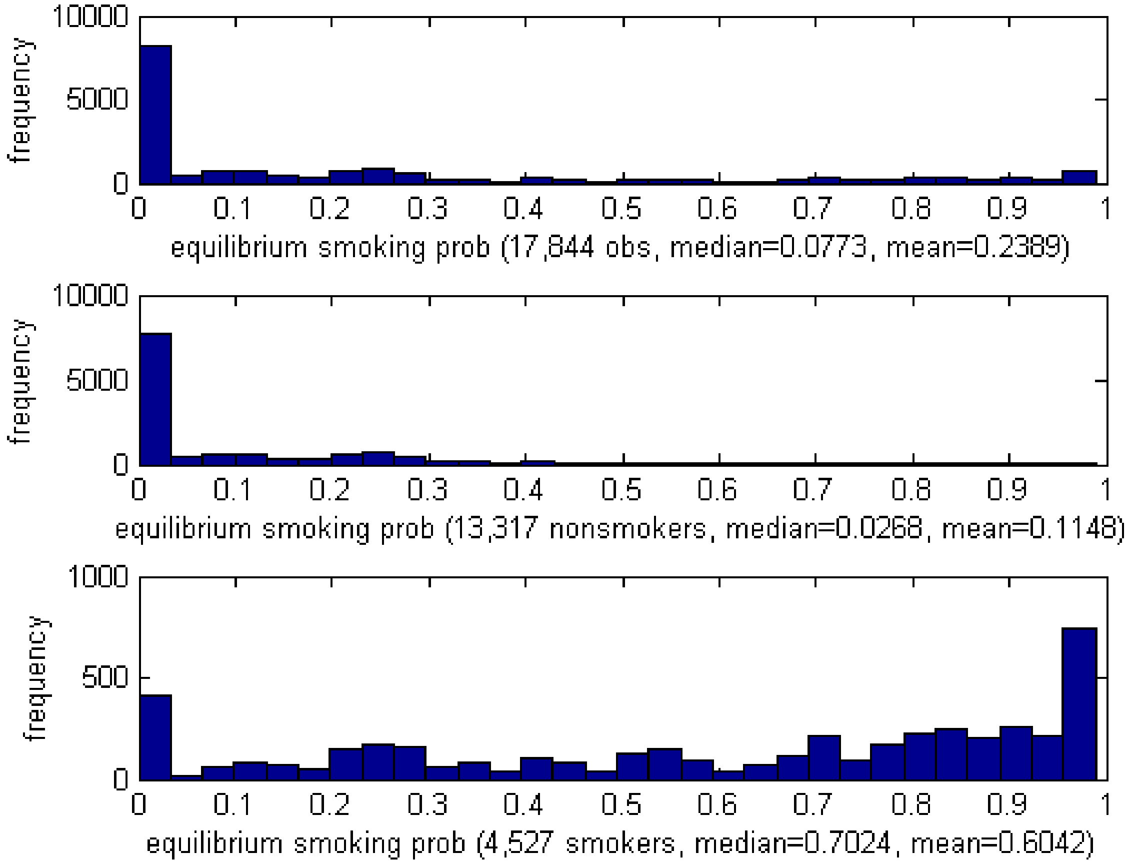

Figure 5: ROC Comparison of Bagged Tree and Logit Smoking Predictions

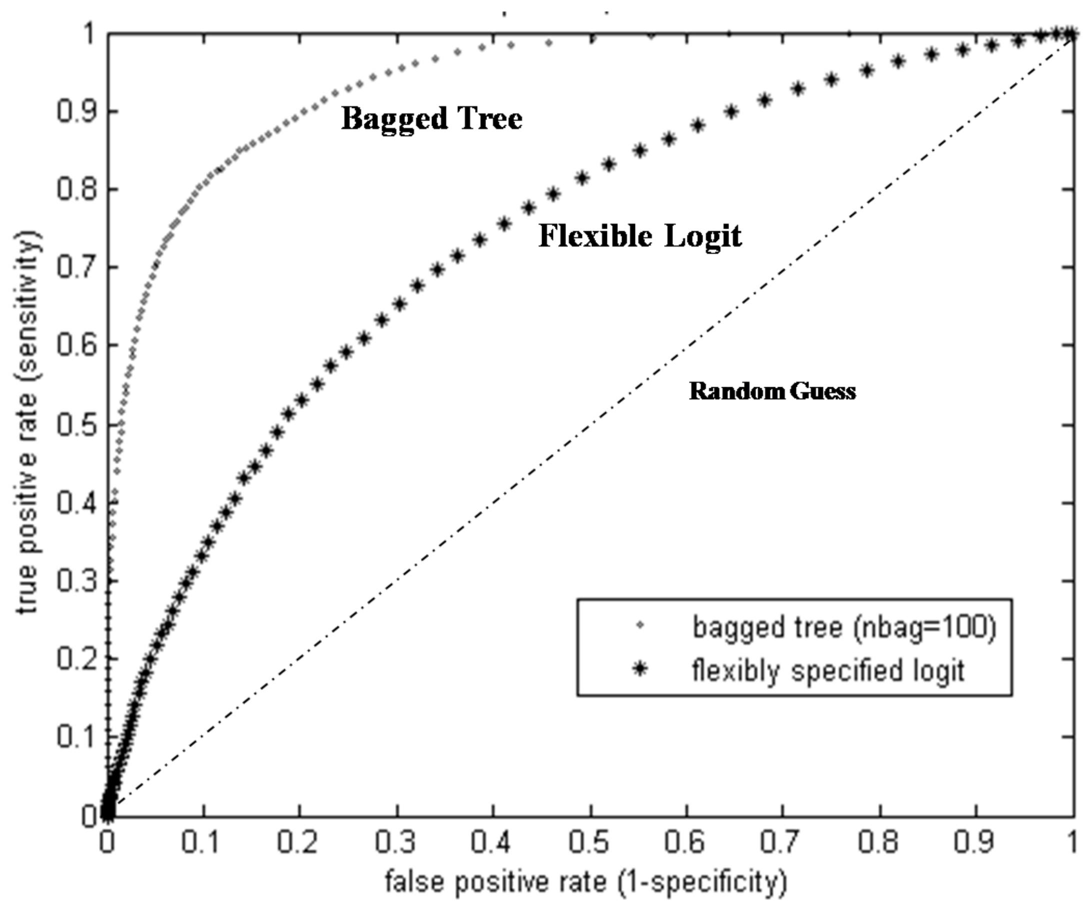

Figure 6: Reliability Comparison of Bagged Tree and Logit Smoking Predictions

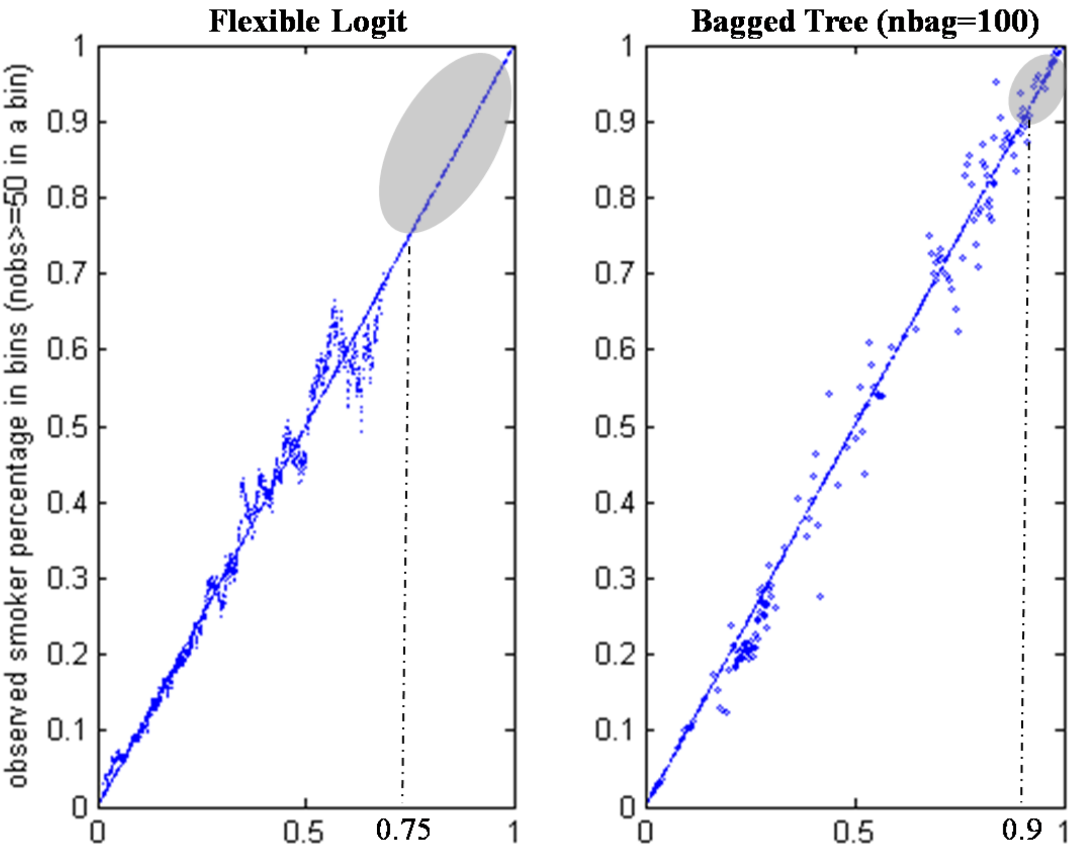

Figure 7: Reliability: Logit Directional Friendship

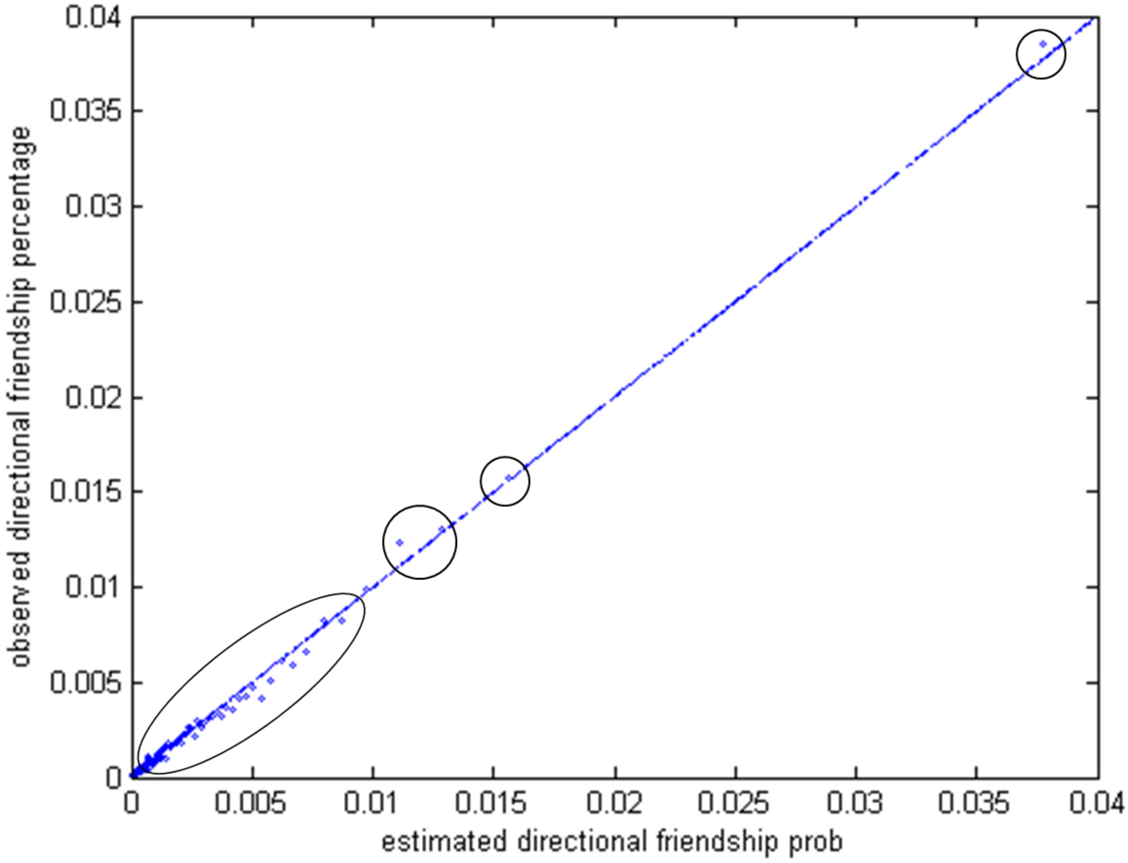

Figure 8: Trend Plot: IV individual Smoking Probability vs. IV Peer Smoking Norm

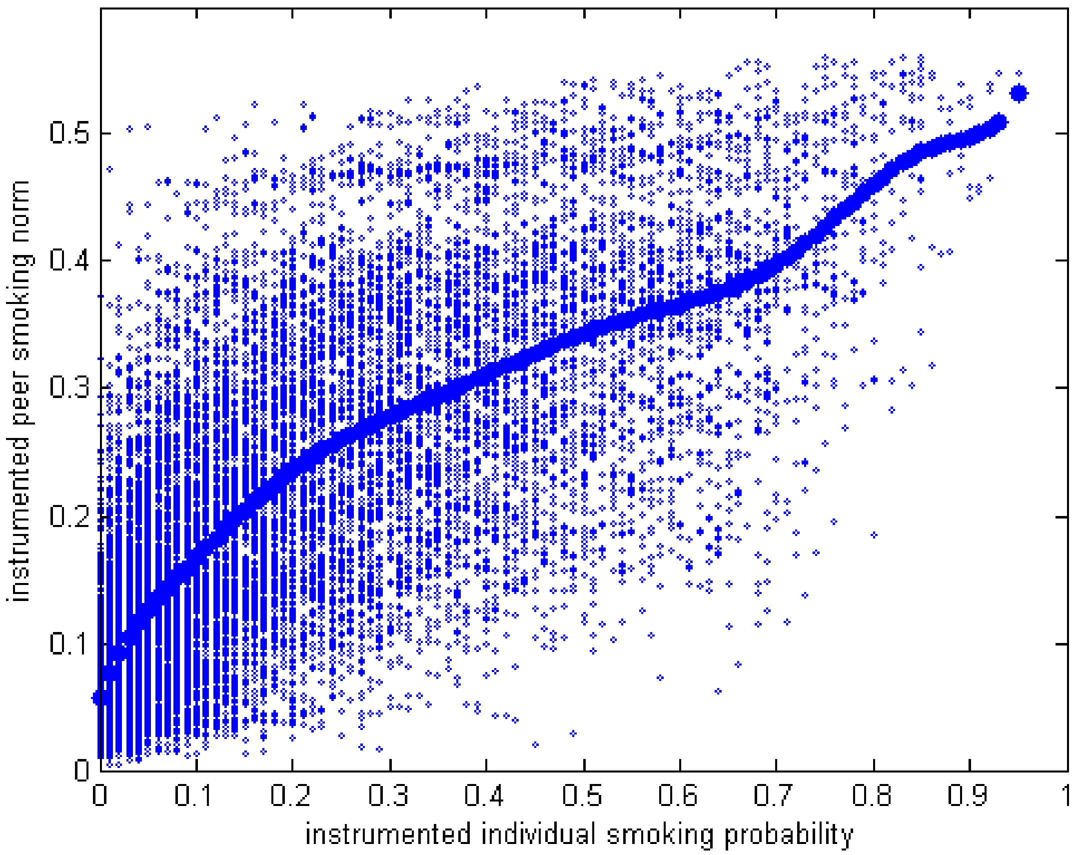

Figure 9: Effects of Cigarette Tax Increases on Equilibrium Smoking Initiation Rates

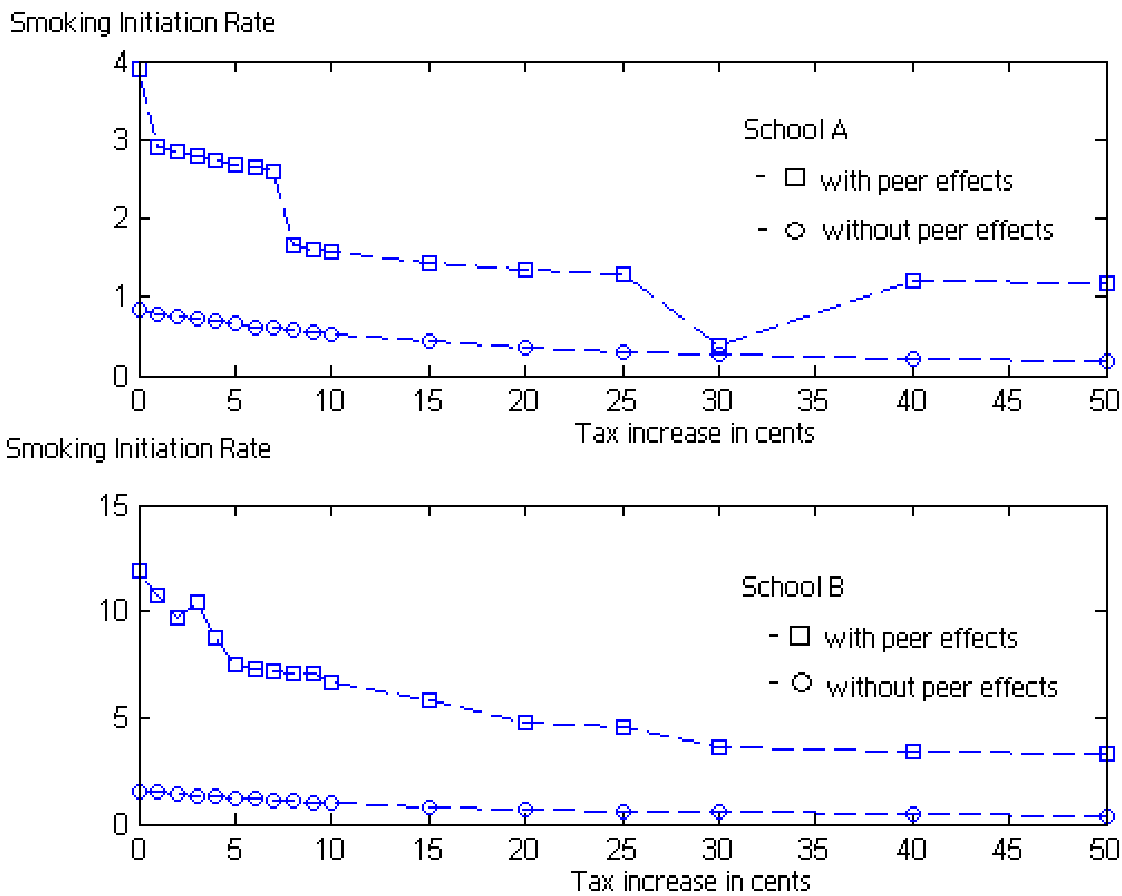

Table 1
Construction of Research Samples used in Analysis

Sample number	Selection criterion	Sample size
1	wave I respondent	20,745
2	with at least one parent	19,903
3	with family income information	18,245
4	with state cigarette tax information	17,844
4.1	those asked to nominate 5 schoolmate friends of each gender[a]	5,774
4.2	those who report at least one schoolmate friend	4,268
4.3	those who never smoked regularly before wave I[b]	13,924

a Using the 5,774 observations, we generate 2,987,761 directional pairwise observations to model friendship formation.

b Samples 4.1 and 4.3 are independent subsets of Sample 4; Sample 4.2 is a subset of Sample 4.1.

Table 2
Summary Statistics

Variables	Sample 4 Full research sample Mean	Std. Dev.	Sample 4.1 those asked to nominate 10 friends Mean	Std. Dev.	Sample 4.3 those with no smoking history Mean	Std. Dev.
Smoke	0.253	(0.435)	0.277	(0.448)	0.117	(0.320)
State cigarette tax (in 10 cents)	3.273	(1.631)	3.437	(1.459)	3.240	(1.595)
Age	16.114	(1.701)	16.379	(1.672)	16.028	(1.729)
Grade level						
7^{th}	0.134	(0.341)	0.105	(0.306)	0.149	(0.356)
8^{th}	0.135	(0.342)	0.107	(0.309)	0.139	(0.346)
9^{th}	0.181	(0.385)	0.154	(0.362)	0.179	(0.383)
10^{th}	0.196	(0.398)	0.225	(0.420)	0.194	(0.396)
11^{th}	0.187	(0.390)	0.213	(0.409)	0.180	(0.384)
12^{th}	0.165	(0.371)	0.195	(0.396)	0.158	(0.364)
Female	0.502	(0.500)	0.480	(0.500)	0.504	(0.500)
Race						
White	0.626	(0.484)	0.624	(0.484)	0.582	(0.493)
Black	0.229	(0.420)	0.183	(0.387)	0.266	(0.442)
Asian	0.071	(0.257)	0.103	(0.303)	0.076	(0.265)
Other	0.074	(0.262)	0.090	(0.286)	0.076	(0.265)
Family income (in 10K dollars)	3.477	(4.854)	3.398	(4.654)	3.455	(4.857)
Parents' highest education						
College degree	0.333	(0.471)	0.323	(0.468)	0.334	(0.472)
High school degree	0.497	(0.500)	0.501	(0.500)	0.499	(0.500)
Less than high school degree	0.170	(0.375)	0.176	(0.381)	0.167	(0.373)
Parent's smoking behavior						
No smokers	0.357	(0.479)	0.360	(0.480)	0.384	(0.487)
One smoker	0.244	(0.429)	0.245	(0.430)	0.250	(0.433)
Two smokers	0.399	(0.490)	0.395	(0.488)	0.367	(0.482)
Religious importance						
Very important	0.419	(0.493)	0.404	(0.491)	0.457	(0.498)
Important	0.352	(0.478)	0.360	(0.480)	0.340	(0.474)
Unimportant	0.061	(0.240)	0.062	(0.241)	0.056	(0.229)
Not religious	0.168	(0.374)	0.174	(0.379)	0.148	(0.335)
Sample Size	17,844		5,774		13,924	

Table 3
Estimation Results from Friendship Model

Variables[a]	Coefficient	Standard Error[b]	
Age deviation (omitted: $age_i = age_j$)			
(age_i-$age_j > 0$) × ($age_i <= 14$)	-0.484	(0.141)	***
(age_i-$age_j < 0$) × ($age_i <= 14$)	-0.457	(0.057)	***
(age_i-$age_j > 0$) × ($14 < age_i <= 17$)	-0.407	(0.038)	***
(age_i-$age_j < 0$) × ($14 < age_i <= 17$)	-0.350	(0.024)	***
(age_i-$age_j > 0$) × ($age_i > 17$)	-0.417	(0.031)	***
(age_i-$age_j < 0$) × ($age_i > 17$)	-0.283	(0.064)	***
Grade level deviation (omitted: 7_i^{th}, 7_j^{th})			
7_i^{th}, upper $grader_j$	-1.872	(0.149)	***
8_i^{th}, 8_j^{th}	0.025	(0.106)	
8_i^{th}, upper $grader_j$	-1.460	(0.178)	***
8_i^{th}, lower $grader_j$	-1.791	(0.169)	***
9_i^{th}, 9_j^{th}	-0.414	(0.151)	***
9_i^{th}, upper $grader_j$	-1.951	(0.162)	***
9_i^{th}, lower $grader_j$	-1.252	(0.267)	***
10_i^{th}, 10_j^{th}	-0.525	(0.149)	***
10_i^{th}, upper $grader_j$	-1.814	(0.155)	***
10_i^{th}, lower $grader_j$	-1.985	(0.174)	***
11_i^{th}, 11_j^{th}	-0.477	(0.149)	***
11_i^{th}, upper $grader_j$	-1.708	(0.160)	***
11_i^{th}, lower $grader_j$	-1.872	(0.159)	***
12_i^{th}, 12_j^{th}	-0.472	(0.150)	***
12_i^{th}, lower $grader_j$	-1.980	(0.163)	***
Gender deviation (omitted: $male_i$, $male_j$)			
$male_i$, $female_j$	-0.320	(0.027)	***
$female_i$, $male_j$	-0.381	(0.037)	***
$female_i$, $female_j$	0.107	(0.031)	***
Race deviation (omitted: $white_i$, $white_j$)			
$white_i$, $black_j$	-1.842	(0.130)	***
$white_i$, $asian_j$	-1.399	(0.111)	***
$white_i$, $other_j$	-0.357	(0.076)	***
$black_i$, $white_j$	-1.957	(0.165)	***
$black_i$, $black_j$	0.107	(0.064)	*
$black_i$, $asian_j$	-2.850	(0.251)	***
$black_i$, $other_j$	-2.000	(0.209)	***

Table 3 - continued
Estimation Results from Friendship Model

Variables	Coefficient	Standard Error	
Race deviation (omitted: $white_i$, $white_j$)			
$asian_i$, $white_j$	-1.420	(0.153)	***
$asian_i$, $black_j$	-3.118	(0.239)	***
$asian_i$, $asian_j$	0.589	(0.066)	***
$asian_i$, $other_j$	-1.507	(0.129)	***
$other_i$, $white_j$	-0.389	(0.081)	***
$white_i$, $black_j$	-1.916	(0.206)	***
$other_i$, $asian_j$	-1.238	(0.161)	***
$other_i$, $other_j$	0.023	(0.088)	
Family income deviation (omitted: $income_i = income_j$)			
($income_i$-$income_j > 0$) × ($income_i \leq$ 25K)	-0.0073	(0.0037)	***
($income_i$-$income_j < 0$) × ($income_i \leq$ 25K)	-0.0026	(0.0008)	
($income_i$-$income_j > 0$) × (25K$<income_i \leq$ 50K)	-0.0037	(0.0012)	
($income_i$-$income_j < 0$) × (25K$<income_i \leq$ 50K)	0.0003	(0.0006)	
($income_i$-$income_j > 0$) × (50K$<income_i \leq$ 75K)	-0.0010	(0.0009)	
($income_i$-$income_j < 0$) × (50K$<income_i \leq$ 75K)	-0.0003	(0.0008)	
($income_i$-$income_j > 0$) × ($income_i >$ 75K)	-0.0005	(0.0004)	
($income_i$-$income_j < 0$) × ($income_i >$ 75K)	0.0001	(0.0009)	
Smoking probability deviation (omitted: $\hat{p}(smoke_i)$-$\hat{p}(smoke_j)$=0)			
($\hat{p}(smoke_i)$-$\hat{p}(smoke_j) > 0$)	-0.641	(0.115)	***
($\hat{p}(smoke_i)$-$\hat{p}(smoke_j) < 0$)	-0.297	(0.096)	***
School size			
(size$\leq$ 100)	0.0654	(0.0084)	***
(size$\leq$ 100)^2	-0.0006	(0.0001)	***
(100 $<$size$\leq$ 200)	0.0304	(0.0037)	***
(100 $<$size$\leq$ 200)^2	-0.0001	(0.0000)	***
(size$>$ 200)	0.0028	(0.0006)	***
(size$>$ 200)^2	-0.0000	(0.0000)	
Constant	-4.306	(0.287)	
N=2,987,761 created from 5,774 students in Sample 4.1			

a Subscript i indicates variable for individual i, while subscript j indicates variable for potential friend j.
b Standard errors are clustered at the person level.
*** indicates significance at the 1% level; **, 5% level; *, 10% level.

Table 4
Marginal Effects on Smoking Initiation Rate - Specifications 1-3

Variables	Specification 1 No peer effect			Specification 2 School norm			Specification 3 IV School norm		
Average peer smoking effect	NA			0.18	(0.03)	***	0.16	(0.03)	***
(unit: 0.01)		-		11.78	(0.27		11.79	(0.27)	
Peer effect by grade level									
grade 7		-		0.15	(0.07)	**	0.13	(0.09)	
		-		9.35	(0.64)		9.36	(0.64)	
grade 8		-		-0.02	(0.09)		-0.08	(0.11)	
		-		12.70	(0.76)		12.72	(0.75)	
grade 9		-		0.18	(0.06)	***	0.21	(0.06)	***
		-		12.22	(0.65)		12.23	(0.66)	
grade 10		-		0.19	(0.06)	***	0.18	(0.06)	***
		-		11.87	(0.61)		11.86	(0.61)	
grade 11		-		0.20	(0.07)	***	0.17	(0.07)	**
		-		12.84	(0.66)		12.86	(0.66)	
grade 12		-		0.35	(0.07)	***	0.28	(0.07)	***
		-		11.49	(0.68)		11.49	(0.68)	
Average tax effect	-3.86	(1.81)	**	-0.35	(0.17)	**	-0.35	(0.17)	**
(unit: 10 cents)	12.10	(0.28)		11.78	(0.27)		11.79	(0.27)	
Tax effect by grade level									
grade 7	-4.01	(1.34)	***	-0.91	(0.35)	***	-0.86	(0.36)	**
	9.75	(0.65)		9.35	(0.64)		9.36	(0.64)	
grade 8	-4.64	(1.90)	**	-0.19	(0.47)		-0.16	(0.46)	
	13.16	(0.76)		12.70	(0.76)		12.72	(0.75)	
grade 9	-4.27	(1.83)	**	-1.14	(0.37)	***	-1.11	(0.36)	***
	12.49	(0.66)		12.22	(0.65)		12.23	(0.66)	
grade 10	-3.30	(1.95)	*	-0.03	(0.38)		-0.02	(0.38)	
	12.11	(0.62)		11.87	(0.61)		11.86	(0.61)	
grade 11	-3.66	(2.05)	*	-0.27	(0.40)		0.30	(0.42)	
	13.09	(0.65)		12.84	(0.66)		12.86	(0.66)	
grade 12	-2.96	(1.93)	*	0.41	(0.44)		0.38	(0.43)	
	11.72	(0.68)		11.49	(0.68)		11.49	(0.68)	
Age	0.79	(0.44)		0.69	(0.42)		0.65	(0.42)	
	12.10	(0.28)		11.78	(0.27)		11.79	(0.27)	
Grade level									
grade 7 (base)	12.40	(1.62)		11.68	(1.42)		11.56	(1.47)	
grade 8	3.30	(1.22)	***	2.07	(1.35)		1.76	(1.47)	
grade 9	0.13	(1.67)		0.34	(1.39)		0.28	(1.43)	
grade 10	-0.92	(1.91)		-0.43	(1.60)		-0.27	(1.66)	
grade 11	-0.37	(2.20)		0.19	(1.91)		0.39	(1.96)	
grade 12	2.40	(2.36)		-2.05	(2.08)		-1.78	(2.13)	
Gender									
Male (base)	13.77	(0.43)		12.02	(0.39)		12.02	(0.40)	
Female	-0.53	(0.55)		-0.46	(0.54)		-0.46	(0.55)	

Table 4 - continued
Marginal Effects on Smoking Initiation Rate - Specifications 1-3

Variables	Specification 1 No peer effect			Specification 2 School norm[a]			Specification 3 IV School norm[a,b]		
Race									
White (base)	13.76	(0.43)		12.96	(0.38)		12.95	(0.38)	
Black	-4.86	(0.81)	***	-3.80	(0.68)	***	-3.74	(0.69)	***
Asian	-4.97	(1.14)	***	-3.83	(1.05)	***	-3.79	(1.06)	***
Other	-0.97	(1.25)		-0.13	(1.15)		0.11	(1.15)	
Average family income effect	-0.03	(0.06)		-0.01	(0.06)		-0.02	(0.06)	
(unit: 10K dollars)	12.10	(0.28)		11.78	(0.27)		11.79	(0.27)	
Income effect by grade level									
grade 7	-0.59	(0.23)	***	-4.46	(1.24)		-0.63	(0.22)	***
	9.75	(0.65)		9.35	(0.64)		9.36	(0.64)	
grade 8	0.01	(0.15)		0.14	(1.47)		0.02	(0.15)	
	13.16	(0.76)		12.70	(0.76)		12.72	(0.75)	
grade 9	0.07	(0.16)		1.19	(1.64)		0.09	(0.15)	
	12.49	(0.66)		12.22	(0.65)		12.23	(0.66)	
grade 10	-0.02	(0.10)		-0.10	(0.95)		-0.01	(0.10)	
	12.11	(0.62)		11.87	(0.61)		11.86	(0.61)	
grade 11	0.13	(0.12)		1.98	(1.33)		0.19	(0.12)	
	13.09	(0.65)		12.84	(0.66)		12.86	(0.66)	
grade 12	0.18	(0.13)		1.92	(1.44)		0.18	(0.13)	
	11.72	(0.68)		11.49	(0.68)		11.49	(0.68)	
Parents' highest education									
College degree (base)	11.19	(0.46)		10.95	(0.46)		10.95	(0.45)	
High school degree	1.25	(0.60)	**	1.13	(0.60)	*	1.14	(0.59)	*
Less than high school degree	1.69	(0.82)	**	1.59	(0.83)	*	1.58	(0.81)	**
Parents' smoking behavior									
No smokers (base)	10.15	(0.42)		9.84	(0.41)		9.83	(0.42)	
One smoker	2.52	(0.71)	***	2.50	(0.68)	***	2.49	(0.70)	***
Two smokers	3.41	(0.65)	***	3.41	(0.64)	***	3.47	(0.65)	***
Religious importance									
Not religious (base)	13.62	(0.75)		13.36	(0.72)		13.42	(0.73)	
Very important	-3.84	(0.86)	***	-3.88	(0.83)	***	-3.96	(0.84)	***
Important	-0.10	(0.89)		-0.17	(0.86)		-0.22	(0.88)	
Not important	2.83	(1.46)	**	2.64	(1.44)	*	2.61	(1.46)	*
Sample Size	13,924			13,924			13,924		

Table 5
Marginal Effects on Smoking Initiation Rate - Specifications 4-6

Variables	Specification 4 Peer norm			Specification 5 IV peer norm[c]			Specification 6 Preferred[b]		
Average peer smoking effect	0.11	(0.01)	***	0.19	(0.04)	***	1.07	(0.09)	***
(unit: 0.01)	13.47	(0.51)		13.68	(1.41)		12.10	(0.28)	
Peer effect by grade level									
grade 7	0.11	(0.03)	***	0.30	(0.12)	**	1.12	(0.16)	***
	12.61	(1.45)		12.61	(1.44)		9.78	(0.64)	
grade 8	0.11	(0.04)	***	0.14	(0.10)		1.07	(0.16)	***
	16.78	(1.57)		16.87	(1.60)		13.19	(0.74)	
grade 9	0.16	(0.03)	***	0.18	(0.08)	**	1.08	(0.12)	***
	14.64	(1.34)		14.68	(1.34)		12.51	(0.65)	
grade 10	0.12	(0.03)	***	0.25	(0.06)	***	1.11	(0.12)	***
	12.26	(1.02)		12.24	(1.02)		12.11	(0.62)	
grade 11	0.08	(0.03)	***	0.12	(0.06)	**	1.02	(0.11)	***
	12.85	(1.06)		12.83	(1.07)		13.07	(0.67)	
grade 12	0.11	(0.03)	***	0.19	(0.07)	***	1.06	(0.11)	***
	13.10	(1.23)		13.15	(1.24)		11.69	(0.66)	
Average tax effect	-4.70	(3.82)		-4.27	(3.93)		-3.68	(2.29)	*
(unit: 10 cents)	13.47	(0.51)		13.68	(1.41)		12.22	(0.32)	
Tax effect by grade level									
grade 7	-6.16	(3.07)	**	-5.86	(3.13)	*	-3.24	(1.93)	*
	12.61	(1.45)		12.61	(1.44)		9.83	(0.69)	
grade 8	-7.68	(4.11)	*	-7.40	(4.15)	*	-4.08	(2.53)	*
	16.78	(1.57)		16.87	(1.60)		13.28	(0.81)	
grade 9	-5.32	(4.07)		-4.85	(4.28)		-3.74	(2.39)	
	14.64	(1.34)		14.68	(1.34)		12.57	(0.68)	
grade 10	-3.36	(3.94)		-2.74	(4.02)		-3.35	(2.43)	
	12.26	(1.02)		12.24	(1.02)		12.16	(0.64)	
grade 11	-3.98	(4.02)		-3.52	(4.09)		-3.35	(2.43)	
	12.85	(1.06)		12.83	(1.07)		13.14	(0.69)	
grade 12	-3.87	(4.14)		-3.30	(4.21)		-3.86	(2.17)	*
	13.10	(1.23)		13.15	(1.24)		11.77	(0.67)	
Age	1.87	(0.93)	**	1.68	(0.93)	*	0.11	(0.43)	
	13.47	(0.51)		13.68	(1.41)		12.22	(0.32)	
Grade level									
grade 7 (base)	18.91	(3.96)		23.83	(4.78)		26.99	(3.65)	
grade 8	2.28	(2.75)		-1.89	(3.85)		-8.27	(3.22)	***
grade 9	-4.10	(3.85)		-9.23	(4.56)	**	-14.71	(3.58)	***
grade 10	-6.41	(4.21)		-11.81	(4.93)	**	-15.85	(3.71)	***
grade 11	-7.31	(4.69)		-12.34	(5.40)	**	-17.09	(3.94)	***
grade 12	-8.98	(5.14)	*	-13.99	(5.81)	**	-18.17	(4.00)	***
Gender									
Male (base)	13.33	(0.74)		13.26	(0.74)		12.31	(0.41)	
Female	0.26	(1.03)		0.44	(1.02)		-0.29	(0.55)	

Table 5 - continued
Marginal Effects on Smoking Initiation Rate - Specifications 4-6

Variables	Specification 4 Peer norm			Specification 5 IV peer norm[c]			Specification 6 Preferred[b]		
Race									
White (base)	15.67	(0.82)		15.55	(0.82)		11.25	(0.37)	
Black	-6.78	(1.62)	***	-6.11	(1.75)	***	7.03	(1.55)	***
Asian	-5.21	(2.07)	**	-5.20	(2.10)	***	-1.42	(1.17)	
Other	-4.38	(2.26)	**	-4.62	(2.19)	**	0.71	(1.17)	
Average family income effect	-0.09	(0.12)		-0.09	(0.12)		-0.01	(0.07)	
(unit: 10K dollars)	13.47	(0.51)		13.68	(1.41)		12.22	(0.32)	
Income effect by grade level									
grade 7	-3.33	(3.69)		3.10	(3.72)		-0.46	(0.23)	
	12.61	(1.45)		12.61	(1.44)		9.83	(0.69)	
grade 8	1.64	(2.29)		1.80	(2.34)		0.07	(0.16)	
	16.78	(1.57)		16.87	(1.60)		13.28	(0.81)	
grade 9	-1.54	(3.49)		1.44	(3.65)		0.13	(0.20)	
	14.64	(1.34)		14.68	(1.34)		12.57	(0.68)	
grade 10	-2.01	(1.75)		-2.58	(1.80)		-0.01	(0.13)	
	12.26	(1.02)		12.24	(1.02)		12.16	(0.64)	
grade 11	1.85	(1.92)		1.57	(1.91)		0.06	(0.14)	
	12.85	(1.06)		12.83	(1.07)		13.14	(0.69)	
grade 12	-0.38	(2.41)		-0.17	(2.43)		0.09	(0.15)	
	13.10	(1.23)		13.15	(1.24)		11.77	(0.67)	
Parents' highest education									
College degree (base)	11.72	(0.85)		11.65	(0.84)		11.22	(0.46)	
High school degree	2.65	(1.11)	**	2.77	(1.14)	**	1.26	(0.62)	**
Less than high school degree	2.51	(1.55)	*	2.69	(1.53)	*	1.86	(0.85)	**
Parents' smoking behavior									
No smokers (base)	11.16	(0.82)		11.00	(0.82)		10.40	(0.44)	
One smoker	3.41	(1.29)	***	3.73	(1.31)	***	2.47	(0.71)	***
Two smokers	3.71	(1.25)	***	4.02	(1.25)	***	2.93	(0.67)	***
Religious importance									
Not religious (base)	13.47	(1.39)		13.75	(1.44)		13.67	(0.78)	
Very important	-1.87	(1.65)		-2.08	(1.67)		-3.75	(0.89)	***
Important	1.51	(1.66)		1.22	(1.64)		-0.14	(0.94)	
Not important	5.12	(2.66)	**	4.30	(2.68)	*	2.54	(1.47)	*
Sample Size	4,268			4,268			13,924		

Table A1
Estimation Results from Smoking Initiation Specifications

Variables	Specification 1 No peer effect			Specification 2 School norm			Specification 3 IV School norm		
Smoking norm (unit: 0.01) and grade interactions									
norm × grade 7	-			1.750	(0.889)	*	1.541	(1.004)	
norm × grade 8	-			-0.206	(0.822)		-0.761	(0.948)	
norm × grade 9	-			1.726	(0.591)	***	2.002	(0.609)	***
norm × grade 10	-			1.849	(0.618)	***	1.740	(0.572)	***
norm × grade 11	-			1.780	(0.624)	***	1.519	(0.599)	***
norm × grade 12	-			3.450	(0.691)	***	2.861	(0.577)	***
State cigarette tax (unit: 10 cents) and grade interactions									
tax × grade 7	-0.620	(0.248)	***	-0.115	(0.046)	**			
tax × grade 8	-0.532	(0.246)	**	-0.018	(0.043)				
tax × grade 9	-0.509	(0.244)	**	-0.113	(0.038)	***			
tax × grade 10	-0.392	(0.243)	*	-0.004	(0.037)				
tax × grade 11	-0.407	(0.243)	*	-0.027	(0.039)				
tax × grade 12	-0.361	(0.245)		0.041	(0.043)				
Age	0.074	(0.041)	*	0.066	(0.040)	*	0.063	(0.038)	*
Grade level (comparison: 7^{th})									
grade 8	-0.214	(0.235)		0.156	(0.316)		-0.057	(0.276)	
grade 9	-0.581	(0.319)	*	-0.234	(0.330)		-0.330	(0.298)	
grade 10	-1.030	(0.341)	***	-0.650	(0.348)	*	-6.393	(0.301)	**
grade 11	-0.979	(0.358)	***	-0.564	(0.361)		-0.521	(0.317)	*
grade 12	-1.377	(0.382)	***	-1.517	(.408)	***	-1.252	(0.363)	***
Female	0.052	(0.055)		0.045	(0.054)		0.046	(0.054)	
Race (comparison: white)									
Black	-0.510	(0.091)	***	-0.399	(0.075)	***	-0.391	(0.076)	***
Asian	-0.529	(0.138)	***	-0.407	(0.125)	***	-0.405	(0.127)	***
Other	-0.092	(0.115)		-0.014	(0.106)	***	0.014	(0.107)	
Family income (unit: 10K dollars) and grade interactions									
income × grade 7	-0.072	(0.028)	***	-0.077	(0.028)	***	-0.078	(0.028)	***
income × grade 8	0.001	(0.014)		0.001	(0.013)		0.001	(0.013)	
income × grade 9	0.067	(0.015)		0.010	(0.015)		0.009	(0.015)	
income × grade 10	-0.002	(0.010)		-0.002	(0.010)		-0.001	(0.010)	
income × grade 11	0.001	(0.011)		0.017	(0.011)		0.017	(0.011)	
income × grade 12	0.018	(0.013)		0.018	(0.013)		0.018	(0.013)	
Parents' highest education (comparison: college degree)									
High school degree	0.127	(0.061)	*	0.115	(0.061)	*	0.115	(0.061)	*
Less than high school degree	0.169	(0.080)	*	0.158	(0.079)	**	0.155	(0.079)	**
Parents' smoking behavior (comparison: no smokers)									
One smoker	0.260	(0.071)	***	0.259	(0.071)	***	0.258	(0.072)	***
Two smokers	0.342	(0.065)	***	0.343	(0.064)	***	0.347	(0.068)	***
Religious importance (comparison: not religious)									
Very important	-0.389	(0.084)	***	-0.393	(0.080)	***	-0.402	(0.081)	***
Important	-0.0085	(0.080)		-0.015	(0.077)		-0.020	(0.079)	
Unimportant	0.230	(0.117)	*	0.215	(0.115)	*	0.210	(0.118)	*
School fixed effects	yes			no			no		
Sample size	13,924			13,924			13,924		

Table A1 - continued
Estimation Results from Smoking Initiation Specifications

Variables	Specification 4 Peer norm			Specification 5 IV peer norm			Specification 6 Preferred		
Smoking norm (unit: 0.01) and grade interactions									
norm × grade 7	1.163	(0.373)	***	3.150	(1.254)	***	12.932	(1.699)	***
norm × grade 8	0.915	(0.315)	***	1.144	(0.821)		9.684	(1.424)	***
norm × grade 9	1.483	(0.315)	***	1577	(0.695)	**	10.145	(1.096)	***
norm × grade 10	1.289	(0.268)	***	2.589	(0.634)	***	10.485	(1.111)	***
norm × grade 11	0.797	(0.263)	***	1.194	(0.556)	**	9.179	(0.924)	***
norm × grade 12	1.091	(0.291)	***	1.817	(0.654)	***	10.676	(1.001)	***
State cigarette tax (unit: 10 cents) and grade interactions									
tax × grade 7	-0.939	(0.546)	*	-0.874	(0.528)		-0.505	(0.328)	*
tax × grade 8	-0.891	(0.544)	*	-0.841	(0.524)		-0.478	(0.321)	
tax × grade 9	-0.674	(0.537)		-0.598	(0.517)		-0.456	(0.313)	
tax × grade 10	-0.489	(0.535)		-0.416	(0.516)		-0.414	(0.314)	
tax × grade 11	-0.552	(0.535)		-0.493	(0.516)		-0.448	(0.314)	
tax × grade 12	-0.526	(0.534)		-0.454	(0.515)		-0.509	(0.314)	*
Age	0.171	(0.082)	*	0.139	(0.079)	*	0.010	(0.043)	
Grade level (comparison: 7^{th})									
grade 8	-0.088	(0.480)		0.028	(0.491)		-0.190	(0.283)	
grade 9	-1.350	(0.654)	**	-1.285	(0.667)		-0.922	(0.393)	**
grade 10	-2.073	(0.679)	***	-2.328	(0.685)	***	-1.201	(0.447)	***
grade 11	-1.953	(0.713)	***	-1.953	(0.714)	*	-0.971	(0.458)	**
grade 12	-2.263	(0.776)	***	-2.372	(0.777)	**	-1.261	(0.471)	***
Female	-0.026	(0.101)		-0.044	(0.100)		0.028	(0.056)	
Race (comparison: white)									
Black	-0.725	(0.193)	***	-0.631	(0.195)	**	0.621	(0.124)	***
Asian	-0.522	(0.230)	**	-0.526	(0.228)	**	-0.168	(0.138)	
Other	-0.432	(0.236)		-0.458	(0.223)		0.074	(0.119)	
Family income (unit: 10K dollars) and grade interactions									
income × grade 7	-0.046	(0.048)		-0.874	(0.528)	*	-0.058	(0.028)	**
income × grade 8	0.013	(0.018)		-0.841	(0.525)	*	0.007	(0.015)	
income × grade 9	-0.018	(0.035)		-0.598	(0.517)		0.013	(0.019)	
income × grade 10	-0.018	(0.035)		-0.416	(0.516)		-0.001	(0.013)	
income × grade 11	-0.025	(0.021)		-0.493	(0.516)		0.006	(0.013)	
income × grade 12	0.017	(0.010)		-0.454	(0.515)		0.009	(0.016)	
Parents' highest education (comparison: college degree)									
High school degree	0.268	(0.113)	**	0.278	(0.112)	**	0.129	(0.063)	**
Less than high school degree	0.254	(0.148)	*	0.262	(0.147)	*	0.185	(0.083)	**
Parents' smoking behavior (comparison: no smokers)									
One smoker	0.347	(0.130)	**	0.370	(0.128)	***	0.255	(0.074)	***
Two smokers	0.373	(0.125)	***	0.397	(0.123)	***	0.298	(0.068)	***
Religious importance (comparison: not religious)									
Very important	-0.191	(0.165)		-0.208	(0.163)		-0.382	(0.086)	***
Important	0.144	(0.155)		0.112	(0.154)		-0.012	(0.083)	
Unimportant	0.435	(0.219)	**	0.361	(0.217)	*	0.210	(0.123)	*
School fixed effects	yes			yes			yes		
Sample size	4,268			4,268			13,924		

Appendix: Characterization of Allowed Decision Space

Let $B_i(d_{-i,t})$ denote the set of player i's decisions allowed by her individual budget constraint and other $n_t - 1$ players' decisions $(d_{-i,t})$.

$$B_i(d_{-i,t}) = \{d_{i,t} | C(d_{i,t}; d_{-i,t}, z_t) - Y_{i,t} \leq 0\}$$

If a collection of n_t individual decision sets $\{\widetilde{D}_1, ..., \widetilde{D}_i, ..., \widetilde{D}_{n_t}\}$ $(\widetilde{D}_i \subseteq D_i \ \forall i \in N_t)$ in the game satisfy

$$\widetilde{D}_i = \bigcup_{\tilde{d}_{-i,t} \in \widetilde{D}_{-i}} B_i(\tilde{d}_{-i,t}) \ \forall i \in N_t \tag{26}$$

then the Cartesian product $\prod_i \widetilde{D}_i$ is an ADS of the game.

where $\bigcup_{\tilde{d}_{-i,t} \in \widetilde{D}_{-i}} B_i(\tilde{d}_{-i,t})$ is the union of those player i's decision sets corresponding to all different combinations of $n_t - 1$ players' decisions. $\widetilde{D}_{-i}$ is a Cartesian product defined as $\widetilde{D}_{-i} = \prod_{j \in N_t \backslash \{i\}} \widetilde{D}_j$

Let $K(\prod_i \widetilde{D}_i)$ denote the number of elements in $\prod_i \widetilde{D}_i$.

$\forall k = 1, 2, ..., K(\prod_i \widetilde{D}_i)$ and $\forall i \in N_t$, let $[\tilde{d}^k_{i,t} \ \tilde{d}^k_{-i,t}] = \tilde{d}^k_t$ denote the bi-decomposition of an element in $\prod_i \widetilde{D}_i$, which decomposes $\tilde{d}^k_t$ into the decision made by player i $(\tilde{d}^k_{i,t})$ and the decisions made by the rest $n_t - 1$ players $(\tilde{d}^k_{-i,t})$.

We note, $\tilde{d}^k_t \in \prod_i \widetilde{D}_i \Rightarrow \tilde{d}^k_{-i,t} \in \widetilde{D}_{-i}$. In turn, $\tilde{d}^k_{-i,t} \in \widetilde{D}_{-i}$ and $\widetilde{D}_i = \bigcup_{\tilde{d}_{-i,t} \in \widetilde{D}_{-i}} B_i(\tilde{d}_{-i,t}) \Rightarrow B_i(\tilde{d}^k_{-i,t}) \subseteq \widetilde{D}_i$. This says $\forall i \in N_t$ and $\forall k = 1, 2, ..., K(\prod_i \widetilde{D}_i)$, the set of player i's decisions allowed by $d^k_{-i,t}$ and her individual budget constraint is a subset of the $\widetilde{D}_i$. Recall $\prod_{i \in N_t} \widetilde{D}_i$ is a Cartesian product of $\widetilde{D}_i$s, therefore, $\forall i \in N_t$ and $\forall k = 1, 2, ..., K(\prod_i \widetilde{D}_i)$, $\tilde{d}_{i,t} \in \prod_i \widetilde{D}_i$ satisfies n_t budget constraints simultaneously. **Q.E.D.**